



**UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE FILOSOFIA E CIÊNCIAS HUMANAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GEOGRAFIA**

RAFAEL SILVA DOS ANJOS

**MODELAGEM ESPACIAL DE DOENÇAS E SUAS RELAÇÕES COM FATORES
SOCIOECONÔMICOS, MORFOCLIMÁTICOS E INFRAESTRUTURAIS**

Recife
2021

RAFAEL SILVA DOS ANJOS

**MODELAGEM ESPACIAL DE DOENÇAS E SUAS RELAÇÕES COM FATORES
SOCIOECONÔMICOS, MORFOCLIMÁTICOS E INFRAESTRUTURAS**

Tese apresentada ao Programa de Pós-Graduação em Geografia da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de doutor em geografia. Área de concentração: regionalização e análise regional.

Orientador: Prof. Dr. Ranyére Silva Nóbrega

Coorientador: Prof. Dr. Nuno de Sousa Neves

Recife
2021

Catálogo na fonte
Bibliotecária Maria do Carmo de Paiva, CRB4-1291

A599m Anjos, Rafael Silva dos.
Modelagem espacial de doenças e suas relações com fatores socioeconômicos, morfoclimáticos e infraestruturais / Rafael Silva dos Anjos. – 2021.
136 f. : il. ; 30 cm.

Orientador: Prof. Dr. Ranyére Silva Nóbrega.
Coorientador: Prof. Dr. Nuno de Sousa Neves.
Tese (Doutorado) - Universidade Federal de Pernambuco, CFCH. Programa de Pós-Graduação em Geografia, Recife, 2021.
Inclui referências.

1. Geografia. 2. Análise espacial. 3. Saúde pública. 4. Doenças. I. Nóbrega, Ranyére Silva (Orientador). II. Neves, Nuno de Sousa (Coorientador). III. Título.

910 CDD (22. ed.)

UFPE (BCFCH2021-126)

RAFAEL SILVA DOS ANJOS

**MODELAGEM ESPACIAL DE DOENÇAS E SUAS RELAÇÕES COM FATORES
SOCIOECONÔMICOS, MORFOCLIMÁTICOS E INFRAESTRUTURAIIS**

Tese apresentada ao Programa de Pós-Graduação em Geografia da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de doutor em geografia. Área de concentração: regionalização e análise regional.

Aprovada em: 25/03/2021.

BANCA EXAMINADORA

Prof. Dr. Ranyére Silva Nóbrega (Orientador – Examinador Interno)
Universidade Federal de Pernambuco

Profa. Dra. Amanda Priscila de Santana Cabral Silva (Examinadora Externa)
Universidade Federal de Pernambuco

Profa. Dra. Ana Lúcia Bezerra Candeias (Examinadora Externa)
Universidade Federal de Pernambuco

Prof. Dr. Anselmo Cesar Vasconcelos Bezerra (Examinador Externo)
Instituto Federal de Pernambuco

Profa. Dra. Cristiana Coutinho Duarte (Examinadora Externa)
Universidade Federal de Pernambuco

AGRADECIMENTOS

Agradeço primeiramente a Deus, por ter sempre iluminado os meus caminhos durante toda essa minha jornada.

Agradeço à minha mãe, Alice, e meu pai, Riva (*in memoriam*), que sempre estarão presentes em minha vida através dos seus ensinamentos, valores estes que carregarei para sempre.

Agradeço à minha irmã Aline, minha eterna amiga, por ter sempre me acompanhado nos bons e maus momentos, fazendo-me sentir o verdadeiro amor fraternal.

Agradeço ao Professor Ranyére Nóbrega por ter sempre me acompanhado desde o início da minha carreira universitária até o fim desse ciclo, sendo um dos principais alicerces de toda a minha trajetória acadêmica.

Agradeço ao Professor Nuno Neves, a quem tenho bastante estima, não sendo somente meu orientador de pesquisas, mas um grande amigo que me passou valores que serão eternamente lembrados.

Agradeço a todos os colegas do TropoClima por todas as boas discussões e momentos felizes que pude compartilhar ao longo desses mais de 10 anos.

Agradeço aos funcionários da secretaria do PPGeo, Eduardo e Pablo, por serem sempre solícitos.

Agradeço a Fundação de Amparo a Ciência e Tecnologia de Pernambuco (FACEPE) pelo financiamento à pesquisa.

Agradeço a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro pelo doutorado sanduíche na Universidade de Évora.

*Mas não é só junto ao rio
que o Recife está plantado,
hoje a cidade se estende
por sítios nunca pensados,
dos subúrbios coloridos
aos horizontes molhados.
Horizontes onde habitam
homens de pouco falar,
noturnos como convém
à fúria grave do mar.
Que comem fel de crustáceos
e que vivem do precário
desequilíbrio dos peixes.
Nesse lugar, as mulheres
cultivam brancos silêncios
e nas ausências mais longas,
pousam os olhos no chão,
saem do fundo da noite,
tiram a angústia do bolso
e a contemplam na mão.
Só os velhos adormecem,
lembrando o tempo que foi,
vazios como o vazio
e fácil sono de um boi.
(PENA FILHO, 1999. p.129)*

RESUMO

O presente estudo teve como objetivo principal propor uma abordagem teórico metodológica sobre o processo de modelagem espacial de doenças, baseado na relação entre a distribuição espacial da Zika, Chikungunya, Tuberculose e Leptospirose com condicionantes socioeconômicos, morfoclimáticos e infraestruturais, em Recife. Foram criados modelos espaciais para estimar a distribuição espacial dessas doenças, diferenciando-se de acordo com a combinação do tipo de unidade espacial (setor censitário ou quadrículas de 100 metros), cálculo da variável dependente (número de casos por setor ou média da densidade de Kernel) e método de estimativa da doença (*Ordinary Least Square-OLS* ou *Geographically Weighted Regression-GWR*). Para analisar a relação de cada condicionante com a distribuição espacial de cada doença foram utilizados os coeficientes, globais e locais, dos modelos que apresentaram as melhores eficiências em suas estimativas. Foi identificado que as melhores combinações para uma melhor eficiência nas estimativas dos modelos envolviam a adoção de quadrículas de 100 metros associada a médias de densidades de Kernel, como cálculo da variável dependente. O método de estimativa de GWR apresentou maiores correlações com os dados observados se comparados com o método OLS, embora cada método apresente suas potencialidades de acordo com a proposta da pesquisa. Nesse estudo foi verificado que todas as condicionantes selecionadas para o estudo apresentaram significância estatística com a distribuição espacial de todas as doenças analisadas. Outro ponto a ser destacado é que a adoção da distribuição espacial de algumas doenças como dados de entrada para o modelo, aumentava a eficiência nas estimativas da distribuição espacial de uma determinada doença. Tal abordagem na presente tese pode auxiliar nas estimativas de distribuição espacial de várias doenças, principalmente em casos onde não existe uma confiabilidade total nos dados, ou até mesmo a ausência deles. Outro aspecto a ser levantado é que em casos do surgimento de novas epidemias, essa proposta auxilia em prognósticos nas áreas a serem mais impactadas, baseando-se na distribuição espacial de doenças já conhecidas.

Palavras-chave: análise espacial; saúde pública; doenças; GWR; OLS.

ABSTRACT

The main objective of the present study was to propose a theoretical-methodological approach on the spatial modelling process of diseases, based on the relationship between the spatial distribution of Zika, Chikungunya, Tuberculosis, and Leptospirosis, and socioeconomic, morphoclimatic and infrastructural conditions in Recife. Some spatial models were created to estimate the spatial distribution of those diseases, and they were distinguished according to the combination of the spatial unit type (census sector or 100-meter grids), dependent variable calculation (number of cases per sector or mean kernel density), and disease estimation method (*Ordinary Least Square-OLS* or *Geographically Weighted Regression-GWR*). To analyze the relationship of each conditioning factor with the spatial distribution of each disease, the local and global coefficients of the models that presented the best efficiency in their estimation were used. It was identified that the best combinations for a better efficiency in the models estimation involved the adoption of 100-meter grids associated with kernel density averages, as a dependent variable calculus. The GWR estimation method presented higher correlation with the observed data when compared with the OLS method, although each of them shows their potentialities according to what is proposed in this research. During this study, it was verified that selected conditioning factors presented statistical significance with the spatial distribution of all analyzed diseases. It is important to emphasize that the adoption of spatial distribution of some diseases as input for the model made the estimate efficiency of a certain disease grow. Such an approach in the present thesis can support the spatial distribution estimate of many diseases, mainly in cases where the data is not completely reliable, or even when they are missing. Another important aspect is that in the case of potential epidemics arising, this proposal assists prognosis in the areas that could be more affected, based on the spatial distribution of known diseases.

Keywords: spatial analysis; public health; diseases; GWR; OLS.

LISTA DE FIGURAS

Figura 1- Localização do município de Recife e seus bairros, de acordo com o Estado e país.....	31
Figura 2- Distribuição da renda média mensal do chefe familiar, em Recife.	33
Figura 3 - Média de moradores por domicílio no município de Recife.	34
Figura 4 - Densidade habitacional no município de Recife (hab/km ²).....	36
Figura 5 - Mapa hipsométrico do município de Recife.	38
Figura 6 - Mapa de declividade no município de Recife.....	40
Figura 7 - Porcentagem de áreas não construídas por setor censitário no município de Recife.....	41
Figura 8 - Média da precipitação anual entre os anos 2015 e 2017 no município de Recife.	43
Figura 9 - Proporção de áreas verdes em relação às áreas totais dos bairros no município de Recife.....	44
Figura 10 - Bairro de Boa Viagem (a), zona norte das colinas em Recife (b) e zona norte da planície e das colinas de Recife (c).....	46
Figura 11 - Unidades geomorfológicas orientadas de acordo com o perfil da renda do chefe familiar no município de Recife.	48
Figura 12 - Fluxograma da metodologia da tese.	49
Figura 13- Variáveis utilizadas no modelo, de acordo com sua categoria, e estudos que relacionam as respectivas condicionantes para cada doença.....	51
Figura 14 - Procedimentos para extrair as informações de áreas não construídas.	52
Figura 15 - Unidades espaciais utilizadas nos modelos espaciais.	53
Figura 16 - Organização dos modelos de acordo com o método da estimativa, bandwidth do GWR, unidade espacial e tipo de cálculo da variável dependente.	55
Figura 17 - Casos notificados de Zika e Chikungunya entre 2015 e 2017, de acordo com os bairros de Recife.....	59
Figura 18 - Modelos categorizados de acordo com o tipo de unidade espacial, cálculo da variável dependente e método para estimar os casos de Zika e Chikungunya. O campo para comparação possui o R ² ajustado, a soma do valor absoluto dos resíduos e o critério de informação de Akaike (AICC).....	60
Figura 19 - Densidade da estimativa (método GWR e OLS) e casos observados de Zika e Chikungunya em Recife.	63

Figura 20 - Distribuição espacial do R^2 ajustado no método GWR para a Zika e Chikungunya.	66
Figura 21 - Distribuição espacial dos coeficientes do método GWR para cada variável inserida no modelo para estimar a Zika e a Chikungunya.	67
Figura 22 - Densidade espacial da Zika/Chikungunya, índice de áreas não construídas e coeficiente do GWR, na zona sul, oeste e centro do Recife.	69
Figura 23 - Densidade espacial da Zika/Chikungunya, renda média mensal do chefe familiar e coeficiente do GWR, na zona norte, oeste e no centro do Recife.	71
Figura 24 - Densidade espacial da Zika/Chikungunya, declividade média e coeficiente do GWR, na zona norte, oeste e no centro do Recife.	73
Figura 25 - Casos notificados de Leptospirose entre 2015 e 2017, de acordo com os bairros de Recife.	77
Figura 26 - Modelos categorizados de acordo com o tipo de unidade espacial, cálculo da variável dependente e método para estimar os casos de Leptospirose. O campo para comparação possui o R^2 ajustado, a soma do valor absoluto dos resíduos e o Critério de Informação de Akaike	78
Figura 27 - Densidade da estimativa (método GWR e OLS) e casos observados da Leptospirose em Recife.	80
Figura 28 - Distribuição espacial do R^2 ajustado no método GWR para a Leptospirose.	83
Figura 29 - Distribuição espacial dos coeficientes do método GWR para cada variável inserida no modelo para estimar a Leptospirose.	85
Figura 30 - Densidade espacial da Leptospirose, índice de áreas não construídas e coeficiente do GWR, na zona norte e oeste do Recife.	86
Figura 31 - Densidade espacial da Leptospirose, renda média mensal do chefe familiar e coeficiente do GWR, na zona oeste do Recife.	89
Figura 32 - Densidade espacial da Leptospirose, declividade média e coeficiente do GWR, na zona norte do Recife.	91
Figura 33 - Casos notificados de Tuberculose entre 2015 e 2017, de acordo com os bairros em Recife.	93
Figura 34 - Modelos categorizados de acordo com o tipo de unidade espacial, cálculo da variável dependente e método para estimar os casos de Tuberculose. O campo para comparação possui o R^2 ajustado, a soma do valor absoluto dos resíduos e o critério de informação Akaike.	94

Figura 35 - Densidade da estimativa (método GWR e OLS) e casos observados da Tuberculose em Recife.	96
Figura 36 - Distribuição espacial do R ² ajustado no método GWR para a Tuberculose.	100
Figura 37 - Distribuição espacial dos coeficientes do método GWR para cada variável inserida no modelo para estimar a Tuberculose.	101
Figura 38 - Densidade espacial da Tuberculose, moradores por domicílios e coeficiente do GWR, na zona norte do Recife.	104
Figura 39 - Densidade espacial da Tuberculose, índice de áreas verdes e coeficiente do GWR, na zona oeste do Recife.	107
Figura 40 - Modelos que apresentaram melhores eficiências nas suas estimativas, de acordo com suas respectivas doenças (Zika/Chikungunya, Leptospirose e Tuberculose).	109
Figura 41 - Variáveis utilizadas e os coeficientes do método OLS dos modelos que apresentaram melhor eficiência, organizados de acordo com a relação positiva (vermelho) ou negativa (azul) com suas respectivas doenças estudadas.	111
Figura 42 - A densidade habitacional por bairros e os aspectos levantados sobre a distribuição espacial das doenças analisadas.	114
Figura 43- A renda média mensal do chefe familiar por bairros e os aspectos levantados sobre a distribuição espacial das doenças analisadas.	115
Figura 44 - Proposta teórica para modelagem espacial de doenças, adicionando outras doenças como variável independente.	118
Figura 45 – R ² dos Modelos 4 (Tuberculose), 5 (Zika/Chikununya) e 6 (Leptospirose) com método de estimativa OLS, e suas diferenças quando utilizada outras doenças como variáveis independentes nos seus respectivos modelos.	119
Figura 46 - Densidade espacial observada e estimada da Tuberculose com o método OLS sem as doenças e com as doenças como variáveis independentes.	121
Figura 47 - Densidade espacial observada e estimada da Zika/Chikungunya com o método OLS sem as doenças e com as doenças como variáveis independentes.	123
Figura 48 - Densidade espacial observada e estimada da Leptospirose com o método OLS sem as doenças e com as doenças como variáveis independentes.	125

LISTA DE TABELAS

Tabela 1 - Os 10 bairros com maiores notificações de Zika e Chikungunya, entre 2015 e 2017	58
Tabela 2 - Variáveis independentes utilizadas e suas relações com a Zika e Chikungunya de acordo com o coeficiente, <i>Robust Probability</i> e VIF no método OLS.	64
Tabela 3 - Os 10 bairros com maiores registros de Leptospirose entre 2015 e 2017.....	75
Tabela 4 - Variáveis independentes utilizadas e suas relações com a Leptospirose de acordo com o coeficiente, <i>Robust Probability</i> e VIF no método OLS.	81
Tabela 5 - Os 10 bairros com maiores registros de Tuberculose entre 2015 e 2017.....	92
Tabela 6 - Variáveis independentes utilizadas e suas relações com a Tuberculose de acordo com o coeficiente, <i>Robust Probability</i> e VIF no método OLS.	98
Tabela 7 - Variáveis independentes utilizadas no Modelo 4 adicionando as densidades espaciais da Zika/Chikungunya e Leptospirose e suas relações com a Tuberculose de acordo com o coeficiente, <i>Robust Probability</i> e VIF, no método OLS.	120
Tabela 8 - Variáveis independentes utilizadas no Modelo 5 adicionando as densidades espaciais da Tuberculose e Leptospirose e suas relações com a Zika/Chikungunya de acordo com o coeficiente, <i>Robust Probability</i> e VIF, no método OLS.	122
Tabela 9 - Variáveis independentes utilizadas no Modelo 6 adicionando as densidades espaciais da Tuberculose e Zika/Chikungunya e suas relações com a Leptospirose de acordo com o coeficiente, <i>Robust Probability</i> e VIF, no método OLS.	124

LISTA DE SIGLAS

AICC	<i>Akaike Information Criterion Correction</i>
AVC	Acidente Vascular Cerebral
CHIKV	Chikungunya Vírus
DTN	Doenças Tropicais Negligenciadas (DTN)
EP	Embolia Pulmonar
GWR	<i>Geographically Weighted Regression</i>
IBGE	Instituto Brasileiro de Geografia e Estatística
LIDAR	<i>Light Detection and Ranging (LIDAR)</i>
NDVI	<i>Normalized Difference Vegetation Index</i>
OLS	<i>Ordinary Least Square</i>
TB	Tuberculose
VIF	<i>Variance Inflation Factor</i>
ZEIS	Zonas Especiais de Interesse Social
ZIKV	Zika Vírus

SUMÁRIO

1	INTRODUÇÃO	15
2	OBJETIVOS	18
2.1	OBJETIVOS ESPECÍFICOS	18
3	REFERENCIAL TEÓRICO	19
3.1	A RELAÇÃO DAS DOENÇAS COM FATORES NATURAIS E SOCIOECONÔMICOS NO ESPAÇO GEOGRÁFICO	19
3.2	ASPECTOS EPIDEMIOLÓGICOS DA ZIKA, CHIKUNGUNYA, TUBERCULOSE E LEPTOSPIROSE	21
3.2.1	Zika e Chikungunya	21
3.2.2	Tuberculose	23
3.2.3	Leptospirose	25
3.3	MÉTODOS DE REGRESSÃO LINEAR GLOBAL E LOCAL	26
3.3.1	<i>Ordinary Least Square (OLS)</i>	26
3.3.2	<i>Geographical Weighted Regression (GWR)</i>	27
4	MATERIAIS E MÉTODOS	31
4.1	DESCRIÇÃO DA ÁREA DE ESTUDO	31
4.1.1	Aspectos socioeconômicos da área estudada	32
4.1.2	Aspectos físicos e estruturais da área estudada	37
4.2	METODOLOGIA	49
4.2.1	Coleta de dados e distribuição espacial das doenças	50
4.2.2	Coleta e definição dos condicionantes espaciais a serem relacionados com as doenças	50
4.2.3	Análise e definição do desempenho do modelo de estimativa de distribuição das doenças analisadas	52
4.2.4	Desenvolvimento da proposta teórico metodológica da tese	57
5	RESULTADOS E DISCUSSÕES	58
5.1	ZIKA E CHIKUNGUNYA	58
5.1.1	Avaliação da eficiência dos modelos	60
5.1.2	Análise dos resultados do método OLS	64
5.1.3	Análise dos resultados do método GWR	65
5.2	LEPTOSPIROSE	75
5.2.1	Avaliação da eficiência dos modelos	78

5.2.2	Análise dos resultados do método OLS	81
5.2.3	Análise dos resultados do método GWR	82
5.3	TUBERCULOSE	92
5.3.1	Avaliação da eficiência dos modelos	94
5.3.2	Análise dos resultados do método OLS	97
5.3.3	Análise dos resultados do método GWR	98
5.4	DESENVOLVIMENTO TEÓRICO SOBRE A MODELAGEM ESPACIAL DE DOENÇAS.....	108
6	CONSIDERAÇÕES FINAIS.....	127
	REFERÊNCIAS	129

1 INTRODUÇÃO

A paisagem urbana é considerada como a expressão das relações e ações que propiciam uma investigação sobre a cidade aos olhos do geógrafo. Sob o ponto de vista geográfico, a paisagem liga-se ao plano imediato, que estaria inerente ao conceito de produção do espaço, sendo esse analisado como produto das transformações que a sociedade humana realiza a partir da natureza em um determinado momento do desenvolvimento das forças produtivas sob múltiplas formas de uso, podendo ser construções para moradia, lazer e atividades relacionadas ao trabalho (CARLOS, 2007).

Diante da complexidade das interações das atividades humanas no espaço geográfico, é imprescindível ao geógrafo identificar os problemas inerentes aos fatores relacionados a urbanização. Um desses problemas, estão relacionados à proliferação de doenças por fatores antrópicos e/ou naturais.

Considerando as condições socioeconômica dos países, George (1989) afirma que há uma correlação da permanência de determinadas endemias em países que apresentam más condições de vida. Ainda segundo o autor, a Geografia da Saúde permitiu avaliar que embora alguns países possam ter todo aparato de equipamentos médicos com recursos tecnológicos avançados, a segurança em termos de saúde pode variar dentro do país de acordo com os locais e as condições de residência, formas de trabalho, nível de instrução e renda.

Sob o ponto de vista das condicionantes naturais, essas podem ser integradas com os fatores socioeconômicos para uma análise mais completa da distribuição espacial de uma determinada doença. Como exemplo disso, Ayoade (2010) considera que as funções fisiológicas do homem respondem às mudanças no tempo atmosférico. Segundo ele, certas doenças são induzidas pelo clima em tempos diferentes, condicionadas pela configuração climática e as estações do ano.

As variáveis socioeconômicas e ambientais, por apresentarem possibilidades de espacialização, podem indicar correlações com a concentração espacial de determinadas enfermidades, auxiliando na identificação de quais fatores estabelecidos no espaço geográfico podem interferir em sua maior densidade espacial. Tais procedimentos começaram já na metade do século XIX, por meio da análise espacial com a contribuição do médico John Snow, identificando a relação espacial dos casos de cólera e as águas contaminadas oriundas de poços, em Londres (SHIODE *et al.*, 2015).

Com essas características, algumas doenças se distribuem dentro dos centros urbanos apresentando uma alta correlação espacial com condicionantes naturais e socioeconômicos. Dentre essas doenças, podemos destacar as arboviroses, Leptospirose e Tuberculose.

No caso das arboviroses, a Zika e a Chikungunya têm como principal vetor o mosquito *Aedes aegypti* que, de acordo com Mendonça *et al.* (2009), sua proliferação está condicionada ao crescimento urbano e as condições atmosféricas, visto que o mosquito tende a se reproduzir em ambientes com características climáticas tropicais. Já a Leptospirose, pode apresentar uma relação com altos índices pluviométricos associados a más condições de infraestrutura urbana. A tuberculose, considerada um problema de saúde prioritário no Brasil, tende a predominar em áreas de grande concentração populacional e precárias condições socioeconômicas e sanitárias (QUEIROGA *et al.*, 2012).

Embora tenha um número crescente de trabalhos acadêmicos que busquem identificar os fatores que contribuem na distribuição espacial de determinadas doenças, de modo geral, os resultados podem mostrar análises pontuais que não levam em consideração o contexto socioeconômico e natural, a partir de uma análise integrada. Associada a essa lacuna, a modelagem espacial surge com o intuito de prever onde essas doenças podem se manifestar de acordo com alguns condicionantes pré-estabelecidas sendo uma metodologia a ser melhor desenvolvida com esforços interdisciplinares envolvendo vários campos do conhecimento.

Nesse sentido, o Sistema de Informação Geográfica (SIG) pode auxiliar na identificação de áreas suscetíveis a disseminação de doenças que estão ligadas não somente a fatores socioambientais, mas relacionados a cobertura e acesso aos serviços de saúde (regulação, vigilância e assistência à saúde).

As técnicas relacionadas a modelagem espacial permitem investigar onde e quão forte são as relações entre as condicionantes geográficas com a distribuição espacial de doenças. Desse modo, torna-se possível a criação de modelos preditivos que possam estimar espacialmente uma determinada enfermidade. Além disso, esses modelos espaciais podem auxiliar na produção de possíveis cenários de acordo com as variações climáticas.

Nos casos de doenças emergentes em determinados países, é possível identificar em cenários preditivos, em um modelo espacial eficiente, onde e quando pode ocorrer a maior concentração espacial de uma determinada enfermidade. Esses modelos espaciais, por exemplo, poderiam ser utilizados em casos de pandemia para estimar onde irão ocorrer as maiores notificações da doença, mesmo quando a sua disseminação encontrasse ainda na fase inicial. Esses resultados auxiliariam num planejamento de atendimento de saúde mais eficiente em áreas com maior precariedade dos serviços públicos de saúde.

Destarte, partem alguns questionamentos acerca do tema proposto, com intuito de nortear o direcionamento da pesquisa, dentre eles:

- Como e quais os fatores interferem na distribuição espacial das doenças?
- Existem condicionantes que possuem correlações espaciais em comum entre algumas doenças?
- Quais metodologias são aconselháveis para estimar espacialmente algumas doenças?
- É possível criar uma proposta teórica acerca da modelagem espacial de doenças?

A partir dessas análises, deve-se considerar que para um conjunto de ações preventivas efetivas é necessária a compreensão de quais fatores influenciam em determinadas doenças no espaço geográfico e como, a partir da resposta, deve ser direcionado o foco do planejamento dos órgãos públicos, ressaltando o papel do cidadão na prevenção de epidemias. Partindo desse princípio, considerando o objeto de estudo da geografia, torna-se necessária a interpretação de como a sociedade e o meio físico influenciam na distribuição espacial das doenças, tornando trabalhos científicos nesse âmbito instrumentos de utilidade pública, contribuindo para um planejamento urbano mais efetivo e abrangente no que concerne à saúde pública.

Diante do exposto acima, a presente tese buscou investigar quais métodos e técnicas são mais eficientes nas estimativas da distribuição espacial da Zika, Chikungunya, Leptospirose e Tuberculose, levantando propostas para a modelagem espacial de doenças. Outro aspecto levantado foi analisar a relação de condicionantes geográficas (renda média mensal, densidade populacional, presença de áreas verdes, precipitação, áreas não edificadas, morador por domicílio e declividade) com as doenças já mencionadas.

2 OBJETIVOS

Propor uma abordagem teórico-metodológica para a modelagem espacial de doenças, utilizando como base as relações entre a distribuição espacial da Zika, Chikungunya, Tuberculose e Leptospirose com os condicionantes socioeconômicos, morfoclimáticos e infraestruturais, em Recife.

2.1 OBJETIVOS ESPECÍFICOS

- Identificar o modelo que apresente melhor desempenho para estimar a distribuição espacial das doenças analisadas a partir da combinação do tipo de unidade espacial, cálculo da variável dependente e método de estimativa.
- Verificar quais e como as condicionantes morfoclimáticas, socioeconômicas e estruturais influenciam na distribuição espacial da Zika, Chikungunya, Leptospirose e Tuberculose.
- Elaborar uma proposta teórica e metodológica acerca do processo de modelagem espacial de doenças de acordo com os resultados encontrados.

3 REFERENCIAL TEÓRICO

Nesta seção será apresentado o referencial teórico da pesquisa.

3.1 A RELAÇÃO DAS DOENÇAS COM FATORES NATURAIS E SOCIOECONÔMICOS NO ESPAÇO GEOGRÁFICO

Historicamente, observações acerca do diferencial de doenças de acordo com o local, vem desde a Grécia Antiga, com Hipócrates (SILVA, 1997). No entanto, os trabalhos desenvolvidos por John Snow, no século XIX, foram pioneiros nos procedimentos metodológicos para investigação epidemiológica (SHIODE *et al.*, 2015). Algumas teorias que se debruçaram sobre a relação do espaço geográfico com a disseminação de doenças merecem destaque. Por exemplo, a teoria desenvolvida pelo parasitologista russo Eugene Pavlovsky, em 1939, denominada de teoria dos focos naturais, abordava que algumas doenças estavam associadas a determinadas paisagens. De acordo com o parasitologista, alguns agentes patógenos circulam livremente em ambientes ausentes da presença humana, sendo consideradas como foco natural da doença. Naturalmente, os indivíduos que circulassem nessas áreas entrariam em contato com os agentes patógenos e, conseqüentemente infectariam-se pela doença (MALKHAZOVA *et al.*, 2014).

De modo geral, a junção dos fatores ambientais e socioeconômicos podem influenciar numa maior incidência de determinadas doenças, e alguns geógrafos vêm se destacando com estudos que mostram essa relação. Medeiros (2003) identificou que as atividades no polo gesso localizado no município de Araripina, em Pernambuco, influenciam na ocorrência de doenças respiratórias. É válido ressaltar que o autor destaca que a exposição ao gesso, alta incidência em jovens com baixa escolaridade, interferem nos números de complicações respiratórias. Além disso, a influência climática pode ser considerada, devido as partículas suspensas na atmosfera estarem condicionadas a precipitação, direção dos ventos e umidade.

Sobre a leptospirose, alguns estudos evidenciam que ela está relacionada a fatores sociais, econômicos e ambientais. Vasconcelos *et al.* (2012) identificaram que no Estado de Pernambuco, os casos de leptospirose localizaram-se predominantemente em áreas urbanas (87%), sendo o ambiente domiciliar onde ocorre a infecção, na maioria dos casos. Os autores ainda identificaram que as maiores concentrações espaciais da doença ocorrem onde existe a pior situação de renda, destino inadequado do lixo e maior densidade da rede de drenagem na cidade. Tal fato evidencia que a infraestrutura do entorno urbanístico dos domicílios pode influenciar diretamente numa maior exposição das pessoas à doença.

Dentro do espaço urbano, Anjos et al. (2013) identificaram que a quantidade precipitada associada a áreas de precária infraestrutura urbana favorece a ocorrência de casos de leptospirose. Os autores identificaram que os bairros que não tinham Zonas Especiais de Interesse Social (ZEIS) eram os que possuíam os menores índices de notificações, concomitantemente a isso, verificaram que os anos que mais choviam tendiam a ter maiores quantidades de casos notificados da doença. Ora sendo as ZEIS regiões com condições precárias de habitação, o trabalho dos autores pôde levantar que até certo ponto a regionalização pode auxiliar numa parametrização entre a leptospirose e condições subnormais de saúde humana. Desse modo, como proposta para minimização dos danos causados pela leptospirose, Oliveira (2009) considera como pilares fundamentais o saneamento ambiental, a educação, a informação e a comunicação.

Relacionada a condições socioeconômicas mais precárias, a Tuberculose é considerada como um problema de saúde mundial. A sua distribuição espacial no mundo está associada a fatores geográficos de relevância. San Pedro e Oliveira (2013) verificaram que países com menores Índices de Desenvolvimento Humano (IDH), Produto Interno Bruto (PIB), áreas sem saneamento básico e com alta taxa de mortalidade infantil apresentavam as maiores taxas de incidência de tuberculose.

Ferreira (2018) identificou que os municípios no Estado de Pernambuco que apresentavam baixos índices de escolaridade e renda e altas taxas de desemprego apresentavam maiores taxas de pessoas infectadas com a Tuberculose. Sob o ponto de vista climático, um estudo desenvolvido no Distrito Federal identificou que no inverno concentravam as maiores taxas de pessoas infectadas com tuberculose (FERNANDES et al., 2017). Esse fator, segundo os autores, pode estar relacionado a uma menor incidência de raios ultravioletas nessa época e conseqüentemente menor exposição à vitamina D, uma das vitaminas responsáveis pelo aumento da resistência do corpo à bactéria.

No que concerne as arboviroses emergentes, Zika e Chikungunya, Lima-Camara (2016) destaca que o ambiente está relacionado a maior proliferação dos vetores dessas doenças. Ainda segundo a autora, o crescimento desordenado das cidades, associada a poluição dos rios, formação de valas disponibiliza sítios de oviposição artificiais para a proliferação e disseminação dos mosquitos. Além disso, uma maior frequência das chuvas resulta em acúmulo de água em recipientes, aumentando a oferta de criadouros, por outro lado, em períodos de seca em determinadas regiões, obriga as pessoas a armazenarem água em tonéis e outros reservatórios artificiais (LIMA-CAMARA, 2016). Em Recife, Ferreira (2016) identificou que

maiores temperaturas associada a alta umidade relativa do ar na cidade, promovem durante todo o ano condições ideais de reprodução e dispersão do vetor da dengue.

Considerando os aspectos de má infraestrutura de habitação e condições sociais precárias, Lins (2019) verificou que as maiores concentrações da Zika e Chikungunya em Recife concentram-se espacialmente em áreas de aglomerados subnormais. Outro estudo elaborado por Silva et al., (2021) no Distrito Sanitário III do município de Recife verificou que os bairros que possuíam Zonas Especiais de Interesse Social eram os que apresentavam as maiores incidências de Dengue, Zika e Chikungunya.

Para Valle et al. (2016) a falha na prevenção da dengue é devido a fatores que extrapolam o setor da Saúde. Para os autores, merecem destaques aspectos macroestruturais, socioeconômicos e ambientais, historicamente ignorados em prol de intervenções meramente biomédicas ou tecnológicas.

Sob o ponto de vista climático, Ferreira et al. (1998) verificaram que a precipitação possui relação com os casos notificados de Hepatite A. No entanto as condições socioeconômicas podem influenciar na distribuição espacial da doença. Segundo os autores, esses mesmos fatores estão relacionados com o aumento da Dengue e Leptospirose. A partir dessas considerações, podemos inferir que a abordagem acerca das condições socioespaciais, atreladas às condições climáticas, podem responder-nos sobre os riscos que a população vive, mesmo que algumas doenças tenham diferentes vetores, meios de transmissão e agentes etiológicos.

Então, poder-se-ia pensar numa análise mais global da distribuição geral de doenças? Uma análise que levasse em consideração as condições socioambientais e considerassem as doenças como uma consequência desses tais fatores? Pressupõe-se que sim, dado ao fato que uma gama diversificada de enfermidades depende de condições insalubres para sua proliferação no espaço urbano.

3.2 ASPECTOS EPIDEMIOLÓGICOS DA ZIKA, CHIKUNGUNYA, TUBERCULOSE E LEPTOSPIROSE

3.2.1 Zika e Chikungunya

Uma das arboviroses, a Zika é uma doença exantemática, que está relacionada a dengue, febre amarela e febre do Nilo Ocidental. Esta infecção pode durar 1 semana, com sintomas similares à Chikungunya e Dengue, incluindo febre leve, artralgia, artrite, mialgia, enxaqueca, conjuntivite e edema. Casos severos de hospitalização são incomuns, assim como casos de óbitos (PAIXÃO et al, 2016).

Embora o Zika Vírus (ZIKV) não acarrete sintomas severos em um indivíduo infectado, ele está relacionado a outras morbidades graves. Por exemplo, a Zika está associada com o aumento da incidência da Síndrome do Guillain-Barré, um tipo de paralisia crônica que provoca complicadas infecções. Além disso, ela está associada a doenças durante o período de gestação, como aborto espontâneo, morte fetal e a Síndrome Congênita da Zika (CARABALI et al., 2018).

Embora no Brasil seja uma doença com histórico recente, no mundo ela é endêmica em algumas regiões. Os primeiros registros do vírus datam da década de 40, do século XX, no continente africano (SAMPAIO et al., 2019).

A introdução da Zika no Brasil foi confirmada por pesquisadores em março de 2015. Rapidamente ela foi caracterizada como epidemia pelas autoridades públicas brasileiras, a partir da ocorrência de um surto por doença exantemática (que apresenta erupções na pele que coçam), especialmente na região Nordeste.

Suspeita-se que o número de casos suspeitos foi de 300.000 a 440.000 pessoas infectadas, até o fim de 2015. Existem várias hipóteses para a introdução da Zika no Brasil: (a) a Copa Mundial sediada no Brasil em 2014; (b) durante o Campeonato Mundial de Canoagem em 2014, no Rio de Janeiro; (c) Copa das Confederações, que ocorreu em junho de 2013 (CALVET et al., 2016). Todos esses fatores estão relacionados ao fluxo de turistas dos mais variados países para eventos esportivos mundiais.

Outra hipótese levantada por Campos et al., (2018) destaca a intensa migração de haitianos para o Brasil, desde o terremoto de 2010, que pode ter desencadeado numa maior proliferação da doença.

Essa doença mostrou-se mais preocupante quando, no final de 2015, as autoridades públicas de saúde brasileiras comprovaram, pela primeira vez na história, a associação entre o contágio de mulheres pelo vírus Zika durante a gravidez e o nascimento de bebês com microcefalia (MINISTÉRIO DA SAÚDE, 2016). Dentro desse contexto, a microcefalia está dentro de um conjunto de alterações na formação craniofacial, associada a calcificações predominando na junção córtico-subcortical, malformações do desenvolvimento cortical, ventriculomegalia e anormalidades na formação do corpo caloso. Todos esses fatores estão associados à Síndrome Congênita pelo ZIKV (RIBEIRO, et al., 2017). É importante considerar que a maioria das mães com bebês com microcefalia em Recife estão na linha de pobreza vivendo em locais com condições precárias de habitação (SOUZA et al., 2018).

Existe um alto potencial da Zika ser emergente em centros urbanos que tem a presença de vetores, tais como os mosquitos *Aedes aegypti* a *Aedes albopictus* (Song et al., 2017; Musso

and Gubler, 2016). É importante destacar que pesquisadores identificaram que outro mosquito, conhecido como *Culex quinquefasciatus* - que é mais comum nos centros urbanos brasileiros - está suscetível a infecção do ZIKV durante alguns experimentos de alimentação de sangue contaminado (GUEDES et al, 2017; LIU et al, 2017).

Pustiglione (2016) cita a heterogeneidade econômica, social e cultural como condicionantes a restrição do acesso a informações para populações menos favorecidas, criando condições para exposição a doenças. A relação da Zika e microcefalia reflete numa situação em que o a proximidade do vetor de mulheres que não tem acesso a informação ou recurso para evitar a gravidez, podem potencializar o nascimento de crianças com microcefalia.

A Chikungunya é outro arbovírus causado pelo Chikungunya vírus (CHIKV) que pertence ao gênero *Alphavirus da família Togaviridae*. É transmitido principalmente pelos mosquitos *Aedes aegypti* e *Aedes albopictus*; de toda forma, em algumas áreas, transmissões pelo *Culex*, *Mansonia* e *Anopheles* têm sido observadas (TAUBITZ et al., 2007). Cerca de 50 a 97% de pessoas infectadas com o CHIKV desenvolvem febre com artralgia. Essa infecção está associada com início súbito febril ($>38.9C^{\circ}$) (92% dos pacientes), artralgia (87% dos pacientes), dor nas costas (67% dos pacientes), enxaqueca (62% dos pacientes) e fadiga (SILVA et al., 2018).

Em relação a Zika e Chikungunya no Brasil, cerca de 223.230 casos prováveis foram diagnosticados entre 2013 e 2015, sendo 48% desses casos confirmados (AGUIAR, 2018). Em 2016, a distribuição espacial do ZIKV e CHIKV foram similares. Contudo, existiram maiores áreas de contaminação em 2016, se comparado com o ano de 2015.

Por serem consideradas doenças emergentes com um histórico recente de infectados no Brasil e com um número de pesquisas menor, se comparada com a Dengue, a presente tese buscou analisar como os casos de Zika e Chikungunya se relacionam com as condicionantes geográficas, de acordo com a metodologia proposta. Considerando as suas condições similares de proliferação, tais como os mesmos vetores de transmissão, os casos dessas duas doenças foram agrupados para a presente análise.

3.2.2 Tuberculose

A tuberculose (TB) pode ser causada por qualquer espécie de bactérias que integram o complexo *Mycobacterium tuberculosis*: *M. tuberculosis*, *M. bovis*, *M. africanum*, *M. canetti*, *M. microti*, *M. pinnipedi* e *M. caprae* (MINISTÉRIO DA SAÚDE, 2019). Contudo, para a saúde pública, a espécie que é mais considerada é a *M. tuberculosis*, conhecida também como bacilo de Koch (MINISTÉRIO DA SAÚDE, 2019).

A tuberculose se divide em dois grupos principais: pulmonar e extrapulmonar. A Tuberculose Extrapulmonar tem seus sinais e sintomas dependentes dos órgãos ou sistemas acometidos. Seus tipos são: pleural, pleural tuberculoso, ganglionar periférica, meningoencefálica, pericárdica e óssea.

A tuberculose pulmonar – considerada no presente estudo – pode ser primária e secundária. A primária ocorre após o primeiro contato de uma pessoa com o bacilo e é mais comum em crianças. Seus sintomas muitas vezes não são bem compreendidos pelo exame físico pois nem sempre a tosse está presente. A tuberculose secundária pode ocorrer em qualquer idade, sendo mais comum em jovens e adultos e tem como característica principal a tosse seca e prolongada (MINISTÉRIO DA SAÚDE, 2019).

O modo de transmissão da *M.tuberculosis* é por via aérea com tuberculose pulmonar ou laríngea entre pessoas através da tosse, fala ou espirro. Outras formas de transmissão (pele e placenta) são raras. Alguns fatores exógenos são importantes na probabilidade de uma pessoa ser infectada, tais como: infectividade do caso-fonte, duração do contato e o tipo de ambiente compartilhado (MINISTÉRIO DA SAÚDE, 2019). Após a infecção pelo *M. tuberculosis*, transcorrem, em média, 4 a 12 semanas para a detecção das lesões primárias (BRASIL, 2010).

Algumas estimativas mostram que uma pessoa infectada transmita a tuberculose para 10 a 15 pessoas em média, numa comunidade, durante um ano (MINISTÉRIO DA SAÚDE, 2019). É importante ressaltar que pessoas com cultura de escarro negativa e as com tuberculose extrapulmonar exclusivamente são desprovidas de infectividade.

A prevalência observada é maior em áreas de grande concentração populacional, e precárias condições socioeconômicas e sanitárias. A distribuição da doença é mundial, com tendência decrescente da morbidade e mortalidade nos países desenvolvidos (BRASIL, 2010).

No que concerne aos dados estatísticos, no ano de 2017, os coeficientes de incidência variam de 10 a 74,7 casos por 100 mil habitantes entre as Unidades Federadas (UF). Em 2016, foram registrados 4.483 óbitos por TB, correspondendo a um índice de mortalidade de 2,2 óbitos a cada 100 mil habitantes (MINISTÉRIO DA SAÚDE, 2019).

No município de Olinda, no Estado de Pernambuco, Silva et al. (2018) verificaram que as taxas de tuberculoses estavam associadas a taxas de analfabetismo das pessoas infectadas, densidade domiciliar e presença de domicílios sem abastecimento de água.

3.2.3 Leptospirose

A leptospirose é uma doença infecciosa causada por bactérias patogênicas do gênero *leptospiras*, que são transmitidas direta ou indiretamente de animais para humanos (WORLD HEALTH ORGANIZATION, 2003). Vários animais sinantrópicos, domésticos e selvagens servem como reservatórios, porém no meio urbano, os roedores, como o rato de esgoto são os mais comuns. É importante destacar que outras espécies de mamíferos podem ser reservatórios da doença como os suínos, bovinos, equinos, ovinos e cães (BRASIL, 2014).

O homem, hospedeiro terminal e acidental da doença, infecta-se ao entrar em contato com a urina de animais infectados de modo direto ou indireto, por meio do contato com água, lama ou solo contaminados (BRASIL, 2014). A penetração do microrganismo ocorre através da pele com lesões, pele íntegra quando imersa em água por longo tempo ou mucosas. A transmissão inter-humana é muito rara e de pouca relevância epidemiológica (BRASIL, 2014).

Seu espectro clínico pode variar desde a existência de poucos sintomas até formas graves. Alguns sinais, como calafrios, cefaléia, mialgia e febre semelhante, são observados durante essa fase. Mais raramente, conjuntivite e erupção cutânea maculopapular podem ser observados. Após 5 a 7 dias, os sinais podem recuar e desaparecer. Em alguns casos, a doença pode regredir para uma forma subclínica, mas geralmente, após 1 a 3 dias de aparente remissão, a segunda fase da doença começa. Esta é uma consequência da transferência da *leptospiras* dos vasos sanguíneos para os órgãos (WASIŃSKI e DUTKIEWICZ, 2013).

A síndrome de Weil, comumente descrita como icterícia, insuficiência renal e hemorragias é caracterizada pela manifestação mais grave da leptospirose, embora a síndrome de hemorragia pulmonar vem sendo reconhecida como uma forma grave e emergente da doença. A letalidade de formas graves de leptospirose é de aproximadamente 10% e chega a 50% quando ocorre a síndrome de hemorragia pulmonar (BRASIL, 2014).

A leptospirose ocorre em todo o mundo, mas é mais comum em áreas tropicais e subtropicais com alta pluviosidade. A doença é encontrada principalmente onde os seres humanos são expostos à urina de animais infectados ou a um ambiente poluído pela urina (WORLD HEALTH ORGANIZATION, 2003). Mesmo que apresente registros no Brasil durante todos os meses do ano, as epidemias urbanas anuais estão associadas a pós-enchentes e inundações, principalmente em comunidades carentes (BRASIL, 2014).

Alguns dados estatísticos no Brasil merecem ser considerados. Aproximadamente 3.000 casos são notificados por ano com uma letalidade próxima de 11%. A maioria da faixa etária dos infectados estão entre 20 e 49 anos, com média de internações de pacientes com cerca de 75%.

Para Oliveira *et al.* (2009), para a compreensão da leptospirose é preciso considerar quatro categorias de fatores que determinam na sua proliferação: fatores de reservatório (animais domésticos, taxas de reprodução, dispersão); fatores ambientais (pluviosidade, tipo de solo, hidrografia); fatores sócio-demográficos (densidade populacional, renda, moradia) e outros fatores (tempo e repetição de exposição, resposta imunológica). Outro aspecto relevante é que a exposição laboral também deve ser considerada, podendo ser uma das causas da predominância da leptospirose na população adulta, jovem e masculina.

3.3 MÉTODOS DE REGRESSÃO LINEAR GLOBAL E LOCAL

Nesta seção será apresentado os métodos geoestatísticos de regressão global e local.

3.3.1 *Ordinary Least Square* (OLS)

O método dos mínimos quadrados é usado para estimar os valores numéricos dos parâmetros para ajustar uma função a um conjunto de dados e para caracterizar as propriedades estatísticas das estimativas. Existe diversas variações desse método, como o *Ordinary Least Squares* (OLS) (ABDI, 2007). A equação OLS pode ser descrita como:

$$Y_i = \beta_0 + \beta_1 X_i + u_i \quad (1)$$

Onde para cada i , o Y é a variável dependente e X é a variável independente. Aqui, u_i é o termo de erro para a observação i , uma vez que contém todos os fatores que afetam Y_i além de X_i . O fator β_1 pode ser descrito como:

$$\beta_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad (2)$$

Onde $\bar{Y} = n^{-1} \sum_{i=1}^n Y_i$ é a média da amostra do Y_i e da mesma forma para \bar{X} . Assim, a equação acima é a covariância da amostra entre x e y dividida pela variância da amostra de x (WOOLDRIDGE, 2003). O último fator (β_0) é descrito como:

$$B_0 = Y_i - \beta_1 X_i - u_i \quad (3)$$

É importante destacar que o nome “Mínimos Quadrados Ordinários” deriva do fato de que essas estimativas minimizam a soma dos resíduos quadrados (WOOLDRIDGE, 2003).

3.3.2 Geographical Weighted Regression (GWR)

O GWR é uma derivação de uma equação tradicional de regressão linear, no entanto, ela permite que variações locais de uma variável no espaço geográfico possam alterar a variável dependente que se queira estimar. Dessa forma, a equação do GWR é apresentada da seguinte maneira:

$$Y_i = \sum X_{ij} \beta_j (p_i) + \varepsilon_i \quad (4)$$

Onde o X é uma matriz que contém as variáveis independentes, Y é a variável dependente que se deseja estimar num ponto i , e ε é um erro aleatório do qual a distribuição é $N(0, \sigma^2 I)$. β_j é um vetor do coeficiente de regressão de uma variável j , e p_i a localização geográfica, de maneira que a combinação dessas duas variáveis é a principal diferença de uma regressão linear comum (BRUNDON; FOTHERIGHAN; CHARLTON, 1998).

Apesar dessa equação parecer uma simples extensão de uma equação de regressão simples, a calibração desse modelo ajustado ao espaço geográfico, o torna mais complexo. Nesse modelo, torna-se atraente a ideia de que a base da estimativa de $\beta_j (p_i)$ seja variável ao longo de uma região geográfica. Destarte, os pesos atribuídos aos dados observados estariam relacionados a sua localização geográfica, de modo que quanto mais próximo de um ponto determinado, seu peso aumentaria (BRUNDON; FOTHERIGHAN; CHARLTON, 1996)

Portanto, o objetivo do GWR é estimar a variável Y, a partir de cada variável independente (X) em uma determinada localização geográfica (i). Assim, os procedimentos para essa estimativa baseiam-se dessa forma: (a) determinar a área de influência do *bandwidth* (raio ou número de observações em torno de um ponto, controlando a distância de decaimento dos pesos); (b) calcular um peso para cada vizinho com dados observados, com base na sua distância até o centro do dado observado determinado; (c) estimar o modelo de coeficiente usando a regressão do peso dos mínimos quadrados, como mostra a equação a seguir:

$$\beta_i = (x^t w_i x)^{-1} x^t w_i y \quad (5)$$

Onde o w_i é um peso geográfico para o centro i , de modo que $w_i = f(d_i, h)$ – onde $f()$ é uma função do Kernel, d_i é uma distância do vetor entre o centro i e todos os vizinhos, e h é um parâmetro de decaimento ou *bandwidth*. O T sobrescrito refere-se à transposição de uma matriz. Nota-se que nesse método encontra-se semelhanças com o método de regressão de Kernel e a

estimativa de densidade de Kernel, cuja principal diferença é que a atribuição de pesos do GWR leva em consideração a localização (BRUNDON; FOTHERIGHAN; CHARLTON, 1996).

Existem duas maneiras de determinar o *bandwidth* ou a função de decaimento: pelo Kernel fixo e o Kernel adaptativo. O Kernel fixo usa um *bandwidth* que não varia de acordo com a densidade dos dados, enquanto no caso do Kernel adaptativo, a *bandwidth* terá uma variação de acordo com a densidade dos dados observados (YOO, 2012).

Essa diferença permite que o kernel fixo atribua pesos para todos os dados observados, de modo que os pesos nunca terão um valor nulo, por mais que os dados observados estejam distantes uns dos outros (GUO; MA; ZHANG, 2008). Tal fato resulta num enfraquecimento dos resultados finais de correlação entre as variáveis, principalmente em modelos cujo os dados observados estejam concentrados espacialmente. Diante disso, o Kernel adaptativo pode ser uma alternativa de modo que os pesos sejam atribuídos até uma determinada região para aquele determinado dado observado. Essa área de atribuição de pesos pode ser considerada pela função do *bandwidth* associada a distância. Sendo assim, no Kernel adaptativo a função dos pesos apresenta-se da seguinte maneira:

$$w_{ik} = [1 - (d_{ik}/h_i)^2]^2 \text{ quando } d_{ik} \leq h_i \quad (6)$$

$$\text{ou } w_{ik} = 0 \text{ quando } d_{ik} > h_i$$

Onde $w_{ik} = 1$ no centro i , ($d_{ik} = 0$) e $w_{ik} = 0$, quando a distância iguala a distância ou número de vizinhos determinada pelo *bandwidth*, o peso é 0. O *bandwidth* é selecionado de maneira que cada número de observações com valores de pesos não nulos é o mesmo para cada localização i ao longo da área de estudo.

O *bandwidth* pode ser pré-determinado pela distância (através do Kernel Fixo) ou pelo número de vizinhos (Kernel Adaptativo). Contudo, alguns métodos permitem que se estimem o melhor *bandwidth* para que o modelo de previsão da variável Y , apresente menores erros. Dentre esses procedimentos, existem o método da validação cruzada e o *Akaike's Information Criterion* (AIC).

O método de Validação Cruzada (VC), para estimar o *bandwidth* é determinado na seguinte equação:

$$\sum_{i=1,n} [Y_i - Y_{\neq i}(\beta)]^2 \quad (7)$$

Onde o $Y_{\neq i}$ é o valor ajustado de Y_i , com as observações do ponto i omitidos nos processos de calibração. Essa metodologia tende a considerar apenas os pontos que estão

próximos de i , evitando que o próprio valor de i interfira na calibração do modelo. Avaliar o resultado da Validação Cruzada (VC) associado com diferentes parâmetros, permitiria identificar qual *bandwidth* responderia melhor ao parâmetro proposto diante do menor valor da VC (BRUNDON; FOTHERIGHAN; CHARLTON, 1996). É importante mencionar que, uma vez identificada o *bandwidth*, a distribuição espacial de (β) teria sua distribuição espacial baseada nos modelos de Kernel, como já mencionado.

Em relação ao *Akaike's Information Criterion* (AIC), não existe muita diferenciação da VC, de modo que os seus resultados tendem a ser parecidos (GUO et al, 2008). Segundo (GUO; MA; ZHANG, 2008), a vantagem de utilizar o AIC é que esse método considera que o grau de liberdade deve variar entre os modelos, para diferentes localizações.

Esse método pode ser usado na investigação geostatísticas nas mais variadas áreas, como exemplo, Ikefuti (2016) considerou que a renda per capita, proporção de pessoas negras, porcentagem de áreas verdes poderiam ser fatores que contribuiriam para o risco de mortalidade por Acidente Vascular Cerebral (AVC) e Embolia Pulmonar (EP). A partir do método GWR, a autora identificou que esses fatores em conjunto resultavam num coeficiente de correlação ajustado (R^2) de 0,59 e 0,60 para AVC e EP, respectivamente.

No distrito de Swat, no Paquistão, Atique et al., (2017) verificaram que a atuação em conjunto da densidade da população, altitude, distância de cursos d'água conseguiram responder em 49% dos casos notificados de Dengue. Segundo os autores, o fator que apresentou maiores correlações espaciais foi a densidade populacional.

Gwarda (2018) utilizou o GWR para correlacionar com algumas variáveis independentes com o desemprego na Polónia, dentre algumas, podemos citar: o número de empresas a cada 10.000 hab, renda *per capita*, nível de imigração, média do salário mínimo, porcentagem de mulheres. A autora percebeu que a associação desses fatores poderia responder em 69% das taxas de desemprego, embora as correlações espaciais variassem nos setores estudados.

Melo (2015) associou fatores - que dentre alguns, estão a taxa de analfabetismo, renda per capita, taxa de urbanização, migração, saneamento inadequado, variáveis climáticas - com a taxa de incidência de pessoas infectadas por Doenças Tropicais Negligenciadas (DTN) em todos os municípios brasileiros. O autor identificou que o conjunto dessas variáveis apresentaram coeficiente de correlação ajustado (R^2) de 0,84. Foi verificado que os fatores tendem a responder de maneira diferenciada a distribuição da correlação espacial do coeficiente de correlação local.

No Brasil, Cunha et al., (2015) identificou que os casos de hanseníase no Rio de Janeiro associados a fatores sociais e de serviços apresentaram um coeficiente de correlação (R^2) de 0,36, através do método de GWR.

Com o intuito de identificar variações no R^2 utilizando diferentes *bandwidths*, Weber (2018) simulou casos de Zika, em Porto Rico, determinando variáveis independentes a porcentagem de pessoas na pobreza, porcentagem de pessoas sem seguro de saúde, densidade populacional, área urbanizadas e temperatura. Foi identificado que o modelo adaptativo do GWR tende a ter maiores correlações espaciais (0,78) que o método do Kernel Fixo (0,72), embora o último apresentasse um AICc menor. Além disso, o autor identificou que 50 vizinhos correspondia as maiores correlações espaciais e previsibilidade do modelo.

Em Taiwan, Lin e Wen (2011) correlacionaram a densidade populacional e número de focos da Dengue com o número de casos notificados de Dengue. Os autores apresentaram um coeficiente de correlação de 0,59. Além disso, foram comparados os resultados entre o modelo de regressão global OLS (*Ordinary Least Squares*) e o local GWR, sendo o último com melhor desempenho na previsibilidade do modelo.

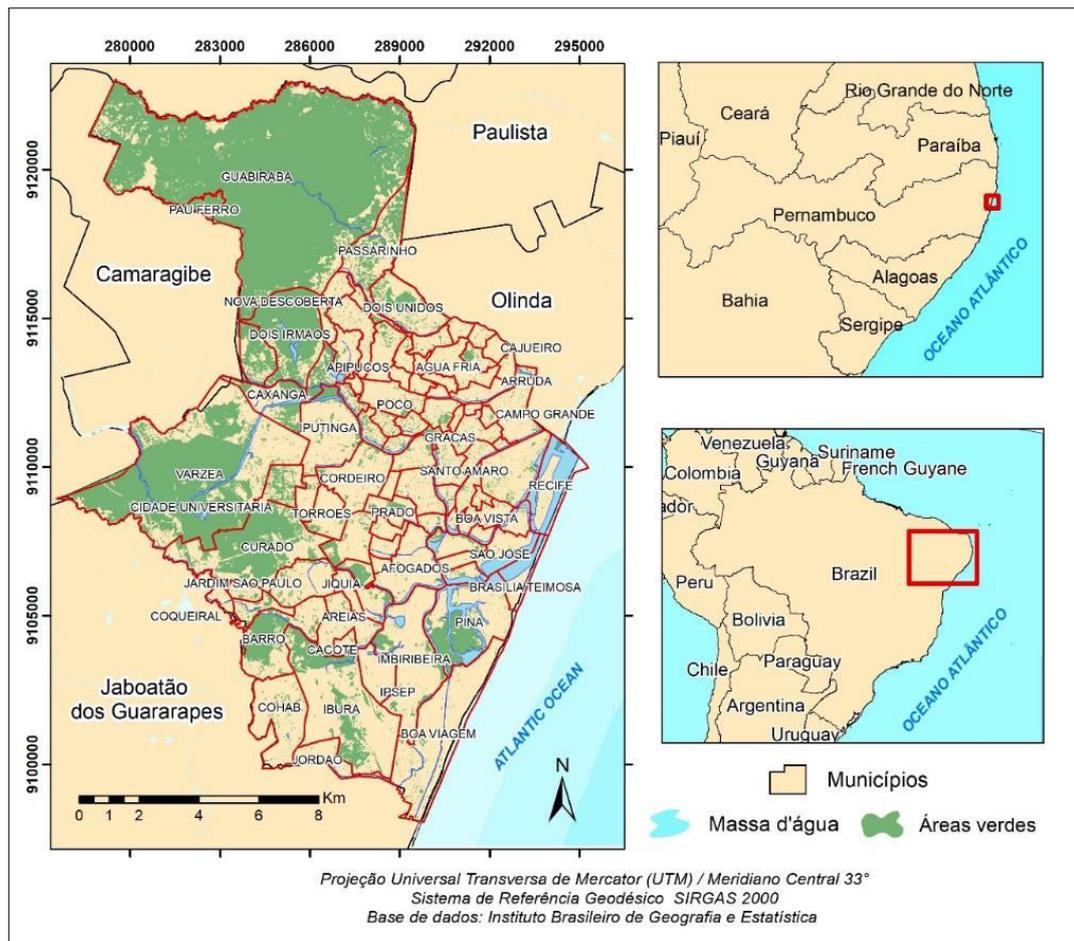
Grillet *et al*, (2010) apresentaram correlações significativas entre densidade populacional, declividade, altitude, distância da avenida principal com a incidência de Malária no nordeste da Venezuela. É válido ressaltar que o R^2 apresentava variações espaço-temporais, variando de 0,47 a 0,83 entre os anos de 2001 e 2007.

4 MATERIAIS E MÉTODOS

4.1 DESCRIÇÃO DA ÁREA DE ESTUDO

O município de Recife é a capital do Estado de Pernambuco e está situada no nordeste brasileiro (Figura 1). A sua área é de 218,4 km² com uma população de aproximadamente 1.653.461 habitantes (IBGE, 2020). A densidade habitacional do município é de aproximadamente 7.333 habitantes por km². Ao todo, Recife possui 94 bairros e 1.854 setores censitários. É válido destacar que os bairros recifenses apresentam uma diversidade social, cultural e econômica, evidenciando contrastes, mesmo em pequenas distâncias. Tal fato, reflete em condições de vida diferenciadas e distribuídas ao longo do seu espaço, condicionando a distribuição espacial de algumas doenças a seguirem a mesma lógica espacial.

Figura 1- Localização do município de Recife e seus bairros, de acordo com o Estado e país.



Fonte dos dados: IBGE (2010). As áreas verdes foram extraídas do Satélite *Sentinel* datada no ano de 2016.

Destarte, os assentamentos precários e as áreas mais valorizadas do município, em muitos casos são vizinhas. Diante de tais características, os setores censitários oferecem uma desagregação do espaço de modo que cada setor possua uma determinada homogeneidade sob vários aspectos: seja por infraestrutura, condições de habitação, características naturais e número de domicílios.

4.1.1 Aspectos socioeconômicos da área estudada

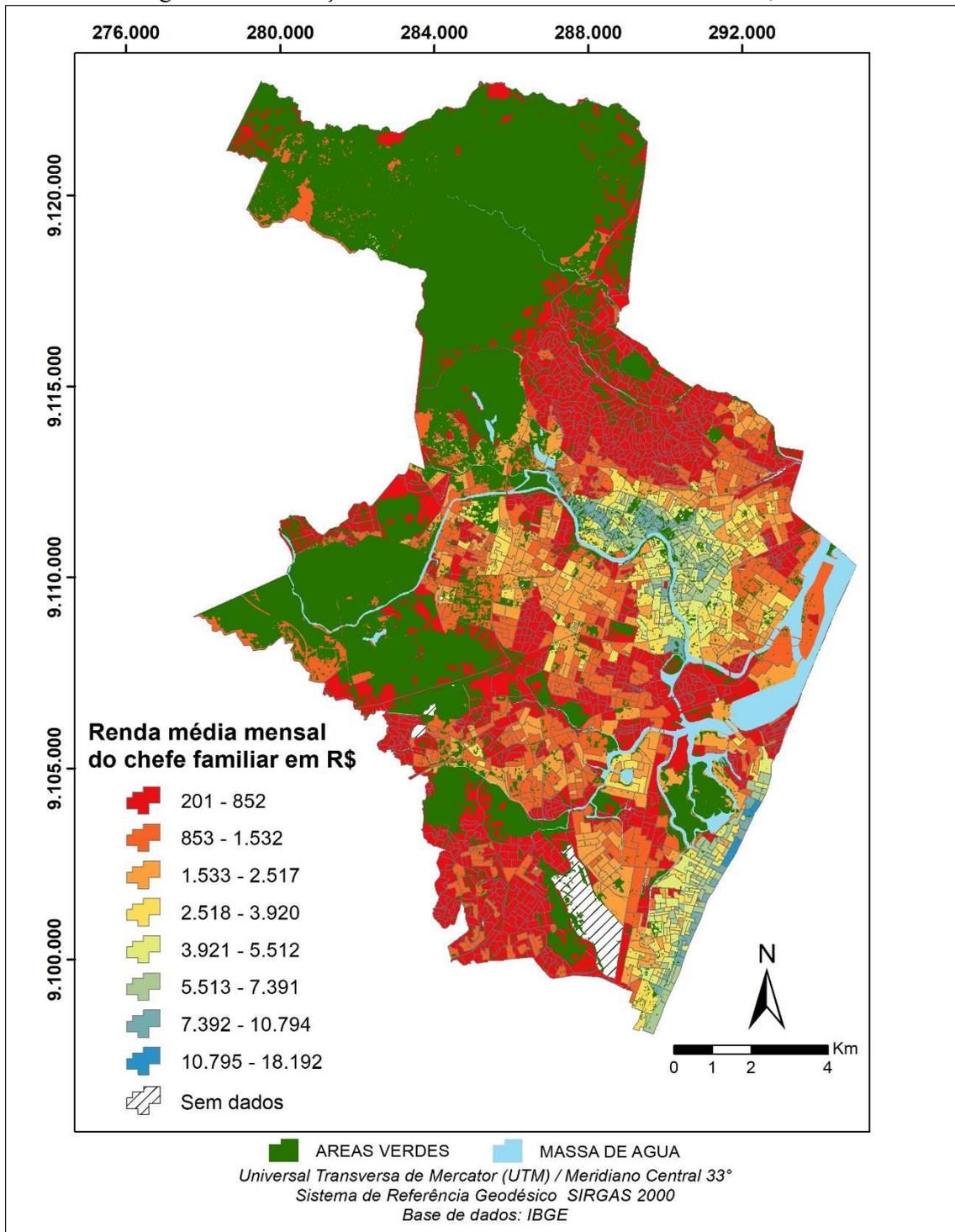
É importante salientar que as variáveis socioeconômicas estão relacionadas entre si, de modo que a distribuição espacial de alguns aspectos possua contrastes similares. Por isso, não se pode negligenciar que a distribuição de renda, densidade habitacional e condições de infraestrutura façam parte de uma mesma lógica espacial. No que se refere a distribuição espacial da renda média mensal dos chefes de família (Figura 2), observa-se que existe uma alta desigualdade na distribuição da renda, que varia de R\$ 201,00 a R\$ 18.192,00. É válido destacar que, durante o censo de 2010, o valor do salário mínimo era de R\$ 510,00, segundo a Lei 12.255/2010. De toda forma, embora os valores sejam alterados anualmente, os padrões espaciais de concentração de renda, tendem a permanecer ao longo dos anos.

Partindo desse princípio, é importante ressaltar algumas peculiaridades em sua distribuição. As áreas onde são apresentadas as maiores rendas do município estão concentradas em duas áreas principais: no litoral - constituído pelos bairros de Boa Viagem e Pina - e a área popularmente conhecida como a *zona norte recifense*, situada na planície flúvio-marinha da margem esquerda do Rio Capibaribe.

Historicamente, as duas áreas tiveram processos de ocupação diferenciados. A *zona norte da planície*, que teve seu início de ocupação relacionada à antiga aristocracia recifense, continuou ao longo do seu processo histórico sendo habitada pelas famílias mais abastadas de Recife. A área litorânea, ocupada de maneira mais acentuada a partir da década de 60 do século passado, continua a ser ocupada pelas famílias com maior poder aquisitivo onde a especulação imobiliária tem sido cada vez mais frequente nos últimos anos.

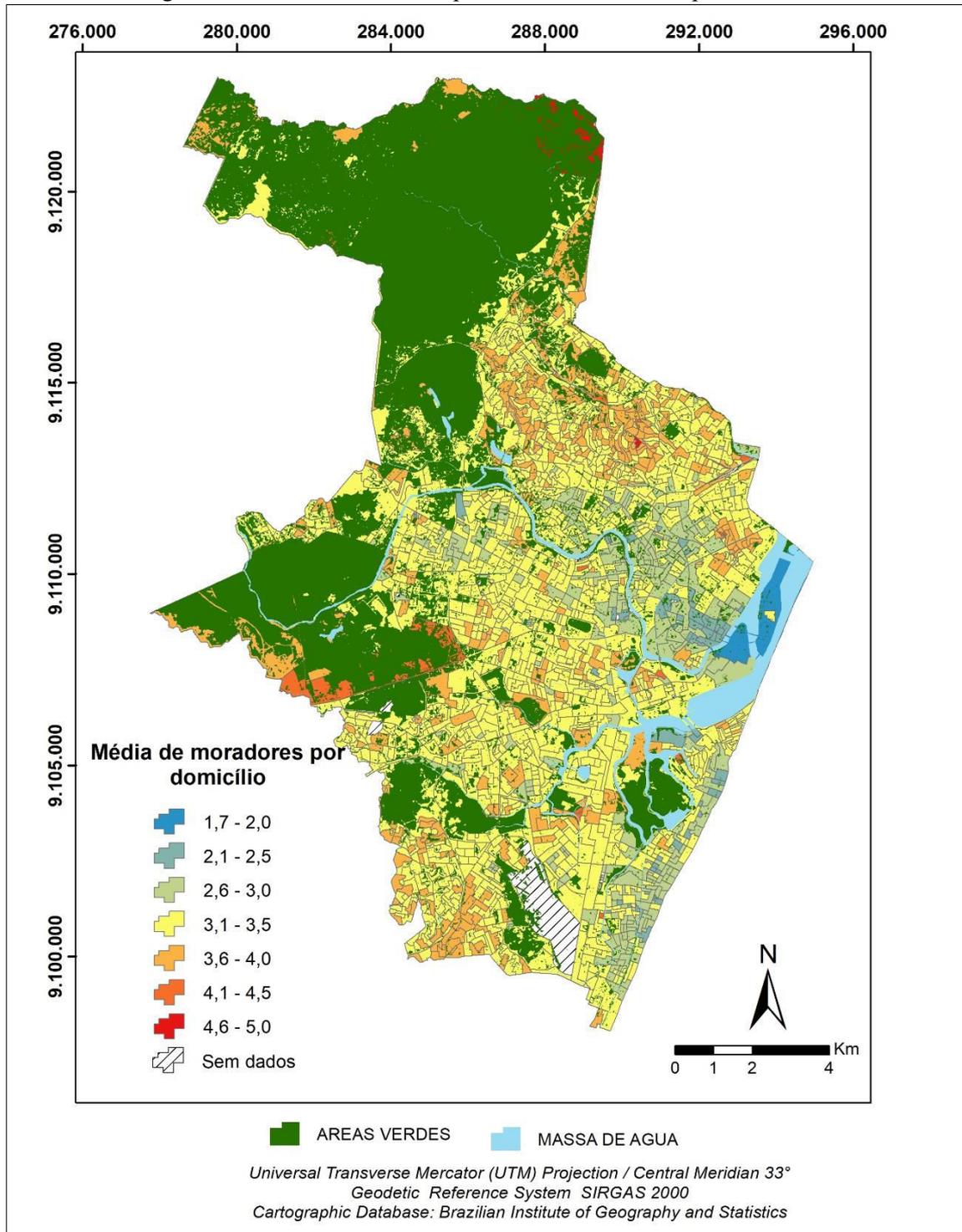
As áreas que apresentam as menores rendas, estão associadas a aglomerações subnormais. Contudo, existem duas regiões onde existe uma maior homogeneidade dos valores de baixa renda: a *zona norte das colinas* do município, que é caracterizada pelo alto número de aglomerações subnormais em seus morros, e a região sudoeste do município, onde também predominam as áreas mais elevadas, em termos de altitude, do município.

Figura 2- Distribuição da renda média mensal do chefe familiar, em Recife.



No que se refere a média de moradores por município (Figura 3), as variações espaciais tendem a ser semelhantes com a distribuição espacial da renda. Em média, os setores que apresentaram as maiores rendas médias mensais, foram as que apresentaram o menor número de moradores por domicílio, até 2,5 moradores por residência, aproximadamente.

Figura 3 - Média de moradores por domicílio no município de Recife.



As áreas onde apresentaram os maiores números de pessoas por domicílio, foram as áreas que apresentaram as menores rendas mensais. Considerando as áreas com precária infraestrutura urbana, os valores variaram de 3,6 a 5,0 pessoas por residência. É importante considerar que essa variável pode estar proporcionalmente relacionada a quantidade de pessoas

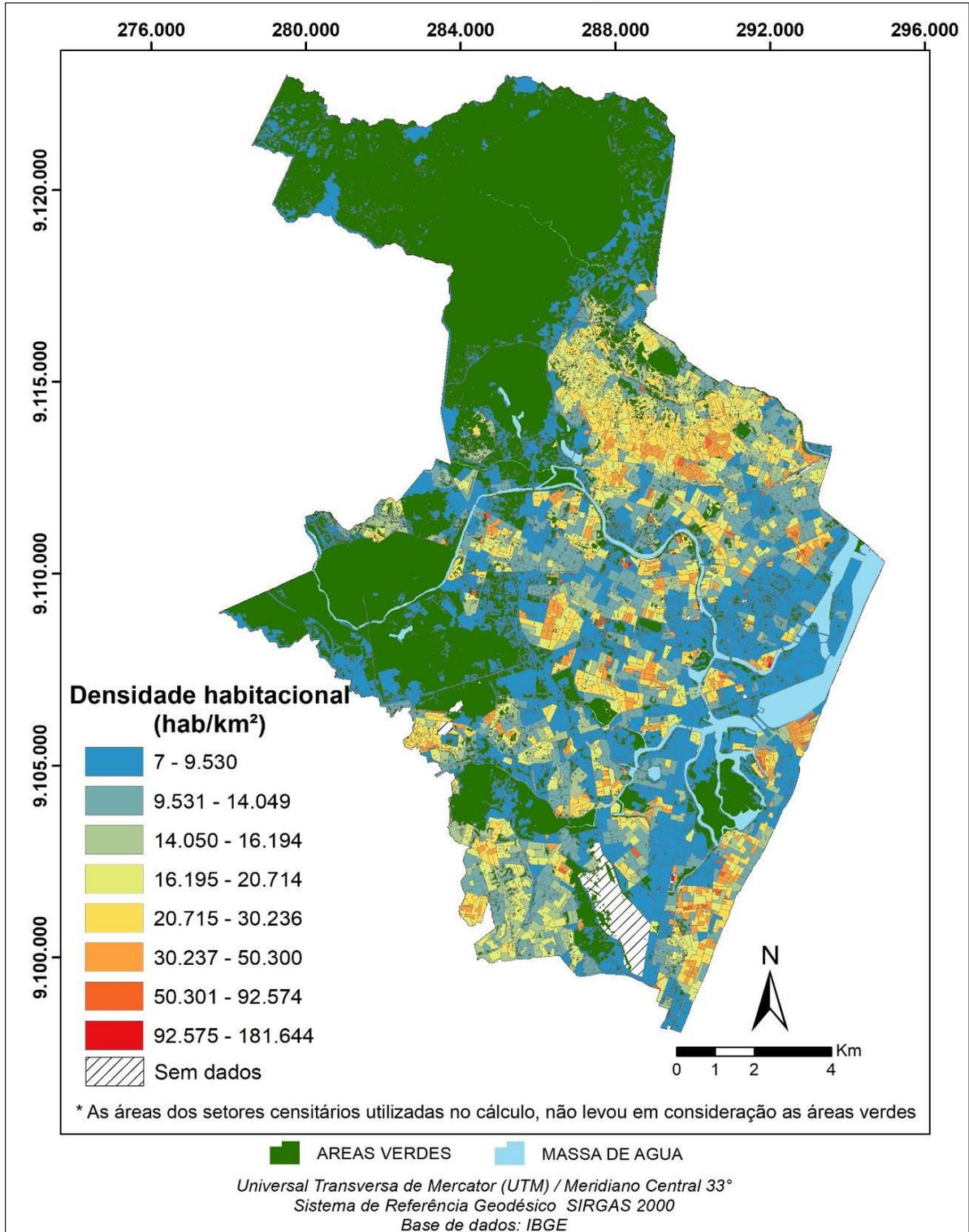
infectadas por doenças com alto índice de contaminação. Nesse sentido, é importante considerar essa variável nos modelos, uma vez que o contato de maior frequência de uma pessoa infectada, na maioria dos casos, encontra-se dentro de seu domicílio.

A densidade populacional em Recife (Figura 4) apresenta uma heterogeneidade que reflete em seus valores por setor censitário: de 7 a 181.644 habitantes por km². Alguns setores mostram semelhanças mesmo com características de infraestrutura e econômicas diferentes. Alguns setores caracterizados pela grande concentração de aglomerações subnormais, com os menores níveis de renda encontrados no município, apresentam valores semelhantes de densidade habitacional com as áreas mais valorizadas de Recife. Podemos exemplificar ao analisar o litoral recifense, onde alguns setores do bairro de Boa Viagem – um dos bairros mais valorizados de Recife possuem os mesmos valores de densidade habitacional dos setores localizados no bairro de Brasília Teimosa, que é caracterizado pelos baixos níveis de qualidade de vida.

Correlacionando a distribuição espacial da média de moradores por domicílio e da densidade habitacional, percebemos que a primeira variável possui uma similaridade maior com a distribuição espacial de renda. Tal fato, evidencia que os processos de ocupação do espaço, possam resultar em similaridades na densidade habitacional, contudo, os mecanismos de ocupação dão-se de maneira diferenciada.

Nos setores censitários onde predominam as áreas mais valorizadas com alta densidade populacional, o mecanismo que promove essa alta procura seria a própria especulação imobiliária, onde a valorização esteja condicionada a condições culturais, existência de melhores serviços e infraestrutura, sendo caracterizada por serem áreas de intensa verticalização. Por outro lado, setores censitários onde predominam uma precária infraestrutura urbana associada a alta densidade habitacional, refletem a ausência de condições financeiras dessas populações em habitarem espaços que ofereçam melhor qualidade de vida e, conseqüentemente mais caros.

Por isso, o que poderia aparentar semelhanças entre a média de moradores por domicílio com a densidade populacional, evidencia que a primeira variável possa destacar uma maior correlação com a disseminação de doenças, seja pelo fator de proximidade ser maior, seja por essa condicionante estar relacionada à renda e, conseqüentemente com melhores ou más condições de vida.

Figura 4 - Densidade habitacional no município de Recife (hab/km²).

Outro aspecto a ser considerado é que áreas com baixos níveis de qualidade de vida apresentaram a mesma densidade populacional que os bairros de maior infraestrutura e poder aquisitivo, mesmo aquelas compostas predominantemente por residências unifamiliares

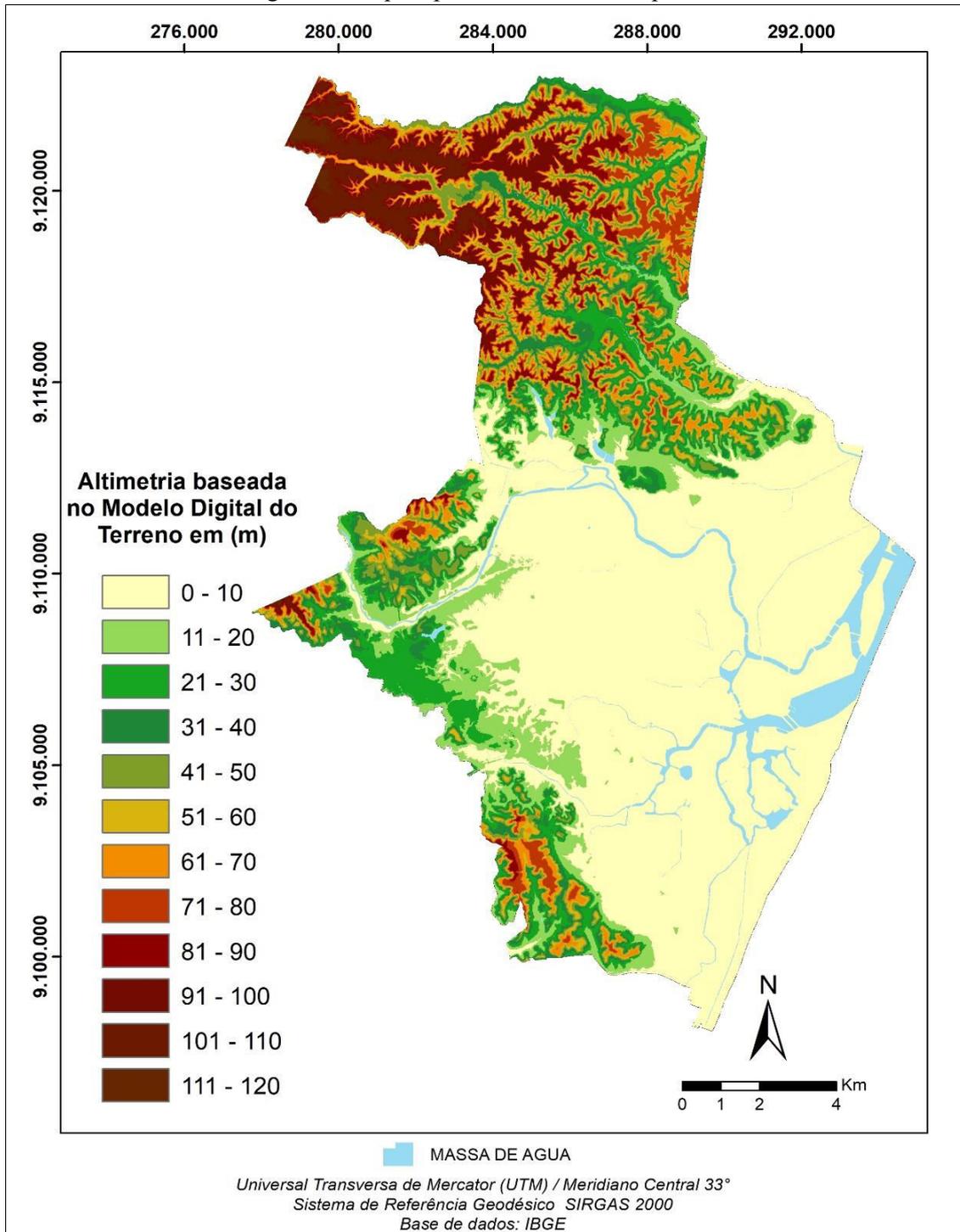
(casas). Contudo, é importante considerar a densidade populacional em análises da distribuição espacial de doenças, visto que essa condicionante pode refletir em taxas de pessoas infectadas em áreas mais nobres.

Alguns produtos sínteses da associação de variáveis socioeconômicas e infraestruturais merecem destaques para o planejamento urbano de Recife. Podemos exemplificar o mapeamento proposto pelo Zoneamento Especial de Interesse Social (ZEIS), na década de 80 do século XX, que são caracterizadas por serem áreas de assentamento habitacional de população de baixa renda, alta densidade de residências, ter carência ou ausência de serviços de infraestrutura básica. Destarte, por serem criadas pelo Poder Público, as ZEIS além de ter o reconhecimento dos territórios conquistados pelos moradores, esse zoneamento tornou-se uma forma de viabilizar intervenções com padrões urbanísticos específicos mais próximos às condições dos assentamentos e sendo um parâmetro regulador contra a pressão da especulação imobiliária (MORAES, 2019).

4.1.2 Aspectos físicos e estruturais da área estudada

Grande parte do município de Recife é constituída pela planície flúvio-marinha, formadas a partir da intensa erosão fluvial e marinha durante as transgressões e regressões marinhas. Porém, as áreas mais elevadas do município, podem chegar até aproximadamente 120 metros acima do mar, que são – em sua maioria - constituídas por tabuleiros costeiros dissecados e áreas de embasamento cristalino soerguidas que circundam toda a planície (Figura 5).

Figura 5 - Mapa hipsométrico do município de Recife.



Com essas características geomorfológicas, o processo de ocupação do município deu-se primeiramente nas áreas mais planas da cidade. Sob forte influência holandesa, os primeiros núcleos urbanos foram se formando próximo aos córregos, riachos e em toda a planície fluvial formada pelos afluentes do Capibaribe. Esses núcleos, ao longo dos séculos foram se

adensando, de tal forma que as populações menos favorecidas foram forçadas a se habitarem nas áreas mais elevadas do município, caracterizadas por serem menos valorizadas devido ao seu alto custo para construção ou pelos riscos inerentes a deslizamento de terras (CASTRO, 2013).

É importante considerar que as populações menos abastadas, antes habitavam as áreas de manguezais e regiões de inundação que margeavam os cursos d'água. Contudo, a partir dos anos 1930, os morros começaram a ser habitados até a metade da década de 1970, onde as aglomerações subnormais aumentaram, a partir da valorização das áreas mais planas da cidade, desde então (SANDHOLZ, 2017). Todo esse processo histórico, reflete até hoje os padrões de habitações que são encontrados nos morros recifenses, onde as más condições de vida, ainda permanecem.

Quando analisado o mapa de declividade (Figura 6), cruzando-o com as variáveis socioeconômicas, percebe-se que quando não há presença de vegetação nas áreas de maior declive, geralmente são habitadas pelas populações com baixo nível de renda, alta densidade populacional e elevado número de moradores por domicílio. É importante ressaltar que as variações de declividade dentro do município recifense variam do plano, até 82°. Essas áreas de alta declividade, em sua maioria, estão situadas ao norte, oeste e sudoeste do município.

Em relação a porcentagem de áreas não construídas (sem considerar as áreas verdes) por setor censitário (Figura 7), podemos identificar onde as áreas de maior densidade de habitações concentram-se espacialmente. De maneira implícita, essa variável pode inferir que uma menor densidade de habitações está condicionada a vias de maior extensão, existência de praças, maiores distâncias entre as residências e, conseqüentemente uma melhor infraestrutura. Mesmo considerando uma densidade considerável das habitações em setores mais nobres da cidade devido a especulação imobiliária, percebe-se que os valores encontrados em regiões onde a renda média mensal do chefe de família é baixa são encontradas áreas de densidade de edificações com valores superiores.

Figura 6 - Mapa de declividade no município de Recife.

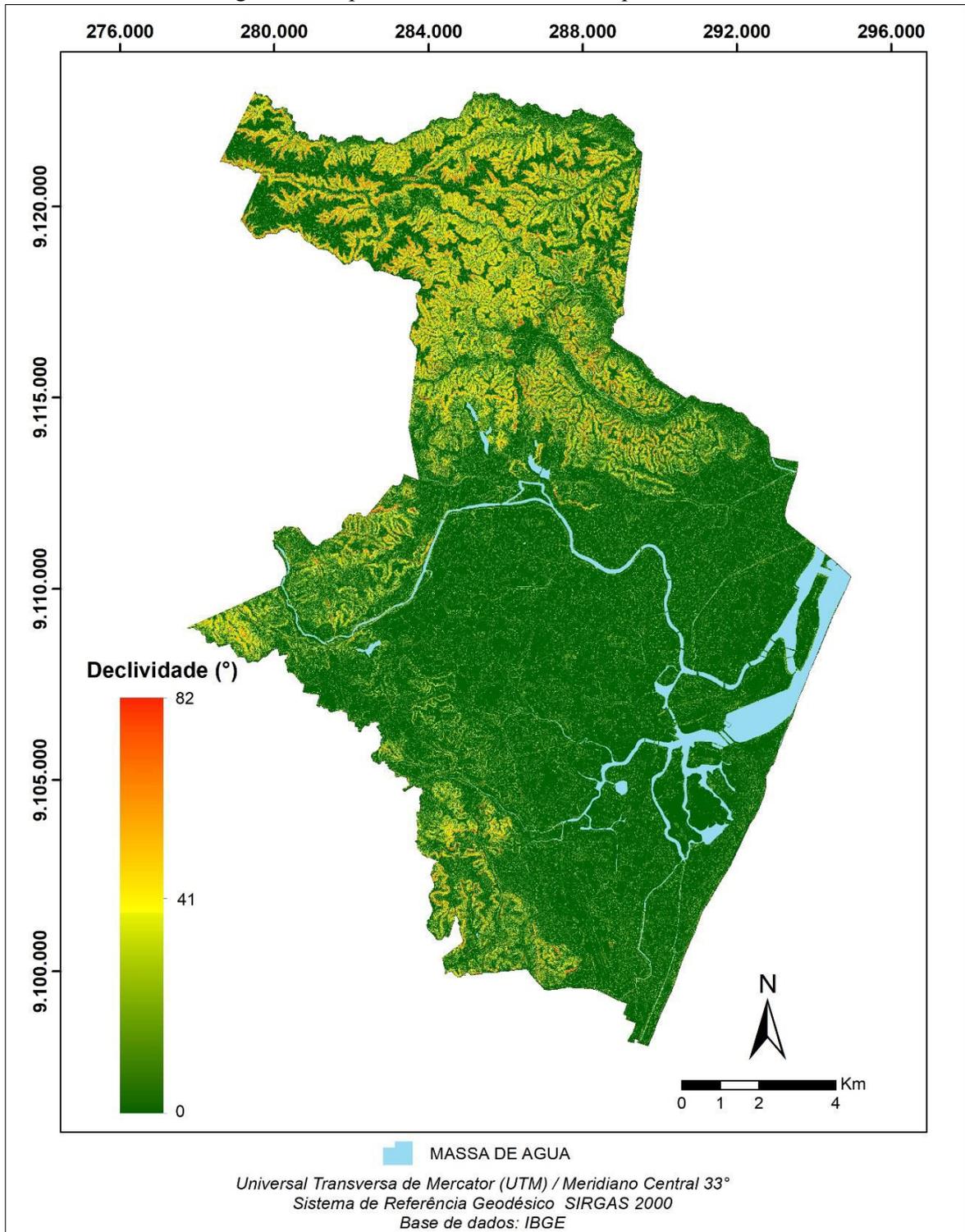
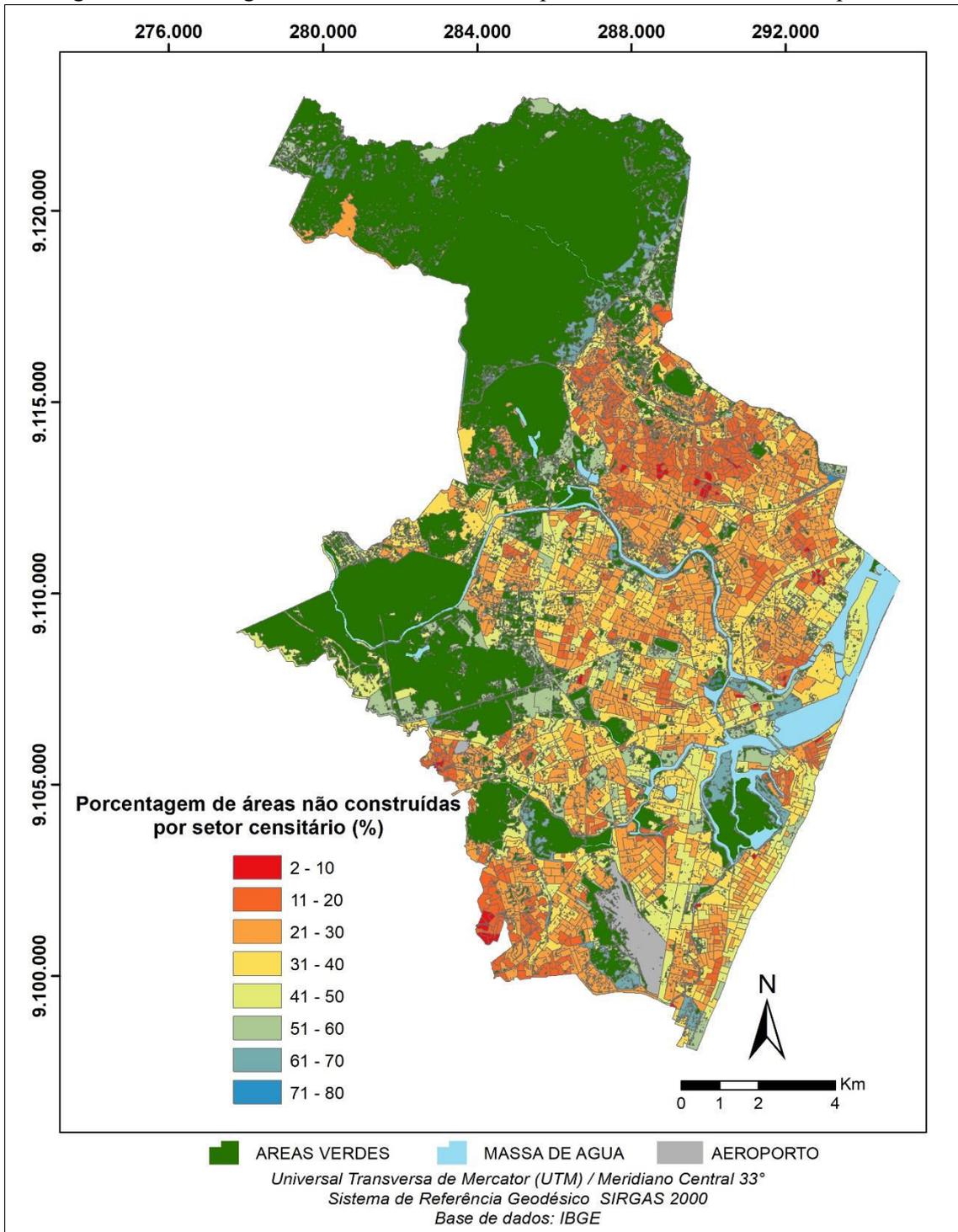


Figura 7 - Porcentagem de áreas não construídas por setor censitário no município de Recife.



A distribuição espacial da precipitação em Recife apresenta algumas particularidades (Figura 8). De modo geral, a região litorânea apresenta menor quantidade de precipitação anual que o oeste do município. Considerando a média de precipitação entre os anos de 2015 a 2017, a quantidade precipitada variou de 1.500 mm a 2.000 mm, com um gradiente espacial de aproximadamente 500 mm dentro do município.

É importante considerar que alguns fatores podem estar condicionados a distribuição espacial das chuvas em Recife. Anjos *et al.* (2020) e Nascimento *et al.* (2018) verificaram que a direção dos ventos associada a rugosidade do relevo do município de Recife pode estar relacionada à distribuição espacial da precipitação. Partindo desse princípio, as áreas cujas vertentes estejam voltadas para sudeste, recebem maior umidade dos ventos alísios de sudeste, que é predominante durante todo o ano. Os autores discutem que é importante considerar que a turbulência pode provocar uma convecção forçada no sopé das encostas, favorecendo uma maior quantidade precipitada nessa região, se comparada com áreas de maior altitude do município.

Ainda segundo os autores, as áreas dos morros recifenses apresentam menor quantidade precipitada que as áreas de planície próximas ao sopé das encostas, devido a sua altitude ser mais elevada e, conseqüentemente ter menor atividade convectiva. Por isso, é válido considerar que, mesmo não existindo uma relação de causa-efeito, possa existir uma relação entre condicionantes socioeconômicas e pluviométricas em Recife.

Outra condicionante abordada na presente tese são as áreas verdes

Considerando os bairros e suas respectivas porcentagens de áreas verdes (Figura 9) podemos evidenciar que alguns que estão localizados ao norte e oeste do município apresentam as maiores porcentagens de áreas verdes em relação a área total do município. Com aproximadamente 95% de áreas verdes, Pau Ferro apresenta a maior porcentagem de áreas verdes em Recife, seguido da Guabiraba (91%), Dois Irmãos (86%), Várzea (67%) e Apipucos (59%).

Em contrapartida, bairros como Brasília Teimosa, Mangueira, Morro da Conceição, Mustardinha e Alto José do Pinho apresentam menos de 2% de áreas verdes em relação a área total do bairro. Considerando os aspectos socioeconômicos desses bairros é evidente que são caracterizados por uma alta densidade populacional e rendas médias mensais do chefe familiar mais baixos.

Figura 8 - Média da precipitação anual entre os anos 2015 e 2017 no município de Recife.

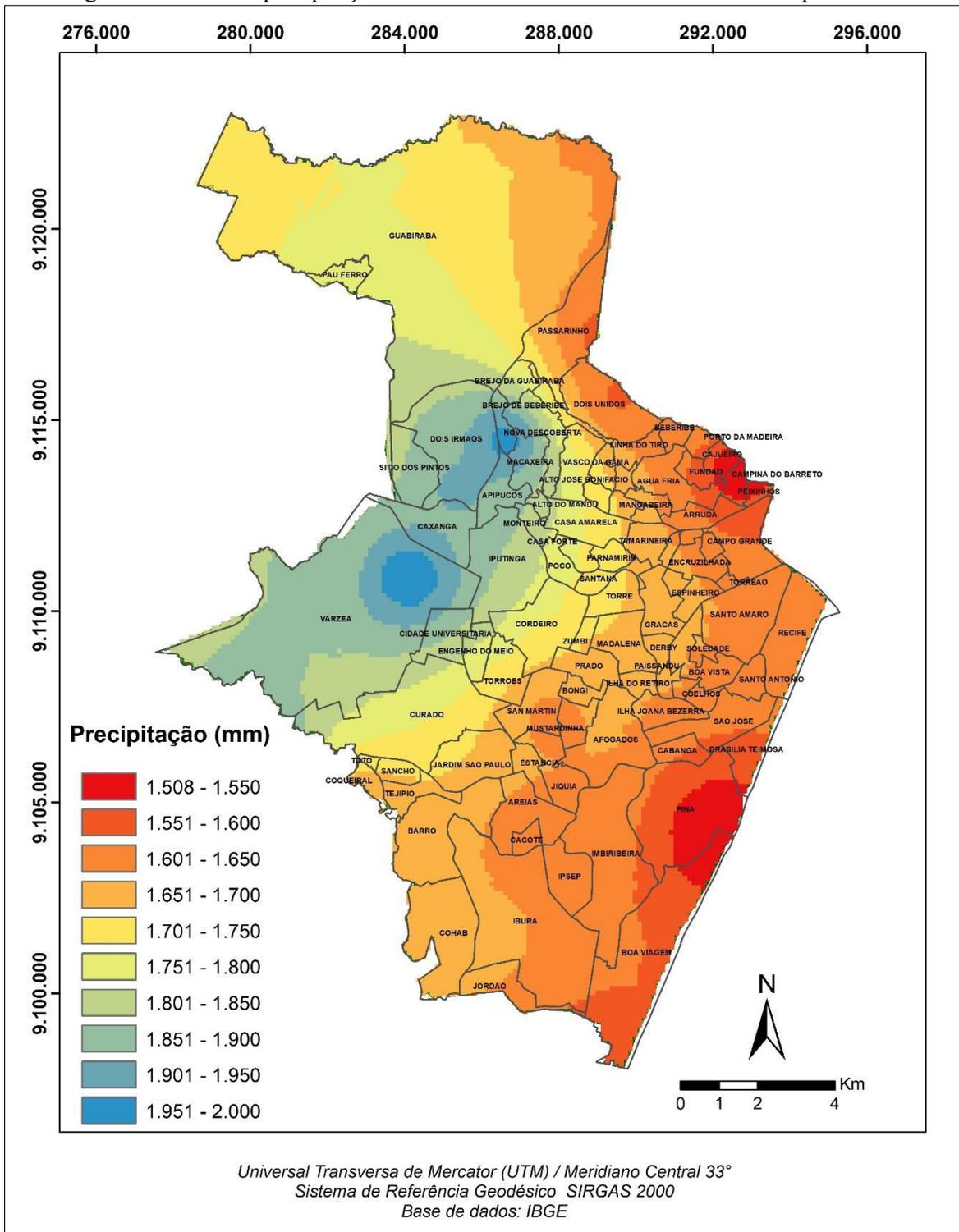
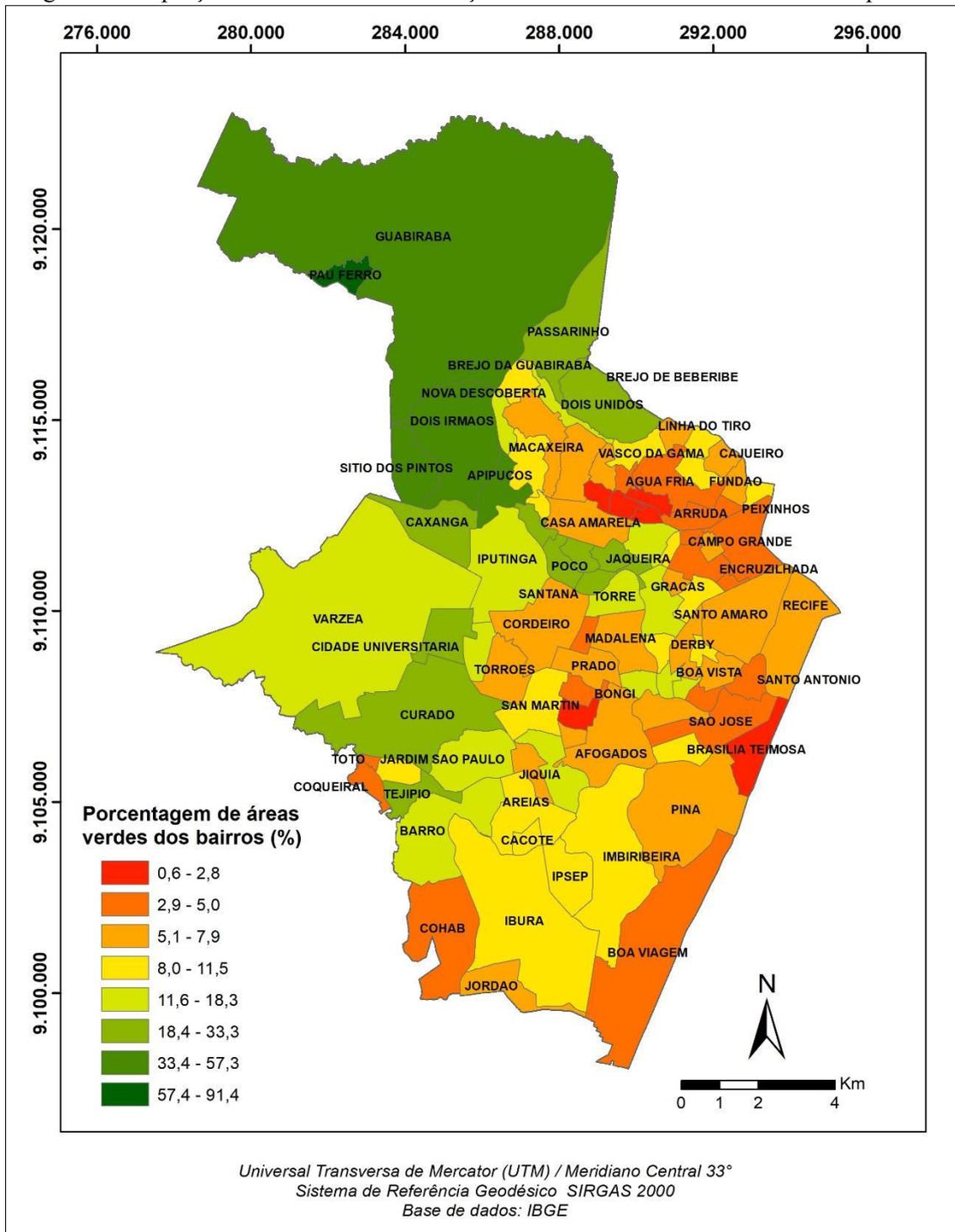


Figura 9 - Proporção de áreas verdes em relação às áreas totais dos bairros no município de Recife.



Quando é considerada a análise da paisagem, percebe-se que as áreas mais nobres de Recife apresentam alta densidade de edificações associadas a uma intensa verticalização. É perceptível essas características quando observamos na Figura 10 (a) o bairro de Boa Viagem, localizado na zona sul do município. Tais elementos paisagísticos nas áreas nobres da cidade apresentam uma boa infraestrutura urbana, associada a existência de vias pavimentadas e mais largas, iluminação pública, saneamento básico e presença de parques e praças. Outro ponto a ser destacado é que a própria verticalização evidencia uma intensa valorização dessas áreas, resultando em densidades habitacionais mais elevadas.

Caracterizada também pela alta densidade de edificações, porém com precária infraestrutura urbana, as áreas de maior vulnerabilidade social concentram-se nas margens alagáveis dos cursos d'água e nas áreas de morros em Recife, como observa-se na Figura 10 (b). Essas áreas apresentam vielas com acesso limitado a determinados veículos, ausência de saneamento básico e áreas de lazer, como parques e praças. É importante ressaltar que essas áreas ao longo dos anos foram negligenciadas pela população abastada da cidade, devido à dificuldade de construção que essas áreas ofereciam, sejam por apresentarem terrenos mais íngremes ou apresentarem suscetibilidade a alagamentos. Por isso, essas áreas foram ocupadas pela parcela mais pobre da população recifense, acarretando numa vulnerabilidade social crescente ao longo dos anos, relacionada às limitações impostas pelo terreno e habitações improvisadas devido à falta de recurso financeiro dessas famílias.

Tais aspectos, merecem destaque quando analisada a paisagem de maneira conjunta, sendo perceptível as discrepâncias sociais na paisagem recifense, onde a planície - caracterizada pela sua melhor infraestrutura urbana - contrastando com as áreas de maior vulnerabilidade social, sendo nesse caso representado pelos morros do Recife, na Figura 10 (c).

Figura 10 - Bairro de Boa Viagem (a), zona norte das colinas em Recife (b) e zona norte da planície e das colinas de Recife (c)

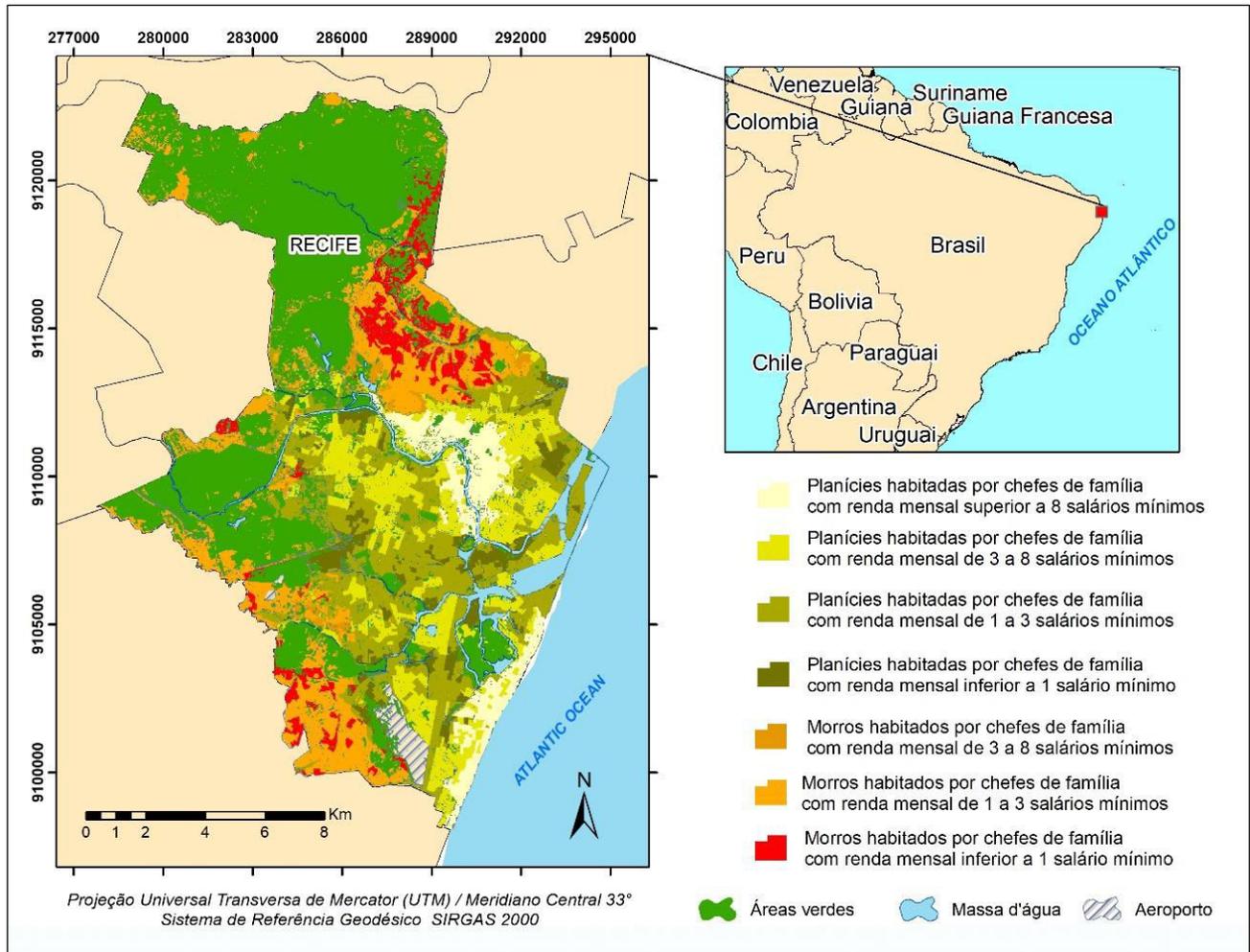


Foto: Rafael Anjos (2021)

Quando cruzadas as informações geomorfológicas com a renda média mensal do chefe familiar percebemos algumas particularidades (Figura 11). A classe que apresenta maior área em relação ao município são as planícies que habitam chefes de família com renda entre 1 e 3 salários mínimos, com cerca de 34 km². As áreas de morro com renda média mensal entre 1 e 3 salários mínimos é a segunda classe com maior extensão territorial, com aproximadamente 26 km². As planícies que são habitadas por chefes de família com renda média mensal entre 3 e 8, e acima de 8 salários mínimos possuem a 3^o e 4^o maior extensão territorial do município com 25km² e 11 km², respectivamente. Com renda média mensal de até 1 salário mínimo, as planícies e as áreas de morro, ocupam a 5^o e 6^o maior extensão territorial com 8 e 9 km², respectivamente. As áreas de morros habitadas por chefes de família com salário mínimo de 3 a 8, ocupa o menor território, com cerca de 2 km². Apesar das pequenas dimensões territoriais das áreas com a população de baixa renda, é válido ressaltar que são caracterizadas por apresentarem alta densidade populacional.

A partir da análise desses aspectos, fica evidente que os aspectos físicos, sociais e econômicos estão relacionados entre si no município de Recife. Destarte, é válido considerar que as relações de algumas dessas variáveis demonstram que existiram processos históricos que condicionaram a combinação dessas variáveis, por outro lado, alguns fatores devem ser levados em consideração o fato de não existir uma relação clara de causa e efeito, como a distribuição espacial da precipitação.

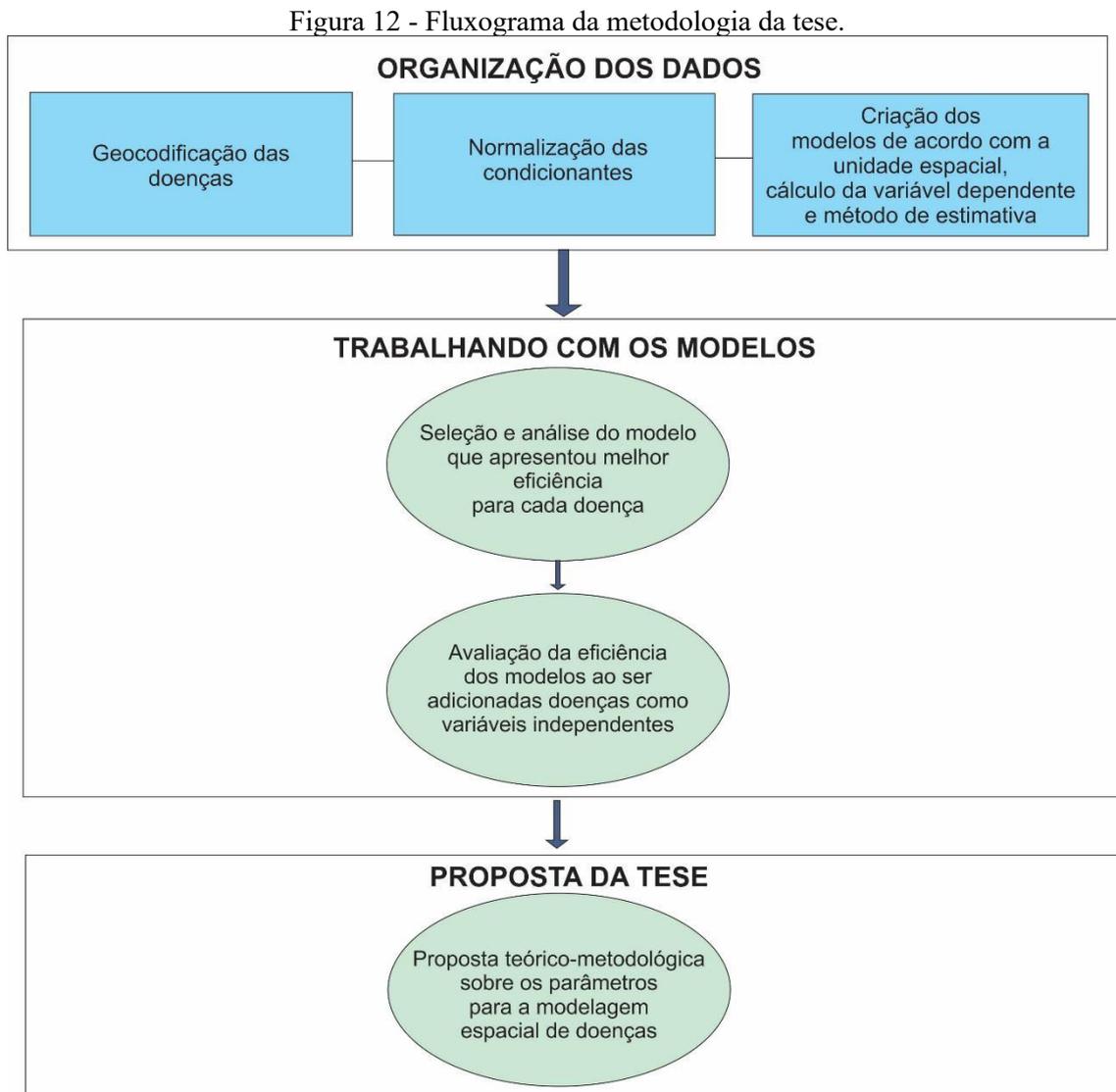
Figura 11 - Unidades geomorfológicas orientadas de acordo com o perfil da renda do chefe familiar no município de Recife.



4.2 METODOLOGIA

A metodologia da presente tese baseou-se em três partes principais, conforme o fluxograma na

Figura 12. Foram elas: organização dos dados, análise dos modelos e a proposta teórico-metodológica sobre a modelagem espacial de doenças, de acordo com os resultados encontrados.



4.2.1 Coleta de dados e distribuição espacial das doenças

A escolha de cada doença teve como princípio a relação entre sua distribuição espacial com aspectos socioambientais, de acordo com a bibliografia encontrada.

Diante disso, foram considerados 3 grupos de doenças a serem analisados: 1- (Zika e Chikungunya); 2- (Leptospirose); 3- (Tuberculose). Todos os dados de casos notificados foram disponibilizados diretamente pela Secretaria de Saúde do Recife entre os anos de 2015 e 2017. Para cada conjunto da doença, os anos foram agrupados e cada notificação geocodificada, de maneira que os casos em penitenciárias e sem endereço foram desconsiderados na presente pesquisa. Todos os casos foram geocodificados utilizando o software Google Earth, disponibilizado gratuitamente.

Ao todo, foram 5.755 casos registrados de Zika e Chikungunya desconsiderando casos sem endereço e notificações em penitenciárias. Desse total, 4.861 casos foram geocodificados com 15% dos casos não geocodificados por inconsistência do endereço.

Em relação aos casos de Leptospirose, 306 casos (sem registros em penitenciárias e ausência de endereço) foram submetidos a geocodificação. Desse total, 282 foram geocodificados, com aproximadamente 8% dos casos submetidos não geocodificados devido a inconsistência do endereço.

Já em relação a Tuberculose, foram submetidos 4.781 casos (sem registros em penitenciárias e ausência de endereço). Desse total, 4.265 casos foram geocodificados, com 11% dos casos submetidos não geocodificados devido a inconsistência do endereço.

4.2.2 Coleta e definição dos condicionantes espaciais a serem relacionados com as doenças

Com o intuito de investigar como se relacionam algumas condicionantes geográficas para cada doença descrita, foram levantadas quatro categorias de variáveis a serem utilizadas nos modelos espaciais (Figura 13).

Desse modo, justifica-se que as condicionantes geográficas diferenciassem entre si, de acordo com suas respectivas categorias, para que buscasse entender qual categoria poderia apresentar uma maior relação com a distribuição espacial para cada doença analisada. Foram selecionadas as condicionantes geográficas que apresentassem algum tipo de relação com todas as doenças e que já tivessem sido levantadas por estudos anteriores.

Ao todo, foram sete condicionantes geográficas utilizadas como dados de entrada para cada modelo espacial. A disponibilidade dos dados foi um dos critérios considerados.

Figura 13 - Variáveis utilizadas no modelo, de acordo com sua categoria, e estudos que relacionam as respectivas condicionantes para cada doença.

Variáveis e suas categorias		Estudos relacionados para cada doença		
		Arboviroses	Leptospirose	Tuberculose
VARIÁVEIS SOCIOECONÔMICAS	Moradores por domicílio	ORTIZ et al., 2017	SOARES et al., 2010	LOBO et al., 2016
	Renda do chefe familiar	ALMEIDA et al., 2007	OLIVEIRA, 2009	MAGALHÃES, 2014
	Densidade populacional	QUINTERO et al., 2014	BARCELLOS e SABROZA, 2001	SILVA, 2016
VARIÁVEIS FÍSICAS	Áreas verdes	FEITOSA, 2016	GUTIÉRREZ et al., 2019	LIU et al., 2019
	Declividade	FEITOSA, 2016	FERREIRA e FERREIRA, 2016	MEDEIROS, 2013
VARIÁVEIS DE INFRAESTRUTURA	Áreas não construídas	LESSER e KITRON, 2016	PINTO FERREIRA et al., (2019)	CLARCK et al., 2002
VARIÁVEIS CLIMÁTICAS	Precipitação	FULLER et al., 2017	MOHAMMADINIA et al., 2019	KUDDUS et al., 2019

Elaborado pelo autor (2021)

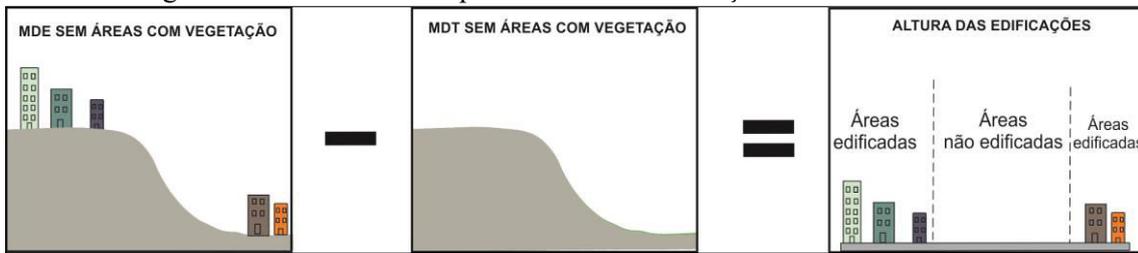
As variáveis socioeconômicas foram obtidas a partir do banco de dados do Instituto Brasileiro de Geografia e Estatística (IBGE), de acordo com os setores censitários existentes no município de Recife.

A variável de áreas verdes, categorizadas como variável física, foi produzida a partir da análise do *Normalized Difference Vegetation Index* (NDVI) do satélite Sentinel, com resolução espacial de 10 metros, associada a fotointerpretação de imagens de satélite disponibilizadas pelo Google Earth. A imagem para produção do NDVI, foi do dia 05 de setembro de 2016, por apresentar baixa nebulosidade.

Os dados de precipitação do município de Recife, entre os anos de 2015 e 2017, foram do Centro Nacional de Monitoramento e Alertas de Desastres Naturais (CEMADEN).

As variáveis declividade e porcentagem de áreas não construídas por unidade espacial, foram obtidas a partir das imagens do *Light Detection and Ranging* (LIDAR) com resolução espacial de 1 metro. Para extrair a informação de áreas não construídas, foi necessária a remoção das áreas verdes e massa d'água do modelo digital de elevação (MDE), para que ficassem apenas a altimetria das áreas urbanizadas de Recife. Após esse procedimento, foram consideradas as áreas edificadas que apresentassem uma elevação superior a 2 metros, conforme a Figura 14.

Figura 14 - Procedimentos para extrair as informações de áreas não construídas.



Todas as variáveis foram calculadas para que elas fossem utilizadas nos modelos, por meio de índices. A equação para normalizar os valores das variáveis, é mostrada a seguir:

$$\text{Índice das variáveis} = (V_i - V_a) / (V_{std}) \quad (8)$$

Onde V_i é a média do valor da variável em uma determinada unidade espacial i , V_a é a média geral da variável para todo o município de Recife e V_{std} é o desvio padrão da variável em todo o território recifense.

A utilização dessa variável justifica-se dado ao fato que podem ser identificadas as áreas com alta densidade de habitações e, conseqüentemente menor taxa de áreas construídas. Outro aspecto é que áreas com altas taxas de áreas não construídas podem ser não somente áreas com presença de parques, praças e largas avenidas, mas locais com presença de terrenos baldios, por exemplo.

4.2.3 Análise e definição do desempenho do modelo de estimativa de distribuição das doenças analisadas

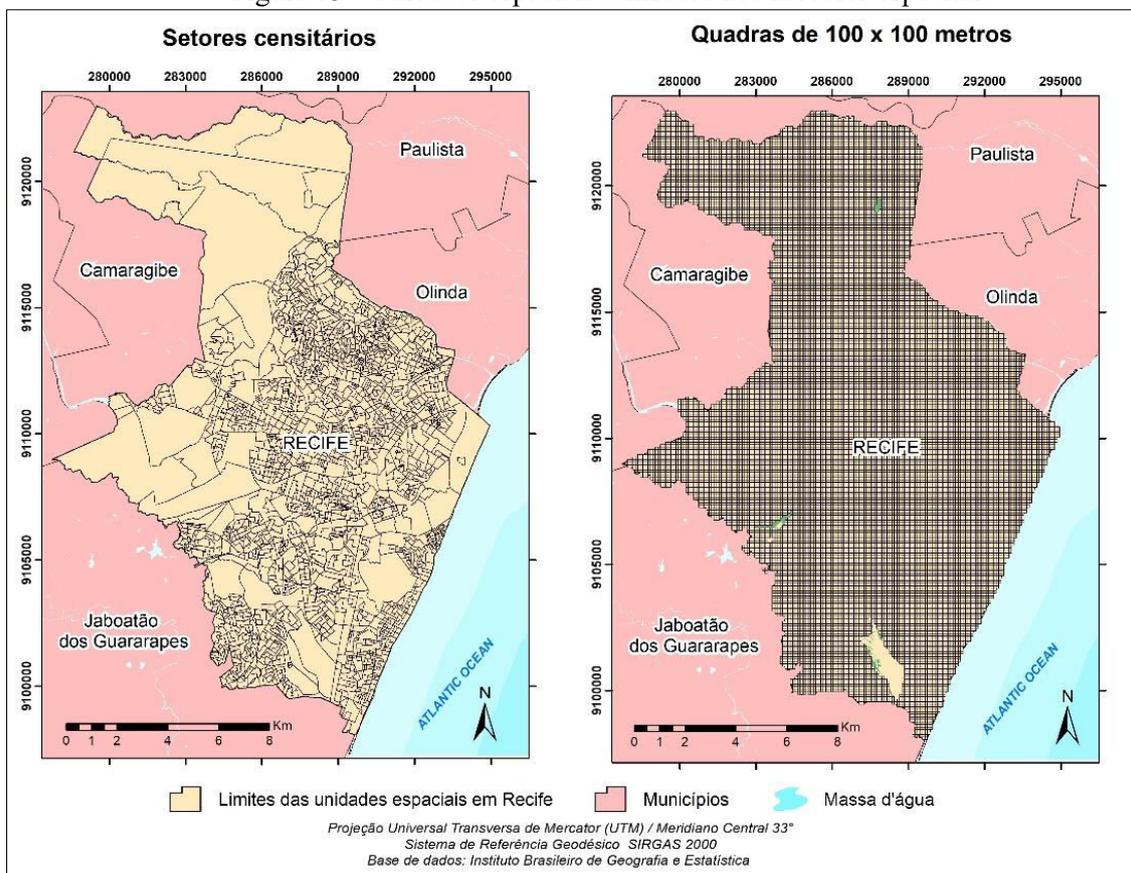
Todos os procedimentos a serem descritos a seguir, foram elaborados dentro da plataforma do software ArcGis 10.2, cuja licença foi disponibilizada gratuitamente pelo Laboratório de Informática, da Universidade de Évora – Portugal.

Assim, foram criados 6 modelos para estimar a distribuição espacial das doenças descritas. Esses modelos diferenciavam-se de acordo com a combinação de alguns fatores, dentre eles:

1- **Unidade espacial** (Setores censitários, Setores Censitário sem áreas verdes e quadrículas de 100x100 metros). A lógica da utilização desses diferentes tipos de unidades foi para analisar se uma maior desagregação das unidades espaciais poderia interferir numa maior eficiência na estimativa das doenças (Figura 15). Esse fato está atrelado porque algumas variáveis oriundas de dados matriciais (declividade, áreas verdes, precipitação e porcentagem de áreas não construídas) poderiam oferecer um maior detalhamento da correlação espacial entre a distribuição espacial da doença e essas variáveis, mesmo que algumas condicionantes -

como as oriundas do Censo Demográfico de 2010 – permanecessem estáticas com a desagregação espacial. Um outro fator analisado é se a utilização de setores sem áreas verdes resultaria num cálculo de densidade habitacional por setores censitários mais próximo da realidade, supondo assim, uma melhoria no resultado final da estimativa do modelo. Ao todo, foram 1831 setores censitários utilizados para os modelos. Foram retirados das unidades espaciais as áreas que não apresentaram as informações devido à ausência ou baixa quantidade de habitantes, além dos setores com existência de penitenciárias.

Figura 15 - Unidades espaciais utilizadas nos modelos espaciais.



2- Cálculo para a distribuição espacial da variável dependente (número de casos por unidade espacial ou média da densidade da doença, utilizando o método de Kernel e seus respectivos raios de atuação). A análise do cálculo da variável dependente está relacionada como os métodos de estimativa irão trabalhar a distribuição espacial da doença, correlacionando-as com as condicionantes analisadas.

A proposta de utilizar número de casos por uma determinada unidade espacial permite identificar com exatidão a ocorrência dos casos, embora as variações bruscas de valores entre

as unidades espaciais possam resultar numa relação não muito clara com as variações espaciais de algumas condicionantes utilizadas no modelo. Por outro lado, utilizar a distribuição espacial da doença utilizando a média da densidade de Kernel por unidade espacial possa mostrar uma suavidade maior nas variações espaciais entre as unidades espaciais, facilitando numa interpretação mais clara com algumas condicionantes. É importante ressaltar que a densidade de Kernel é baseada em dois parâmetros: o raio de busca adotado e as unidades de medidas, como km^2 ou m^2 . Exemplificando: se foi adotado um raio de busca de 500 metros e foram encontradas 3 notificações dentro desse raio, abrangendo essa proporção para um raio de 1000 metros (1 km), teríamos 6 casos por km^2 registrados dentro da área de atuação daquele raio. Na presente pesquisa foram utilizadas 3 distâncias de raios: 500, 700 e 900 metros.

3- Método de estimativa da doença (OLS ou GWR). Embora ambos os métodos já estivessem integrados em cada modelo, eles diferenciavam-se entre si a partir do *bandwidth* utilizado no método GWR. Por padronização, os *bandwidths* adotados foram selecionados de acordo com o raio de busca do método de densidade do Kernel escolhido para cada modelo estudado. Portanto, embora cada modelo apresentasse os dois métodos descritos, eles diferenciavam-se de acordo com os *bandwidths* adotados no GWR (500, 700 e 900 metros).

O resultado final da combinação desses fatores para cada modelo, encontram-se na Figura 16.

Figura 16 -Organização dos modelos de acordo com o método da estimativa, bandwidth do GWR, unidade espacial e tipo de cálculo da variável dependente.

Modelo 1		Variável dependente
Método	Bandwidth (m)	
OLS		casos por setor censitário
GWR	500	casos por setor censitário
Modelo 2		Variável dependente
Método	Bandwidth (m)	
OLS		casos por setor censitário - sem áreas verdes
GWR	500	casos por setor censitário - sem áreas verdes
Modelo 3		Variável dependente
Método	Bandwidth (m)	
OLS		casos por quadrícula de 100 m
GWR	500	casos por quadrícula de 100 m
Modelo 4		Variável dependente
Método	Bandwidth (m)	
OLS		densidade de Kernel com raio de 500m por quadrícula de 100m
GWR	500	densidade de Kernel com raio de 500m por quadrícula de 100m
Modelo 5		Variável dependente
Método	Bandwidth (m)	
OLS		densidade de Kernel com raio de 700m por quadrícula de 100m
GWR	700	densidade de Kernel com raio de 700m por quadrícula de 100m
Modelo 6		Variável dependente
Método	Bandwidth (m)	
OLS		densidade de Kernel com raio de 900m por quadrícula de 100m
GWR	900	densidade de Kernel com raio de 900m por quadrícula de 100m

Para comparação da eficiência de cada modelo foi considerado o coeficiente de correlação (R^2 ajustado), enquanto para a comparação da qualidade das estimativas entre os métodos dos modelos (GWR e OLS), foram adotadas as somas dos residuais e o *Corrected Akaike Information Criterion* (AICC).

O R^2 ajustado tem seus valores que variam de 0 a 1, em que os altos valores evidenciam uma maior eficiência do modelo. O ajuste para o R é aplicado porque quando uma variável explanatória é adicionada ao modelo, o denominador não é alterado, mas o numerador sim. Cálculos para o R^2 ajustado normalizam o numerador e o denominador pelo grau de liberdade, resolvendo o problema da alteração.

A soma dos residuais é a soma absoluta dos resíduos ao quadrado, derivado pela diferença entre os valores estimados e observados. O *Akaike Information Criterion Correction* é uma medida da performance do método de regressão dos modelos.

Após a avaliação e escolha dos modelos que apresentaram a melhor eficiência em suas estimativas, foram avaliados os coeficientes do método de regressão global (OLS) e local

(GWR) para que analisassem as relações entre as variáveis utilizadas e as doenças analisadas. Além disso, foram avaliados os mapas das estimativas de ambos os métodos, comparando-os com os dados observados.

Três parâmetros foram utilizados para analisar as relações globais entre as variáveis utilizadas e as doenças no método OLS: *Robust Probability*, *Variance Inflation Factor* (VIF) e *Koenker Statistics*.

O *Robust Probability* indica se um coeficiente possui significância estatística ($p < 0,01$). Esse indicador, também chamado de p-valor, apresenta-nos se a relação entre uma determinada condicionante e a doença podem ser relevantes, uma vez rejeitada a hipótese nula – que é a não existência de uma relação entre a variável e a doença em questão - (quando o valor é inferior que 0,01) (ESRI, 2020).

O VIF é uma medida que visa analisar a multicolinearidade de determinado conjunto de variáveis. Para cada condicionante utilizada no modelo, o VIF avaliará se as relações entre as condicionantes apresentam similaridades entre si, evitando que variáveis inseridas no modelo sejam redundantes (quando os valores ultrapassam 7,5) (ESRI, 2020). O VIF pode ser descrita na seguinte equação:

$$VIF_i = 1 / (1 - R^2_i) \quad (9)$$

Onde o R^2 representa o coeficiente de determinação não ajustado para a regressão da variável independente i das restantes variáveis independentes.

Já o *Koenker Statistics* é um teste para determinar se a variável independente do modelo se comporta de maneira uniforme no espaço geográfico (ESRI 2020). Esse índice, se for inferior a 0,01, indica que as relações entre a condicionante geográfica e a doença podem ser diretamente ou inversamente proporcional de acordo com a unidade espacial analisada.

4.2.4 Desenvolvimento da proposta teórico metodológica da tese

Com o intuito de englobar os resultados, o último capítulo dessa tese abordou como as variáveis inseridas para estimar cada doença inter-relacionavam-se entre os modelos que apresentaram as melhores eficiências. Partindo desse pressuposto, foram respondidos quais as variáveis inseridas apresentavam correlações espaciais com todas as doenças e como eram as relações (diretamente ou inversamente proporcional).

A partir desta investigação, foram gerados outros modelos para cada doença, pelo método OLS, considerando além das variáveis inseridas, as densidades espaciais das doenças que não faziam parte da variável dependente do modelo. Como exemplo, para o modelo que apresentou a melhor eficiência para estimar a Tuberculose, foram inseridas, além das variáveis independentes já descritas, as densidades espaciais da Leptospirose e Zika/Chikungunya. O intuito desse método foi avaliar se a inserção dessas doenças espacializadas, melhorariam as estimativas dos modelos até então, estudados.

Ao fim de todos os procedimentos no âmbito da avaliação dos modelos espaciais, foi desenvolvida uma proposta teórico-metodológica sobre os parâmetros que merecem destaque no processo de se estimar as doenças espacialmente. Além disso, foi avaliado como as doenças podem estar relacionadas entre si, seguindo uma lógica espacial, e como essas relações podem auxiliar na criação de modelos espaciais.

5 RESULTADOS E DISCUSSÕES

Nesta seção será apresentado os resultados e as discussões da presente tese.

5.1 ZIKA E CHIKUNGUNYA

Ao todo, foram geocodificados 4.861 casos de Zika e Chikungunya em Recife, entre os anos de 2015 e 2017. Considerando os 10 bairros com maior número de casos notificados de Zika e Chikungunya (Tabela 1), sete deles concentram-se na porção norte do município. É importante destacar que é nessa área do município que concentram os maiores índices de densidade habitacional, assim como os menores valores de renda média mensal por chefe de domicílio.

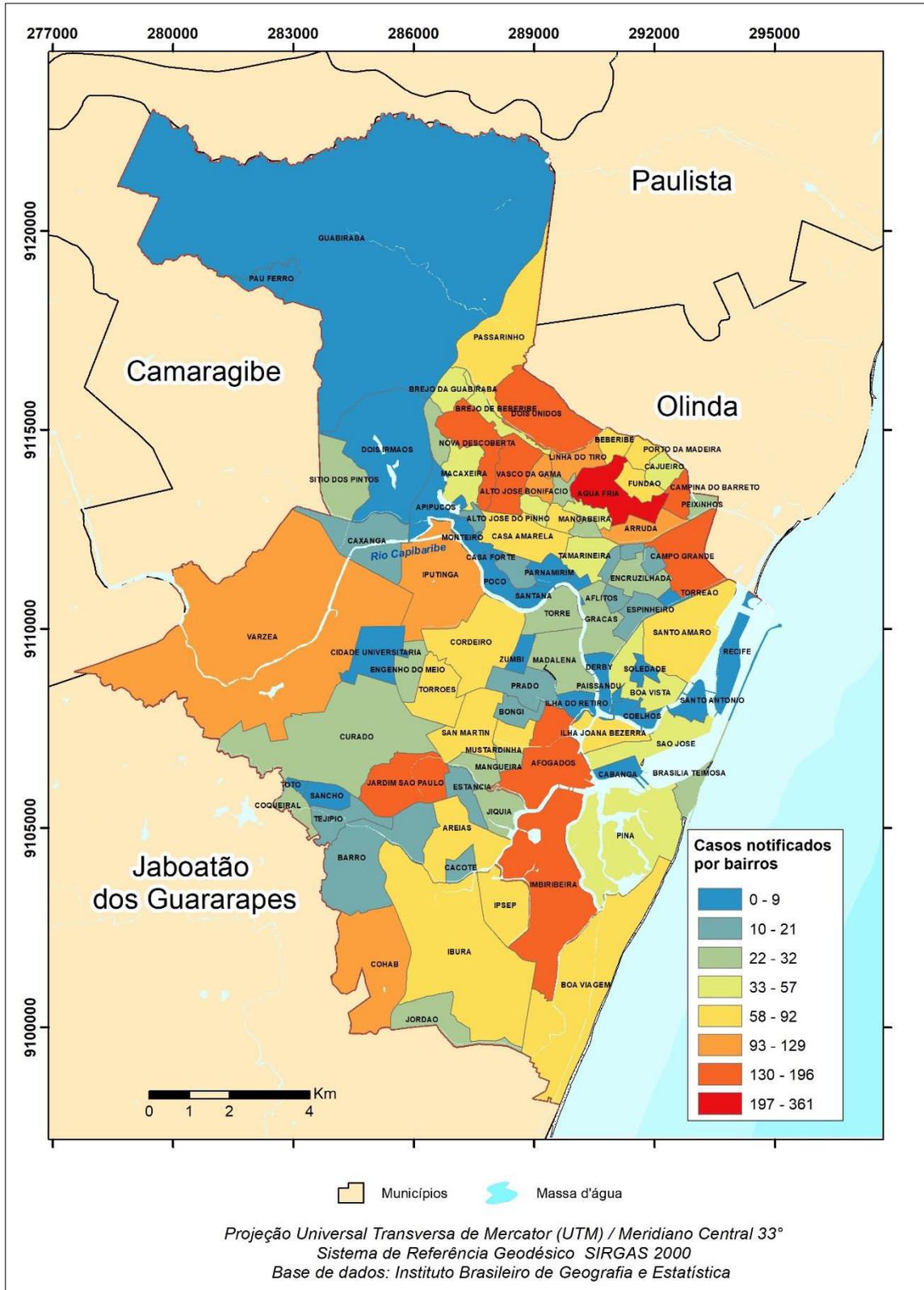
Tabela 1 - Os 10 bairros com maiores notificações de Zika e Chikungunya, entre 2015 e 2017

Bairros	Casos notificados de Zika e Chikungunya
Água fria	361
Jardim São Paulo	196
Campo Grande	184
Afogados	179
Dois unidos	171
Nova Descoberta	169
Campina do Barreto	158
Vasco da Gama	158
Imbiribeira	155
Linha do Tiro	129

Considerando a distribuição espacial da doença por bairros (Figura 17), observa-se que os bairros com menor número de casos notificados concentram-se na margem esquerda do Rio Capibaribe (Parnamirim, Poço da Panela, Casa Forte, Santana). De modo geral, são nesses bairros onde encontram-se melhores condições de infraestrutura urbana. Outros bairros com baixas notificações, ocorrem em áreas com alta percentagem de áreas verdes como Guabiraba, Dois Irmãos e Pau Ferro.

Por outro lado, a maior concentração espacial de bairros com maiores notificações localiza-se em áreas mais elevadas dos morros no norte do município, como os bairros de Água Fria, Dois Unidos, Nova Descoberta e Vasco da Gama.

Figura 17 - Casos notificados de Zika e Chikungunya entre 2015 e 2017, de acordo com os bairros de Recife.



5.1.1 Avaliação da eficiência dos modelos

Com base nos resultados da análise da eficiência dos modelos (Figura 18), observou-se que a eficiência dos modelos varia de acordo com a combinação entre a unidade espacial adotada, cálculo da variável dependente e método do modelo.

Figura 18 - Modelos categorizados de acordo com o tipo de unidade espacial, cálculo da variável dependente e método para estimar os casos de Zika e Chikungunya. O campo para comparação possui o R^2 ajustado, a soma do valor absoluto dos resíduos e o critério de informação de Akaike (AICC).

Modelo 1		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		casos por setor censitário	0,092	1.134	5.024
GWR	500	casos por setor censitário	0,47	593	37
Modelo 2		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		casos por setor censitário - sem áreas verdes	0,095	1.133	5.022
GWR	500	casos por setor censitários - sem áreas verdes	0,47	595	47
Modelo 3		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		casos por quadrícula de 100 m	0,11	9.121	52.935
GWR	500	casos por quadrícula de 100 m	0,19	7.280	14.110
Modelo 4		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		densidade de Kernel com raio de 500m por quadrícula de 100m	0,36	9.461	45912
GWR	500	densidade de Kernel com raio de 500m por quadrícula de 100m	0,89	2.631	5135
Modelo 5		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		densidade de Kernel com raio de 700m por quadrícula de 100m	0,367	9.717	45.767
GWR	700	densidade de Kernel com raio de 700m por quadrícula de 100m	0,92	2.743	7067
Modelo 6		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		densidade de Kernel com raio de 900m por quadrícula de 100m	0,359	9.949	45.886
GWR	900	densidade de Kernel com raio de 900m por quadrícula de 100m	0,92	2.855	8.036

Levando em consideração a diferença entre os modelos com diferentes unidades espaciais, os R^2 ajustados apresentaram melhores resultados nos modelos que utilizaram a quadrícula de 100 metros.

É importante ressaltar que, excluindo os espaços verdes dos setores censitários, a diferença não foi considerável entre os métodos, se analisado o coeficiente (R^2 ajustado), a soma absoluta dos resíduos e o AICC. A comparação entre os modelos utilizando esses parâmetros só foi possível porque os modelos 1 e 2 utilizam a origem das mesmas unidades espaciais: setores censitários.

Quando combinado o número de notificações da Zika e Chikungunya de acordo com as unidades espaciais, apenas o método OLS apresentou uma melhora quando a unidade

espacial de quadrículas de 100 metros foi adotada. Essa melhoria no OLS pode estar relacionada à divisão das variáveis contínuas da superfície - como declividade, espaços verdes, áreas não construídas e precipitação - em unidades espaciais menores que as áreas censitárias. Esse fato possibilita mais detalhes das informações do modelo, melhorando as estimativas. Além disso, pode inferir que a desagregação das unidades espaciais, mantendo uma mesma distância pode prejudicar o desempenho do GWR, dada a sua metodologia em trabalhar com os vizinhos das unidades espaciais.

De acordo com o cálculo da variável dependente, foram analisados quatro métodos: casos de Zika e Chikungunya por unidade espacial, média de densidade Kernel por unidade espacial, adotando raios de 500, 700 e 900 metros, respectivamente. Além disso, no método GWR, deu preferência em adotar a mesma distância do *bandwidth* da que foi utilizada como raio para a densidade de Kernel.

Os melhores resultados, analisando o coeficiente (R^2) e a soma absoluta dos resíduos, foram observados no modelo que utilizou a densidade do Kernel com 700 metros de raio. É evidente que as diferenças no cálculo da variável dependente são mais notáveis entre os casos por unidade espacial e densidade de Kernel.

Os R^2 ajustados dos métodos OLS foram bastante próximos nos três raios de busca da densidade de Kernel adotados (500, 700 e 900 metros), embora tenha apresentado a maior soma absoluta de resíduos, quando comparado ao GWR.

Ao analisar as diferenças entre os métodos dos modelos, observou-se que o método o método GWR obteve melhores resultados em todos os modelos, se comparado com o método OLS.

Estudos mostraram que essa eficiência no método GWR é melhor que o método OLS. Por exemplo, a correlação entre alguns fatores e a notificação da dengue no Paquistão foi de 0,37 no método OLS e 0,49 no GWR (ATIQUÉ *et al.*, 2017). Em Taiwan, a relação entre densidade populacional e taxas de dengue apresentou correlação maior no método GWR que no OLS (LIN e WEN, 2011). Outro ponto a considerar é que essas correlações do método GWR podem variar no espaço e no tempo. Por exemplo, a correlação entre densidade populacional, declive, altimetria e malária variou de 0,47 a 0,83, entre 2001 e 2007 (GRILLET *et al.*, 2010).

De qualquer forma, é importante observar que o método GWR é considerado um modelo exploratório. Sendo assim, o método GWR é apropriado para identificar onde o coeficiente é mais significativo, de acordo com o *bandwidth* utilizado.

Analisando a eficiência dos modelos, identificou-se que os melhores resultados estavam no Modelo 5. Este modelo possui a unidade espacial composto por quadrículas de 100 metros

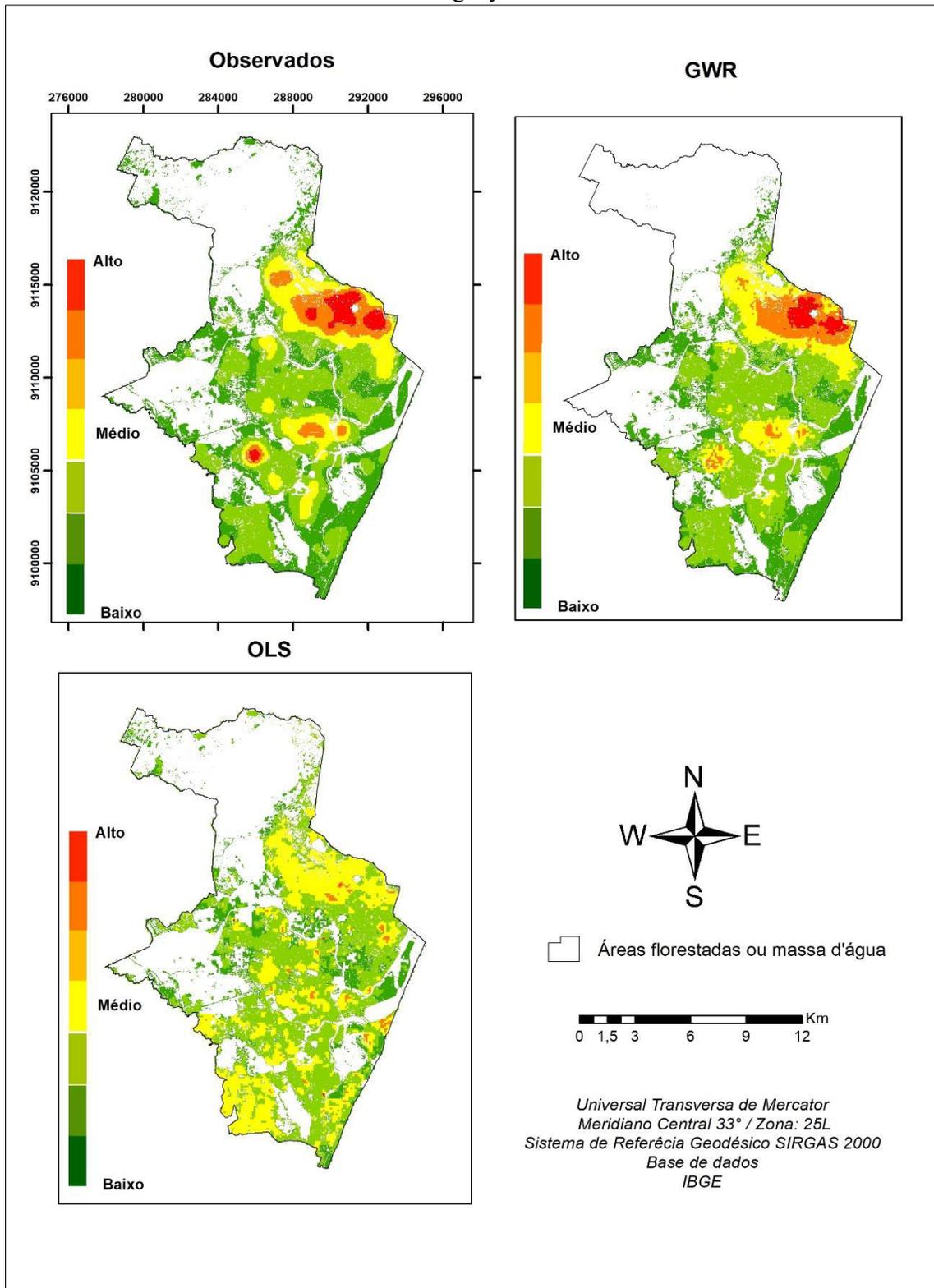
e o cálculo da variável dependente utilizando a densidade de Kernel com raio de busca de 700 metros. É importante destacar que esses parâmetros são adequados à distribuição espacial do Zika e Chikungunya no Recife. De qualquer forma, esse modelo pode ser testado em outras cidades para investigar a eficiência desses parâmetros em alguns métodos, como GWR e OLS.

Ao analisar as diferenças entre os casos observados e estimados do Modelo 5 (Figura 19), algumas particularidades foram analisadas. A espacialização desses casos notificados, utilizando os parâmetros do modelo 5, permitiu identificar que o nordeste do município, concentra os altos índices da doença. No presente caso, o raio de busca adotado para a densidade do Kernel foi de 700 metros e a unidade adotada foi de km^2 . Considerando a proporção dos casos notificados encontrados num raio de 700 metros para uma área de 1 km^2 , os valores máximos observados no município de Recife foram de 236 pessoas infectadas por km^2 , com maiores concentrações desses valores no nordeste do município.

O método OLS apresentou a maior diferença, se comparado aos dados observados. Embora esse método superestime os casos de Zika e Chikungunya em áreas onde ocorreram baixos índices das doenças, de modo geral, sua eficiência é melhor para estimar onde não ocorreram notificações das doenças.

Por outro lado, o modelo GWR mostrou semelhanças com a distribuição espacial das arboviroses. É importante mencionar que o método GWR é considerado como um modelo exploratório para avaliar a relação entre variáveis e as arboviroses. Por isso, apenas o método OLS pode ser utilizado, usando os coeficientes, para estimar o Zika e o Chikungunya em outras cidades. Esse fato não exclui a utilização de métodos de regressão local, como o GWR, para identificar a variabilidade espacial dos coeficientes.

Figura 19 - Densidade da estimativa (método GWR e OLS) e casos observados de Zika e Chikungunya em Recife.



5.1.2 Análise dos resultados do método OLS

Analisando os parâmetros no método OLS, observou-se que a *Variance Inflation Factor* (VIF) era baixo, indicando não haver problemas com multicolinearidade em todas as variáveis. Tal evidência, mostra que as variáveis utilizadas no modelo não são redundantes no modelo.

Outro ponto a considerar é que o *Robust Probability* apresentou autocorrelação espacial significativa para todas as variáveis. Tal fato mostra que todas as variáveis apresentaram correlação com os casos de Zika e Chikungunya.

O valor da estatística *Koenker* para todas as variáveis foi estatisticamente significativo ($p < 0,01$), mostrando que os relacionamentos modelados não são consistentes devido à não estacionariedade. Isso significa que as relações da variável dependente com as outras variáveis ora podem ser positivas, ora negativas, evidenciando que os coeficientes se alternam de acordo com a região estudada. Destarte, métodos de regressão linear geograficamente ponderada, tais como o GWR, podem apresentar essas variações espaciais de coeficientes, assim como as regiões onde a relação das variáveis com a doença pode ser mais evidente, tornando-se uma alternativa a ser considerada.

Três fatores apresentaram coeficientes positivos no modelo OLS: pessoas por domicílio, densidade populacional e declive (Tabela 2). As demais variáveis independentes apresentam coeficientes negativos: espaços verdes, renda do chefe de família, estruturas não construtivas e precipitação.

Tabela 2 - Variáveis independentes utilizadas e suas relações com a Zika e Chikungunya de acordo com o coeficiente, *Robust Probability* e VIF no método OLS.

Variáveis	Coefficiente	<i>Robust Probability</i>	VIF
Pessoas por domicílio	0,327402	0,000000*	1,252564
Renda do chefe familiar	-0,173202	0,000000*	1,154345
Precipitação	-0,125738	0,000000*	1,141473
Áreas não construídas	-0,031793	0,000005*	1,578071
Declividade	0,05937	0,000000*	1,240517
Densidade populacional	0,452072	0,000000*	1,567344
Áreas verdes	-0,239382	0,000000*	2,547128

5.1.3 Análise dos resultados do método GWR

Embora o valor de R^2 ajustado do método GWR tenha apresentado 0,92, existe uma variabilidade espacial da correlação entre as variáveis e as arboviroses estudadas (Figura 20). Nesse modelo, o método adotado na GWR – com 700 metros de *bandwidth* - apresentou bom desempenho neste estudo.

Na maioria das áreas censitárias (60%), o R^2 ajustado foi superior a 0,40. De qualquer forma, não existe um padrão espacial da relação entre o conjunto de variáveis e as arboviroses. Regiões onde haviam altas taxas de Zika e Chikungunya mostraram valores baixos e altos de R^2 ajustado. Assim, por exemplo, o método GWR pode mostrar que algumas variáveis são necessárias para incluir no modelo em áreas onde os valores de R^2 eram baixos. De qualquer forma, a variabilidade espacial do R^2 ajustado na Figura 20 indicou que o modelo tem proximidade com os dados observados na maioria das unidades espaciais, em Recife.

Uma observação a ser feita é que as diferenças entre o R^2 ajustado que se apresentam de maneira espacial e o valor único do R^2 ajustado apresentado nos modelos, é que no último os valores para o seu cálculo dizem respeito a divisão entre a soma quadrada dos resíduos e a soma quadrada dos valores observados de todas as unidades espaciais utilizadas, enquanto no primeiro, os cálculos são feitos apenas com as unidades espaciais e seus vizinhos, definidos de acordo com o *bandwidth* adotado.

Analisando os coeficientes da relação entre cada variável e a distribuição espacial da Zika e Chikungunya no método GWR na Figura 21, é possível identificar que os valores variam de acordo com a localização geográfica. Embora as variáveis independentes mostrassem semelhanças entre os coeficientes dos métodos GWR e OLS, algumas regiões apresentaram relação inversa.

Figura 20 - Distribuição espacial do R^2 ajustado no método GWR para a Zika e Chikungunya.

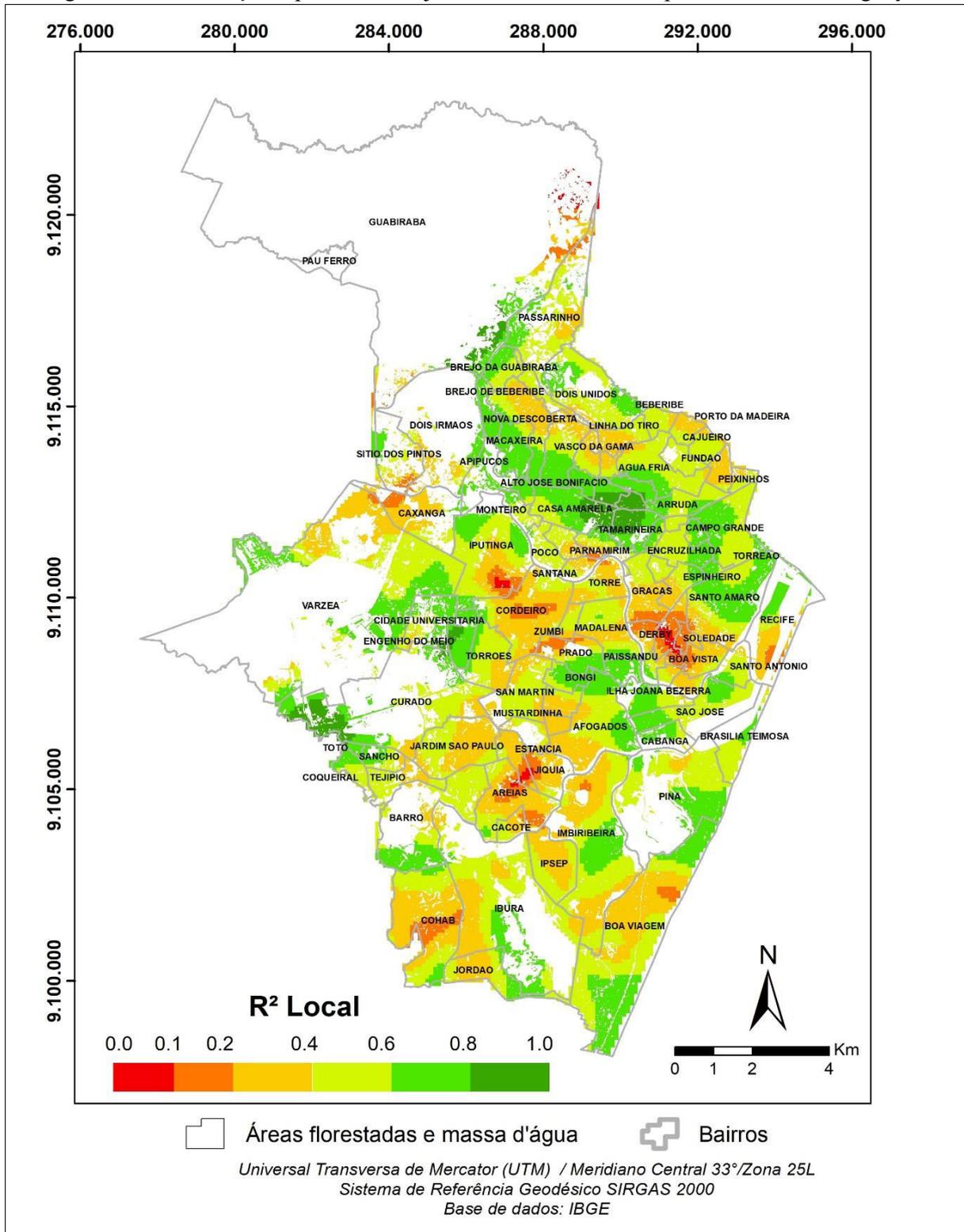
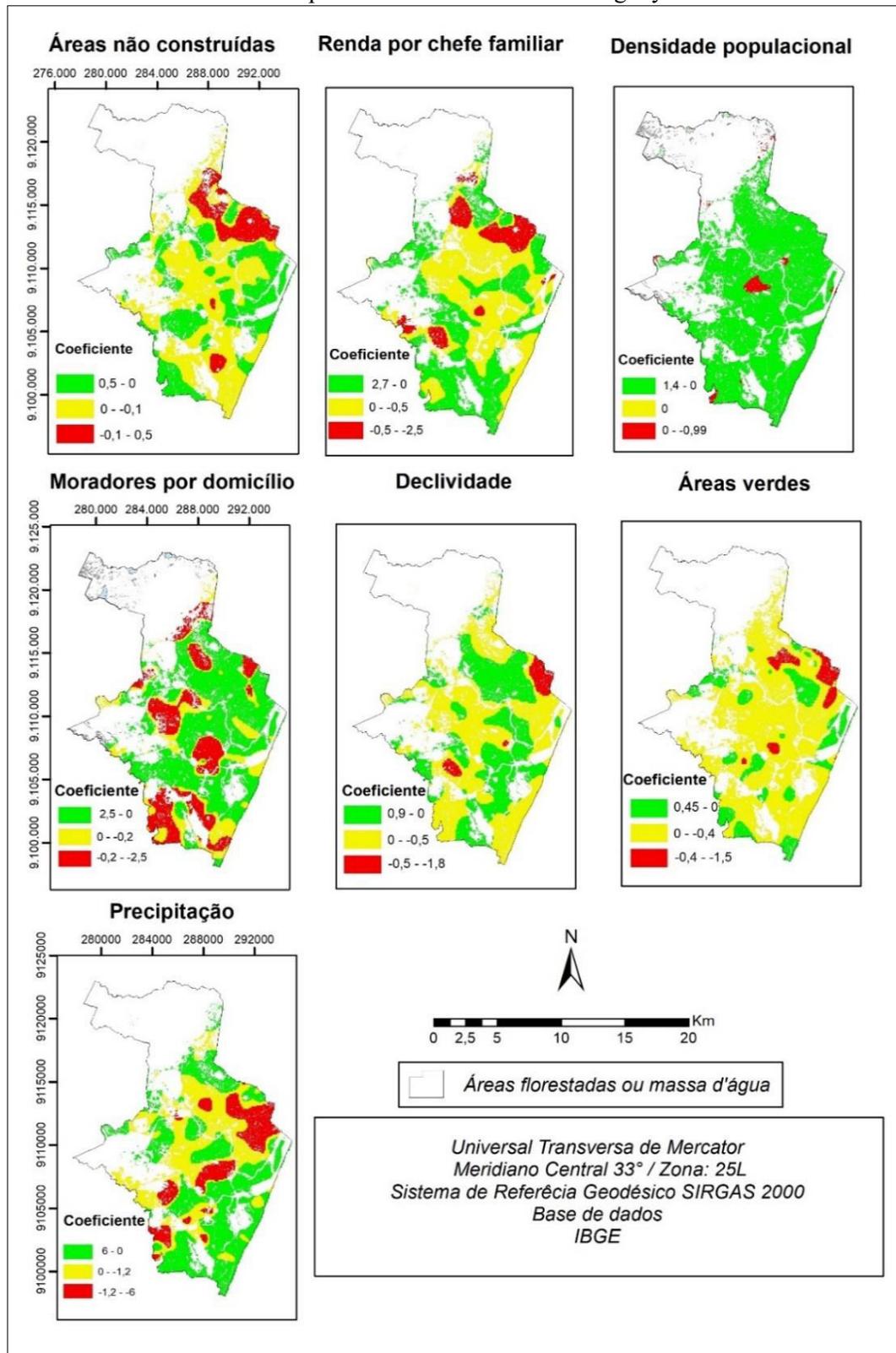


Figura 21 - Distribuição espacial dos coeficientes do método GWR para cada variável inserida no modelo para estimar a Zika e a Chikungunya.



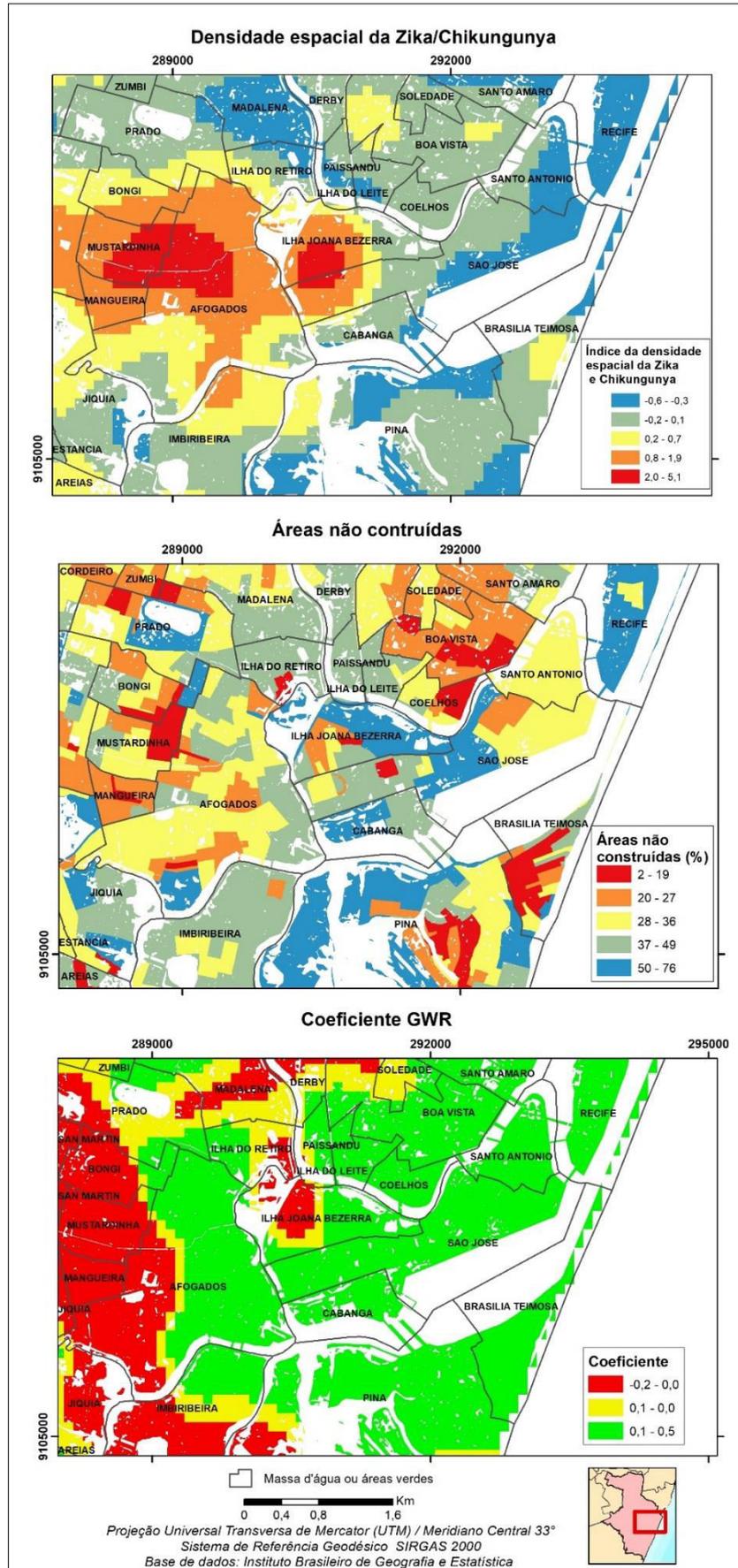
As áreas não construídas apresentaram correlação negativa com a Zika e Chikungunya na maioria das unidades espaciais (55%). É importante notar que algumas regiões (leste e sudoeste) apresentaram valores positivos do coeficiente. Nessas áreas, as taxas de áreas não construídas eram baixas, enquanto as taxas de Zika e Chikungunya tendiam a acompanhar essa tendência de diminuição. Esse fato apresentou que outras variáveis podem evitar a disseminação das doenças, mesmo em áreas onde a densidade de construções é considerada elevada.

Outro ponto a ser levantado nessas áreas de coeficientes positivo é que a presença de terrenos baldios (áreas não construídas) favorece criadouros dos mosquitos, aumentando a probabilidade de mais pessoas serem infectadas no entorno. Esse fato fica mais evidente na Figura 22, onde na Ilha Joana Bezerra as áreas não construídas mostram-se bem próximas dos núcleos de maior notificação da doença, ainda que nesses núcleos a densidade de edificações seja alta. Outro aspecto é que as áreas de coeficientes positivos no centro do Recife (caracterizado pelos bairros do Recife, Boa Vista, Santo Amaro, São José e Santo Antônio) de modo geral, são áreas com densidade populacional baixa, se comparadas com outras regiões do município.

No entanto, os valores negativos do coeficiente podem ser explicados pelo fato de as áreas onde as taxas de áreas não construídas serem baixas, a distância entre os prédios é menor, aumentando as oportunidades do mosquito de infectar mais famílias. Outro ponto a considerar é que as áreas de aglomerações subnormais apresentam as menores taxas de áreas não construídas em Recife, justificando a correlação entre as maiores taxas de Zika e Chikungunya com as menores taxas de áreas não construídas. Lesser e Kitron (2016) perceberam que as altas concentrações espaciais da Zika em São Paulo, estavam relacionadas a uma alta densidade de moradias com estruturas de habitação improvisadas.

É importante observar que a densidade de edificações pode ser um indicador da temperatura devido ao calor latente que é absorvido em áreas com alta concentração de construções. Esse fato pode estar relacionado às altas taxas de Zika e Chikungunya, devido às preferências do *Aedes aegypti* a altas temperaturas.

Figura 22 - Densidade espacial da Zika/Chikungunya, índice de áreas não construídas e coeficiente de GWR, na zona sul, oeste e centro do Recife.



A correlação negativa no modelo OLS, entre a renda familiar e as doenças, foi confirmada nos valores do coeficiente de GWR na maioria das unidades espaciais.

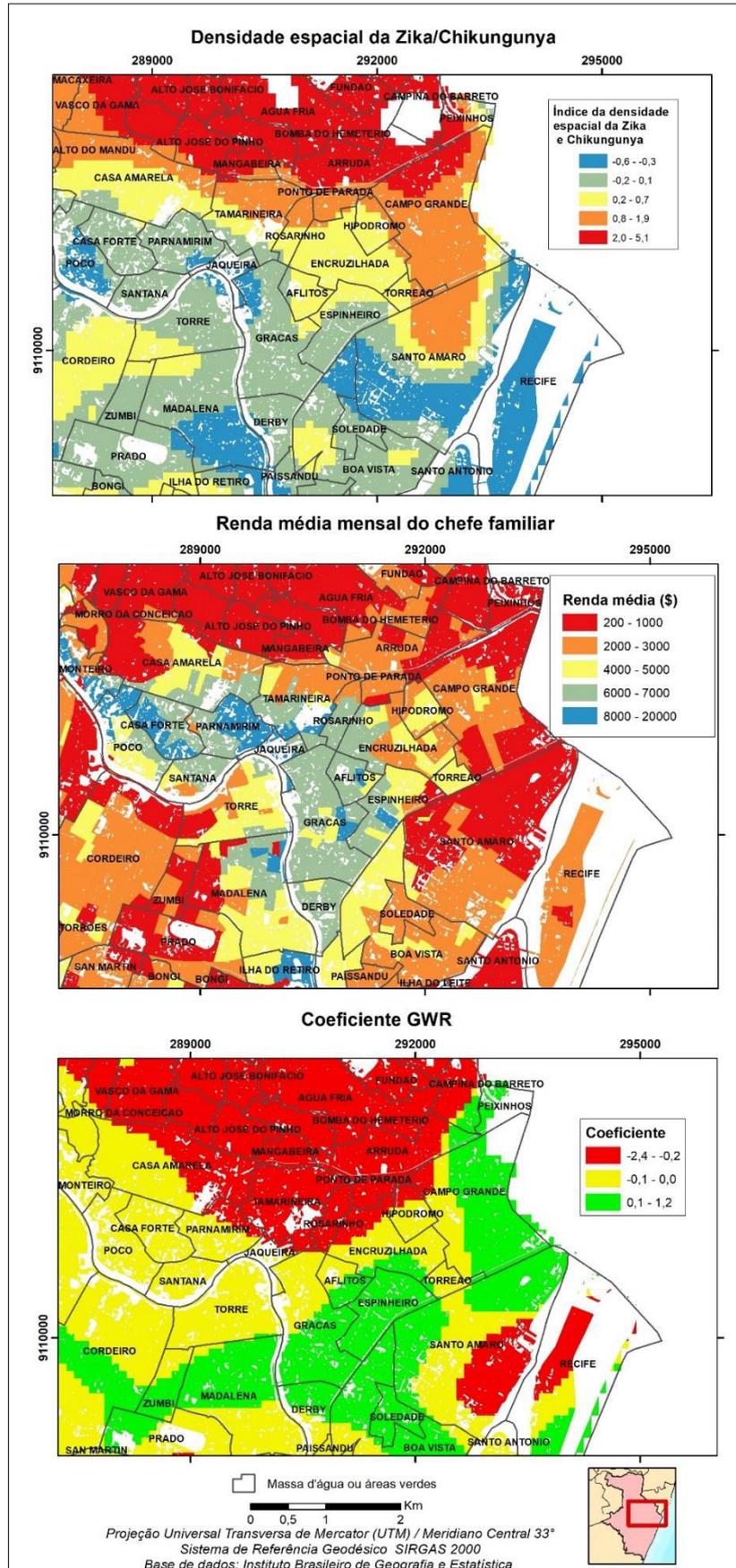
No entanto, a renda não é um fator predominante em algumas unidades espaciais no norte da cidade, se considerarmos os coeficientes positivos da cidade. Esse fato pode estar relacionado à diminuição de casos das arboviroses em áreas onde a população de baixa renda está localizada. Tal fato fica mais evidente quando analisado o bairro de Campo Grande e a porção norte do bairro de Santo Amaro, na Figura 23. Outra hipótese associada a isso é a existência de áreas caracterizadas por baixos índices de renda média mensal com baixo número de notificações de Zika e Chikungunya também pode estar relacionada com a cobertura da atenção básica e da vigilância nessas áreas, por meio da Vigilância ambiental.

A heterogeneidade das variáveis relacionadas à doença arboviral é observada em outros estudos, onde alguns indicadores socioeconômicos não foram estatisticamente significantes na categorização das áreas de risco (RODRIGUES *et al*, 2018). No entanto, estudos mostraram que a taxa de renda é o fator que melhor caracterizou as áreas de risco para arboviroses, como a dengue, por exemplo (ALMEIDA *et al.*, 2007).

Ao analisar a densidade populacional, na maioria das unidades espaciais, o valor do coeficiente no modelo GWR é positivo, principalmente onde as taxas de doença foram altas, confirmando os valores do coeficiente no modelo OLS. É importante notar que, onde ocorreu uma correlação inversa entre a densidade populacional e as notificações de Zika e Chikungunya, mostrou uma taxa de densidade populacional menor que a de seus vizinhos, mesmo com taxas semelhantes das arboviroses. Por isso, é provável que, nesse fato, a correlação tenha sido negativa. Apesar disso, outros estudos mostraram que, onde há densidades populacionais humanas mais altas, existem condições socioeconômicas mais baixas, más condições de água e saneamento e instalações frequentes de cobertura de pneus, mas também com infraestrutura razoável (QUINTERO *et al*, 2014). Por exemplo, em Porto Rico foi identificado que população pobre, densidade populacional, áreas urbanizadas e temperatura apresentaram correlação de 0,78 com a Zika, no modelo GWR (WEBER, 2018). Outro estudo no Paquistão verificou que, dentre as variáveis, a altimetria, distância dos rios e densidade populacional, o último fator apresentou a maior correlação com a notificação da dengue (ATIQUÉ *et al*, 2017).

No que concerne à relação entre Zika e Chikungunya e pessoas por domicílio, na maioria das unidades espaciais, a correlação foi positiva no modelo GWR, concordando com os valores do coeficiente no modelo OLS.

Figura 23 - Densidade espacial da Zika/Chikungunya, renda média mensal do chefe familiar e coeficiente do GWR, na zona norte, oeste e no centro do Recife.



É importante notar que, embora algumas unidades espaciais apresentem altas taxas de pessoas no domicílio, as taxas das ZIKV e CHIKV foram baixas, causando um coeficiente negativo no modelo GWR. Em geral, a correlação entre pessoas por domicílio e as notificações de Zika e Chikungunya em ambos os modelos mostrou uma correlação diretamente proporcional. Corroborando com essa análise, alguns estudos identificaram que a média de pessoas por domicílio pode ser um indicador da disseminação de algumas arboviroses. Por exemplo, foram detectados resultados no Equador, onde a relação entre Zika e número de moradores por domicílio foi estatisticamente significativo (ORTIZ *et al*, 2017). Na Venezuela, observou-se que a relação entre dengue e pessoas por domicílio era diretamente proporcional (VICENTI-GONZALEZ *et al*, 2017).

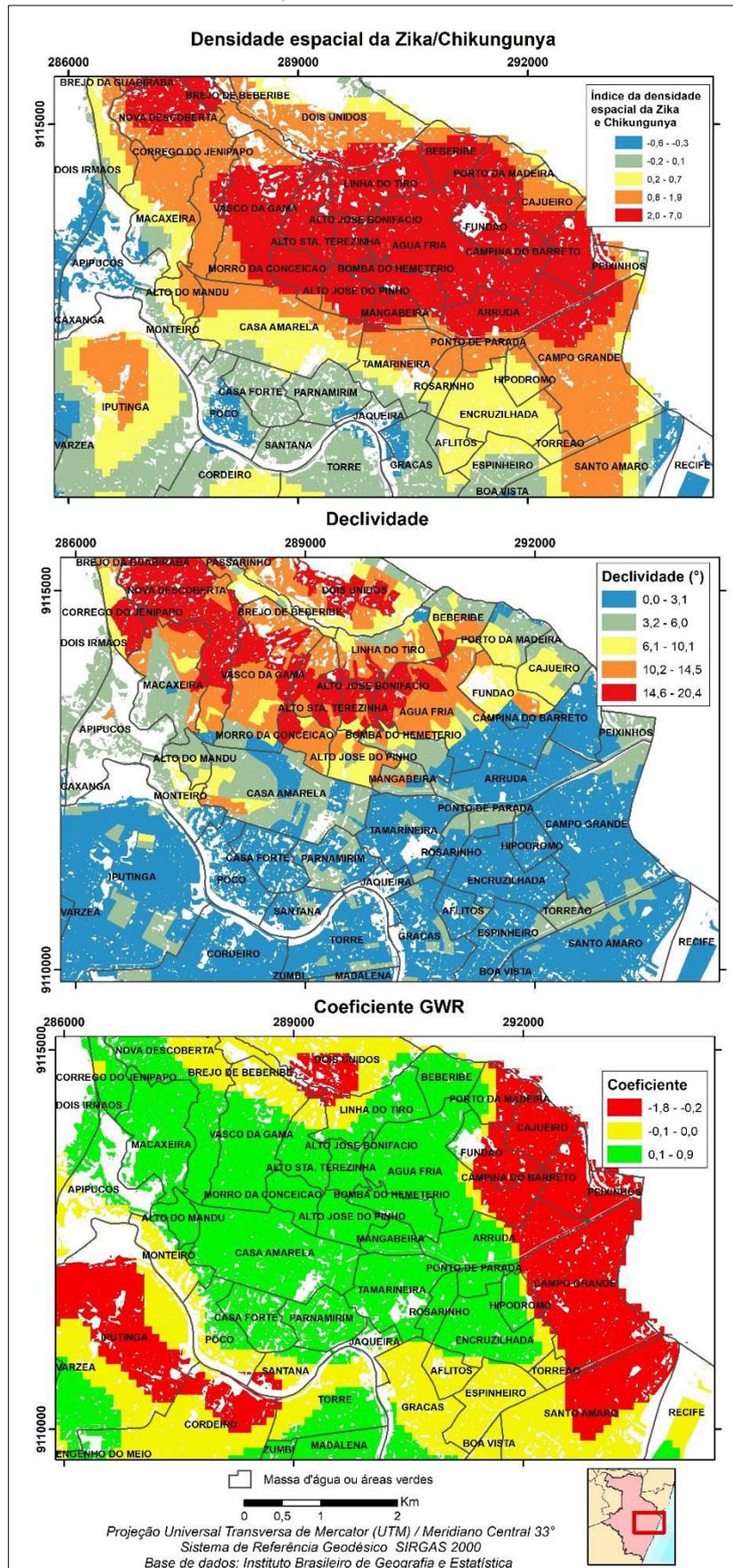
A relação entre declividade e as arboviroses analisadas apresentou correlações positivas e negativas, sem padrão de distribuição espacial. Algumas hipóteses podem ser levantadas nas áreas onde o coeficiente do GWR foi negativo. Feitosa (2016) afirma que em áreas onde o solo é impermeável associada a uma baixa declividade, favorece que o acúmulo de água parada e limpa se acumule e forme criadouros do *Aedes aegypti*.

Os coeficientes positivos podem estar associados a áreas com declividades inferiores as áreas subjacentes, mas que apresentam alta densidade espacial da doença. Esse fato fica mais evidente quando observamos os bairros de Cajueiro, Campina do Barreto e Campo Grande (Figura 24).

Entretanto, os valores do coeficiente no modelo OLS mostraram que a relação entre as densidades espaciais da doença e a declividade é diretamente proporcional.

Esse fator é muito complexo devido às áreas de aglomerações subnormais estarem localizadas em margens de rios, encostas íngremes ou aterros sanitários. Esse fato está relacionado ao baixo poder aquisitivo dos moradores de favelas no mercado formal de terras, quando comparado aos grupos de alta renda (MAHABIR *et al*, 2016).

Figura 24 - Densidade espacial da Zika/Chikungunya, declividade média e coeficiente do GWR, na zona norte, oeste e no centro do Recife.



É importante ressaltar que a variação espacial dos coeficientes obtidos pelo GWR pode estar relacionada à complexa interação entre topografia e população de baixa renda. Em Recife, após o fim da escravidão em 1888, muitos escravos mudaram-se das plantações de cana-de-açúcar para o Recife e se estabeleceram nas áreas de mangue ao longo dos cursos d'água. No entanto, desde o final da década de 1930 até meados da década de 1970, o número de aglomerações subnormais aumentaram significativamente, forçando muitas famílias pobres para as áreas de morros da cidade (SANDHOLZ, 2017). Portanto, pode-se inferir que as notificações de Zika e Chikungunya em aglomerações subnormais localizadas em áreas planas tenham causado a correlação negativa nessas regiões, embora o coeficiente positivo tenha sido identificado em áreas de declives acentuados.

A relação inversamente proporcional entre espaços verdes e as arboviroses encontrada no método OLS também é apresentada na maioria das unidades espaciais. As áreas onde ocorreram os maiores valores de coeficiente positivo (principalmente a nordeste da cidade) apresentaram baixa presença de espaços verdes e baixas taxas de Zika e Chikungunya. No entanto, a presença de espaços verdes pode ser um fator que contribuiu para diminuir a temperatura nas áreas urbanas, inibindo a criação de mosquitos. Esse fato pode ser confirmado com as semelhanças da relação inversamente proporcional entre chuva e as densidades espaciais da Zika e Chikungunya na maioria das unidades espaciais. Vale ressaltar que em regiões tropicais quando há pouca precipitação, a temperatura do ar se eleva, contribuindo para a proliferação do mosquito.

Outro ponto a considerar é que o mosquito *Aedes aegypti* prefere áreas urbanizadas, onde a presença de cobertura vegetal é mais limitada. Assim, os espaços verdes podem ser um fator correlacionado com a urbanização e, conseqüentemente, relacionado à distribuição espacial das arboviroses. Esse resultado é semelhante ao estudo que identificou a maior incidência de Chikungunya nas regiões com maiores porcentagens de áreas urbanizadas no Rio de Janeiro (FULLER *et al*, 2017). Outro fator a destacar é que a disponibilidade de sangue de outros animais diminui à medida que se aumentam áreas urbanizadas, aumentando a preferência do *Aedes aegypti* pelo sangue humano (NDENGA *et al*, 2017). Além disso, é importante destacar que a maioria das áreas de aglomerações subnormais é altamente compacta, apresentando uma cobertura superficial alta com baixa taxa de vegetação (KOHLI *et al*, 2016). Em outras palavras, a presença de vegetação pode estar relacionada a vários fatores que influenciam na distribuição espacial do Zika e Chikungunya.

Em relação à variável chuva, os valores do coeficiente de OLS corroboraram com a maioria das unidades espaciais dos coeficientes do GWR, onde as regiões de menor precipitação

ocorreram as maiores taxas de Zika e Chikungunya. De fato, no nordeste do Recife, onde estão localizadas as maiores taxas de doenças, os valores do coeficiente de GWR são negativos na maior parte da área censitária.

Alguns modelos identificaram que, embora um aumento na temperatura em latitudes mais altas possa aumentar a presença do *Aedes aegypti* para essas regiões, o aumento da precipitação poderia evitar a proliferação desse mosquito (VALDEZ *et al.*, 2017). Na maioria dos estudos que analisaram a relação entre parâmetros climáticos e as arboviroses, foi identificado que as arboviroses estão relacionadas à distribuição temporal, desconsiderando a distribuição espacial devido à falta de estações meteorológicas bem distribuídas nas cidades. No entanto, as análises temporais indicam que tempestades extremamente intensas que produzem um alto volume de precipitação em poucas horas podem afetar a sobrevivência das larvas, levando a diminuição de vetores e a transmissão de arboviroses (FULLER *et. al*, 2017). Essas observações podem sugerir que as regiões de baixo volume de precipitação são melhores para a proliferação de mosquitos.

5.2 LEPTOSPIROSE

Ao todo, foram geocodificados 222 casos de leptospirose em Recife, entre os anos de 2015 e 2017 (Figura 25). De todos os casos notificados, 37% concentravam-se nos 10 bairros com maiores notificações da doença (Tabela 3). Liderando a lista, o bairro de Campo Grande registrou o maior número pessoas infectadas, sendo 12 ao total.

Tabela 3 - Os 10 bairros com maiores registros de Leptospirose entre 2015 e 2017.

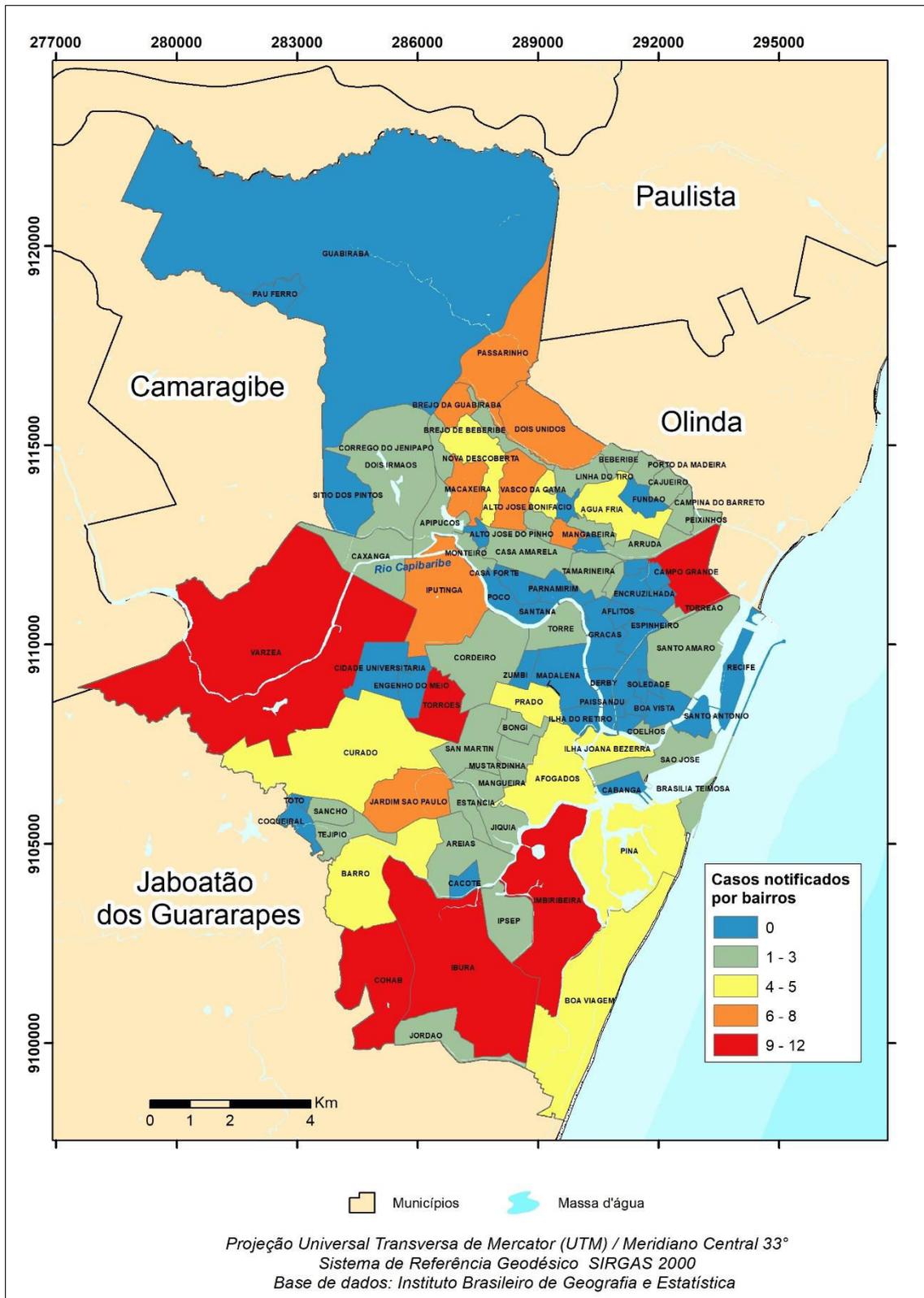
Bairros	Casos notificados de Leptospirose
Campo Grande	12
Ibura	11
Várzea	10
Imbiribeira	9
Torrões	9
Cohab	9
Macaxeira	8
Iputinga	8
Dois Unidos	7
Passarinho	7

De modo geral, é importante ressaltar o núcleo de bairros caracterizados por boas condições de infraestrutura e habitação, como os bairros de Casa Forte, Santana, Poço da Panela e Parnamirim apresentarem menor número de casos notificados. Contudo, alguns bairros mais

populares na zona oeste de Recife, como Cordeiro, Zumbi, Bongi e San Martin também mostraram poucos registros da doença. Por isso, é importante salientar que a distribuição espacial da leptospirose levanta a complexidade de condicionantes associados na sua proliferação, assim como a existência de fatores locais da doença onde necessita uma unidade espacial de análise menor que o bairro para ser identificada.

Outro aspecto a ser ressaltado é que os bairros com maiores notificações da doença apresentam as suas áreas habitadas localizadas na planície flúvio-marinha do município, como a Várzea, Campo Grande e Imbiribeira. Esse aspecto pode estar intimamente relacionado a áreas de inundação e, conseqüentemente maior exposição dos moradores dessas áreas à água contaminada.

Figura 25 - Casos notificados de Leptospirose entre 2015 e 2017, de acordo com os bairros de Recife.



5.2.1 Avaliação da eficiência dos modelos

Analisando os resultados da Figura 26, observou-se que a eficiência dos modelos varia de acordo com a combinação entre a unidade espacial adotada, cálculo da variável dependente e método do modelo.

Considerando os três primeiros modelos, percebe-se que o tipo de unidade espacial adotada não apresenta diferenças significativas no método OLS. Contudo, no método GWR – que apresentou os melhores resultados que o método OLS, se observados a soma dos residuais e o AICC – demonstrou um maior desempenho no Modelo 2 com um R^2 ajustado de 0,86, cuja unidade espacial são os setores censitários sem áreas verdes. Esse valor do R^2 pode estar relacionado ao fato de que 14 setores censitários foram desconsiderados do cálculo, pois não possuíam um vizinho num raio de 500 metros (que foi o *bandwidth* adotado), contados a partir do centróide do polígono. Tal fato, compromete o cálculo geral e, conseqüentemente, as estimativas das doenças para os setores, o que nesse caso melhorou os coeficientes em relação aos outros modelos.

Figura 26 - Modelos categorizados de acordo com o tipo de unidade espacial, cálculo da variável dependente e método para estimar os casos de Leptospirose. O campo para comparação possui o R^2 ajustado, a soma do valor absoluto dos resíduos e o Critério de Informação de Akaike

Modelo 1		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		casos por setor censitário	0,02	1.042	5.165
GWR	500	casos por setor censitário	0,12	630	43
Modelo 2		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		casos por setor censitário - sem áreas verdes	0,02	1.030	5.165
GWR	500	casos por setor censitários - sem áreas verdes	0,86	643,278985	-3
Modelo 3		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		casos por quadrícula de 100 m	0,01	4.325	55215
GWR	500	casos por quadrícula de 100 m	0,03	4.300	17.209
Modelo 4		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		densidade de Kernel com raio de 500m por quadrícula de 100m	0,24	11.311	50921
GWR	500	densidade de Kernel com raio de 500m por quadrícula de 100m	0,75	5.294	11081
Modelo 5		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		densidade de Kernel com raio de 700m por quadrícula de 100m	0,31	9.073	40.147
GWR	700	densidade de Kernel com raio de 700m por quadrícula de 100m	0,78	4.183	12595
Modelo 6		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		densidade de Kernel com raio de 900m por quadrícula de 100m	0,36	15.751	58.775
GWR	900	densidade de Kernel com raio de 900m por quadrícula de 100m	0,8	7.058	25.262

É importante destacar que o método GWR é bastante sensível ao *bandwidth* adotado, assim como a unidade espacial que vai ser inserida a variável dependente. Destarte, isso pode evidenciar que cada unidade espacial possa ter um número de *bandwidth* mais adequado para que os coeficientes sejam satisfatórios no cálculo dos vizinhos.

Quando observados os diferentes tipos de cálculo da variável dependente, percebe-se que os melhores resultados dos modelos tendem a estarem relacionados ao método da densidade de Kernel, se comparado com o número de notificações por unidade espacial. De maneira mais precisa, ao ser utilizada a densidade de Kernel com raio de 900 metros, foi observada uma maior aproximação entre os coeficientes dos métodos de OLS e GWR, apresentando 0,36 e 0,8, respectivamente. Uma das hipóteses para essas melhoras consideráveis, é que as variações espaciais, quando utilizadas o método de densidade de Kernel são menos abruptas entre as unidades espaciais – se comparadas com o número de casos por unidade espacial- o que tendem a acompanhar as flutuações espaciais das condicionantes de maneira mais aproximada a variação espacial da leptospirose.

No que concerne ao método de estimativa das doenças, analisando o AICC e a soma dos residuais, percebe-se que o GWR apresentou os melhores resultados em todos os modelos. É importante frisar que a própria particularidade do método GWR permite que seus resultados sejam mais satisfatórios que o método OLS, uma vez que o último trabalha com estacionariedade espacial, deixando os coeficientes menos flexíveis.

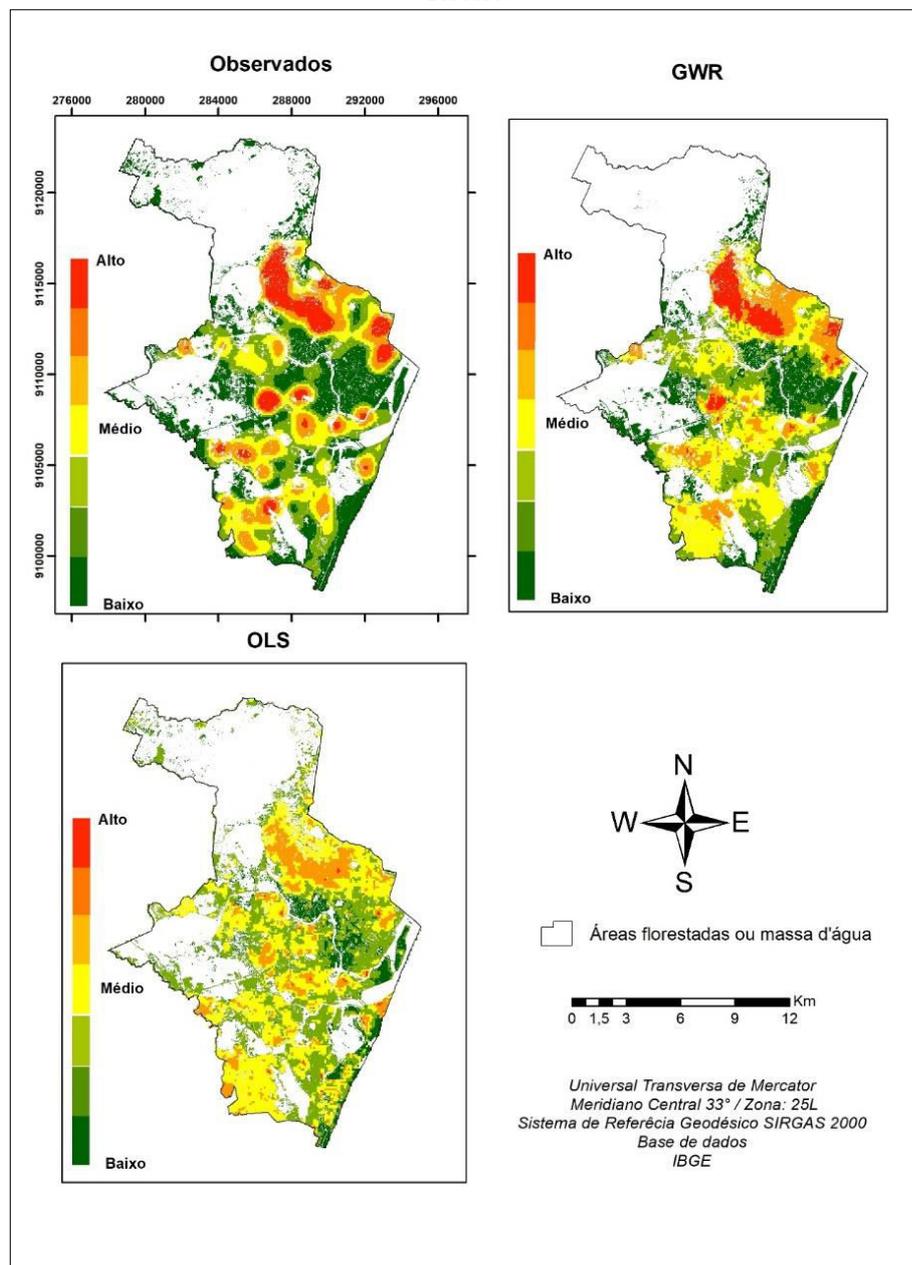
Considerando a eficiência de todos os modelos, observando os valores de R^2 , o Modelo 6 apresentou resultados mais satisfatórios, tanto para o método OLS, como para o GWR. Tal modelo é caracterizado por adotar um *bandwidth* de 900 metros no método GWR, com unidade espacial de quadrículas 100x100 metros e cálculo da variável dependente utilizando a densidade do Kernel com 900 metros de raio.

A partir da espacialização dessas notificações utilizando o método da densidade de Kernel com um raio de 900 metros, podemos identificar que a maior concentração dos casos se localiza ao extremo norte do município (Figura 27). No entanto é importante ressaltar que alguns núcleos de notificações se fazem presentes em algumas localidades do município. As menores notificações de leptospirose apresentam-se no litoral recifense e na zona norte da planície recifense, situada na margem esquerda do rio Capibaribe, e ao sul dos morros do norte de Recife, sendo caracterizada pela presença dos bairros onde concentram a área nobre da cidade.

Em termos espaciais, os maiores índices atingiram 9 pessoas por km^2 , porém essas variações apresentam distinções entre os dados observados e os dados estimados pelos métodos

OLS e GWR. As estimativas apresentadas pelo método GWR são as que apresentam maiores semelhanças com os dados observados. Contudo, cabe uma crítica a ser ressaltada: por mais que existam similaridades entre o método GWR com os dados observados, os coeficientes que foram utilizados para produzir as estimativas não podem ser replicados em outras localidades, pois eles são únicos para aquela localização geográfica. Por isso, é válido ressaltar a importância de métodos de regressão linear global, como o OLS, que apresentam valores únicos de coeficiente para cada condicionante e sua respectiva relação com a doença.

Figura 27 - Densidade da estimativa (método GWR e OLS) e casos observados da Leptospirose em Recife.



5.2.2 Análise dos resultados do método OLS

Analisando os parâmetros no método OLS, observou-se que a *Variance Inflation Factor* (VIF) era baixo, indicando não haver problemas com multicolinearidade em todas as variáveis, ou seja, não existem variáveis que sejam redundantes no modelo.

Outro ponto a considerar é que a *Robust Probability* apresentou autocorrelação espacial significativa para todas as variáveis. Tal fato mostra que todas as variáveis apresentaram correlação com os casos de Leptospirose.

O valor da estatística *Koenker* para todas as variáveis foi estatisticamente significativo ($p < 0,01$), mostrando que os relacionamentos modelados não são consistentes devido à não estacionariedade. Isso significa que as relações da variável dependente com as outras variáveis ora podem ser positivas, ora negativas, evidenciando que os coeficientes se alternam de acordo com a região estudada. Assim, evidencia-se que métodos de regressão geográfica ponderada, como o GWR, podem ser utilizadas na estimação dos casos de leptospirose.

Três fatores apresentaram coeficientes positivos no modelo OLS: pessoas por domicílio, densidade populacional, áreas não construídas, precipitação e declividade (Tabela 4). As demais variáveis independentes apresentaram coeficientes negativos: áreas verdes e renda do chefe de família.

Tabela 4 - Variáveis independentes utilizadas e suas relações com a Leptospirose de acordo com o coeficiente, *Robust Probability* e VIF no método OLS.

Variáveis	Coefficiente	<i>Robust Probability</i>	VIF
Pessoas por domicílio	0,329912	0,000000*	1,251669
Renda do chefe familiar	-0,359291	0,000000*	1,154417
Precipitação	0,033156	0,000924*	1,141215
Áreas não construídas	0,028745	0,007892*	1,572145
Declividade	0,143386	0,000000*	1,240199
Densidade populacional	0,601906	0,000000*	1,562560
Áreas verdes	-0,419618	0,000000*	2,527576

Cada coeficiente mostrará como cada variável incluída no modelo irá se comportar com a Leptospirose. Destarte, os valores dos coeficientes representam a mudança na densidade espacial da Leptospirose para cada mudança em uma unidade da variável independente, mantendo todas as variáveis constantes. De todas as variáveis analisadas, percebe-se que a densidade populacional apresenta o maior valor do coeficiente, evidenciando que a doença

tende a acompanhar de maneira mais aproximada as variações espaciais da densidade populacional, considerando o conjunto de variáveis utilizadas no modelo. Por outro lado, a variável independente que demonstrou pouca alteração em relação a Leptospirose foi áreas não construídas, evidenciando que a alteração em uma unidade do índice dessa variável, aumentará apenas 0,02 no índice espacial da doença.

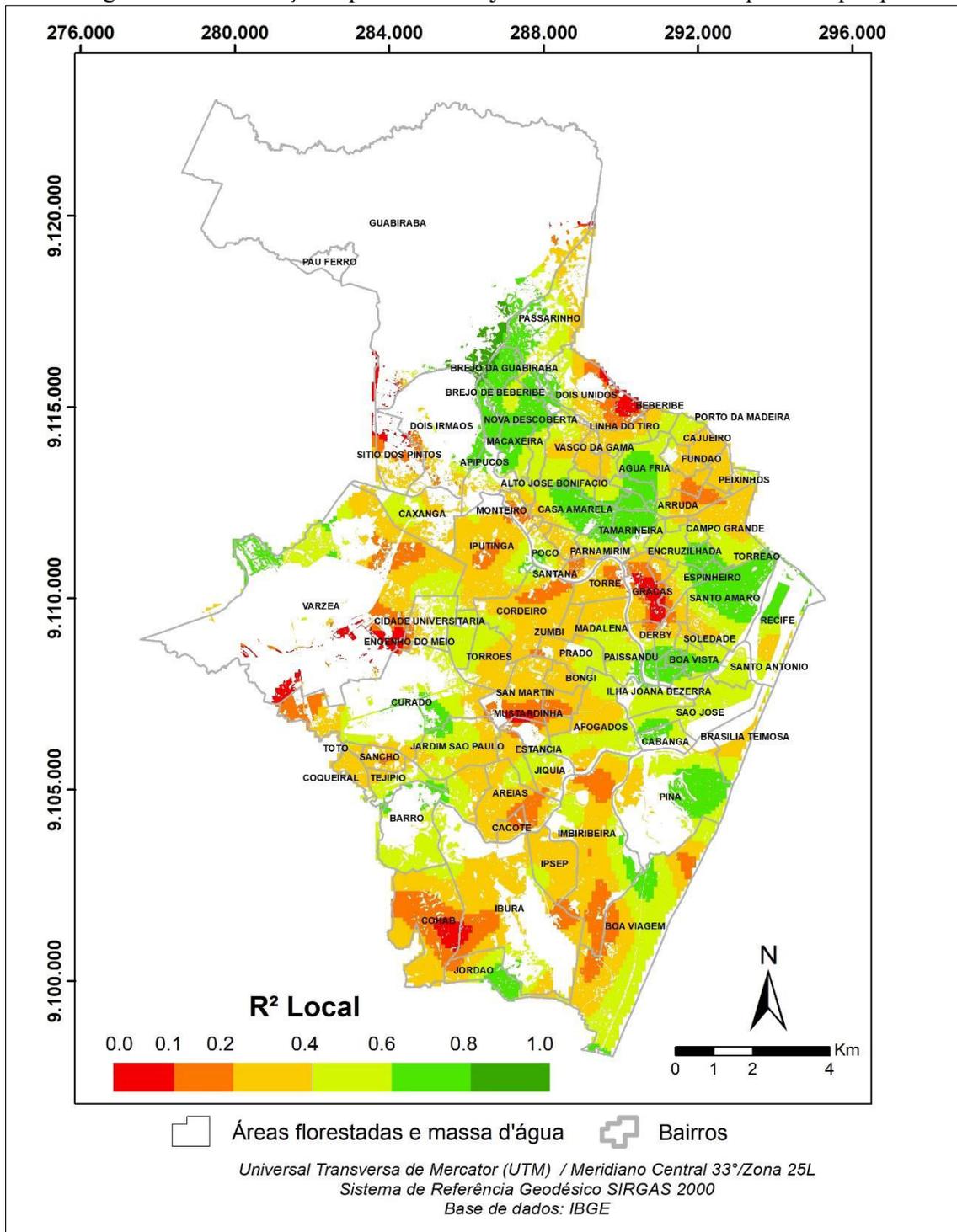
5.2.3 Análise dos resultados do método GWR

Apesar do valor de R^2 ajustado do método GWR tenha apresentado 0,80, os valores do coeficiente variaram de acordo com sua localização espacial (Figura 28). Essas variações apresentaram valores de 0 a 0,87, com maiores coeficientes localizados nas áreas onde houveram a maior concentração da doença, na zona norte do Recife. Dessas unidades espaciais, aproximadamente 42% de todas as unidades espaciais apresentaram coeficientes maiores que 0,4. Outro aspecto a ser considerado é que aproximadamente 10% das unidades espaciais tiveram valores nulos devido às unidades espaciais não terem um número mínimo de vizinhos aceitáveis para o cálculo (acontecendo esse fato próximo das bordas do município). É importante levar tal fato em consideração, visto que essas unidades espaciais são desconsideradas nos valores dos coeficientes, mostrando a alta sensibilidade desse método em relação ao *bandwidth* adotado.

Um dos pontos fortes em analisar a variação espacial dos coeficientes é identificar onde outros fatores precisariam ser incluídos. Desse modo, essa representação espacial dos coeficientes (R^2) pode servir de base para futuras investigações em campo com o intuito de identificar e compreender quais outras condicionantes ocorrem somente nessas regiões que, ou evitam a disseminação da leptospirose, ou a favorece.

Uma observação a ser feita é que as diferenças entre o R^2 que se apresentam de maneira espacial e o R^2 ajustado apresentado nos modelos, é que no último os valores para o seu cálculo dizem respeito a divisão entre a soma quadrada dos resíduos e a soma quadrada dos valores observados de todas as unidades espaciais utilizadas, enquanto no primeiro, essa mesma equação é feita apenas em uma determinada unidade espacial e o número de vizinhos, definidos de acordo com o *bandwidth* adotado.

Figura 28 - Distribuição espacial do R² ajustado no método GWR para a Leptospirose.



Em relação a variação espacial dos coeficientes, no método GWR, referentes as relações da leptospirose com cada condicionante, é possível identificar que as relações analisadas no método OLS, podem se inverter em algumas regiões do município (Figura 29). É importante salientar que algumas condicionantes podem apresentar alta variabilidade espacial em seus coeficientes, outras não. Trazendo essa relação entre os coeficientes apresentados espacialmente no método GWR com os apresentados, podemos levantar algumas abordagens.

No que concerne a relação da leptospirose com áreas não construídas, aproximadamente 46% das unidades espaciais apresentaram coeficientes negativos, 44% apresentaram coeficientes positivos e aproximadamente 10% apresentaram valores nulos. Comparando esses valores no método OLS, percebe-se que a maioria das unidades espaciais no método GWR apresentaram um coeficiente negativo, enquanto no método de regressão linear apresentou um valor positivo. Essa relação entre áreas não construídas e casos de leptospirose ainda é pouco discutida nos trabalhos publicados.

Na Figura 30 pode-se observar que as áreas que apresentaram coeficientes negativos estavam nos bairros que configuravam baixa densidade espacial da doença associada à taxas de áreas não edificadas maior que os bairros vizinhos. Por outro lado, os coeficientes positivos predominaram na maior parte da porção norte do município.

Figura 29 -Distribuição espacial dos coeficientes do método GWR para cada variável inserida no modelo para estimar a Leptospirose.

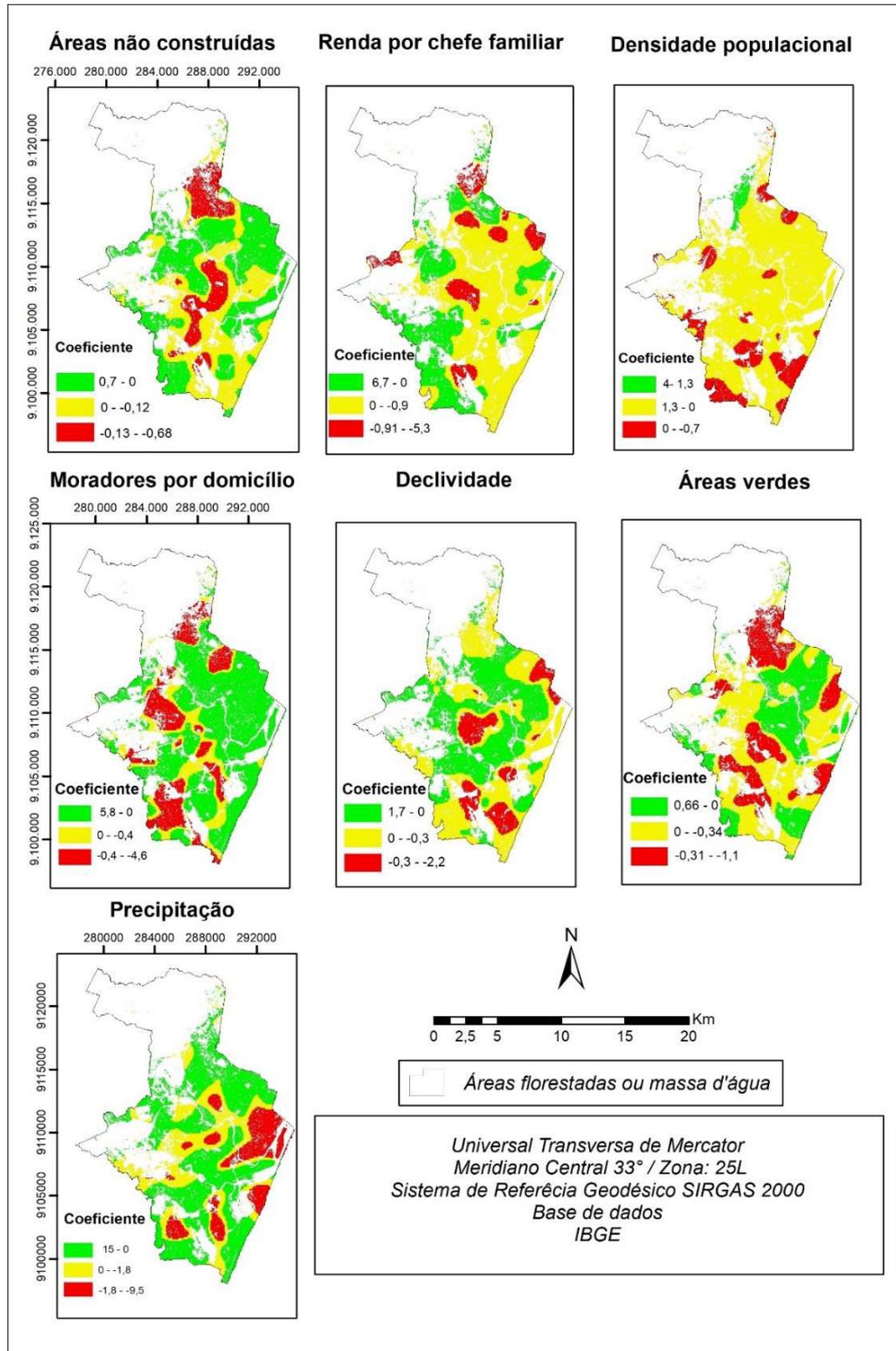
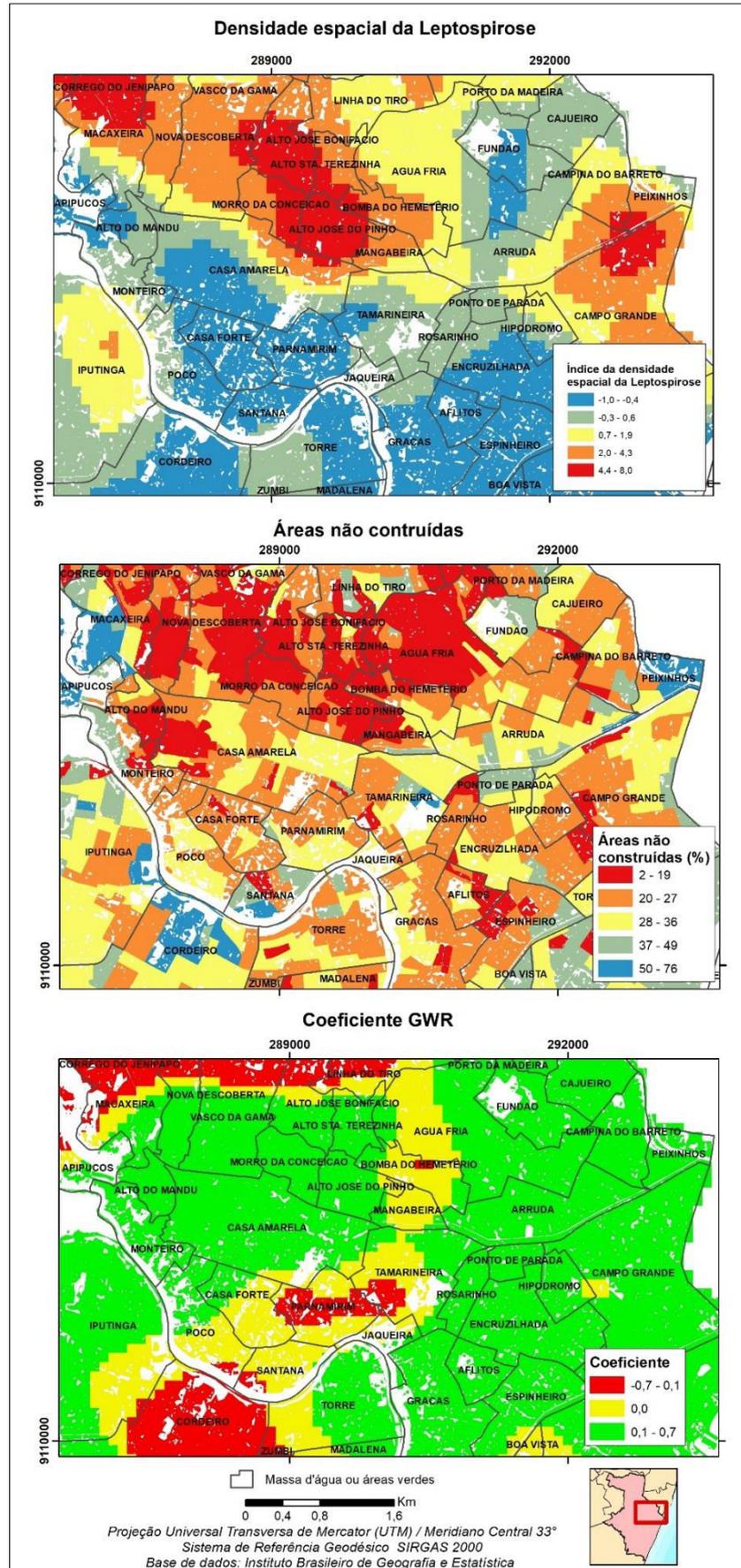


Figura 30 - Densidade espacial da Leptospirose, índice de áreas não construídas e coeficiente de GWR, na zona norte e oeste do Recife.



Algumas considerações podem ser levantadas como a de Pinto-Ferreira *et al.* (2019) que verificaram que as áreas onde houveram as maiores concentrações de edificações (e menor concentração de áreas não construídas) apresentaram um maior número de notificações de leptospirose em cachorros. Mesmo que não tenha sido um trabalho a partir de pessoas infectadas, esse trabalho pode-se tomar por base dado ao fato de poder existir a transmissão de leptospirose de outros mamíferos, como o cachorro, para humanos, como abordam Polachini e Fujimori (2015).

Embora as áreas com menor proporção de áreas não construídas tenham registrado alta incidência da leptospirose, outros núcleos apresentaram alta incidência porém maiores proporções de áreas não construídas, se comparado com os vizinhos (eixo central do município). Por ser uma doença que demanda fluxo de água para sua disseminação, geralmente as áreas próximas aos cursos d'água são mais vulneráveis a proliferação da leptospirose. Uma das hipóteses é que como a instalação de domicílios próximo dos cursos d'água torna-se um grande empecilho - devido às constantes variações diárias e sazonais dos níveis de águas fluviais e marinhas - a densidade de edificações nessas áreas de planície de inundação é menor que as áreas dos morros recifenses, com presença de áreas não construídas. É importante ressaltar que as áreas nobres do município, também possui baixo índice de áreas não construídas edificações, porém apresentam baixos índices da doença.

No que se refere a relação entre a renda do chefe familiar e a taxa de leptospirose, foi identificado que 60% das unidades espaciais apresentaram coeficientes negativos, apresentando coerência com o coeficiente apresentado no método OLS. Não muito diferente de outras doenças, a renda pode ser um dos fatores que apresentam maior relação com a distribuição espacial de algumas morbidades. Oliveira (2009) identificou que a renda apresentava uma forte correlação com a distribuição espacial da doença em Recife, evidenciando que esse fator era um dos mais importantes para a modelagem espacial da doença no município.

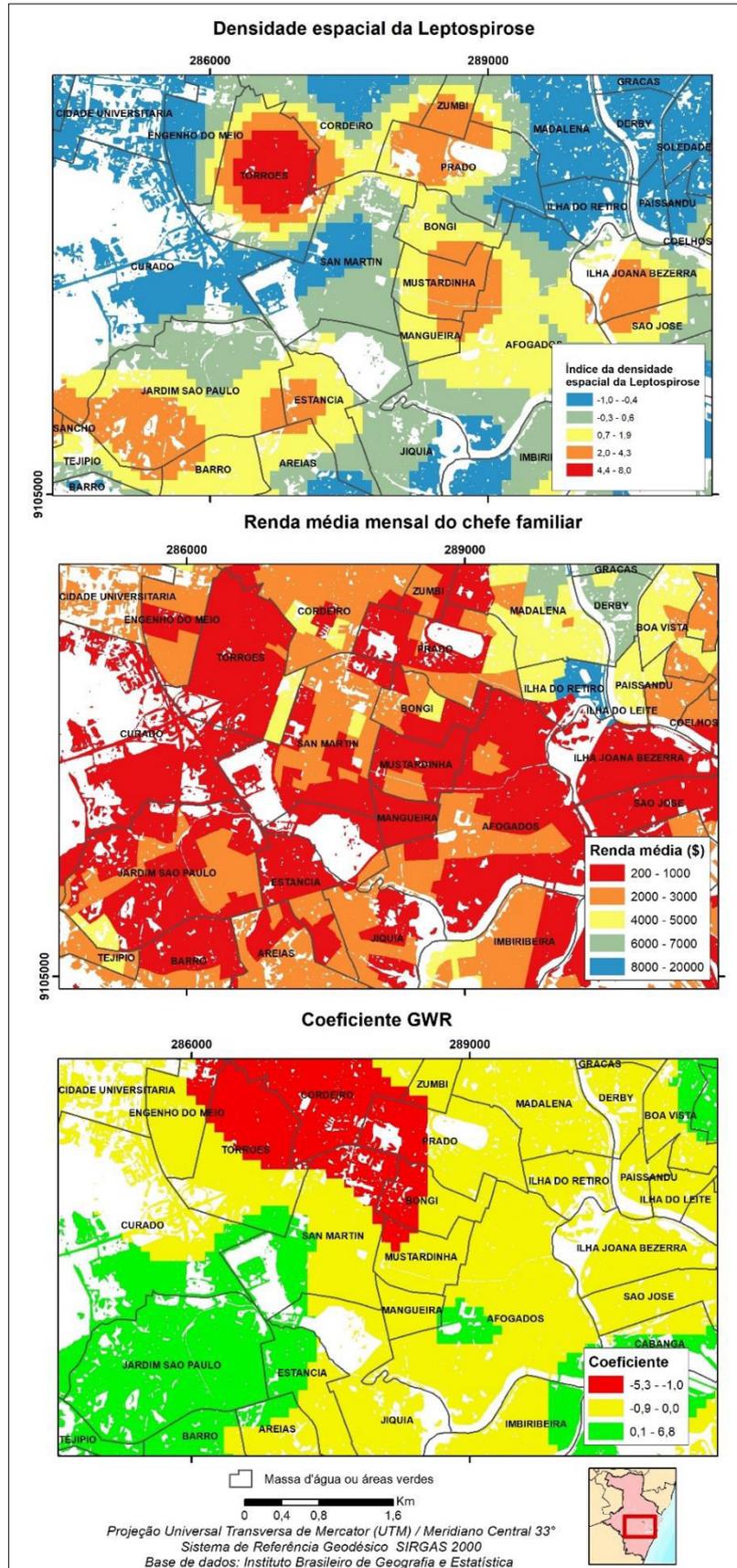
Reis *et al.* (2008) identificaram que as menores rendas *per capita* foi um dos fatores que influenciaram numa maior concentração espacial da leptospirose em Salvador. Também em Salvador, Dias *et al.* (2007) identificaram que o número de pessoas infectadas estava relacionado às suas condições salariais. Os autores levantaram alguns questionamentos acerca de um certo “endemismo” da doença em regiões específicas da cidade, que inevitavelmente apresentam baixas condições de infraestrutura urbana e socioeconômicas. Considerando outras cidades de outros países, Briskin (2016) também identificou que em Porto Rico a renda está associada a distribuição espacial da doença.

É válido ressaltar que algumas regiões apresentaram coeficientes positivos com a leptospirose. Dentre essas regiões, no oeste do município, algumas unidades espaciais que não apresentaram rendas inferiores que seus vizinhos apresentaram maior concentração da doença. Outros casos de coeficientes positivos encontrados podem estar relacionados ao fato de que embora as unidades espaciais apresentem níveis de renda abaixo que seus vizinhos, o índice de concentração da doença foi menor que as unidades adjacentes, como em alguns núcleos nos bairros de Jardim São Paulo, Barro e Estância (Figura 31).

Dentre as variáveis analisadas, a que apresentou maior uniformidade e coerência com os resultados do método OLS foi a densidade populacional. Exceto numa pequena área ao norte do município, as demais unidades espaciais apresentaram coeficientes positivos, evidenciando que quanto maior a densidade populacional da unidade, maior a concentração da doença. De fato, Barcellos e Sabroza (2001) abordaram sobre esse fator estar intrinsecamente ligado à população de ratos, uma vez que há maior disponibilidade de comida para os roedores. Considerando as variações temporais da leptospirose, Gracie et al. (2014) identificaram que a densidade populacional é um fator que está correlacionado tanto em períodos endêmicos e epidêmicos no Rio de Janeiro.

Em relação ao número de pessoas por domicílio e sua relação com a leptospirose, podemos identificar que na maior parte do município, as unidades espaciais mostraram valores positivos, como foi apresentado no método OLS. É importante ressaltar que praticamente nas unidades espaciais onde existiram coeficientes negativos, houve algum tipo de notificação da doença, mesmo com um índice de pessoas por domicílio menor que as unidades espaciais vizinhas. Soares *et al.* (2010) identificou que esse fator em São Paulo apresentou uma relação diretamente proporcional com a leptospirose. Esse é um dos fatores curiosos a serem abordados. Embora a leptospirose não seja infectada de pessoa para pessoa, um maior número de pessoas num domicílio que possuem características propícias para manifestação da doença, favorece que todas as pessoas possam ser infectadas. Podendo assim, ser um fator de relação indireta.

Figura 31 - Densidade espacial da Leptospirose, renda média mensal do chefe familiar e coeficiente do GWR, na zona oeste do Recife.



No que concerne a relação entre a declividade e a leptospirose, podemos identificar algumas particularidades na relação entre declividade e leptospirose. Na porção norte do município (Figura 32) alguns bairros situados nos morros com alta declividade (Alto José Bonifácio, Alto José do Pinho, Morro da Conceição) apresentaram altas notificações da doença, resultando em coeficientes positivos nessas áreas. Por outro lado, setores censitários nos bairros de Campina do Barreto, Campo Grande e Peixinhos, apresentaram alta densidade espacial da doença associada a áreas de baixa declividade, resultando em coeficientes negativos.

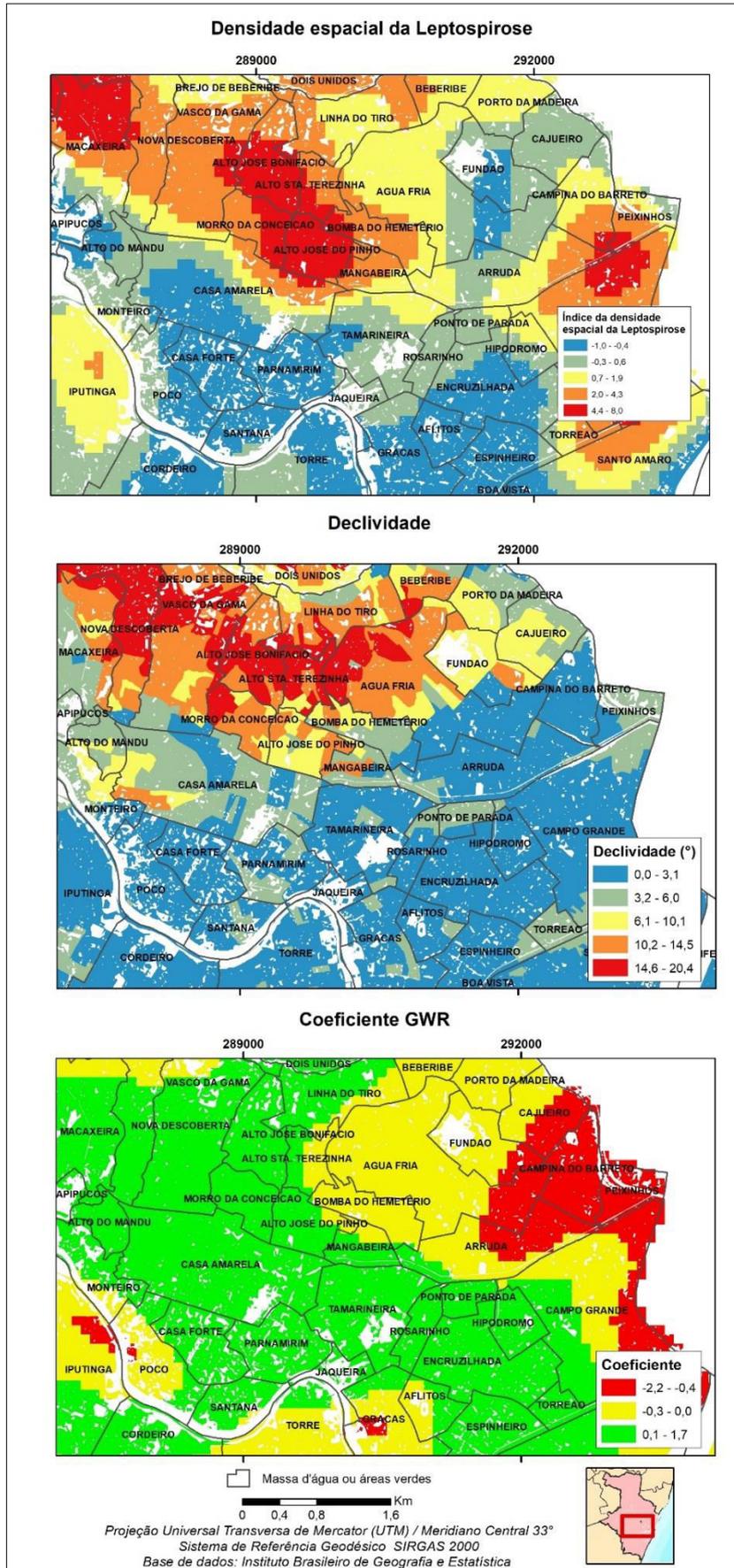
Levando em consideração todas as unidades espaciais, 48% apresentaram valores menor que 0, enquanto 41% de valores positivos e 11% de valores nulos. Isso evidencia que a maioria das unidades espaciais apresentam correlação com o resultado apresentado no método OLS. De modo geral, os valores positivos dos coeficientes estavam em áreas planas com índices de notificação da leptospirose baixas ou não existentes. Essa relação entre declividade e a distribuição espacial da leptospirose é discutida em alguns trabalhos. Ferreira e Ferreira (2016) apresentaram que as maiores densidades espacial da doença, concentravam em áreas com menor declividade, em São Paulo.

Em Recife, embora as características das aglomerações subnormais possuam abrangências que se situam tanto em áreas de morros como em regiões mais planas, como os manguezais, a concentração da leptospirose tende a ocorrer nas aglomerações subnormais com maior susceptibilidade a alagamentos, ou seja, em áreas mais planas.

As áreas verdes e sua associação com os casos de leptospirose possuem semelhanças entre os métodos OLS e GWR, uma vez que 62% das unidades espaciais dos coeficientes apresentaram valores negativos. Esse fator tende a se comportar de maneira diferenciada de acordo com a localização espacial. Gutiérrez et al. (2019) mostraram que existiu uma relação positiva entre cobertura vegetal e casos de leptospirose na Colômbia. Segundo os autores, a existência da cobertura vegetal favorece a uma produção de chuvas maior e, conseqüentemente maior probabilidade de contaminação, devido a maior disponibilidade de água.

No caso de Recife, essa relação pode ter se mostrado com uma relação inversa devido ao fato das áreas de aglomeração subnormal apresentarem índices de vegetação baixa ou nulas. Isso significa que as áreas verdes podem estar associadas de maneira indireta com os casos de leptospirose.

Figura 32 - Densidade espacial da Leptospirose, declividade média e coeficiente do GWR, na zona norte do Recife.



Considerando a precipitação e sua relação com a leptospirose, no método GWR, apenas 38% das unidades espaciais apresentaram valores negativos, enquanto 51% valores positivos. A maioria das unidades espaciais está em concordância com o método OLS.

Essa influência da precipitação e casos de leptospirose é amplamente discutida nos trabalhos e pesquisas acadêmicas. Mohammadinia *et al.* (2019) identificou, através do método GWR, que os maiores valores dos coeficientes – referidos a precipitação e leptospirose – eram positivos, no Irã. Chaiblich *et al.* (2017) também identificaram que a sazonalidade da doença estava intrinsecamente ligada as variações de precipitação no Rio de Janeiro.

Como a leptospirose é uma doença de veiculação hídrica, sua disseminação está associada a áreas de inundação, assim como regiões onde existem maior quantidade precipitada. Tal fato justifica os valores positivos dos coeficientes encontrados, tanto no método OLS e GWR.

5.3 TUBERCULOSE

Ao todo foram geocodificados 4.265 casos de Tuberculose distribuídos espacialmente no município de Recife entre os anos de 2015 e 2017 (Figura 33). Considerando a Tabela 5, os 10 bairros com maiores notificações concentram 35% de todos os casos notificados no município. O bairro de Água Fria registrou o maior caso de notificações, com 276 pessoas infectadas com a doença, com diferença superior a 100 casos em relação ao segundo bairro com maior registro (Boa Viagem).

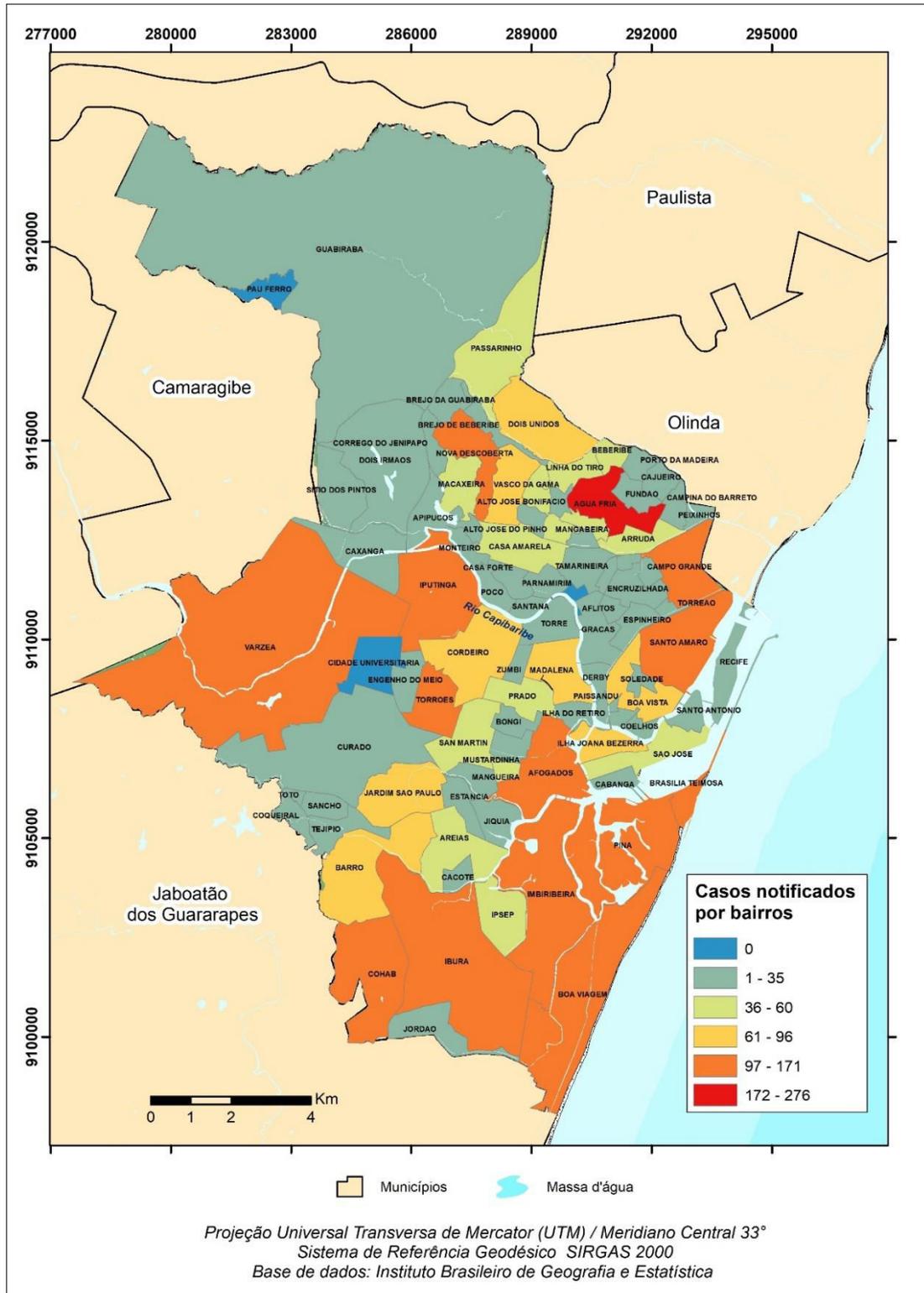
Tabela 5 -Os 10 bairros com maiores registros de Tuberculose entre 2015 e 2017.

Bairros	Casos notificados de Tuberculose
Água Fria	276
Boa Viagem	171
Afogados	161
Campo Grande	147
Santo Amaro	142
Ibura	137
Cohab	134
Brasília Teimosa	121
Várzea	121
Iputinga	118

A distribuição espacial da tuberculose por bairros, em Recife, não apresenta núcleos de alta concentração de casos notificados de maneira concentrada (Figura 33). Por outro lado, alguns bairros situados na margem esquerda do Capibaribe apresentam poucas notificações de

peças infectadas. Tal aspecto é válido destacar, visto que esses mesmos bairros (Santana, Poço da Panela, Parnamirim, Casa Forte) apresentaram poucos registros de notificação das outras doenças analisadas (Zika, Chikungunya e Leptospirose).

Figura 33 - Casos notificados de Tuberculose entre 2015 e 2017, de acordo com os bairros em Recife.



5.3.1 Avaliação da eficiência dos modelos

Quando analisados os modelos, algumas particularidades merecem ser destacadas em relação ao tipo de cálculo da variável dependente, tipo de unidade espacial e método para estimar os casos de Tuberculose (Figura 34).

Considerando o R^2 ajustado, podemos observar dentre os modelos 1 e 3, o modelo 2 apresentou o melhor desempenho, cuja unidade espacial foram os setores censitários sem áreas verdes. É válido ressaltar que as variações entre o R^2 não apresentaram valores significativos entre os setores censitários sem áreas verdes e quadrícula de 100 metros, apresentando 0,09 e 0,08 (OLS); 0,77 e 0,75 (GWR), respectivamente.

Figura 34 - Modelos categorizados de acordo com o tipo de unidade espacial, cálculo da variável dependente e método para estimar os casos de Tuberculose. O campo para comparação possui o R^2 ajustado, a soma do valor absoluto dos resíduos e o critério de informação Akaike.

Modelo 1		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		casos por setor censitário	0,04	938	3.697
GWR	500	casos por setor censitário	0,52	758	36
Modelo 2		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		casos por setor censitário - sem áreas verdes	0,09	1.269	5.038
GWR	500	casos por setor censitários - sem áreas verdes	0,77	763	-0,52
Modelo 3		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		casos por quadrícula de 100 m	0,08	9.217	53.520
GWR	500	casos por quadrícula de 100 m	0,75	2787	2.796
Modelo 4		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		densidade de Kernel com raio de 500m por quadrícula de 100m	0,51	8.379	40.774
GWR	500	densidade de Kernel com raio de 500m por quadrícula de 100m	0,89	3.260	4.267
Modelo 5		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		densidade de Kernel com raio de 700m por quadrícula de 100m	0,50	8.595	40.700
GWR	700	densidade de Kernel com raio de 700m por quadrícula de 100m	0,9	3.364	7.000
Modelo 6		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		densidade de Kernel com raio de 900m por quadrícula de 100m	0,48	8.873	41.038
GWR	900	densidade de Kernel com raio de 900m por quadrícula de 100m	0,89	3.525	10.340

Observando o Modelo 1, percebemos que o método de unidade espacial associado ao número de casos de pessoas infectadas por Tuberculose tende a ter um coeficiente de determinação (R^2) baixo, (0,04). Magalhães (2014) apresentou resultados semelhantes de coeficiente de determinação de 0,04, utilizando variáveis independentes, como: proporção de

responsável com renda maior que 1 salário-mínimo e menor que 2 salários-mínimos; proporção de analfabetos; proporção de domicílios com pessoas que moram sozinhas e renda média do responsável, no Rio de Janeiro.

Observando os diferentes tipos de cálculos da variável dependente e os *bandwidths* utilizados nos métodos do GWR, é evidente que os melhores resultados se encontram quando o método de densidade de Kernel foi adotado. Considerando os Modelos 4, 5, e 6, podemos identificar que as diferenças entre os parâmetros não são consideráveis. Tal fato evidencia que existem dois grupos que se dividem de acordo com a eficiência em estimar os casos e o método de cálculo da variável dependente adotado. Os Modelos 1, 2 e 3, cujo cálculo da variável dependente foi o número de casos por unidade espacial e com menor eficiência em estimar as notificações que o grupo composto pelos Modelos 4, 5 e 6 – caracterizados por utilizarem a densidade de casos notificados de Tuberculose por unidade espacial.

A suavidade espacial promovida pela espacialização dos casos de Tuberculose por meio da densidade de Kernel tende a acompanhar as variações espaciais das variáveis independentes utilizadas nos modelos, resultando numa melhor eficiência dos modelos em estimar as maiores concentrações espaciais da doença.

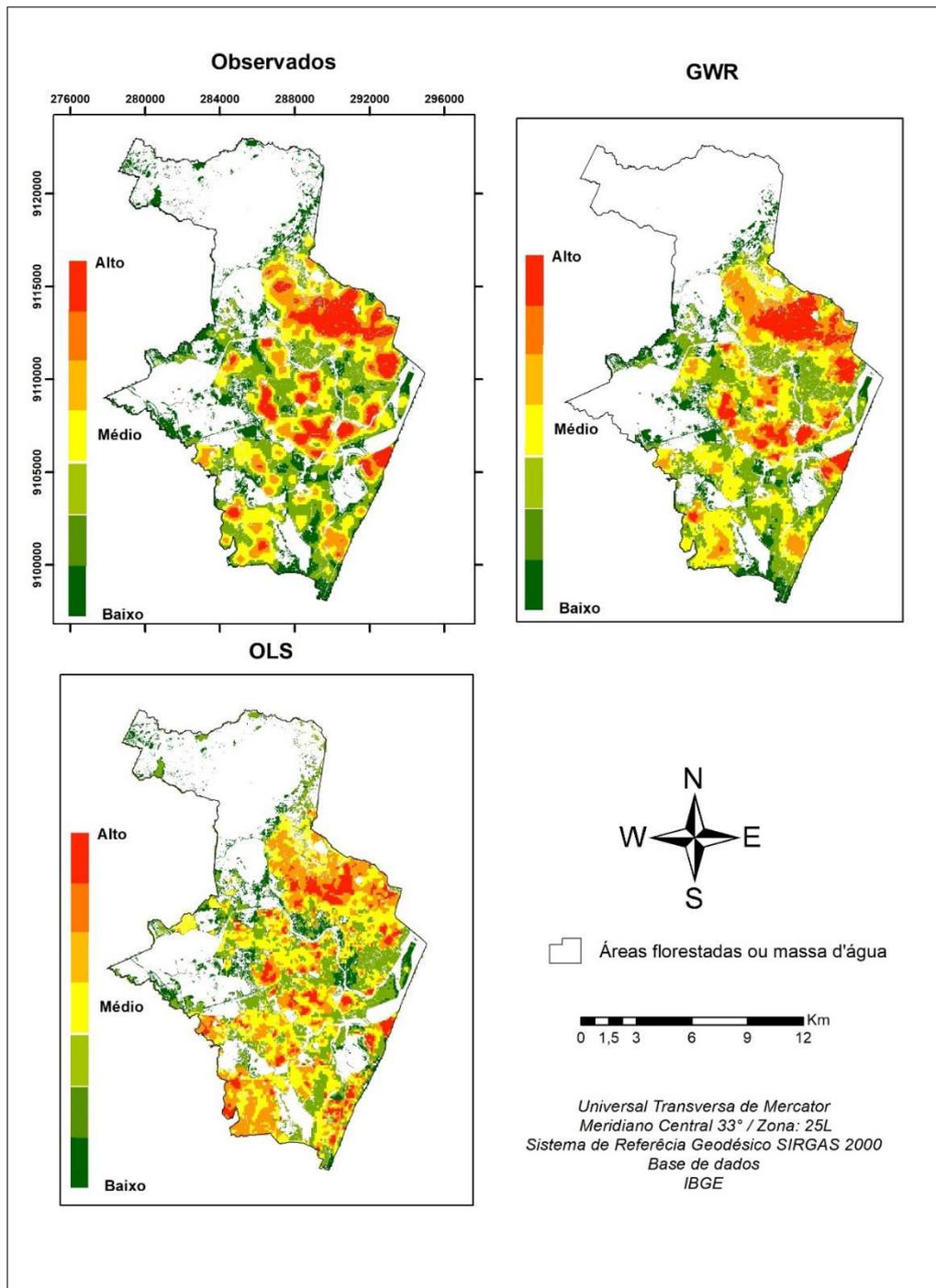
Analisando as somas residuais e o AICC para cada modelo, percebe-se que o método GWR apresenta melhor eficiência em todos os modelos analisados. É importante ressaltar que os dois métodos são extremamente importantes para uma análise do comportamento espacial da doença. O GWR, é caracterizado por utilizar regressões para um *bandwidth* determinado (seja ele por uma distância previamente estabelecida ou por uma quantidade determinada de unidades espaciais). O OLS tende a apresentar as correlações gerais entre as variáveis utilizadas no modelo com a doença a ser estudada.

Observando a Figura 34, podemos identificar que as variações no *bandwidth*, não alteraram de maneira significativa os coeficientes (R^2) dos Modelos 4, 5 e 6. Devido às semelhanças dos parâmetros dos modelos 4, 5 e 6, foi definido que o modelo que apresentasse o maior R^2 do método OLS, fosse utilizado para as análises posteriores. No caso presente, foi utilizado o Modelo 4.

Através da análise da distribuição espacial dos casos observados e os casos estimados de Tuberculose percebe-se em todos os mapas uma alta concentração dos casos de Tuberculose no nordeste do município de Recife (Figura 35). Considerando os dados observados e o método de densidade de Kernel adotado com raio de busca de 500 metros, as maiores concentrações indicam 282 casos notificados por km^2 , nos morros situados na zona norte recifense.

Outro aspecto a ser levantado é que o método OLS apresentou as maiores diferenças, quando comparado com o método GWR. É importante ressaltar que ambos os modelos tendem a identificar os menores índices das notificações da doença no município de Recife. Os núcleos na porção central do município (zona norte recifense) tendem a concentrar-se as menores densidades da doença, como pode ser observado nos métodos OLS e GWR.

Figura 35 - Densidade da estimativa (método GWR e OLS) e casos observados da Tuberculose em Recife.



Por outro lado, algumas regiões – principalmente na porção norte do município – o método OLS subestimou as notificações de Tuberculose. De toda forma, ambos os métodos mostraram um bom desempenho em estimar as variações espaciais da doença. É importante ressaltar que o método GWR, por ser um método de regressão local, é utilizado apenas como um modelo exploratório, uma vez que os coeficientes não podem ser utilizados para outras cidades, devido as suas variações espaciais. Já o método OLS, embora apresente maiores diferenças, oferece coeficientes que pode ser utilizado para estimar a Tuberculose em outros municípios, por exemplo.

5.3.2 Análise dos resultados do método OLS

De acordo com os parâmetros do método OLS, a *Variance Inflation Factor* (VIF) para todas as variáveis independentes utilizadas no Modelo 4 foram consideradas baixas (< 7). Tal fato evidencia que os parâmetros utilizados não são redundantes no modelo. O *Robust Probability* com valores próximos de 0, mostrou que houve autocorrelação espacial para todas as variáveis e existe uma relação com os casos de Tuberculose.

O valor da estatística Koenker para todas as variáveis foi estatisticamente significativo ($p < 0,01$), mostrando que os relacionamentos modelados não são consistentes devido à não estacionariedade. Isso significa que as relações da variável dependente com as outras variáveis ora podem ser positivas, ora negativas, evidenciando que os coeficientes se alternam de acordo com a região estudada. Assim, evidencia-se que métodos de regressão geográfica ponderada, como o GWR, podem ser utilizadas na estimação dos casos de Tuberculose.

Duas variáveis mostraram coeficientes positivos com os casos de Tuberculose: pessoas por domicílio e densidade populacional (Tabela 6). Cinco variáveis mostraram coeficientes negativos: renda do chefe familiar, precipitação, áreas não construídas, declividade e áreas verdes.

Tabela 6 - Variáveis independentes utilizadas e suas relações com a Tuberculose de acordo com o coeficiente, *Robust Probability* e VIF no método OLS.

Variáveis	Coeficiente	Robust Probability	VIF
Pessoas por domicílio	0,114124	0,000000*	1,251669
Renda do chefe familiar	-0,155030	0,000000*	1,154417
Precipitação	-0,084784	0,000000*	1,141215
Áreas não construídas	-0,049166	0,000000*	1,572145
Declividade	-0,030622	0,000000*	1,240199
Densidade populacional	0,623429	0,000000*	1,562560
Áreas verdes	-0,185464	0,000000*	2,527576

O comportamento de cada coeficiente irá corresponder de que maneira cada variável independente influenciará no aumento da doença. Partindo desse princípio, percebemos que a distribuição espacial da Tuberculose é mais sensível à densidade populacional, enquanto a variável declividade é a que oferece menos impacto nos casos de Tuberculose. Apesar dessas variações é imprescindível avaliar o comportamento, mesmo que sutil, da relação dessas variáveis com a doença.

5.3.3 Análise dos resultados do método GWR

As relações entre o conjunto das variáveis independentes com os casos de Tuberculose podem variar de acordo com a sua localização geográfica. Como o princípio do método GWR é estimar as correlações para uma região pré-determinada por um *bandwidth*, é possível identificar quais regiões tiveram uma correlação maior do conjunto das variáveis independentes com a variável dependente.

É importante destacar que os valores do R^2 distribuídos espacialmente diferem-se do R^2 ajustado apresentado na Tabela 6 devido às metodologias diferentes de ambos. No último, os valores para o seu cálculo dizem respeito a divisão entre a soma quadrada dos resíduos e a soma quadrada dos valores observados de todas as unidades espaciais utilizadas, enquanto no primeiro, os cálculos são feitos apenas com as unidades espaciais e seus vizinhos, definidos de acordo com o *bandwidth* adotado.

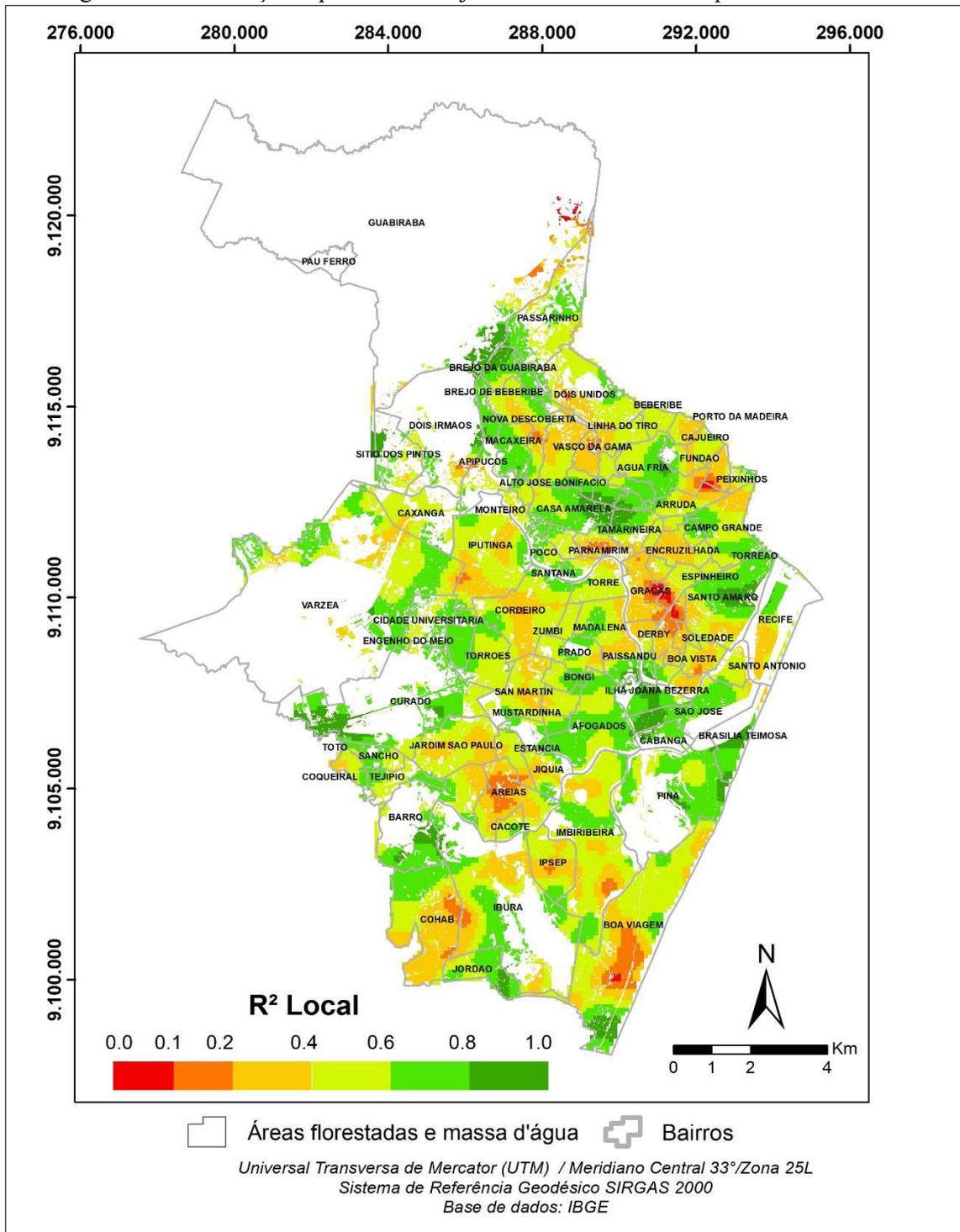
As variações espaciais do R^2 não apresentam um padrão espacial significativo. 2.865 unidades espaciais não foram consideradas no cálculo devido a impossibilidade dos cálculos do método do GWR não terem vizinhos suficientes para gerar os coeficientes de relação entre as

variáveis e os casos de Tuberculose. 15.655 unidades espaciais apresentaram algum valor de R^2 referentes ao *bandwidth* utilizado (500 metros). Dentre essas unidades espaciais, 12.018 apresentaram coeficientes iguais ou superiores a 0.4, representando 77% de todas as unidades espaciais que apresentaram algum valor.

É válido destacar que nas áreas onde foram apresentadas as maiores densidades da doença, houveram diferenças nos valores dos coeficientes espaciais. Tal fato pode inferir que outros condicionantes, além dos que foram utilizados nos modelos, podem estar influenciando na distribuição espacial da doença nessas localidades.

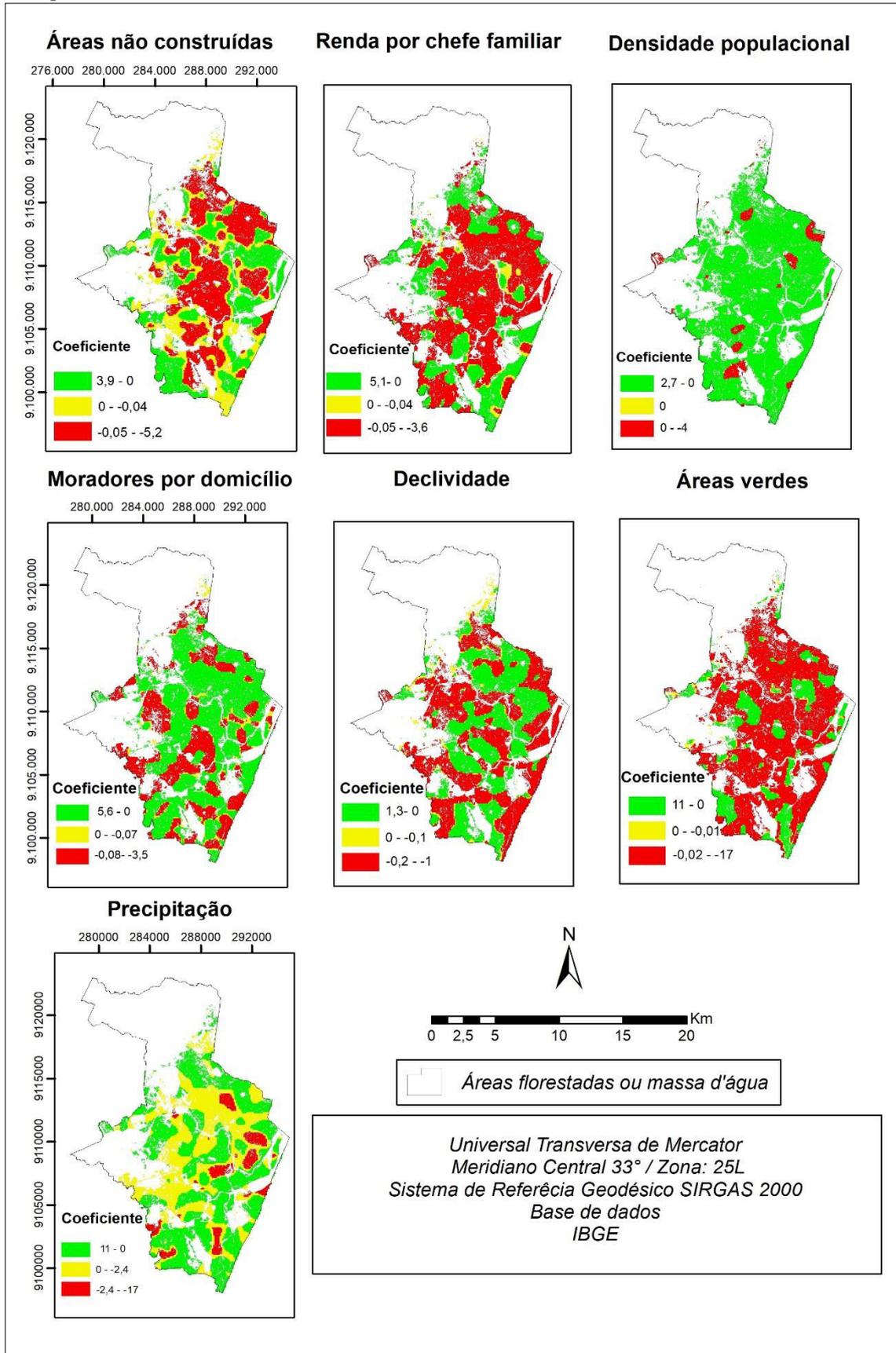
Uma observação a ser considerada é que, de modo geral, as regiões de maiores densidades de casos de Tuberculose apresentaram os maiores coeficientes espaciais. Partindo desse pressuposto, é importante identificar quais fatores não foram considerados no modelo que impedem a disseminação da doença nessas áreas, resultando em coeficientes menores.

Figura 36 -Distribuição espacial do R² ajustado no método GWR para a Tuberculose.



A análise dos coeficientes originadas pelo método GWR para cada variável, permite que sejam analisados os comportamentos espaciais de cada fator e sua relação local com a distribuição espacial da doença. Podemos identificar na Figura 37 que as variações locais dos coeficientes para cada variável independente podem ser complexas e variáveis de acordo com a região estudada porém pode apresentar em sua maioria, convergências com os coeficientes do método OLS.

Figura 37 - Distribuição espacial dos coeficientes do método GWR para cada variável inserida no modelo para estimar a Tuberculose.



Em relação às áreas não construídas, aproximadamente 65% das unidades espaciais apresentaram valores negativos em seus coeficientes no método GWR. De modo geral, as áreas onde predominam os coeficientes negativos concentram-se nas áreas onde a notificação de Tuberculose apresentou alta densidade. Esses dados corroboram com os coeficientes apresentados pelo método OLS, onde as áreas que concentram as maiores densidades de edificações, apresentam as maiores taxas da doença. Partindo dessa análise, Clarck et al., (2002) identificaram que alta densidade de casas pode ocasionar numa maior susceptibilidade a contrair Tuberculose. Outro ponto levantado por PC et al., (2013) é que um maior distanciamento entre as casas promove maior circulação do ar e maior exposição à luz solar, melhorando a qualidade do ar. Em regiões onde a densidade de edificações é alta, a qualidade do ar tende a ser mais poluída, aumentando a susceptibilidade de contaminação de Tuberculose.

Uma das principais causas para que a densidade habitacional esteja relacionada a uma maior probabilidade de disseminação da doença está na proximidade entre uma pessoa infectada e as pessoas que o circundam. De toda forma, a Tuberculose não apresenta alto poder de transmissibilidade, visto que quando uma pessoa saudável é exposta ao bacilo da TB, tem apenas 30% de chance de infectar-se, dependendo do grau de exposição (proximidade, condições do ambiente e tempo de convivência), da infectividade do caso índice (quantidade de bacilos eliminados, presença de caverna na radiografia de tórax) e de fatores imunológicos individuais (MINISTÉRIO DA SAÚDE, 2019).

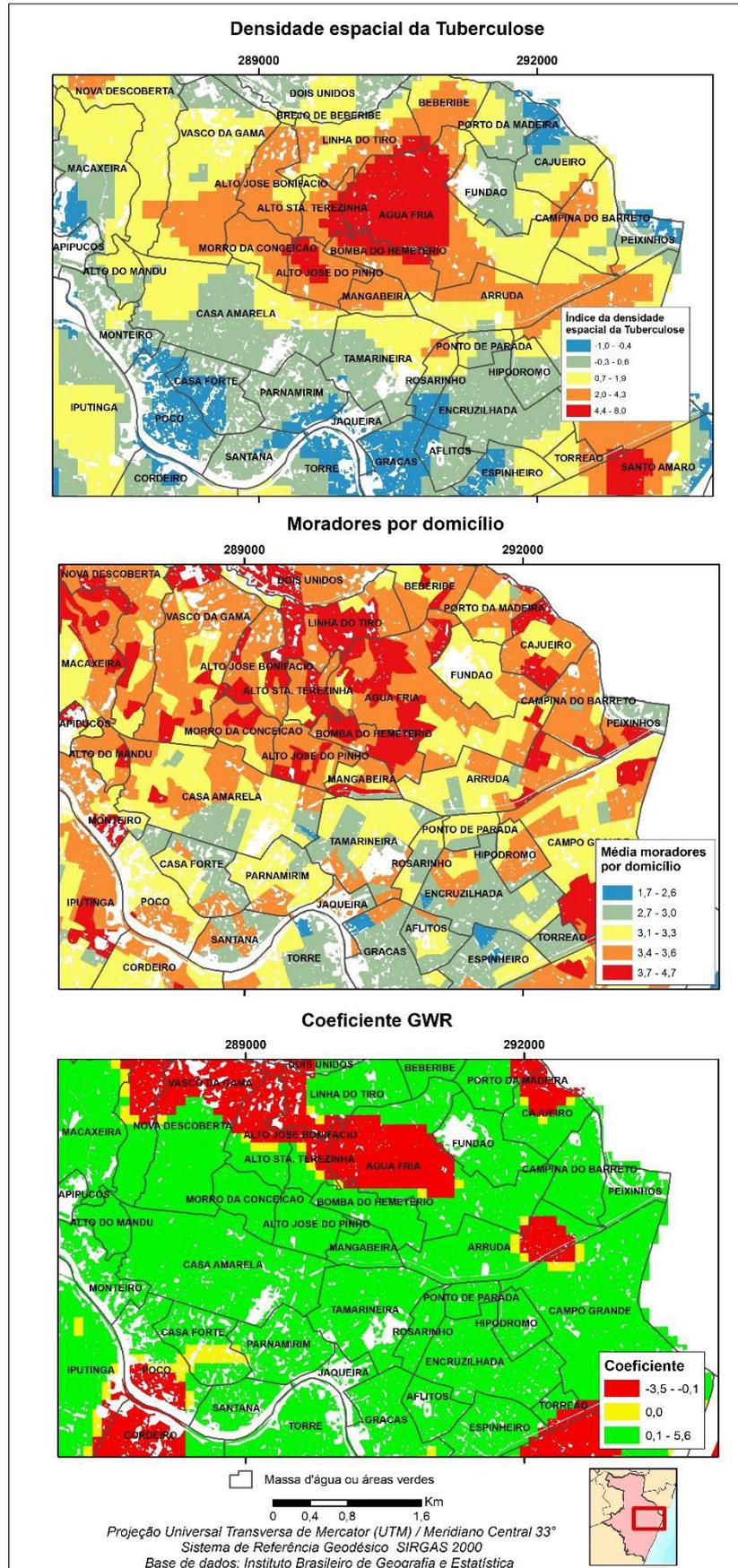
No que concerne a relação da distribuição espacial da Tuberculose com a renda média mensal do chefe familiar é possível identificar algumas particularidades. Das 15.655 unidades espaciais que apresentaram algum valor de coeficiente, 10.417 unidades espaciais apresentaram coeficientes negativos da relação renda x tuberculose, o que resulta em aproximadamente 66% das unidades espaciais. Esse resultado é próximo ao que foi identificado pelo método OLS, que apresentou um coeficiente de -0,15. Magalhães (2014) identificou, pelo método OLS que os casos de Tuberculose, no Rio de Janeiro também apresentaram coeficientes negativos. Utilizando outros modelos para observação, Silva (2016) verificou que o grupo de chefes de família sem renda mensal, elevava em 18% o risco de contaminação da doença, em Olinda.

Baseando-se nos coeficientes espaciais do GWR percebe-se que em grande porção do município de Recife existe uma relação diretamente proporcional entre densidade populacional e densidade de notificações de Tuberculoses. 14.719 das unidades espaciais apresentaram valores positivos em seus coeficientes, o que representa 94% das unidades espaciais calculadas. Esse fator corrobora com os coeficientes apresentados no método OLS. Coincidentemente, foi a condicionante que apresentou maior valor no coeficiente, inferindo que os casos de

Tuberculose são extremamente sensíveis às variações de densidade populacional. Com resultados similares Silva (2016) identificou que as densidades populacionais apresentaram maiores correlações com a densidade de casos de Tuberculose que as variáveis de renda e nível de escolaridade.

Considerando a relação entre moradores por domicílios e densidade espacial da tuberculose algumas particularidades merecem destaque. Alguns coeficientes negativos foram presentes em algumas unidades espaciais. A causa dos coeficientes negativos se mostrarem presentes em alguns bairros (Vasco da Gama, Alto José do Bonifácio e Água Fria) está relacionada ao fato dessas localidades apresentarem densidades espaciais da doença maiores que seus vizinhos, mas com taxas de moradores próximas ou inferiores que as unidades espaciais vizinhas (Figura 38).

Figura 38 - Densidade espacial da Tuberculose, moradores por domicílios e coeficiente do GWR, na zona norte do Recife.



Contudo, os valores positivos contabilizaram 60% de todas as unidades espaciais calculadas. Essa predominância dos coeficientes positivos é perceptível quando analisamos a zona norte de Recife. Não diferente, o método OLS demonstrou valores positivos nessa relação diretamente proporcional. Não se pode negar que a aglomeração de pessoas em um domicílio com alguém infectado favorece a exposição de outras pessoas ao contágio. Contudo, ela não pode ser um fator determinante, se comparada com a densidade populacional por exemplo. Por isso, existe uma heterogeneidade na distribuição dos coeficientes de moradores por domicílios, com algumas regiões apresentando relações inversamente proporcional. Lobo *et al.*, (2016) explicitam que em alguns setores censitários localizados em condomínios residenciais para classe média alta possuem um alto número de pessoas por domicílios, aproximando dos valores encontrados em aglomerados subnormais.

Como alternativa a contrapor esses problemas os autores propõem o número de pessoas por quartos. Embora a discussão seja válida, em Recife, os bairros nobres recifenses possuem menos moradores por domicílio que os aglomerados subnormais.

As unidades espaciais onde apresentaram uma relação inversamente proporcional entre moradores por domicílio e casos notificados de tuberculose (o que representou a minoria em relação ao total das unidades espaciais estudadas), apresentaram as taxas de densidade da doença maior que seus vizinhos, porém com índices de moradores menor que o seu entorno.

Coincidindo com o coeficiente do método OLS, 60% das unidades espaciais apresentaram coeficientes negativos em relação à declividade. Apesar desse fato, ao ser analisada a variação espacial dos coeficientes do GWR, não se percebe um padrão espacial dos valores positivos e negativos, por isso tenha refletido num coeficiente do OLS baixo (-0,03). A declividade é um fator complexo e merece ser analisado sob vários enfoques. Estudos evidenciam que em grandes cidades, regiões de alta declividade são habitadas por comunidades carentes de infraestrutura urbana, alta densidade populacional e baixos níveis de renda e educação. Partindo desse princípio, estudos mostram que existe uma relação diretamente proporcional entre declividade e casos de tuberculose (MEDEIROS, 2013).

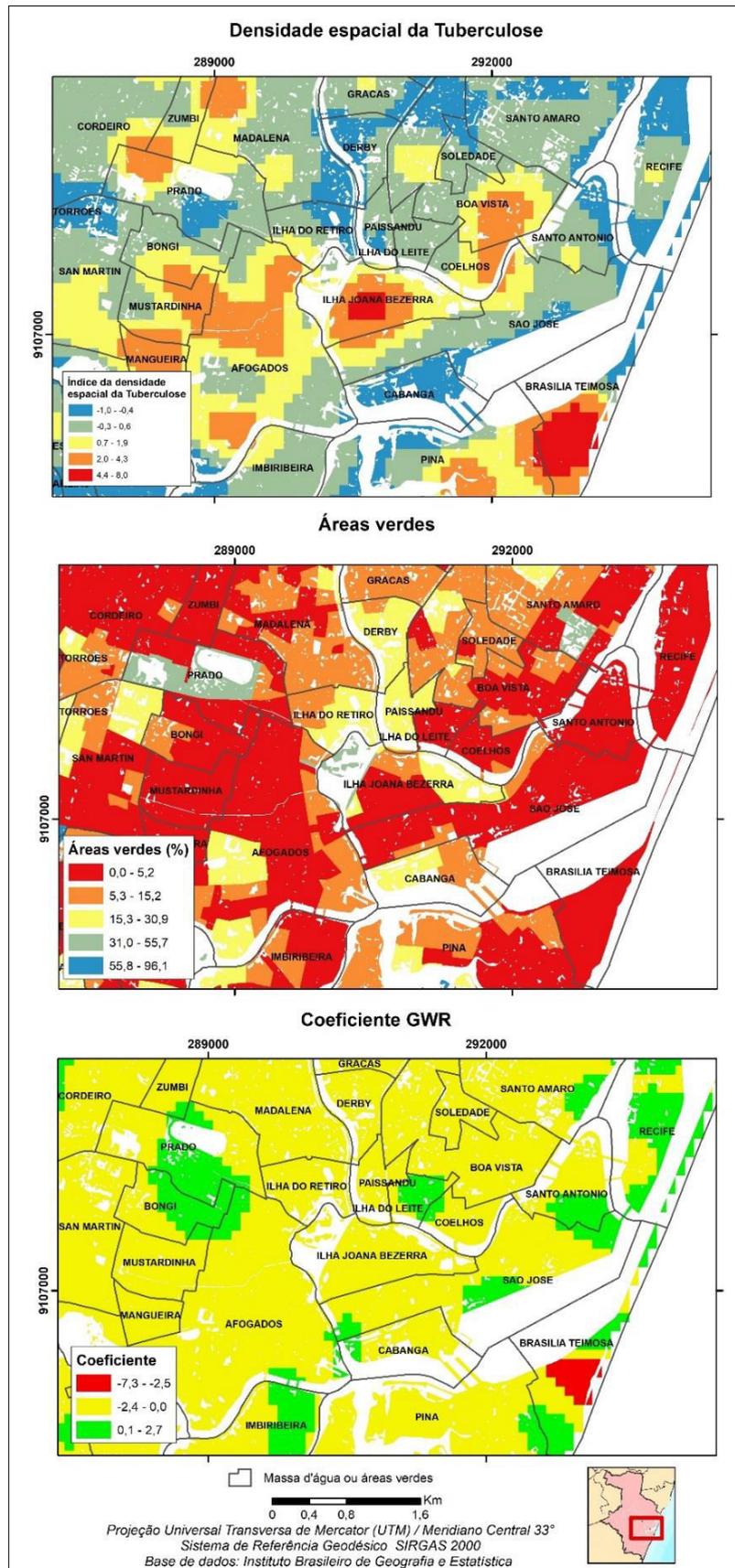
No entanto, como explicar os coeficientes negativos, nos métodos OLS e GWR, no município de Recife? Mansoer *et al.*, (1997) verificou que o aumento da altitude no Quênia, diminui o índice de pessoas infectadas por Tuberculose. Embora seja uma relação indireta entre declividade e altitude, além do baixo gradiente altimétrico que se encontra em Recife (~100 metros), pode-se inferir que a circulação do ar nessas áreas tende a circular mais livremente, melhorando sua qualidade do ar. É preciso considerar estudos futuros que possam corroborar tal afirmativa.

Cerca de 79% das unidades espaciais com coeficiente, apresentaram relações inversamente proporcionais entre a quantidade espaços verdes com a incidência de Tuberculose. Esse fato fica mais evidente quando observa-se na

Figura 39, os bairros situados na zona oeste do município. Os núcleos de alta incidência da doença nos bairros de Ilha Joana Bezerra, Mangueira e Brasília Teimosa apresentaram baixa porcentagem de áreas verdes. Liu et al., (2019) mostraram algumas evidências que comprovam que a existência de áreas verdes pode diminuir a disseminação de casos de Tuberculose. Dentre elas, está no fato de que espaços verdes diminuem a alta densidade populacional, melhora a qualidade do ar e da água (que está relacionada a disseminação de tuberculose) e, consequentemente diminuindo os casos de Tuberculose.

Dentre as variáveis analisadas, a precipitação foi a única variável que diferiu os seus resultados do método OLS, se considerada a maioria das unidades espaciais. 47% das unidades espaciais apresentaram valores negativos, contra 53% dos coeficientes positivos. Os coeficientes negativos oriundos do método GWR localizavam-se próximos de onde as incidências da Tuberculose eram maiores. Considerando as regiões onde os valores positivos foram encontrados (menos precipitação, menos casos encontrados), podem existir outros fatores que são mais determinantes que evitem a disseminação da doença nessas áreas, como o litoral recifense que apresenta as maiores rendas dos chefes de família.

Figura 39 - Densidade espacial da Tuberculose, índice de áreas verdes e coeficiente do GWR, na zona oeste do Recife.



Considerando os coeficientes levantados pelo método OLS, percebe-se que existe uma relação inversamente proporcional entre a precipitação e casos de Tuberculose, em Recife. Kuddus *et al.*, (2019) verificou que maiores quantidades de precipitação estão associadas a diminuição das notificações de Tuberculose em Bangladesh, mesmo com um atraso das notificações de Tuberculose em relação à chuva. Beiravand *et al.*, (2016) verificaram que existe uma relação inversa entre chuva e incidência de Tubérculos, no Irã. Ou seja, existia uma maior concentração espacial de pessoas infectadas com Tuberculose nas regiões mais secas no país estudado.

Xiao *et al.*, (2018) explica que um dos possíveis motivos para essa relação está no fato de que a exposição demasiada ao ar seco reduz a produção do muco protetor do trato respiratório, reduzindo a resistência ao crescimento do *Mycobacterium tuberculosis*. Outra explicação, segundo os autores, é que a umidade favorece a ativação da Tuberculose latente, que é caracterizada por ser assintomática, mas que pode transforma-se em Tuberculose ativa.

5.4 DESENVOLVIMENTO TEÓRICO SOBRE A MODELAGEM ESPACIAL DE DOENÇAS

A abordagem do presente capítulo, está em enfatizar quais parâmetros coincidiram entre as doenças estudadas e como os procedimentos de criação dos modelos apresentaram semelhanças entre si.

Partindo desse princípio podemos analisar que os melhores modelos espaciais para estimar as doenças analisadas (Leptospirose, Zika/Chikungunya e Tuberculose), basearam-se em unidades espaciais regulares organizadas por células de 100x100 metros com cálculos da variável dependente utilizando o método de densidade de Kernel.

Em relação às unidades espaciais, embora o detalhamento de informação não se altere ao desagregar as informações contidas dos setores censitários, no presente caso, as condicionantes oriundas de dados matriciais (áreas verdes, declividade, áreas não construídas e precipitação) fornecem variações espaciais mais detalhadas, que se divididas em unidades espaciais menores, podem acompanhar a distribuição espacial das doenças.

Ao mesmo tempo que a unidade espacial menor pode melhorar o modelo, observando os modelos que demonstraram melhor desempenho, foi perceptível verificar que suas bases de cálculo da variável dependente era o método de densidade de Kernel. Destarte, é evidente que utilizar o número de casos por unidade espacial enrijece o modelo de modo que as variações entre as unidades espaciais são bruscas, diminuindo o coeficiente de determinação. Partindo

desse princípio, utilizar um método de densidade dos casos notificados permite suavizar as variações entre as unidades espaciais utilizadas no modelo.

Figura 40 - Modelos que apresentaram melhores eficiências nas suas estimativas, de acordo com suas respectivas doenças (Zika/Chikungunya, Leptospirose e Tuberculose).

ZIKA/CHIKUNGUNYA					
Modelo 5		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		densidade de Kernel com raio de 700m por quadrícula de 100m	0,367	9.717	45.767
GWR	700	densidade de Kernel com raio de 700m por quadrícula de 100m	0,92	2.743	7067
LEPTOSPIROSE					
Modelo 6		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		densidade de Kernel com raio de 900m por quadrícula de 100m	0,36	15.751	58.775
GWR	900	densidade de Kernel com raio de 900m por quadrícula de 100m	0,8	7.058	25.262
TUBERCULOSE					
Modelo 4		Variável dependente	Parâmetros		
Método	Bandwidth (m)		R2	S. Res	AICC
OLS		densidade de Kernel com raio de 500m por quadrícula de 100m	0,51	8.379	40.774
GWR	500	densidade de Kernel com raio de 500m por quadrícula de 100m	0,89	3.260	4.267

É importante levantar uma discussão acerca do objetivo da modelagem de dados espaciais destinadas a estimar doenças. Talvez, a maior utilidade da modelagem espacial de estimar a densidade de doenças é mostrar as regiões suscetíveis à contaminação de uma determinada doença e não estimar o local exato que uma pessoa foi contaminada. Por isso, não há necessidade - e nem é possível - que os coeficientes de determinação devam ser próximos de 1, e muito menos que as correlações das variáveis com as doenças perfeitamente alinhadas.

Destarte, as abordagens na área da geografia da saúde permitem, sob várias escalas, estimar regiões que estão suscetíveis a contaminação de determinadas doenças e, o fato de transformar os casos pontuais para uma espacialização a partir de densidades de casos, permite identificar quais condicionantes semelhantes permitem a disseminação ou a contenção de determinadas doenças. Por isso, o intuito em criar um modelo espacial de doenças não é prever o local exato da residência de uma pessoa contaminada, mas identificar quais características geográficas promoveram para essa ocorrência e em que dimensão um geógrafo possa determinar as áreas mais suscetíveis ou menos suscetíveis à proliferação de doenças.

Outro aspecto a ser abordado é o método de estimativa das doenças analisadas, seja por regressão global (OLS) ou local (GWR). Em todos os modelos analisados e em todas as doenças, o método GWR apresentou melhor eficiência em estimar as doenças. Isso não significa

que o método OLS possa ser descartado, muito pelo contrário. Ambos os métodos apresentam cálculos diferenciados que permitem analisar a variável dependente sob várias perspectivas.

Uma das características do GWR está no fato de que seus coeficientes variam espacialmente, impedindo que sejam utilizados para outras localidades, já que o coeficiente é um produto da interação entre a condicionante – ou o conjunto delas – e uma determinada doença naquela específica unidade espacial. É essa característica que o torna como um modelo exploratório. Uma vez que o R^2 do modelo GWR estiver mais próximo de 1, é possível analisar os coeficientes espacialmente permitindo identificar onde e qual o grau da relação entre a condicionante e a variável dependente. Por isso, considerando a análise dos resultados dos modelos aqui descritas, é possível afirmar que as doenças não apresentam uma relação homogênea espacialmente com nenhuma condicionante e, provavelmente seja um fenômeno que aconteça com inúmeras doenças.

Por outro lado, o método OLS oferece coeficientes gerados a partir da relação global entre as condicionantes e as doenças. Esses coeficientes permitem que se possa estimar a mesma doença em outras localidades com ausência de notificações, além de identificar qual o tipo de relação entre os condicionantes e as doenças analisadas ocorrem de maneira geral no espaço geográfico.

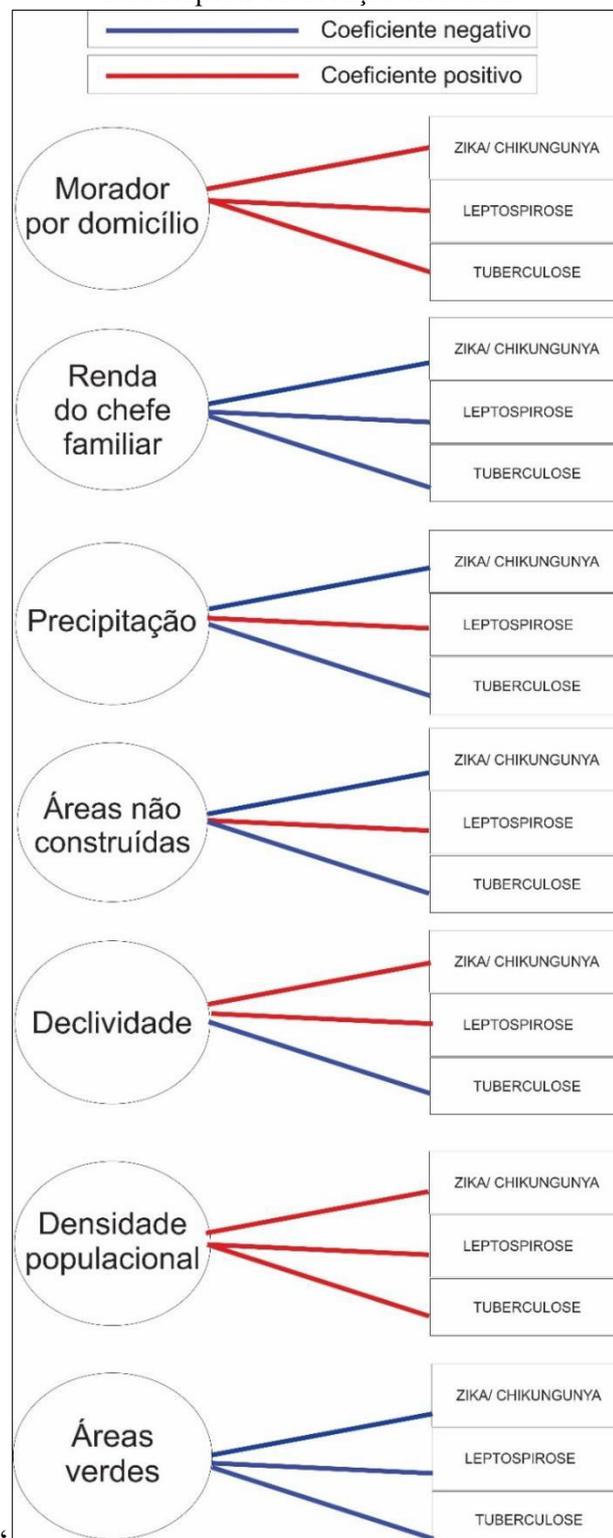
Outro aspecto, não menos importante, é a sensibilidade da qualidade das estimativas dos modelos em relação ao raio da densidade do Kernel adotado e o *bandwidth* adotado no GWR. O raio da densidade do Kernel refletirá no grau de abrangência espacial que aquele determinado fenômeno será representado no mapa. Quanto maior o raio, maior será sua expansão e suavidade visual. Tal efeito será crucial ao ser feito o cruzamento com as outras variáveis independentes do modelo.

O *bandwidth* adotado no método GWR é o parâmetro crucial para o desempenho do modelo. Quando se tem uma alta heterogeneidade espacial entre a doença e as variáveis independentes numa área a ser estudada (ora coeficientes positivos, ora negativos) menor deverá ser o *bandwidth* adotado para uma maior eficiência das estimativas. Isso se dá porque o *bandwidth* pode se adaptar a essas variações espaciais de coeficientes para um número limitado de unidades espaciais, bastando apenas que dentro do *bandwidth* exista uma boa correlação espacial entre a doença e as variáveis, seja o coeficiente positivo ou negativo.

Na Figura 41, podemos identificar as relações entre as condicionantes e as doenças analisadas, a partir dos coeficientes gerados pelo método OLS. Uma das primeiras análises a serem consideradas é que todas as condicionantes apresentaram autocorrelação espacial com

todas as doenças, alterando em alguns casos os tipos de relação entre as mesmas (diretamente ou inversamente proporcional).

Figura 41 - Variáveis utilizadas e os coeficientes do método OLS dos modelos que apresentaram melhor eficiência, organizados de acordo com a relação positiva (vermelho) ou negativa (azul) com suas respectivas doenças estudadas.



Analisando de maneira conjunta, percebe-se que as relações de algumas condicionantes se comportam na mesma direção com as doenças estudadas, mesmo que elas apresentem características epidemiológicas diferentes. Dentre as sete condicionantes utilizadas nos modelos, quatro apresentaram as mesmas relações com todas as doenças: número de moradores por domicílio, renda do chefe familiar, densidade populacional e áreas verdes.

O número de moradores por domicílio mostrou coeficientes positivos para todas as doenças. Dois aspectos dessa condicionante devem ser levados em consideração na modelagem espacial de doenças. A primeira diz respeito à proximidade entre indivíduos dentro de um mesmo domicílio, aumentando a probabilidade de transmissão da doença, seja de maneira indireta (como a Zika e Chikungunya, por meio do mosquito infectado) ou de maneira direta (como a Tuberculose). O segundo aspecto está atrelado às condições de infraestrutura que o domicílio está inserido. Embora a Leptospirose não seja transmitida de pessoa para pessoa, a condição sanitária de uma habitação pode favorecer que os seus moradores tenham os mesmos riscos de contaminação da doença.

A renda do chefe familiar mostrou uma relação inversamente proporcional para todas as doenças. Essa variável está intimamente ligada às condições de infraestrutura domiciliar e urbana, acesso à informação, assistência médica que muitas vezes obedecem a uma lógica espacial. Portanto, embora as notificações dos mais variados tipos de doenças possam fazer parte da realidade de todos os bairros de um município, a suas concentrações espaciais fazem parte da desigualdade social de uma cidade.

A densidade populacional está diretamente ligada ao número de notificações das doenças analisadas. No presente estudo, deu-se preferência à densidade de notificações sem transformá-la em índices por 100.000 habitantes. Existe uma discussão acerca dos melhores métodos para espacialização de fenômenos pontuais que ocorrem no espaço. A preferência pela utilização da densidade de pontos sem dividi-las pela população da unidade espacial evita que algumas unidades possam mostrar uma “falsa” baixa densidade de pontos por não existir população suficiente (WANG, 2014).

É importante salientar que todas as condicionantes devem ser analisadas de maneira conjunta, uma vez que apenas uma condicionante não é o fator determinante para favorecimento de uma determinada doença.

No caso do município de Recife, os maiores níveis de renda do chefe familiar encontram-se na zona norte da planície e no litoral, nos bairros de Boa Viagem e Pina. Todavia, quando considerada a distribuição espacial das doenças até aqui analisadas, os bairros litorâneos

apresentavam mais notificações se comparados com os bairros da zona norte da planície, como nos bairros de Casa Forte, Poço da Panela, Parnamirim, Santana.

Tal aspecto infere que outras condicionantes podem estar relacionadas a uma maior proliferação de doenças, mesmo quando o fator econômico favorece a uma melhor condição de vida. Partindo desse princípio, percebe-se que nos bairros com maiores rendas médias mensais com maior densidade populacional apresentou número significativos de casos notificados, como o bairro de Boa Viagem que apresentou o 2º maior número de casos de Tuberculose (Figura 42 e Figura 43).

Por outro lado, o bairro de Jardim São Paulo apresentou média de densidade populacional inferior que outros bairros e, mesmo assim, registrou o 2º maior número de notificações de Zika e Chikungunya, mesmo com renda média mensal do chefe familiar superior aos bairros vizinhos, como o Curado e Torrões (Figura 42 e Figura 43).

Diante dessas observações, é importante ressaltar que a complexidade espacial socioeconômica do Recife associada aos seus aspectos naturais, resulta em múltiplas interações de acordo com a doença analisada. É salutar considerar que cada unidade espacial tem suas particularidades e infinitas combinações de condicionantes que possam resultar ou evitar a proliferação de cada doença. Por isso, é sugerível que um modelo espacial abarque um conjunto diversificado de categorias de variáveis para que consiga estimar uma determinada doença mais próximo da realidade.

Figura 42 - A densidade habitacional por bairros e os aspectos levantados sobre a distribuição espacial das doenças analisadas.

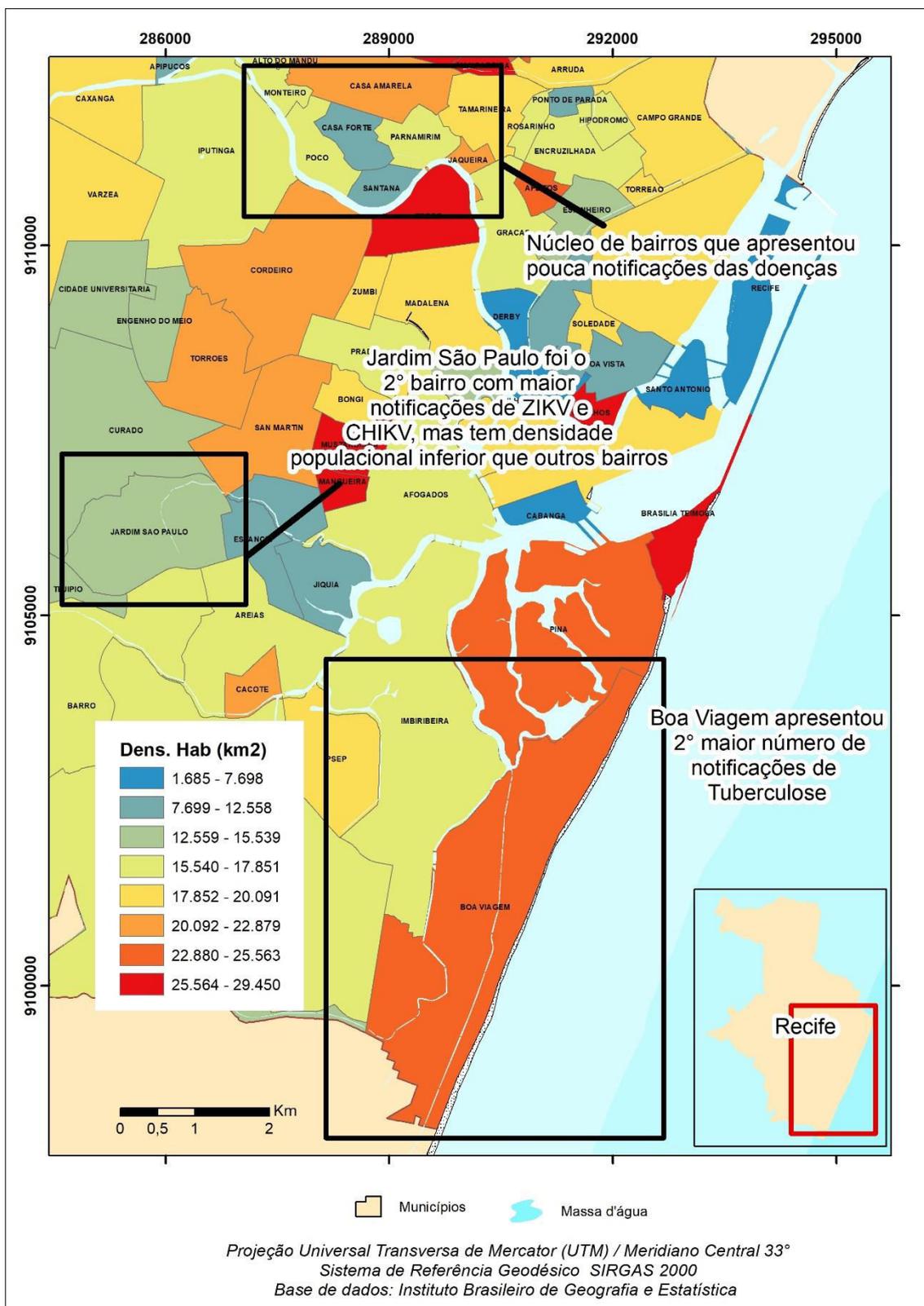
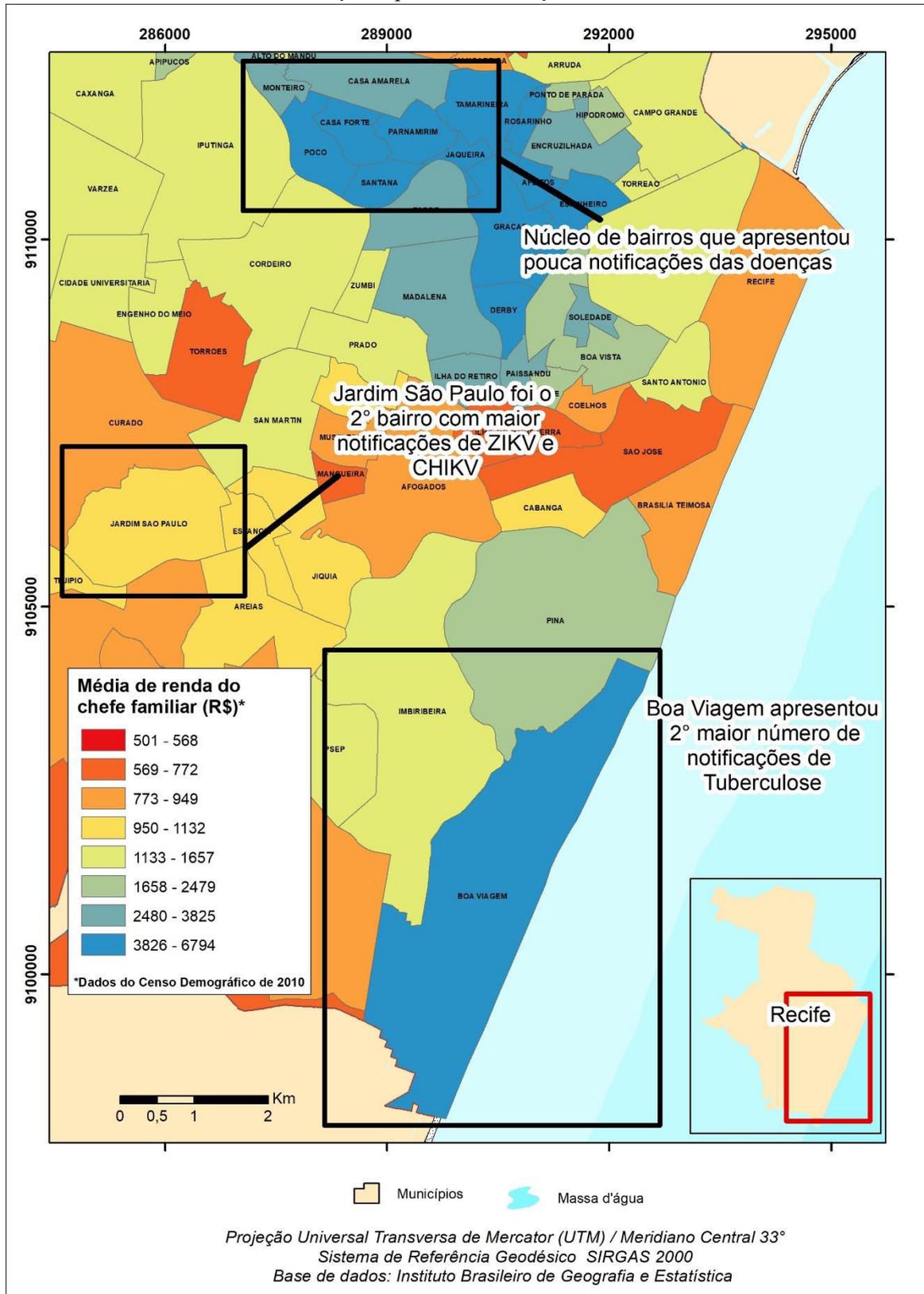


Figura 43- A renda média mensal do chefe familiar por bairros e os aspectos levantados sobre a distribuição espacial das doenças analisadas.



Outro fator que apresentou as mesmas relações inversamente proporcionais com as doenças foi a presença de Áreas Verdes. Apesar do tipo de relação ser a mesma, a presença de áreas verdes influencia diferenciadamente cada doença, como abordado nos capítulos anteriores. É partindo desse princípio que para um planejamento urbano, sob o ponto de vista da saúde pública, as áreas verdes devem ser consideradas como elemento urbanístico crucial. Seja por gerar amenidades microclimáticas, atenuar a densidade populacional, ou até mesmo espaços para que sejam utilizados para exercícios físicos e lazer.

Consideradas as condicionantes que apresentaram as mesmas relações com as doenças (diretamente ou inversamente proporcional), outras comportam-se de maneira diferenciada para cada doença, mesmo que apresentem autocorrelação espacial com as enfermidades analisadas. Esses fatores (Precipitação, Áreas não construídas e Declividade) fazem parte de um grupo a parte de condicionantes que mesmo tendo correlação espacial com algumas notificações podem ter comportamento diferentes, necessitando uma maior observação de sua relação com a doença em questão.

Partindo desses pressupostos, podemos classificar as condicionantes em dois grupos: as que possuem os mesmos tipos de relações, independentemente das doenças; e as que embora apresentem correlação espacial, os seus coeficientes podem ser negativos ou positivos de acordo com a morbidade analisada.

Tal fato, levanta alguns questionamentos acerca da influência de condicionantes que atuam de maneira semelhante na distribuição espacial de doenças e que muitas vezes são tratadas de forma segmentada. Por isso, é preciso considerar que existem fatores que se relacionam de modo similar mesmo com doenças que apresentam agentes etiológicos, períodos de incubação, níveis de transmissão e tipos de microrganismos diferenciados.

Diante dessa discussão acerca da influência de como algumas condicionantes espaciais favorecem ou impedem a maiores proliferações de doenças, fica evidente que o processo de modelagem espacial é complexo, e para ser considerado um modelo espacial mais próximo da realidade pode ser levantado uma equação baseada nos seguintes princípios:

$$Yr_i = \beta \sum_{CSG} + \beta \sum_{CSE} + \beta \sum_{CEED} \quad (10)$$

Em que Yr_i é a densidade real de uma determinada doença para uma localização i . A variável $\beta \sum_{CSG}$ é o somatório dos coeficientes de todas as condicionantes socioambientais gerais que influenciam na distribuição espacial da doença Y , assim como de outras doenças,

podendo ser a densidade populacional, nível de renda e outras variáveis. O $\beta_{\Sigma_{CSE}}$ é o somatório dos coeficientes das condicionantes socioambientais específicas da doença Y, como a precipitação em relação à leptospirose ou a qualidade do ar com doenças que envolvem o trato respiratório. O $\beta_{\Sigma_{CEED}}$ é o conjunto das condicionantes espaciais que estão condicionadas às características epidemiológicas da doença Y, tais como a área de atuação dos vetores, nível de transmissibilidade e potencial de disseminação espacial do agente etiológico. Dentre os conjuntos abordados, esse último tende a ser mais complexo pois não existem instrumentos ou ferramentas para avaliar como essas condicionantes comportam-se espacialmente. Esse fenômeno levanta questionamentos, tais como: qual o raio de atuação de um mosquito *Aedes aegypti*? Qual o potencial de disseminação espacial de uma doença a partir de uma pessoa infectada?

É a partir dessa discussão que fica evidente que o processo de modelagem espacial não pode e nem deve chegar à perfeição de estimar com exatidão as regiões mais suscetíveis a contaminação de uma determinada morbidade. Contudo, a partir da análise dessa equação, pode-se levantar algumas soluções para aproximar os modelos espaciais da realidade.

Uma das principais dificuldades em estimar espacialmente uma doença é identificar o máximo de condicionantes espaciais que possam entrar no modelo espacial para conseguir aproximar-se ao máximo da realidade. Uma das alternativas para adicionar os fatores que não podem ser identificados é adicionar outra doença para que as condicionantes socioambientais gerais que não foram adicionadas, sejam indiretamente colocadas no modelo, seguindo a seguinte equação:

Segundo a maioria dos modelos espaciais para estimar uma determinada doença, algumas variáveis não são colocadas pelos mais variados motivos: dificuldade de obtenção dos dados, limitações em mensurar a variável, série temporal curta, dentre outras. Por isso, para estimar uma doença, geralmente os estudos baseiam-se na seguinte equação:

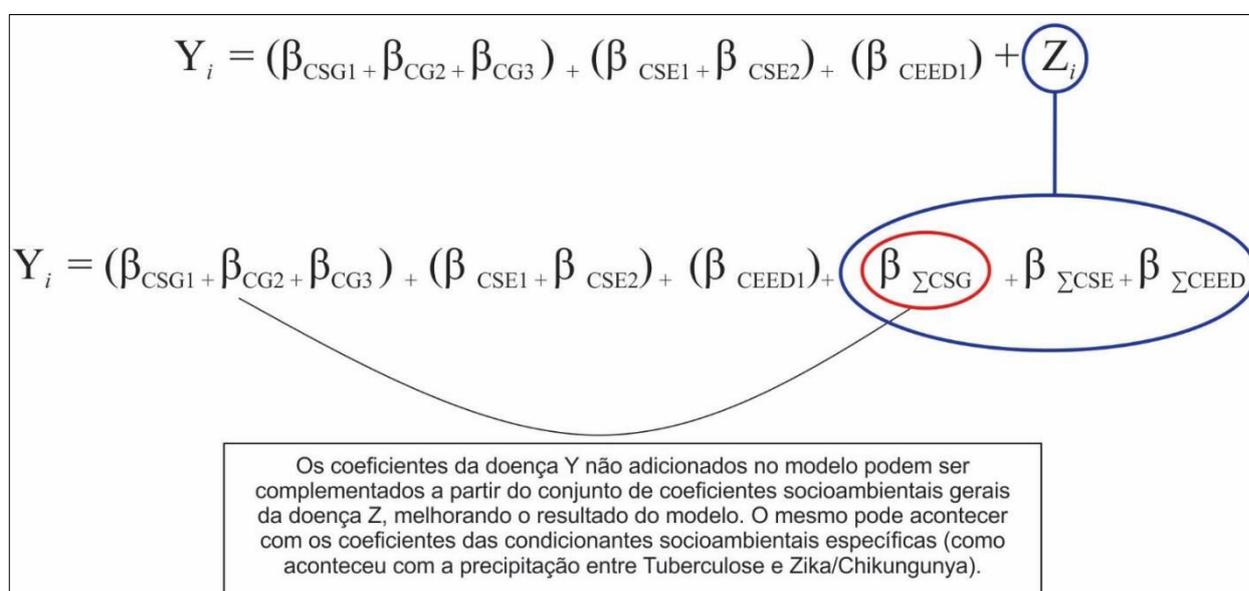
$$Y_i = (\beta_{CSG1} + \beta_{CG2} + \beta_{CG3}) + (\beta_{CSE1} + \beta_{CSE2}) + (\beta_{CEED1}) \quad (11)$$

Em que Y_i é a densidade estimada para a doença numa localização i. Diferente da equação anterior, apenas algumas condicionantes são colocadas no modelo de acordo com sua categoria, ao invés do conjunto (Σ). Sendo assim, os β_{CSG1} , β_{CSG2} e β_{CSG3} são os coeficientes

de cada condicionante socioambiental geral, podendo ser, por exemplo, a renda, densidade populacional e moradores por domicílio. O β_{CSE1} e β_{CSE2} podem ser dois coeficientes de variáveis socioambientais específicas da doença, como a precipitação e áreas verdes para a leptospirose, por exemplo. A última variável β_{CEED1} (condicionante espacial específica de uma doença) é uma das mais complicadas em adicionar no modelo, dada a sua dificuldade de mensurar. Exemplificando, pode ser a quantidade de um determinado vetor de uma doença por unidade espacial, como os registros de ratos por bairros, em um município.

A proposta a seguir (Figura 44), tem como princípio adicionar mais coeficientes socioambientais gerais de modo indireto no modelo espacial, com o intuito de melhorar os resultados das estimativas. O modelo a seguir é composto por todas as variáveis já descritas na equação anterior, com uma variável a mais: a Z_i . Essa variável é a densidade espacial de uma outra doença para a localização i .

Figura 44 - Proposta teórica para modelagem espacial de doenças, adicionando outras doenças como variável independente.



Quando se adiciona a densidade de uma doença (Z) como variável independente de um modelo espacial para estimar uma doença (Y), adicionamos os conjuntos dos coeficientes intrínsecos àquela morbidade. A grande questão é que, como não se consegue descobrir todas as variáveis que influenciam em Y , podemos inferir que alguns coeficientes socioambientais gerais que não foram adicionados, possam estar dentro do conjunto de coeficientes da doença Z . Obviamente que alguns coeficientes que estão relacionados à (Z) não irão apresentar

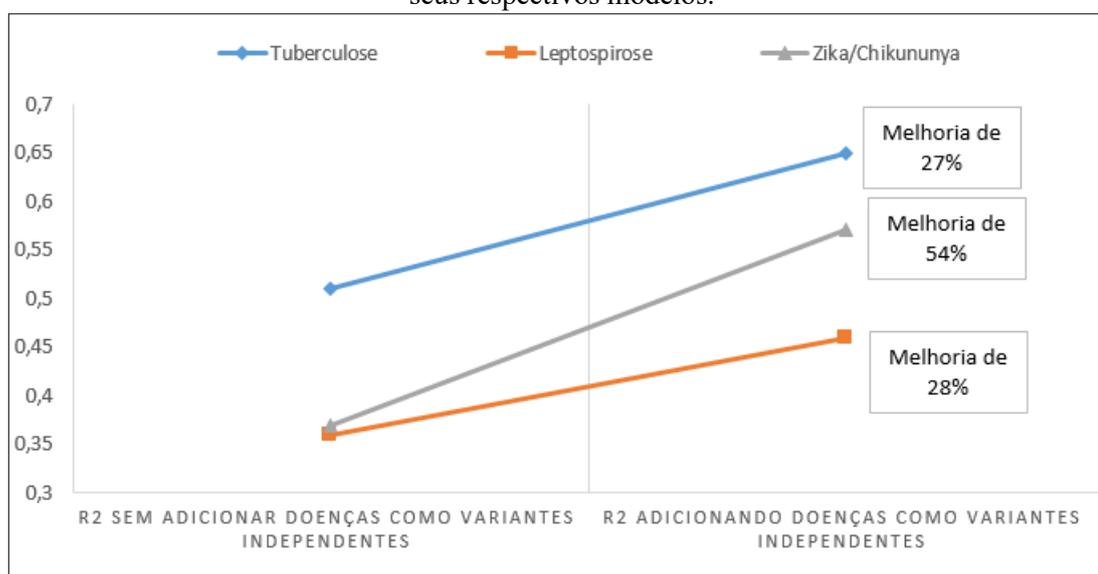
correlação com a (Y), o que não prejudicaria o resultado do modelo, apenas apresentaria seu coeficiente próximo a 0.

Podemos exemplificar essa discussão teórica a partir das doenças analisadas nesse estudo, e verificar as diferenças de coeficientes nas estimativas entre os modelos com e sem as doenças como variável independente.

A Figura 45 mostra o coeficiente de determinação ajustado (R^2) do método OLS dos melhores modelos estudados nos capítulos anteriores para cada doença e a diferença que resultaria quando adicionadas, como variável independente, as doenças que ficaram fora da análise.

No caso da Tuberculose, o Modelo 4 apresentou um coeficiente de determinação de 0,51. Ao serem adicionadas as densidades espaciais da Zika/Chikungunya do Modelo 5 e Leptospirose do Modelo 6 como variáveis independentes, o R^2 ajustado passou a ser 0,65, no método OLS. Isso resultou numa melhora de 27% na qualidade de estimativa do modelo.

Figura 45 – R^2 dos Modelos 4 (Tuberculose), 5 (Zika/Chikungunya) e 6 (Leptospirose) com método de estimativa OLS, e suas diferenças quando utilizada outras doenças como variáveis independentes nos seus respectivos modelos.



Considerando o Modelo 5 da Zika/Chikungunya, com o método OLS, percebe-se que o coeficiente de determinação, quando adicionadas as densidades de Tuberculose e Leptospirose de seus respectivos modelos, passou de 0,37 para 0,57. Tal resultado mostrou que houve uma melhora de 54% nas estimativas.

O mesmo aconteceu com a Leptospirose (Modelo 6), quando adicionadas a Zika/Chikungunya e Tuberculose na doença. O R^2 ajustado no método OLS, foi de 0,36 para 0,46, resultando numa melhoria de 28% das estimativas.

Considerando os novos coeficientes para o novo Modelo 4 de Tuberculose (Tabela 7), percebe-se que as relações se mantiveram as mesmas que o Modelo 4 sem as doenças inclusas como variável independente, exceto a variável de pessoas por domicílios que teve a relação do coeficiente mudada de positivo para negativo. Contudo essa variável, segundo a *Robust Probability*, deixou de apresentar correlação estatística com a Tuberculose, no modelo. Esse fato evidencia que quando é adicionada variáveis com forte correlação com a doença – no caso a densidade de Leptospirose e Zika/Chikungunya - as outras variáveis tem seus valores de coeficientes diminuídos, podendo comprometer a tal ponto, que ela não seja mais útil em ser utilizada no modelo.

De toda forma, ao ser avaliada a a *Variance Inflation Factor* (VIF), percebe-se que não houve problemas de multicolinearidade entre as variáveis.

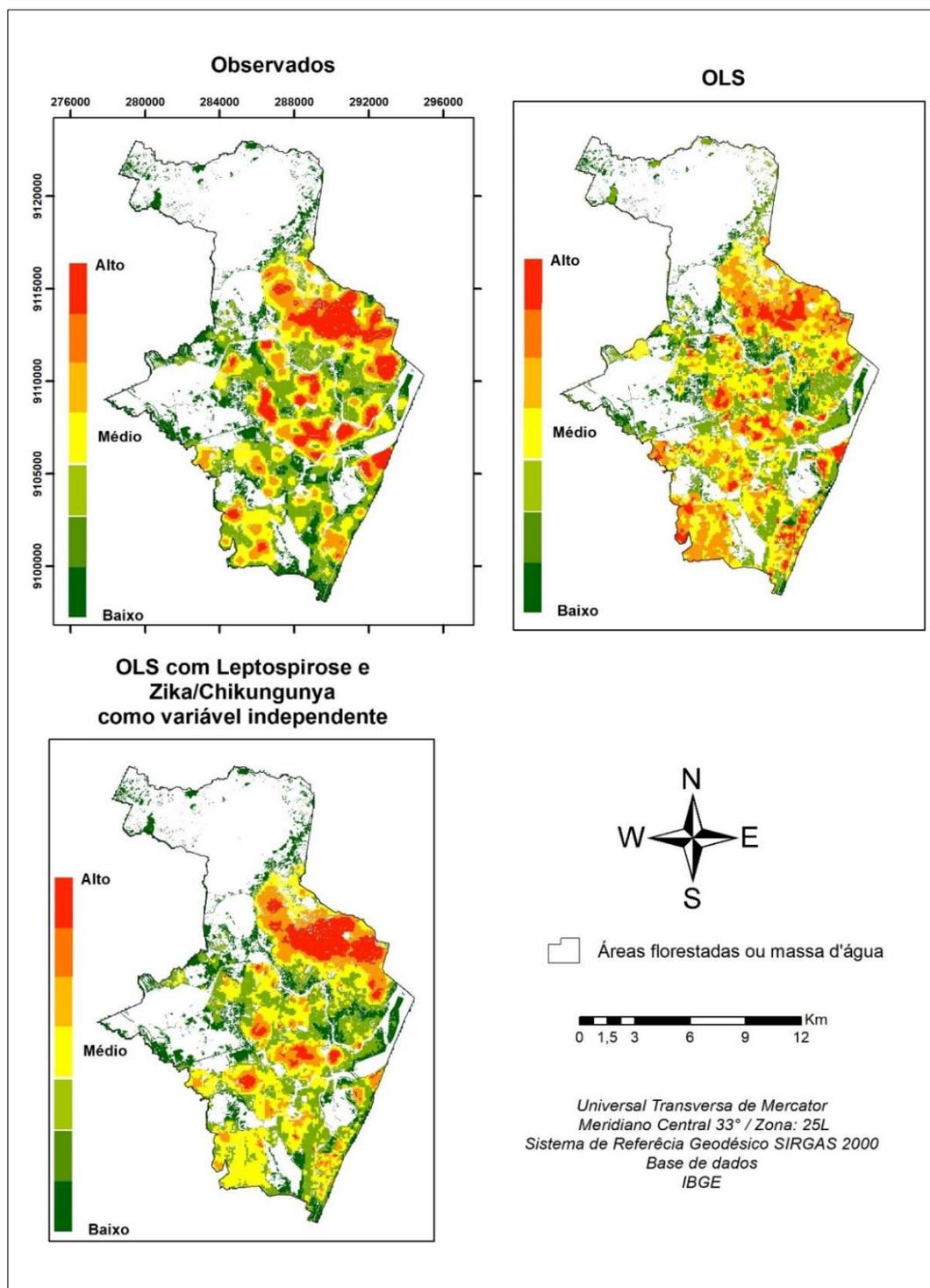
Tabela 7 - Variáveis independentes utilizadas no Modelo 4 adicionando as densidades espaciais da Zika/Chikungunya e Leptospirose e suas relações com a Tuberculose de acordo com o coeficiente, *Robust Probability* e VIF, no método OLS.

Variáveis	Coefficiente	Robust Probability	VIF
Pessoas por domicílio	-0,019783	0,088225	1,270544
Renda do chefe familiar	-0,067234	0,000000*	1,257926
Precipitação	-0,050258	0,000000*	1,170687
Áreas não construídas	-0,037451	0,000000*	1,581369
Declividade	-0,058086	0,000000*	1,253376
Densidade populacional	0,348754	0,000000*	2,000442
Áreas verdes	-0,057482	0,000000*	2,665268
Zika e Chikungunya Kernel 700m	0,401461	0,000000*	1,832466
Leptospirose Kernel 900m	0,114124	0,000000*	1,815459

Ao ser analisado os mapas das estimativas do Modelo 4 (Figura 46), utilizando apenas o método OLS, percebemos que o modelo espacial que foram adicionadas as doenças mostrou

maiores semelhanças com os dados observados da densidade espacial da Tuberculose em Recife. A importância do mapa, nesse sentido, mostra que as variações espaciais obedecem a uma lógica espacial, onde todas as doenças aqui analisadas fazem parte de um mesmo processo de distribuição ao longo do município.

Figura 46 - Densidade espacial observada e estimada da Tuberculose com o método OLS sem as doenças e com as doenças como variáveis independentes.



Considerando os novos coeficientes para o novo Modelo 5 da Zika/Chikungunya, percebe-se que as relações se mantiveram as mesmas ao serem comparadas o modelo sem as doenças inclusivas como variável independente. É válido ressaltar que os coeficientes apresentaram o mesmo comportamento encontrado no modelo anterior, que foi a diminuição dos seus valores, mesmo que as relações estivessem mantidas. Tal fato pode estar relacionado com o que já foi descrito anteriormente.

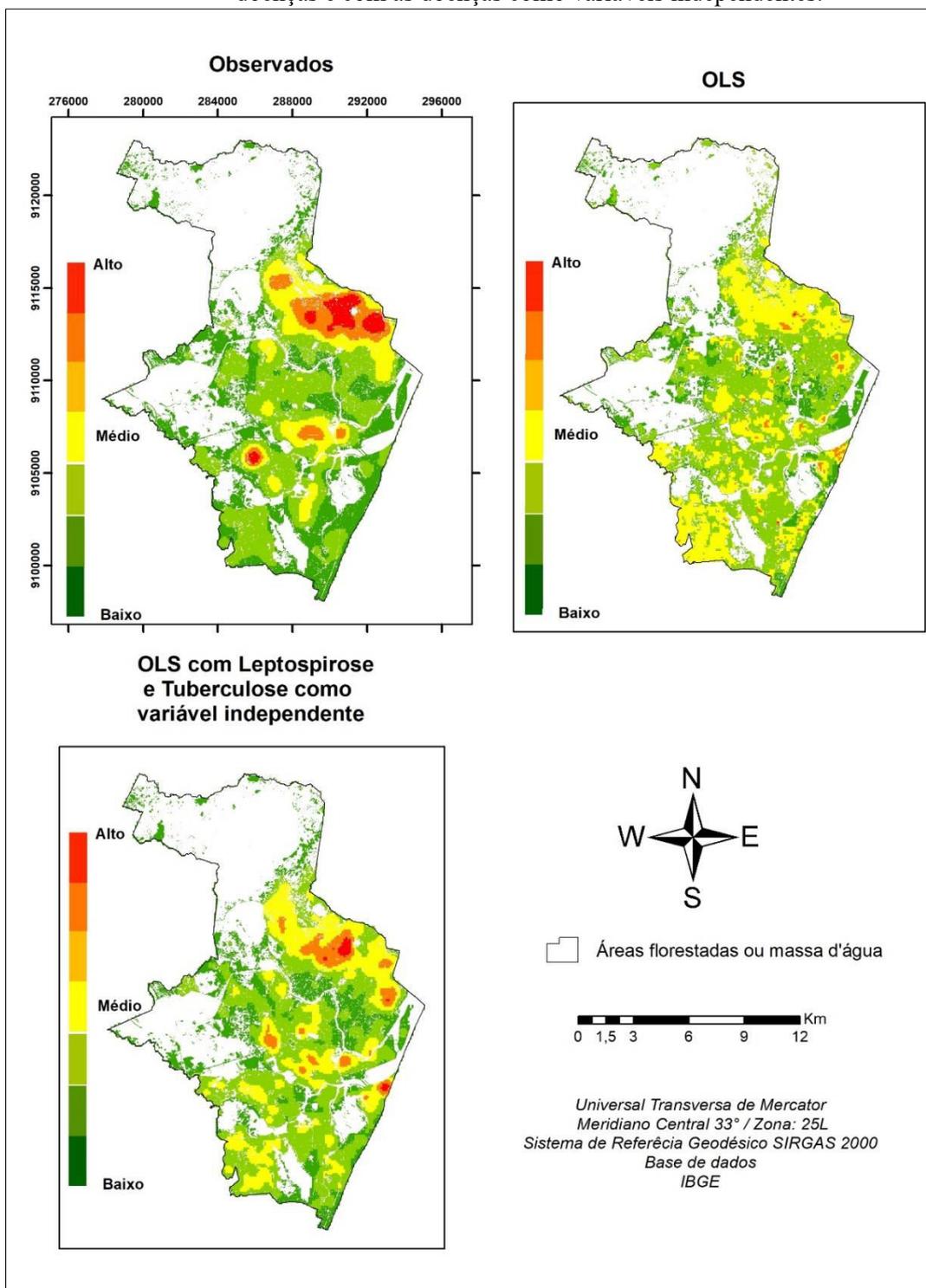
Os baixos valores do *Variance Inflation Factor* (VIF) mostrou que não existiram problemas de multicolinearidade entre as variáveis.

Tabela 8 - Variáveis independentes utilizadas no Modelo 5 adicionando as densidades espaciais da Tuberculose e Leptospirose e suas relações com a Zika/Chikungunya de acordo com o coeficiente, *Robust Probability* e VIF, no método OLS.

Variáveis	Coeficiente	Robust Probability	VIF
Pessoas por domicílio	0,193229	0,000000*	1,262234
Renda do chefe familiar	-0,03008	0,000000*	1,269078
Precipitação	-0,07358	0,000000*	1,166407
Áreas não construídas	-0,01252	0,014293*	1,584954
Declividade	0,048111	0,000000*	1,257379
Densidade populacional	0,052554	0,000000*	2,360879
Áreas verdes	-0,06196	0,000000*	2,666357
Leptospirose Kernel 900m	0,239038	0,000000*	1,730996
Tuberculose Kernel 500m	0,54663	0,000000*	2,256555

Analisando a distribuição espacial da doença no novo Modelo 5, percebe-se que a proximidade dos dados observados é maior onde as estimativas apresentaram maior densidade.

Figura 47 - Densidade espacial observada e estimada da Zika/Chikungunya com o método OLS sem as doenças e com as doenças como variáveis independentes.



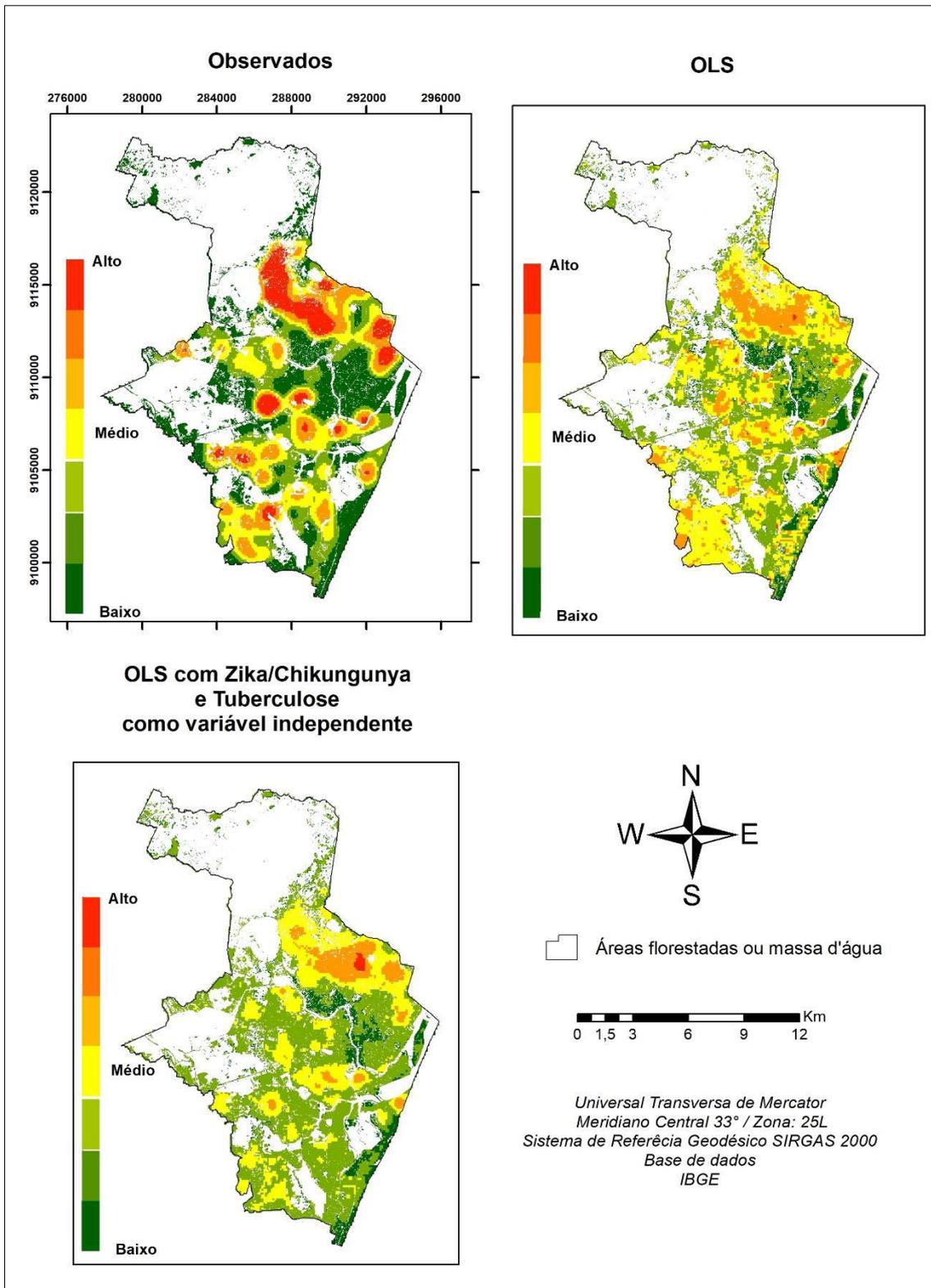
Em relação aos coeficientes do novo Modelo 6 da Leptospirose (Tabela 9), verifica-se que todas as variáveis apresentaram as mesmas relações, além de mostrarem autocorrelação espacial com a densidade da doença, segundo o *Robust Probability*. Os novos valores dos coeficientes apresentaram uma diminuição em relação ao modelo tradicional, exceto as variáveis de Densidade Populacional e Áreas Verdes, que apresentaram aumento. Os valores do VIF não apresentaram problemas de multicolinearidade.

Tabela 9 - Variáveis independentes utilizadas no Modelo 6 adicionando as densidades espaciais da Tuberculose e Zika/Chikungunya e suas relações com a Leptospirose de acordo com o coeficiente, Robust Probability e VIF, no método OLS.

Variáveis	Coeficiente	Robust Probability	VIF
Pessoas por domicílio	0,112234	0,000000*	1,268205
Renda do chefe familiar	-0,16314	0,000000*	1,222113
Precipitação	0,073786	0,000000*	1,167937
Áreas não construídas	0,036048	0,000000*	1,582977
Declividade	0,085715	0,000000*	1,250287
Densidade populacional	0,171003	0,000000*	2,311026
Áreas verdes	-0,19076	0,000000*	2,620713
Zika e Chikungunya Kernel 700m	0,276436	0,000000*	2,192532
Tuberculose Kernel 500m	0,179702	0,000000*	2,831696

Ao observar o mapa do novo Modelo 6 da Leptospirose (Figura 48), percebe-se que houve uma maior aproximação da realidade em relação ao Modelo 6 anterior. Uma das diferenças a serem percebidas é que as regiões superestimadas no modelo anterior (sudoeste de Recife), conseguiram aproximar da realidade no novo modelo. Tal fato também é percebido de maneira isolada em algumas áreas ao longo do município, onde pelo modelo anterior a superestimação era evidente, diminuindo essa ocorrência no novo modelo.

Figura 48 - Densidade espacial observada e estimada da Leptospirose com o método OLS sem as doenças e com as doenças como variáveis independentes.



De modo geral, em todos os novos modelos houveram mudanças significativas em termos estatísticos e espaciais nas estimativas das doenças, quando outras doenças eram colocadas como variáveis adicionais nos modelos estudados. Esse fato corrobora com os pressupostos acima de que algumas variáveis não são colocadas em modelos espaciais pelos mais variados motivos, mas que adicionando outras doenças como variáveis de entrada, elas podem melhorar as estimativas dos modelos. Não obstante essa melhoria nas estimativas, isso evidencia que doenças com diferentes características epidemiológicas fazem parte de uma mesma lógica espacial onde algumas variáveis socioambientais (mesmo que não identificadas) contribuem para a concentração de enfermidades no município de Recife.

A tese até aqui defendida evidencia que em situações que uma doença pouco conhecida passa a se distribuir espacialmente de modo intenso, torna-se possível estimar a sua distribuição espacial a partir de outras doenças já registradas, associadas a alguns condicionantes geográficos.

Tal proposta mostra-se relevante dado ao fenômeno ocorrido pela disseminação do COVID-19, doença causada pelo coronavírus, denominado SARS-CoV-2. Em março de 2020, Recife registrou os primeiros casos de pessoas infectadas. Ao longo dos meses de disseminação dessa doença na cidade, é perceptível que a lógica espacial da distribuição da taxa de letalidade da doença está intimamente relacionada a fatores como renda, taxa de escolaridade e condições de habitação (SOUZA, 2020).

A proposta teórica dessa tese defende, além dessas afirmativas, que exista a possibilidade desse fenômeno não ser somente abordado sob vários enfoques (políticos, culturais, sociais, econômicos e ambientais), mas também sob várias escalas (de um bairro até a análise de vários países).

6 CONSIDERAÇÕES FINAIS

De acordo com os dados levantados na presente tese, podemos considerar os resultados em três partes principais: (a) parâmetros que apresentaram melhores eficiência nas estimativas da doença; (b) análise das relações entre as condicionantes e as doenças estudadas; (c) proposta teórico-metodológica acerca da modelagem espacial de doença.

A combinação dos parâmetros que resultou nas melhores estimativas da distribuição espacial da Zika e Chikungunya foi a unidade espacial de quadrícula de 100 metros, associada a densidade de Kernel com raio de 700 metros e método de estimativa de GWR com 700 metros de *bandwidth*.

Em relação ao modelo para estimar a Leptospirose, a combinação dos parâmetros que resultou nas melhores estimativas foi a unidade espacial da quadrícula de 100 metros, associada a densidade de Kernel com raio de 900 metros e método de estimativa GWR com 900 metros de *bandwidth*.

Sobre a distribuição espacial da Tuberculose, os parâmetros que resultou nas melhores estimativas no modelo foi a unidade espacial da quadrícula de 100 metros, associada a densidade de Kernel com raio de 500 metros e método de estimativa GWR com 500 metros de *bandwidth*.

Um aspecto a ser considerado sobre os parâmetros dos modelos é que embora o método GWR apresente os melhores resultados – se comparado com o método OLS – uma de suas principais utilidades é avaliar como comporta-se as relações espaciais entre uma determinada condicionante e a doença analisada. Tal fato o torna mais como um modelo exploratório que preditivo.

Todas as condicionantes consideradas no presente estudo (socioeconômicas, físicas, infraestruturais e climáticas) apresentaram correlação estatística significativa em todas as doenças estudadas. Baseando-se nos coeficientes do método OLS, algumas delas apresentaram coeficientes positivos para todas as doenças (morador por domicílio e densidade populacional), enquanto a renda do chefe familiar e áreas verdes apresentaram coeficientes negativos para todas as doenças. Alguns condicionantes apresentaram um comportamento específico, com relação direta ou inversamente proporcional de acordo com cada doença.

Esse fenômeno espacial pode levantar algumas propostas para a modelagem espacial de doenças. Destarte, podemos considerar que para prever uma determinada doença devemos considerar três tipos de relação das condicionantes com uma doença num modelo espacial: (a) os **socioambientais gerais**: que são as condicionantes espaciais que apresentam as mesmas relações espaciais com outras doenças; (b) **socioambientais específicos**: que são os

condicionantes socioambientais específicos de uma determinada doença e (c) **condicionantes espaciais específicos da doença**: esse fator caracteriza a fenômenos espaciais que estão atreladas a distribuição espacial da doença, como por exemplo, a área de atuação de um vetor ou sua distribuição espacial.

Partindo desse princípio, o principal desafio de modelar espacialmente uma determinada doença é que a obtenção dos dados que abarque toda essa gama de condicionantes é um fator limitador. Por isso, como proposta teórico-metodológica considera-se que é possível prever a distribuição espacial de uma doença, a partir do registro espacial de outras doenças já conhecidas, uma vez que elas fazem parte de uma mesma lógica espacial.

Corroborando a essas afirmativas, os resultados dessa tese mostraram que modelos espaciais que consideraram a distribuição espacial de patologias como dados de entrada para estimar a distribuição espacial de uma determinada doença apresentaram melhoras em suas estimativas, se comparada com os mesmos modelos espaciais sem as doenças como variáveis independentes.

Infere-se que ao colocar uma determinada doença como variável independente de um modelo, todas as condicionantes intrínsecas a ela também vão junto. Destarte, algumas condicionantes, que até então não tinham sido incluídas no modelo, começam a fazer parte do modelo espacial de maneira indireta, melhorando as suas estimativas.

Tal proposta pode auxiliar nas estimativas de distribuição espacial de várias doenças, principalmente em casos onde não existe uma confiabilidade total nos dados, ou até mesmo a ausência deles. Outro aspecto a ser levantado é que em casos do surgimento de novas epidemias, essa proposta auxilia em prognósticos nas áreas a serem mais impactadas, baseando-se na distribuição espacial de doenças já conhecidas.

REFERÊNCIAS

- ABDI, H. The Method of Least Squares. In: SALKIND, N. **Encyclopedia of Measurement and Statistics**. 1 ed. Thousand Oaks: Sage, 2007. p. 1-7.
- AGUIAR, B. S. *et al.* Potential risks of Zika and chikungunya outbreaks in Brazil: A modeling study. **International Journal of Infectious Diseases**, v.70, n.1, p 20–29, 2018.
- ALMEIDA, M. C. M. *et al.* Spatial Vulnerability to Dengue in a Brazilian Urban Area During a 7-Year Surveillance. **Journal of Urban Health**, v. 84, n.3, p. 334–345, 2007.
- ANJOS, R. S.; WANDERLEY, L. S. A.; NÓBREGA, R. S. Análise espacial da precipitação e possíveis fatores que contribuem para sua espacialização em Recife-PE. **Revista Brasileira de Geografia Física** v.13, p. 18-34, 2020.
- ATIQUE, S. *et al.* Investigating spatio-temporal distribution and diffusion patterns of the dengue outbreak in Swat, Pakistan. **Journal of Infection and Public Health**, London v.11, n.4, p. 1-8. 2017.
- AYOADE, J. O. **Introdução à Climatologia para os Trópicos**. 14 eds. Rio de Janeiro: Bertrand Brasil, 2010. 350 p.
- BARCELLOS, C.; SABROZA, P. C. The place behind the case: leptospirosis risks and associated environmental conditions in a flood-related outbreak in Rio de Janeiro. **Cad. Saúde Pública**, Rio de Janeiro, v. 17, n. 1, p. 59-67, 2001.
- BEIRAVAND, R. *et al.* Correlation Assessment of Climate and Geographic Distribution of Tuberculosis Using Geographical Information System (GIS). **Iranian Journal of Public Health**, v. 45, n.1, p. 86-93, 2016.
- BRASIL. Ministério da Saúde. **Doenças infecciosas e parasitárias: guia de bolso**. 8^a ed. Brasília: Ministério da Saúde, 2010.
- _____. **Vigilância em Saúde: Dengue, Esquistossomose, Hanseníase, Malária, Tracoma e Tuberculose**. 2. ed. rev. - Brasília, 197 p. 2008.
- BRINKIN, E. A. **Seroprevalence And Risk Factors For Leptospirosis In An Urban Community Of Puerto Rico**. Mestrado (Master of Public Health). School of Public Health, Yale University, Yale, 2016.
- BRUNSDON, C.; FOTHERIGHAN, A. E.; CHARLTON, M. E. Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. **Journal Geographical Analysis**, Ohio, v.28, n. 4, p. 281-298, 1996.
- BRUNSDON, C.; FOTHERIGHAN, A. E.; CHARLTON, M. E. Geographically weighted regression-modelling spatial non-stationarity. **Journal of the Royal Statistical Society**, London, v. 47, n. 3, p. 431-443, 1998.

CALVET, G. A. et al. Zika virus infection: epidemiology, clinical manifestations and diagnosis. **Curr Opin Infect Dis**, v.29, n.5, p 1- 8, 2016.

CAMPOS, T. L. et al. Revisiting Key Entry Routes of Human Epidemic Arboviruses into the Mainland Americas through Large-Scale Phylogenomics. **International Journal of Genomic**. v. 2018, p. 1-9, 2018.

CARABALI, M. et al. The Zika epidemic and abortion in Latin America: a scoping review. **Global Health Research and Policy**, v.3, n.15, 2, p.1-9, 2018.

CARLOS, A. F. A. **O Espaço Urbano: Novos Escritos sobre a Cidade**. São Paulo: Labur Edições, 2007. 123p.

CASTRO, J. **Um ensaio de geografia urbana: A cidade do Recife**. Recife: Fundação Joaquim Nabuco, Editora Massangana, 2013. 122 p.

CHAILIB, J. V. et al. Estudo espacial de riscos à leptospirose no município do Rio de Janeiro (RJ). **Saúde e Debate**, Rio de Janeiro, v. 41, n. 1, p.225-240, 2017.

CLARCK, M.; RIBEN, P.; NOWGESIC, E. The association of housing density, isolation and tuberculosis in Canadian First Nations communities. **International Journal of Epidemiology**, v. 31, n. 5, p. 940–945, 2002.

COSTA, A. R. R. F. **Análise do Comportamento Epidemiológico da Tuberculose e suas relações com variáveis socioeconômicas no Estado de Pernambuco**. 2018. 48 p. Trabalho de Conclusão de Curso (Bacharelado em Saúde Coletiva), Universidade Federal de Pernambuco, Vitória de Santo Antão. 2018.

CUNHA, M. D. Geographic weighted regression: applicability to epidemiological studies of leprosy. **Revista da Sociedade Brasileira de Medicina Tropical**. v. 49, n.1, p. 74-82, 2016.

DIAS, J. P. et al. Factors associated with *Leptospira* sp infection in a large urban center in northeastern Brazil. **Rev. Soc. Bras. Med. Trop**, Uberaba, v.40, n.5, p. 499-504, 2007.

ESRI. **Interpreting OLS results**. Disponível em: <
<https://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/interpreting-ols-results.htm>> Acesso em dezembro de 2020.

FEITOSA, F. R. S. **Indicadores de sustentabilidade como subsídio para a prevenção e controle da infestação pelo mosquito *Aedes aegypti* no município de Aracaju-SE**. 2016. Dissertação (Mestrado em Desenvolvimento e Meio Ambiente). Universidade Federal de Sergipe, Aracaju, 2016.

FERNANDES, F. M. C. et al. Relationship between climatic factors and air quality with tuberculosis in the Federal District, Brazil, 2003-2012. **Braz J Infect Dis**. v.21 no.4 Salvador, p. 369-375, 2017.

FERREIRA, H. S. **Clima Urbano e Dengue Em Recife: Influência Climática Sobre a Formação das Epidemias**. Dissertação (Mestrado em Geografia), 121 p. Universidade Federal de Pernambuco, Recife. 2016.

FERREIRA, A. V. *et al.* **Correlacionamento do parâmetro meteorológico precipitação com a taxa de incidência de doenças no Município do Rio de Janeiro**. Trabalho apresentado ao X Congresso Brasileiro de Meteorologia, Brasília, 1998.

FERREIRA, M. C.; FERREIRA, M. F.M. Influence of topographic and hydrographic factors on the spatial distribution of leptospirosis disease in São Paulo county, Brazil: an approach using geospatial techniques and GIS analysis. **International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v 41, n. 1, p.197-201.

FULLER T. L, et al. Behavioral, climatic, and environmental risk factors for Zika and Chikungunya virus infections in Rio de Janeiro, Brazil, 2015-16. **PLoS ONE** 12(11), p 1-15, 2017.

GEORGE, P. **O Homem na Terra: A Geografia em ação**. Rio de Janeiro: Edições 70, 1989, 175 p.

GRILLET, M. E, et al. Disentangling the Effect of Local and Global Spatial Variation on a Mosquito-Borne Infection in a Neotropical Heterogeneous Environment. **Journal Tropical of Medicine and Hygiene**. v. 82, n. 2, p. 194–201, 2010.

GRACIE, R. et al. Geographical Scale Effects on the Analysis of Leptospirosis Determinants. **International Journal of Environmental Research and Public Health**, v. 11, n. 10, p. 10366-10383, 2010.

GUEDES, D. R. et al. Zika virus replication in the mosquito *Culex quinquefasciatus* in Brazil. **Emerging Microbes & Infections**, v.6, n.1, p. 1-11, 2017.

GUO, L; MA, Z; ZHANG, L. Comparison of bandwidth selection in application of geographically weighted regression: a case study. **Journal of Forestry Research**, v. 38, n.1, p. 2526–2534, 2008.

GWARDA, K. L. Geographically Weighted Regression in the Analysis of Unemployment in Poland. **J. Geo-Information**, v. 7, n.1, p. 1-17. 2018.

IBGE. **Cidades**. Disponível em: <<https://www.ibge.gov.br/cidades-e-estados/pe/recife.html>>. Acesso em dezembro de 2020.

IKEFUTI, Priscilla Venâncio. **Associação entre variáveis meteorológicas, Índice Climático, fatores socioeconômicos e mortalidade por doenças do aparelho circulatório (Acidente Vascular Cerebral e Embolia Pulmonar) no município de São Paulo – SP**. Tese (Doutorado em Ciências). Faculdade de Filosofia Letras e Ciências Humanas da Universidade de São Paulo, São Paulo, 2016.

- KUDDUS, M. A.; EMMA, S. M.; ADEGBOYE, O.A. Delay effect and burden of weather-related tuberculosis cases in Rajshahi province, Bangladesh, 2007–2012. **Scientific Reports**, v. 9, n. 12720, p.1-13, 2019.
- LESSER, J.; KITRON, U. A geografia social do zika no Brasil. **Estudos Avançados**, v 30 n. 88, p. 168-175, 2016
- LIMA-CAMARA, T. N. Arboviroses emergentes e novos desafios para a saúde pública no Brasil. **Rev Saude Publica**, v. 50, n.36, p.1-7, 2016.
- LIN, C.H.; WEN, T. H. Using Geographically Weighted Regression (GWR) to Explore Spatial Varying Relationships of Immature Mosquitoes and Human Densities with the Incidence of Dengue. **Journal Environ. Res. Public Health**, v.8, n.1, p. 2798-2815, 2011.
- LINS, T. M. P. **Análise de Locais com potencial risco de transmissão de arboviroses usando técnicas de sensoriamento remoto e geoprocessamento**. 2019, 102 p. Dissertação de Mestrado (Mestrado em Ciências Geodésicas e Tecnologia da Geoinformação), Universidade Federal de Pernambuco, Recife, 2019.
- LIU, L. *et al.* Exploring the Relevance of Green Space and Epidemic Diseases Based on Panel Data in China from 2007 to 2016. **Int. J. Environ. Res. Public Health**, v.16, n.14, p. 1-21, 2019.
- LIU, Z. *et al.* Competence of *Aedes aegypti*, *Ae. albopictus*, and *Culex quinquefasciatus* mosquitoes as Zika virus vectors, China. **Emerg Infect Dis**. v.23, n.1, p.1085-1091, 2017.
- LOBO, M. A. A.; PAIVA NETO, H. C.; BELLO, L.A.L. Spatial distribution of pulmonary tuberculosis in urban areas: a case from Belém, Brazil . **Acta Biomedica Brasiliensia**, v. 7, n. 2, p. 1-14, 2016.
- MAGALHÃES, M. A. F. M. **A tuberculose no espaço urbano: um estudo ecológico utilizando análise espacial no município do Rio de Janeiro nos anos de 2005 a 2008**. Tese (Doutorado em Saúde Coletiva). Instituto de Estudos em Saúde Coletiva da Universidade Federal do Rio de Janeiro, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2014.
- MANSOER, J. R.; KIBUGA, D.K.; BORGDOFF, M.W. Altitude: a determinant for tuberculosis in Kenya? **Int J Tuberc Lung Diseases**, v.3, n.2, p.156-161, 1997.
- MINISTÉRIO DA SAÚDE. **Manual de Recomendações para o controle da tuberculose no Brasil**. Brasília: Ministério da saúde, 2019.
- _____. **Vírus Zika no Brasil**. Disponível em: Brasília: Ministério da saúde, 2016.
- MEDEIROS, S. **Poluição ambiental por exposição à poeira de gesso: impactos na saúde da população**. 2003. 200 p. Dissertação (Mestrado em Saúde Pública), Fundação Oswaldo Cruz, Recife. 2003.

- MEDEIROS, H. A. **Análise da acessibilidade geográfica dos doentes de tuberculose ao serviço de saúde em Campina Grande/PB, 2004-2009**. Dissertação (Mestrado em Saúde Pública). Universidade Estadual da Paraíba, Campina Grande, 2013.
- MAHABIR, R. *et al.* The study of slums as social and physical constructs: challenges and emerging research opportunities. **Regional Studies, Regional Science**, v. 3, n.1, p. 737-757, 2016.
- MOHAMMADINIA, A. *et al.* Prediction mapping of human leptospirosis using ANN, GWR, SVM and GLM approaches. **BMC Infectious Diseases**, v. 19, p. 1-18, 2019.
- MALKHAZOVA, S.M., *et al.* Natural-focal diseases: mapping experience in Russia. **Int J Health Geogr** v 13, p. 21 (2014).
- MORAES, D. Revisitando as ZEIS e o PREZEIS no Recife: entre o reformismo e o Direito à Cidade. **Oculum Ensaios**. v 17, n 1. p. 1-18, 2019.
- MELO, F. R. M. **Mortalidade relacionada às doenças tropicais negligenciadas no Brasil, 2000-2011: magnitude, padrões espaço-temporais e fatores associados**. 2015. Tese (Doutorado em Saúde Coletiva). Faculdade de Medicina, Universidade Estadual do Ceará, Fortaleza, 2015.
- MENDONÇA, F. M.; SOUZA, A. V. DUTRA, D. A. Saúde pública, urbanização e Dengue no Brasil. **Sociedade & Natureza, Uberlândia**, v.21, n.3, p. 257-269, 2009.
- MUSSO, D.; GUBLER, D. J. Zika virus. **Clin. Microbiol. Ver**, v.29, n.1, p. 487-524, 2016.
- NASCIMENTO, M. C. *et al.* Variabilidade espaço-temporal da precipitação em casos de eventos extremos na cidade do Recife. *In: Simpósio Brasileiro de Climatologia*, 13, 2018. **Anais**. Juiz de Fora, 2018. P. 110-119.
- NDENGA, B. A. *et al.* Characteristics of Aedes aegypti adult mosquitoes in rural and urban areas of western and coastal Kenya. **Plos One**. v. 12 n. 12, p. 1-14, 2017.
- OLIVEIRA, D. S. C. **Desigualdades intraurbanas de leptospirose no Recife**. 2009. Tese (Doutorado em Saúde Pública). Centro de Pesquisas Aggeu Magalhães, Fundação Oswaldo Cruz, Recife, 2009.
- OLIVEIRA, D. S. C.; GUIMARÃES, M. J. B.; MEDEIROS, Z. Modelo produtivo para a Leptospirose. **Revista de Patologia Tropical**, vol 38, n. 1, p. 17-26, 2009.
- ORTIZ, M. R. *et al.* Post-earthquake Zika virus surge: Disaster and public health threat amid climatic conduciveness. **Scientific Reports**. vol 7, p. 1-10. Article number: 15408, 2017.
- PAIXÃO, E. S. *et al.* History, Epidemiology, and Clinical Manifestations. of Zika: A Systematic Review. **AJPH**, v. 106, n. 4, p. 606-612, 2016.
- PENA FILHO, Carlos. **Livro Geral Poemas**. In: LEÃO, Tânia Carneiro (org.). Guia prático da Cidade do Recife, Recife: Graf. e Edit. Liceu, 1999. p.129.

SAN PEDRO, A.; OLIVEIRA, R. M.. Tuberculose e indicadores socioeconômicos: revisão sistemática da literatura. **Rev Panam Salud Publica**. v. 34, p.294–301, 2013.

PC, L. et al. Risk of tuberculosis in high-rise and high density dwellings: an exploratory spatial analysis. **Environmental Pollution**, v. 183, n.1, p. 40-45, 2013.

PINTO-FEIRREIRA, F. et al. Epidemiological relevance of dogs for the prevention of *Toxoplasma gondii*, *Neospora caninum* and *Leptospira* spp. **Rev. Bras. Parasitol. Vet**, Jaboticabal, v.28, n.3, p. 383–394, 2019.

POLACHINI, C. O.; FUJIMORI, K. Leptospirose canina e humana, uma possível transmissão conjuntival no Município de São Paulo. **Rev Pan-Amaz Saude**, São Paulo, Brasil, v.6, n.1, p. 59-6, 2015.

PUSTIGLIONE, M. Medicina do Trabalho e doenças emergentes, reemergentes e negligenciadas: a conduta no caso das febres da dengue, do Chikungunya e do Zika vírus. **Revista Brasileira de Medicina do Trabalho**, v. 14, n. 1, p.1-12, 2016.

QUINTERO, J. et al. Ecological, biological and social dimensions of dengue vector breeding in five urban settings of Latin America: a multi-country study. **BMC infectious diseases**, v. 14, n. 38, p. 1-13, 2014.

QUEIROGA, R. P. F. *et al.* **Distribuição espacial da tuberculose e a relação com condições de vida na área urbana do município de Campina Grande – 2004 a 2007.** Revista Brasileira de Epidemiologia, 15 (1). p. 222- 231, 2012.

REIS, R. B. et al. Impact of Environment and Social Gradient on *Leptospira* Infection in Urban Slums. **PLoS Negl Trop**, v. 2, n. 4, p. 1-10, 2008.

RIBEIRO, B. N. F. *et al.* Síndrome congênita pelo vírus Zika e achados de neuroimagem: o que sabemos até o momento? **Radiol Bras**. v. 50n. 5, p.314–322, 2017.

RODRIGUES, N. C. P et al. Risk factors for arbovirus infections in a low-income community of Rio de Janeiro, Brazil, 2015- 2016. **PloS one**, v.13, n.1, p. 1-15, 2018.

SAMPAIO, G. S. et al. Expansion of Zika virus circulation from Africa to the Americas, 1947-2018: a literature review. **Epidemiol. Serv. Saúde**, Brasília, v. 28, n. 2, 2019.

SANDHOLZ, S. **Urban centres in Asia and Latin America heritage and identities in changing urban landscapes**. 1 ed. New York: Springer, 2017.397 p.

SHIODE, N.; SHIODE, S.; ROD-THATCHER, E.; RANA, S.; VINTEN-JOHANSEN, P. The mortality rates and the space-time patterns of John Snows cholera epidemic map. **International Journal of Health Geographics**, vol. 14, n.21, p.14–21, 2015.

SILVA, L. J. O conceito de espaço na epidemiologia das doenças infecciosas. **Cad. Saúde Pública**, Rio de Janeiro , v. 13, n. 4, p. 585-593, Oct. 1997 .

SILVA, M. A. *et al.* Spatial distribution of tuberculosis from 2002 to 2012 in a midsize city in Brazil. **BMC Public Health**, v. 16, n. 912, p. 1-8, 2016.

SILVA, A. P. S C *et al.* Factores asociados a la ocurrencia de tuberculosis y su distribución espacial en una ciudad brasileña, 1991-2010. **Salud Colectiva**. v.14, n.1,p 77-91, 2018.

SILVA, A. P. S. C. **Duas décadas de Tuberculose em uma cidade do Nordeste Brasileiro: Avanços e desafios no tempo e no espaço**. 2016. Tese (Doutorado em Saúde Pública) - Centro de Pesquisas Aggeu Magalhães, Fundação Oswaldo Cruz, 2016.

SILVA, J. V. J. *et al.* A scoping review of Chikungunya virus infection: epidemiology, clinical characteristics, viral co-circulation complications, and control. **Acta Tropica**, v.188, n.1, p. 213-224, 2018.

SILVA, M. B. A. Perfil das arboviroses Dengue, Chikungunya e Zika no Distrito Sanitário III do município de Recife, Pernambuco, Brasil. **Revista Brasileira de Meio Ambiente**, v.9, n.1. p.039-050, 2021.

SOARES, T. S. M. Spatial and seasonal analysis on leptospirosis in the municipality of São Paulo, Southeastern Brazil, 1998 to 2006. **Rev. Saúde Pública**, São Paulo, v.44, n.2, p. 1-9, 2010

SONG, B. H., *et al.* Zika virus: History, epidemiology, transmission, and clinical presentation. **Journal of Neuroimmunology**, v. 308, n.1, p.50– 64, 2017.

SOUZA, A. *et al.* An Environmental GIS-based Variable Analysis on SARS-CoV-2 in the City of Recife, Brazil. **Research Square**. PREPRINT (Version 1), p. 1-24, 2020.

SOUZA, W. V. *et al.* Microcephaly epidemic related to the Zikavirus and living conditions in Recife, Northeast Brazil. **BMC Public Health**, v 18, 130, p.1-7, 2018.

TAUBITZ, W. *et al.* Chikungunya Fever in Travelers: Clinical Presentation and Course. **Clinical Infectious Diseases**, v.45, n.1, p.1-4, 2007.

VALLE, D.; PIMENTA, D.N.; AGUIAR, R. Zika, Dengue e Chikungunya: desafios e questões. **Revista Epidemiológica. Serv. Saude**, Brasília, v.25, n.2, p. 419-422, 2016.

VASCONCELOS, C. H.; FONSECA, F.R.; LISE, M. L. Z.; ARSKY, M. L. N. S. Fatores ambientais e socioeconômicos relacionados à distribuição de casos de leptospirose no Estado de Pernambuco, Brasil, 2001 –2009. **Cad. Saúde Colet**, Rio de Janeiro, v. 20 n.1, p. 49-5, 2012.

VICENTI-GONZALEZ, M. F. *et al.* Spatial Analysis of Dengue Seroprevalence and Modeling of Transmission Risk Factors in a Dengue Hyperendemic City of Venezuela. **PLoS Negl Trop Di**, v.11, n.1, p. 1-21, 2017.

WANG, F. **Quantitative Methods and Socio-Economic Applications in GIS**. 2 ed. New York: Taylor & Francis Group, 2014. 327 p.

WEBER, A. V. Exploring Local Influences on Zika Virus Rates in Puerto Rico Utilizing Geographically Weighted Regression. **Papers in Applied Geography**. v.4, n.1, p 1-14, 2018.

WASIŃSKI, B.; DUTKIEWICZ, J. Leptospirosis - Current risk factors connected with human activity and the environment. **Annals of agricultural and environmental medicine**, v. 20. n.1, p.239-44, 2013.

WORLD HEALTH ORGANIZATION. **Human leptospirosis : guidance for diagnosis, surveillance and control**. 1 ed. World Health Organization, 2003.

WOOLDRIDGE, J. M. **Introductory Econometrics: A Modern Approach**. 5 ed. Ohio: South-Western, 2003. 910 p.

XIAO, Y. *et al.* The influence of meteorological factors on tuberculosis incidence in Southwest China from 2006 to 2015. **Scientific Reports**, vol. 8, n.10053, p. 1-8, 2018.

YOO, D. Height and death in the Antebellum United States: A view through the lens of geographically weighted regression. **Journal of Economics and Human Biology**, Amsterdam v. 10, n.1, p. 43–53, 2012.