

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ
ESCOLA DE ARQUITETURA E *DESIGN*
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO URBANA**

GILBERTO BATISTA

**MINERAÇÃO DE DADOS CLIMATOLÓGICOS E VARIÁVEIS
URBANAS ASSOCIADAS À INCIDÊNCIA DE DENGUE:
ESTUDO DE CASO EM FOZ DO IGUAÇU, PARANÁ.**

**CURITIBA
2018**

GILBERTO BATISTA

**MINERAÇÃO DE DADOS CLIMATOLÓGICOS E VARIÁVEIS
URBANAS ASSOCIADAS À INCIDÊNCIA DE DENGUE:
ESTUDO DE CASO EM FOZ DO IGUAÇU, PARANÁ.**

Dissertação apresentada ao Programa de Pós-Graduação em Gestão Urbana da Pontifícia Universidade Católica do Paraná, Linha de Pesquisa Gestão e Tecnologias Ambientais, como requisito parcial à obtenção do título de Mestre em Gestão Urbana.

Orientador: Prof. Dr. Fábio Teodoro de Souza

**CURITIBA
2018**

Dados da Catalogação na Publicação
Pontifícia Universidade Católica do Paraná
Sistema Integrado de Bibliotecas – SIBI/PUCPR
Biblioteca Central

Batista, Gilberto
B333m Mineração de dados climatológicos e variáveis urbanas associadas à
2018 incidência de dengue: estudo de caso em Foz do Iguaçu, Paraná / Gilberto
Batista ; orientador: Fábio Teodoro de Souza. – 2018.
86 f. : il. ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Paraná,
Curitiba, 2018
Bibliografia: f. 82-86

1. Planejamento urbano. 2. Climatologia - Paraná. 3. Sistemas de
informação geográfica. 4. Dengue. I. Souza, Fábio Teodoro de. II. Pontifícia
Universidade Católica do Paraná. Programa de Pós-Graduação em Gestão
Urbana. III. Título.

CDD 20. ed. – 711.4

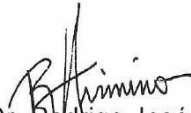
TERMO DE APROVAÇÃO

“MINERAÇÃO DE DADOS CLIMATOLÓGICOS E VARIÁVEIS URBANAS ASSOCIADAS À INCIDÊNCIA DE DENGUE: ESTUDO DE CASO EM FOZ DO IGUAÇU, PARANÁ”

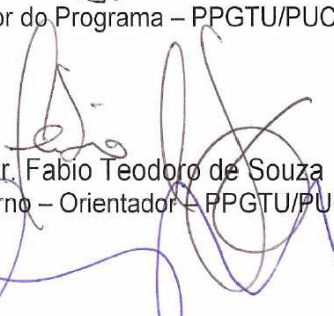
Por

GILBERTO BATISTA

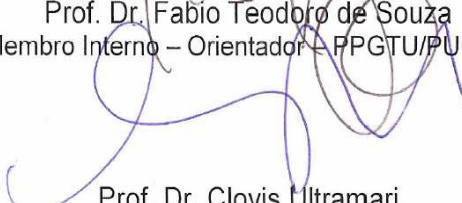
Dissertação aprovada como requisito parcial para obtenção do grau de Mestre no Programa de Pós-Graduação em Gestão Urbana, na modalidade Mestrado Interinstitucional – MINTER, em parceria entre a Pontifícia Universidade Católica do Paraná e o Centro Universitário Dinâmica das Cataratas, área de concentração em Gestão Urbana.



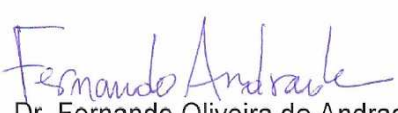
Prof. Dr. Rodrigo José Firmino
Coordenador do Programa – PPGTU/PUCPR



Prof. Dr. Fabio Teodoro de Souza
Membro Interno – Orientador – PPGTU/PUCPR



Prof. Dr. Clovis Ultramari
Membro Interno – PPGTU/PUCPR



Prof. Dr. Fernando Oliveira de Andrade
Membro Externo – UTFPR

Curitiba, 23 de fevereiro de 2018.

DEDICATÓRIA

*A Deus
e aos meus,
amigos e familiares.*

AGRADECIMENTOS

Ao meu orientador, Prof. Dr. Fábio Teodoro de Souza, pela oportunidade, confiança depositada, amizade e valiosas orientações.

Ao professor Carlos Hardt pela acolhida, comprometimento e sábios ensinamentos.

Aos professores membros da banca de qualificação, Clóvis Ultramari e Edilberto Nunes de Moura, pela orientação inicial.

Aos professores membros da banca de defesa, Clóvis Ultramari (PUCPR), Fernando Oliveira de Andrade (UTFPR), Luciene Pimentel da Silva (UERJ) pelas sugestões e acréscimos fundamentais para a finalização deste trabalho.

Aos professores das disciplinas que cursei durante o mestrado, pelas instigações e ensinamentos.

A Pollyana Mara Schlenker (PUCPR) pela acolhida e disponibilidade em alto grau.

As queridas Gisele Batista e Catarina Batista pela amizade incondicional.

Aos meus pais, João Batista e Lizete Batista, por proporcionarem sempre o bater de asas ao infinito, repleto de valores e amor.

Aos meus amores Lilian e Giovana, por todo apoio incondicional e abraços aconchegantes sempre a me esperar.

E a Deus, que sempre me proporciona muito mais do que ousar desejar.

“A sutileza humana nunca irá conceber uma invenção mais bela, mais simples ou mais clara do que a natureza, porque em suas invenções nada falta e nada é supérfluo”.
(Leonardo da Vinci)

RESUMO

A dengue atualmente é uma preocupante arbovirose que acomete os seres humanos na área urbana, em virtude de urbanização acelerada, deficiência de abastecimento de água, precariedade do saneamento e intenso uso de material descartável. Essas condições favorecem a rápida proliferação do principal vetor da doença, o mosquito da espécie *Aedes aegypti*. Nos últimos anos, ocorreu um aumento significativo do número de pessoas contaminadas pela doença. Esses dados são preocupantes e caracterizam um sério problema de saúde pública. Urge a necessidade de medidas preventivas por parte dos órgãos públicos de saúde e uma maior colaboração por parte da sociedade civil com ações que venham a combater o avanço da doença. Este estudo teve como objetivo elaborar um modelo de predição da incidência de dengue baseado em dados obtidos no período de 2010 a 2015 no município de Foz do Iguaçu, Paraná. Os dados foram coletados das fontes de informação da Secretaria Municipal de Saúde, Divisão de Vigilância Epidemiológica, Estações Meteorológicas de Altitude e Superfície do INMET (Instituto Nacional de Meteorologia) e ICEA (Instituto de Controle do Espaço Aéreo). O estudo mostrou um padrão de distribuição da dengue que se repetiu nos anos e nos bairros mais atingidos. O padrão temporal de dengue foi utilizado para prever níveis de incidência da doença nos meses seguintes e elaborar o modelo de predição. Dessa maneira, poderão ser adotadas medidas de vigilância e de mitigação, com estratégia de estratificação das áreas críticas. A análise da pesquisa foi realizada com a utilização de técnicas de mineração de dados e sistemas de informação geográfica. Os resultados apresentados demonstram que o clima do município, as incidências de chuvas, setores populosos e condições geográficas favoreceram a transmissão e evolução da doença. As análises espaciais e a aplicação das técnicas de mineração de dados apresentam potencial para vigilância, pois permitem visualizar os casos de incidência, a densidade de aglomeração dos pontos de vulnerabilidades e outros fatores importantes para o direcionamento das atividades de controle da doença.

Palavras-chave: gestão urbana, dengue, climatologia, mineração de dados, sistemas de informação geográfica, espaço urbano.

ABSTRACT

Dengue fever is currently one of the most critical diseases that affect the human beings in the urban area. It is a serious public health problem due to the accelerated urbanization, lack of water supply and of urban cleaning and intensive use of disposable material. These conditions favor the rapid proliferation of the main vector of the disease, the mosquito of the species *Aedes aegypti*. It has been observed that in Foz do Iguaçu, state of Paraná, Brazil, the number of people affected by the dengue fever has increased in recent years. For this reason, cooperation between governmental authorities and civil society has become essential to prevent the spread of the mosquito and thus to reduce the incidence of this disease. This study aimed the spatial-temporal distribution analysis of cases of dengue fever that occurred during the period of 2010 to 2015 in Foz do Iguaçu by geoprocessing tools and data miner. Data concerning the confirmed number of dengue cases were collected from Municipal Health Secretary, Epidemiological Surveillance Division e Surface Weather Stations of the INMET (National Institute of Meteorology) and ICEA (Institute of Airspace Control) and processed in the geoprocessing with Geographic Information System. The monthly and yearly distributions of cases in the urban area were correlated with pluviometric and elaborate and the prediction model. The presented results showed that the climate conditions of the city, rain, populated areas, flood and geographical conditions favored the development and transmission of dengue fever. The study showed the pattern of spatial distribution of dengue fever that was repeated in the years and in the most affected sectors of the city: west, north and central. The spatial patterns observed can be helpful to predict which sectors may have higher incidence of dengue in the following months. Thus, actions may be adopted for monitoring and control, with the strategy of stratification of areas. One can thus identify the characteristics of the sectors that have a higher risk of contamination and use the most appropriate steps. Techniques of spatial analysis and data miner have the potential for disease surveillance and control of dengue fever, because they can view the cases of incidence, the density of points clustering and other important factors to guide the activities of control.

Keywords: urban management, dengue fever, climatology, data mining, geographical information systems, urban space.

LISTA DE ABREVIATURAS E SIGLAS

AKLP	Associação do Jardim Karla, Laranjeiras e Petrópolis
BD	Banco de Dados
DATASUS	Departamento de Informática do Sistema Único de Saúde
DEN	Dengue
DVE	Departamento de Vigilância Epidemiológica
GIS	Sistema de Informação Geográfica
IA	Inteligência Artificial
ICEA	Instituto de Controle do Espaço Aéreo
INMET	Instituto Nacional de Meteorologia
INPE	Instituto Nacional de Pesquisas Espaciais
IPCC	<i>Intergovernmental Panel on Climate Change</i>
ISO	<i>International Organization for Standardization</i>
KDD	<i>Knowledge Discovery in Databases</i>
MD	Mineração de Dados
MS	Ministério da Saúde
OPAS	Organização Panamericana da Saúde
OMS	Organização Mundial da Saúde
PDM	Plano Diretor Municipal
PMFI	Prefeitura Municipal de Foz do Iguaçu
SIG	Sistema de Informação Geográfica
SMS	Secretaria Municipal da Saúde
SUS	Sistema Único de Saúde
UN	<i>United Nations</i>
WHO	<i>World Health Organization</i>

LISTA DE FIGURAS

FIGURA 1 - LOCALIZAÇÃO DO MUNICÍPIO DE FOZ DO IGUAÇU	21
FIGURA 2 - DIVISÃO TERRITORIAL DO MUNICÍPIO DE FOZ DO IGUAÇU	22
FIGURA 3 - ACESSIBILIDADE DO MUNICÍPIO DE FOZ DO IGUAÇU	25
FIGURA 4 - CRESCIMENTO URBANO DO MUNICÍPIO DE FOZ DO IGUAÇU	27
FIGURA 5 - RISCO GLOBAL DE DENGUE	28
FIGURA 6 - FÊMEA DO MOSQUITO <i>Aedes Aegypti</i>	30
FIGURA 7 - <i>Aedes Aegypti</i> NAS FASES DE DESENVOLVIMENTO	31
FIGURA 8 - RISCO DE DENGUE NO ESTADO DO PARANÁ	33
FIGURA 9 - ETAPAS DO PROCESSO DE DESCOBERTA DO CONHECIMENTO	45
FIGURA 10 - RELAÇÃO ENTRE KDD E MINERAÇÃO DE DADOS.	47
FIGURA 11 - MINERAÇÃO DE DADOS COM UTILIZAÇÃO DE DIFERENTES ÁREAS	48
FIGURA 12 - INTERAÇÃO DA MINERAÇÃO DE DADOS COM O BANCO DE DADOS	50
FIGURA 13 - ESPACIALIZAÇÃO DA INCIDÊNCIA DE DENGUE EM FOZ DO IGUAÇU EM 2010	57
FIGURA 14 - ESPACIALIZAÇÃO DA INCIDÊNCIA DE DENGUE EM FOZ DO IGUAÇU EM 2011	58
FIGURA 15 - ESPACIALIZAÇÃO DA INCIDÊNCIA DE DENGUE EM FOZ DO IGUAÇU EM 2012	59
FIGURA 16 - ESPACIALIZAÇÃO DA INCIDÊNCIA DE DENGUE EM FOZ DO IGUAÇU EM 2013	60
FIGURA 17 - ESPACIALIZAÇÃO DA INCIDÊNCIA DE DENGUE EM FOZ DO IGUAÇU EM 2014	61
FIGURA 18 - ESPACIALIZAÇÃO DA INCIDÊNCIA DE DENGUE EM FOZ DO IGUAÇU EM 2015	62
FIGURA 19 - DECLIVIDADE DAS REGIÕES DE FOZ DO IGUAÇU	63
FIGURA 20 - MAPA HIPSOMÉTRICO DAS REGIÕES DE FOZ DO IGUAÇU	64
FIGURA 21 - ORIENTAÇÃO DAS VERTENTES DAS REGIÕES DE FOZ DO IGUAÇU	65
FIGURA 22 - COMPARAÇÃO DOS MÉTODOS <i>TREE CLUSTERING</i> E <i>K-MEANS</i>	71
FIGURA 23 - MATRIZ CONFUSÃO PARA VARIÁVEL DE PREDIÇÃO T3	72
FIGURA 24 - MATRIZ CONFUSÃO PARA VARIÁVEL DE PREDIÇÃO T2	73
FIGURA 25 - MATRIZ CONFUSÃO PARA VARIÁVEL DE PREDIÇÃO T1	73

LISTA DE QUADROS

QUADRO 1 - CARACTERÍSTICAS DA POPULAÇÃO DE FOZ DO IGUAÇU.....	23
QUADRO 2 - CONJUNTO 1 DE REGRAS DE CLASSIFICAÇÃO.....	77
QUADRO 3 - CONJUNTO 2 DE REGRAS DE CLASSIFICAÇÃO.....	78
QUADRO 4 - CONJUNTO 3 DE REGRAS DE CLASSIFICAÇÃO.....	79
QUADRO 5 - CONJUNTO 4 DE REGRAS DE CLASSIFICAÇÃO.....	80

LISTA DE GRÁFICOS

GRÁFICO 1 - CLIMATOGRAMA DE FOZ DO IGUAÇU.....	66
GRÁFICO 2 - DENDROGRAMA DAS VARIÁVEIS TEMPORAIS	69
GRÁFICO 3 - DENDROGRAMA DAS VARIÁVEIS ESPACIAIS	70

SUMÁRIO

1 INTRODUÇÃO	15
1.1 OBJETIVOS	18
1.2 JUSTIFICATIVA	18
1.3 ESTRUTURA DA DISSERTAÇÃO	19
2 FUNDAMENTAÇÃO TEÓRICA.....	20
2.1 OS DADOS AMBIENTAIS DO MUNICÍPIO DE FOZ DO IGUAÇU	21
2.2 A CLIMATOLOGIA E O URBANO NO CONTEXTO DA DENGUE	28
2.2.1 A dengue sob a perspectiva urbana	28
2.2.2 A influência da dinâmica do clima urbano na dengue	34
2.3 A DIMENSÃO DA URBANIZAÇÃO COM O CLIMA	37
2.3.1 Os fatores climáticos e ambientais para análise urbana	38
2.4 A MINERAÇÃO DE DADOS NA GESTÃO URBANA	40
2.4.1 Análise espacial e o sistema de informação geográfica	40
2.4.2 A descoberta do conhecimento em base de dados	44
2.4.3 A mineração de dados	47
3 METODOLOGIA.....	51
3.1 REVISÃO BIBLIOGRÁFICA.....	51
3.2 COLETA DOS DADOS	52
3.2.1 Dados de variáveis climatológicas.....	52
3.2.2 Dados de variáveis urbanas	52
3.3 PREPARAÇÃO DOS DADOS.....	53
3.3.1 Regionalização dos dados de incidência de dengue	53
3.3.2 Discretização das variáveis para extração das regras	53
3.4 MODELAGEM.....	54
3.4.1 Montagem da matriz dos dados espaciais da incidência de dengue	55
3.4.2 Montagem da matriz dos dados temporais da incidência de dengue.....	55
3.4.3 Geração das regras	55
4 RESULTADOS	56
4.1 ESPACIALIZAÇÃO DA INCIDÊNCIA DE DENGUE EM FOZ DO IGUAÇU.....	56
4.2 GERAÇÃO DOS MAPAS DE PARÂMETROS ESPACIAIS.....	63
4.3 CORRELAÇÕES CLIMATOLÓGICAS E URBANAS COM A DENGUE	66
5 CONCLUSÃO.....	81
REFERÊNCIAS BIBLIOGRÁFICAS.....	82

1 INTRODUÇÃO

A dengue atualmente é uma preocupante arbovirose que acomete os seres humanos na área urbana, em virtude da urbanização acelerada, deficiência de abastecimento de água, precariedade do saneamento e intenso uso de material descartável.

O processo de urbanização das cidades e a complexidade de seus problemas são apontados como os principais desafios na busca de um desenvolvimento sustentável, principalmente ao verificar o impacto causado pelos espaços urbanos, devido à apropriação inadequada dos recursos naturais, degradação e contaminação do meio ambiente.

Efeitos do espaço urbano nos componentes do clima, como temperatura, umidade do ar, radiação solar e vento têm sido documentados no mundo (BESANCENOT, 2012).

O fator climático, em si, desempenha um papel importante para a sobrevivência humana, porque na atualidade as sociedades enfrentam ocorrências freqüentes de eventos extremos, como tempestades, enchentes, secas, e provoca um aumento das situações de vulnerabilidade socioambiental nas cidades (HOGAN; MARANDOLA JR, 2009).

Quanto à incidência de certas doenças que acometem o homem, o fator climático é importante, porque além de afetar a resistência do corpo humano a algumas doenças, influencia o crescimento, a propagação e a difusão de alguns organismos patogênicos ou de seus hospedeiros. Algumas doenças tendem a ser predominantes em certas zonas climáticas, enquanto outras, particularmente as contagiosas, seguem um padrão sazonal na sua incidência (AYOADE, 1998).

Ressalta-se aqui que os processos epidemiológicos da dengue se reproduzem como parte dos fatores de impacto ambiental na cidade. De forma sintética, a Resolução 001 do CONAMA – Conselho Nacional do Meio Ambiente, estabelece que impacto ambiental seja definido como “qualquer alteração das propriedades físicas, químicas e biológicas do meio ambiente, causada por qualquer matéria resultante das atividades humanas” (MENDONÇA, 2000). Estas alterações podem afetar a saúde, segurança e bem-estar da população.

Para Mendonça e Leitão (2008), os impactos socioambientais constituem situações de risco ambiental urbano à medida que, num determinado momento e em determinadas circunstâncias, teriam a capacidade de causar danos diretos à saúde e qualidade de vida dos indivíduos.

O estudo das relações do clima com a saúde é relevante, sobretudo em perspectiva das mudanças climáticas, na previsão de seus prováveis efeitos e vulnerabilidades (SILVA; MARIANO; SCOPEL, 2008).

Do ponto de vista da saúde, o ambiente térmico em áreas urbanas apresenta forte relação entre a termorregulação e a regulação circulatória, que levam à sobrecarga do organismo e ao comprometimento da saúde. A compreensão do clima urbano, em diversas partes do mundo, deve servir para a gestão urbana e adoção de medidas protetoras da saúde (MINAYO, 2002).

As mudanças que ocorrem no espaço urbano com as variáveis climatológicas sobre uma região são vistas como um fator preocupante pelos diversos pesquisadores (MELLO, 2011). O surgimento dos aglomerados urbanos tem contribuído para o desenvolvimento de condições inadequadas de habitação, irregularidade no abastecimento de água e destinação imprópria de resíduos, no qual criam-se ambientes sanitariamente impróprios, uma vez que ocasionam uma variação considerável nas morbidades e vetores de doenças.

Somado a isso, têm-se o crescimento da indústria de materiais não biodegradáveis, clima que se apresenta agravado pelo aquecimento global e o crescente trânsito de pessoas e cargas entre países. Esses fatores são considerados importantes no aumento dos potenciais criadouros do mosquito vetor da dengue (MENDONÇA, 2000).

A propagação do vetor e do vírus conduziu a uma ressurgência global de epidemias de dengue, detectada, principalmente, em muitos centros urbanos. Desse modo, o desafio está em evitar a ocorrência de óbitos pela doença e reduzir os impactos causados. Embora, desde o início do século XX, esforços tenham sido realizados tanto na erradicação da doença, quanto do seu principal vetor, ainda é prevalente no Brasil (WHO, 2016).

As condições ambientais influenciam no processo de desenvolvimento do vetor da dengue, permitem definir o curso da doença, de modo a permitir analisar as características ontológicas da arbovirose em curso no espaço urbano (BARRETO; TEIXEIRA, 2008).

A compreensão do curso da doença a partir da análise espacial, correlações climatológicas e variáveis urbanas são fundamentais para a tomada de decisões acerca da programação das ações de saúde, uma vez que proporciona uma importante estratégia na avaliação da vigilância epidemiológica e na assistência ao paciente, além de propiciar o planejamento e a operacionalização das ações de controle, ao criar abordagens distintas e ações diferenciadas em localidades que apresentam alto risco para a ocorrência da doença (REITER, 2013).

O relatório do IPCC (2015) demonstra que o tipo de estudo que relaciona o clima à saúde possibilita uma visão mais ampla e segura da relação entre as mudanças climáticas e seus impactos sobre a humanidade, desta maneira, o relatório demonstra um nível de confiança de que as mudanças regionais recentes nas temperaturas têm influências consideráveis nos impactos em muitos sistemas físicos e biológicos.

A relação entre clima e transmissão dinâmica de patógenos foi documentada através de pesquisas por todo o mundo. Esta conexão é importante, pois um terço da mortalidade humana é resultado de doenças infecciosas (BESANCENOT, 2012).

Com isso, este trabalho se propõe a analisar espacialmente e temporalmente a incidência da doença, através da elaboração de um modelo de predição que expressa as características entomológicas da doença. A pesquisa utiliza técnicas de mineração de dados e sistemas de informação geográfica para avaliar a notificação dos casos de dengue da Secretaria Municipal de Saúde e Departamento de Vigilância Epidemiológica do Estado do Paraná, de modo a confrontar essa informação com as variáveis climatológicas e urbanas para o município de Foz do Iguaçu.

1.1 OBJETIVOS

O objetivo principal desta pesquisa é elaborar um modelo de predição da incidência de dengue por correlações climatológicas como ferramenta de gestão urbana na cidade de Foz do Iguaçu, Paraná.

Com base nos procedimentos metodológicos adequados, bem como do referencial teórico pertinente ao tema e assuntos relacionados, constituem objetivos específicos:

- Georreferenciar casos de dengue notificados em Foz do Iguaçu;
- Elaborar um referencial teórico sobre a temática dengue associada às variáveis climatológicas e urbanas;
- Realizar análise espacial e temporal dos pontos localizados, com a integração de dados através do sistema de informação geográfica;
- Evidenciar as correlações climatológicas e ambientais com as ocorrências de dengue no espaço urbano.

1.2 JUSTIFICATIVA

Nas duas últimas décadas tem ocorrido um processo de urbanização intensivo e mais de 50% da população mundial vivem em áreas urbanas (BECERRA et al., 2013).

Este fenômeno é ainda mais acentuado na América Latina e em especial no Brasil, onde quase nove em cada dez pessoas vivem em centros urbanos (UN, 2012).

O estudo da relação entre clima e saúde é de suma importância, pois a alteração climática causada pela urbanização ganhou impulso a partir de meados do século XX e com a ampliação do tamanho das cidades (BESANCENOT, 2012).

A dengue se configura como a arbovirose mais relevante do mundo. Estima-se que aproximadamente 42% da população mundial vive em territórios, onde o vírus da doença pode ser transmitido (WHO, 2016).

A ligação entre as condições climatológicas e a transmissão da dengue ainda não é clara em algumas partes do mundo, principalmente no espaço urbano. Na literatura em geral, são encontrados poucos trabalhos que analisam a relação entre as alterações climáticas provocadas pelo fenômeno (HOGAN; MANDAROLA, 2009).

O clima pode influenciar fortemente a dinâmica populacional de doenças vetoriais e conseqüentemente é a componente chave da ecologia dessas doenças (CONFALONIERI, 2003).

Nos últimos anos, ocorreu um aumento significativo do número de pessoas contaminadas pela dengue. Esses dados são preocupantes e caracterizam um sério problema de saúde pública (ALMEIDA, 2013).

A dinâmica que envolve o processo de urbanização é complexa, o tempo entre a causa (padrões atmosféricos) e o efeito (saúde humana) é *não linear*, o que dificulta a aplicação de métodos tradicionais de investigação (RYDIN et al., 2012).

A saúde humana é fortemente influenciada pelo clima através das condições térmicas, de dispersão (ventos e poluição) e de umidade do ar que exercem influência sobre a manifestação de muitas doenças e criam condições favoráveis ao desenvolvimento dos transmissores de doenças contagiosas (MARTINS, 2012).

A alta taxa de crescimento das grandes cidades induz desequilíbrios nos padrões atmosféricos com conseqüências negativas à saúde pública (CEOHATS, 1996).

O entendimento dos conceitos básicos sobre a associação entre as diversas variáveis que envolvem a cidade frente ao espaço urbano valoriza o papel da gestão urbana (ULTRAMARI; FIRMINO, 2010).

1.3 ESTRUTURA DA DISSERTAÇÃO

Este estudo está organizado em cinco capítulos. O primeiro capítulo apresenta uma visão geral da dissertação, introduz o problema do processo de urbanização, as influências das variáveis climáticas e urbanas associadas à incidência de dengue no espaço urbano. Aqui também se descreve o tema, os objetivos e a justificativa da pesquisa.

O segundo capítulo aborda a fundamentação teórica das temáticas envolvidas no objeto da pesquisa, dengue, climatologia e espaço urbano, discorre sobre os sistemas de informações geográficas, a descoberta do conhecimento em base de dados e a mineração de dados.

O terceiro capítulo detalha a metodologia adotada neste trabalho e relata as atividades envolvidas para a escolha das escalas temporais e espaciais, da preparação dos dados e de modelagem.

O quarto capítulo demonstra a espacialização da incidência de dengue e a construção dos mapas de parâmetros espaciais para a análise das correlações climatológicas e urbanas com a ocorrência da dengue, proporciona o modelo de predição de incidência de dengue aplicado à gestão urbana, com a apresentação dos resultados.

No quinto capítulo são apresentadas as considerações finais do trabalho, de maneira a realizar o resgate dos objetivos propostos e as contribuições da pesquisa.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo é destinado aos conceitos investigados por meio da pesquisa bibliográfica e documental, bem como às outras ferramentas de busca utilizadas para obter conceitos adequados e, assim, realizar as análises necessárias dos temas propostos na dissertação.

O capítulo aborda as características da dengue, climatologia e espaço urbano, ao considerar algumas abordagens de classificação e correlações, para discorrer sobre os Sistemas de Informações Geográficas - SIG, bem como a descoberta do conhecimento em base de dados e a mineração de dados.

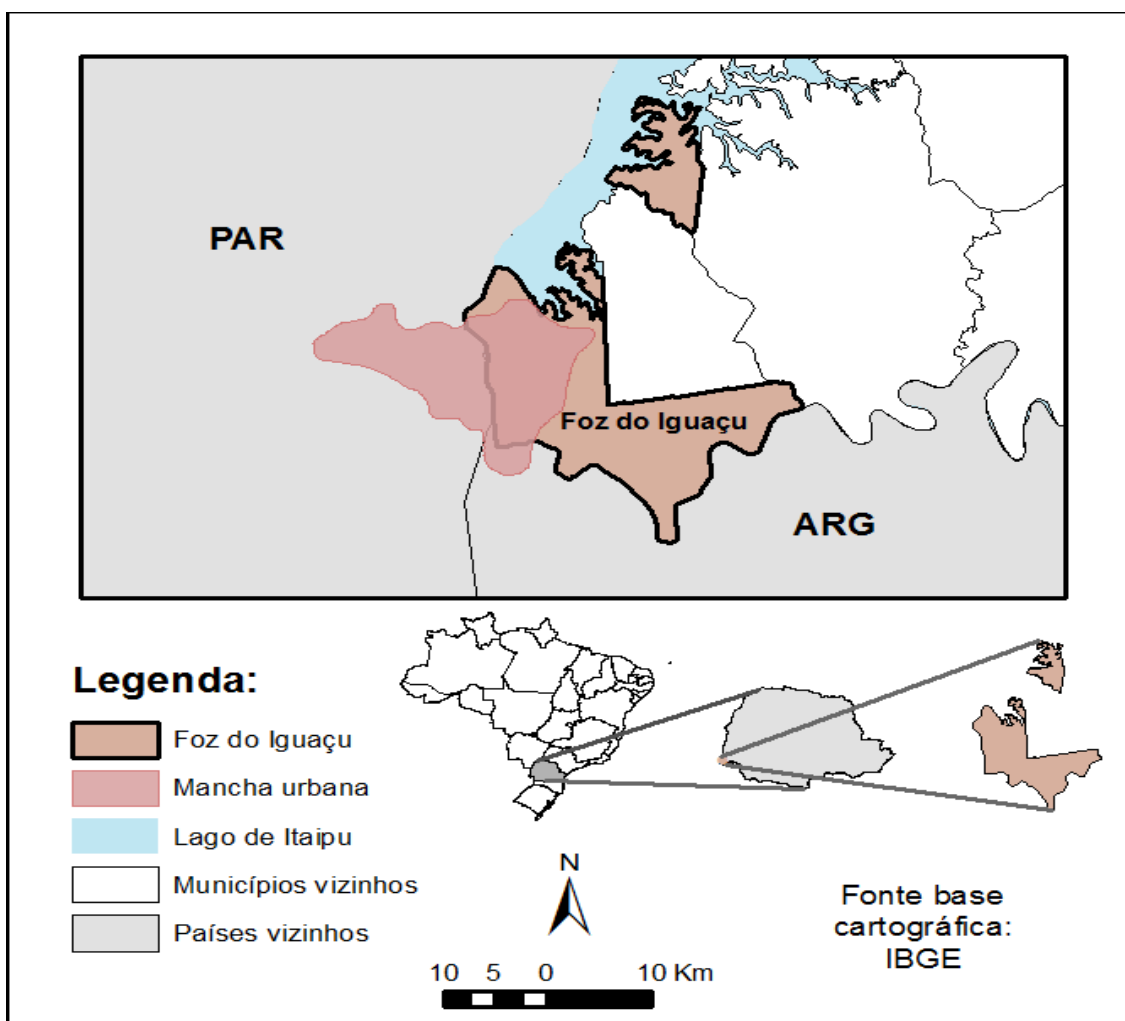
A fundamentação teórica pautou-se primeiramente num percurso sobre o perfil urbano do município de Foz do Iguaçu, as características da dengue, para depois buscar conceitos e bases fundamentais para o entendimento das questões climatológicas e urbanas que propiciam a proliferação da doença. Foram ressaltados referenciais sobre questões ligadas a riscos e vulnerabilidade ambiental, além de condições climáticas e de meio ambiente.

2.1 OS DADOS AMBIENTAIS DO MUNICÍPIO DE FOZ DO IGUAÇU

Foz do Iguaçu é uma cidade que está em pleno desenvolvimento. Além do seu potencial turístico, que atrai visitantes de todo o planeta, atualmente a cidade também acolhe a população acadêmica, atraída pela grande quantidade de instituições de ensino superior.

A figura 1 ilustra a localização do município de Foz do Iguaçu.

FIGURA 1 - LOCALIZAÇÃO DO MUNICÍPIO DE FOZ DO IGUAÇU



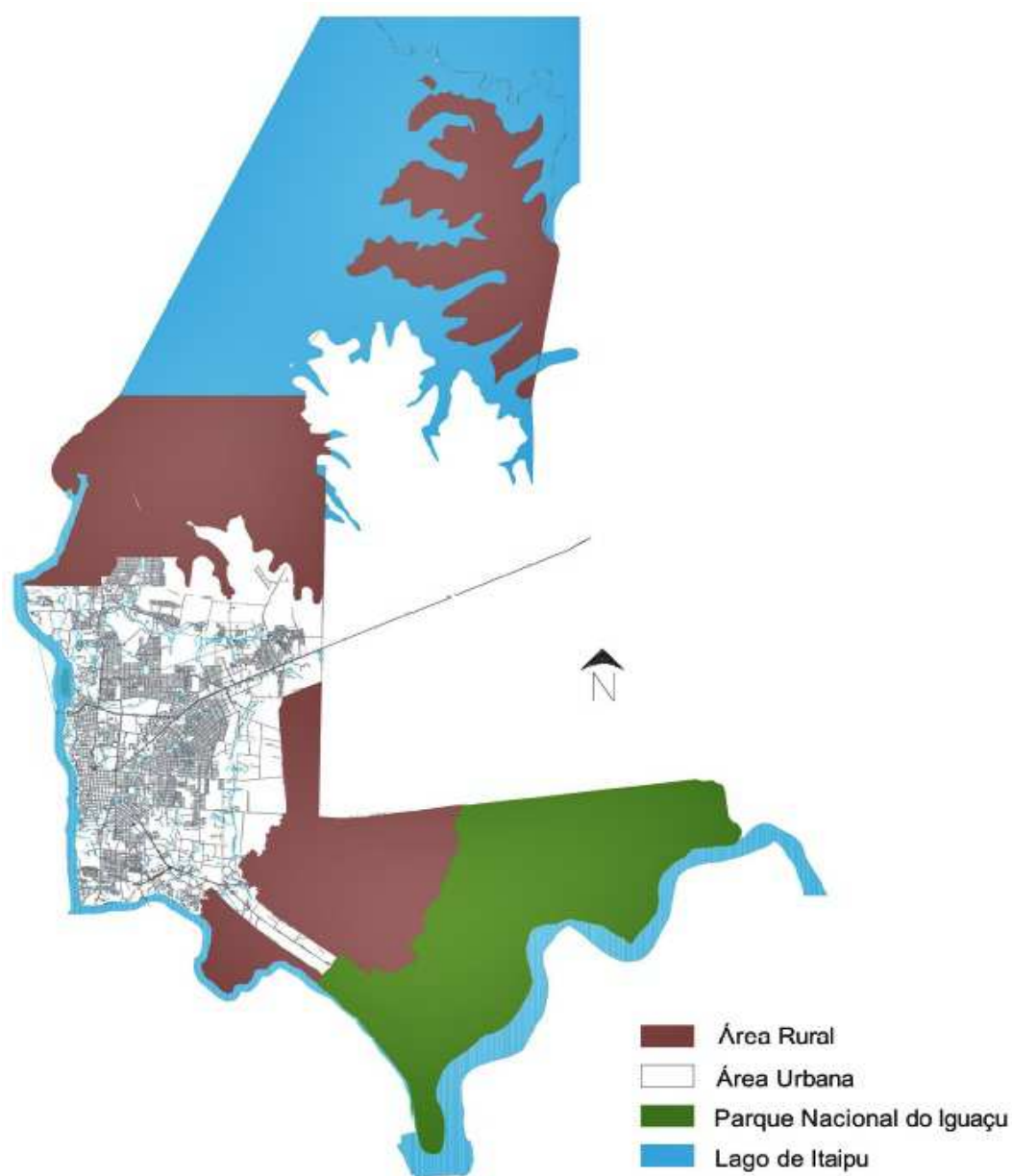
FONTE: PMFI (2015).

A pesquisa foi realizada na cidade de Foz do Iguaçu, Paraná. Conforme informado pela Prefeitura Municipal de Foz do Iguaçu, o município está geograficamente situado a 25° 32' 55" de latitude sul e 54° 35' 17" de longitude oeste, com altitude média de 173 metros, no extremo oeste do estado (PMFI, 2015).

O município possui área territorial totalizada em 617,71 km². A parte do município ocupada pela área urbana totaliza 191,46 km² e corresponde a 31,0% do total, enquanto a área rural representa 138,17 km² e responde por 22,37% do total.

O Parque Nacional do Iguaçu, por sua vez, ocupa 138,6 km², o que representa 22,44% da área total, e a área alagada pela Usina Hidrelétrica de Itaipu corresponde a 149,10 km², que equivale a 24,19% do território do município (PMFI, 2015). A figura 2 ilustra a divisão territorial do município de Foz do Iguaçu.

FIGURA 2 - DIVISÃO TERRITORIAL DO MUNICÍPIO DE FOZ DO IGUAÇU



FONTE: PMFI (2015).

Atualmente, a grande circulação de produtos, pessoas e informações, entre territórios distintos, influenciam culturalmente as populações de muitas cidades e se destacam como os principais motores do desenvolvimento econômico e social de quase todos os países (LEFF, 2002).

Estima-se que a população do município se aproxime dos 400 mil em 2030. Todo esse fluxo de pessoas além de proporcionar o desenvolvimento para a cidade trouxe algumas externalidades, um dos pontos preocupantes é o aumento dos vetores de doenças como a dengue (OPAS, 2013).

As informações estatísticas da população de Foz do Iguaçu, Paraná, são demonstradas no quadro abaixo:

QUADRO 1 - CARACTERÍSTICAS DE POPULAÇÃO DE FOZ DO IGUAÇU

População residente	256.088 pessoas
População residente - Homens	124.218 pessoas
População residente - Mulheres	131.870 pessoas
População residente alfabetizada	221.377 pessoas
População residente que frequentava creche ou escola	84.940 pessoas
População residente, religião católica apostólica romana	149.959 pessoas
População residente, religião espírita	3.604 pessoas
População residente, religião evangélicas	70.420 pessoas
Valor do rendimento nominal mediano mensal per capita dos domicílios particulares permanentes - Rural	502,50 reais
Valor do rendimento nominal mediano mensal per capita dos domicílios particulares permanentes - Urbana	520,00 reais
Valor do rendimento nominal médio mensal dos domicílios particulares permanentes com rendimento domiciliar, por situação do domicílio - Rural	2.759,95 reais
Valor do rendimento nominal médio mensal dos domicílios particulares permanentes com rendimento domiciliar, por situação do domicílio - Urbana	2.693,40 reais

FONTE: PMFI (2015).

Os aspectos geográficos do município de Foz do Iguaçu são apresentados:

- a) Climatologia: clima temperado subtropical úmido, com verões quentes, geadas pouco frequentes e chuvas em todos os meses do ano;
- b) Pluviosidade: precipitação média anual de 1.798,92 mm;
- c) Umidade relativa do ar média de 73,92%;
- d) Relevo: altitude máxima do município de 321 metros (no Parque Nacional do Iguaçu, próximo à divisa com Santa Terezinha de Itaipu); altitude máxima no perímetro urbano de 275 metros (região de Três Lagoas); altitude mínima no perímetro urbano de 100 metros (foz do Rio Iguaçu); altitude da sede de 164 metros;
- e) Vegetação: mata subtropical na região do Parque Nacional do Iguaçu e floresta tropical de várzea nas margens dos rios;
- f) Temperatura média anual máxima: 40° C (janeiro), mínima: 0° C (junho);
- g) Hidrografia: composta de nove microbacias, das quais sete são circunscritas ao perímetro municipal. Principais rios: Paraná, Iguaçu, Tamanduá, São João, Almada, M'Boicy e Monjolo. Afluentes do Rio Paraná: rios M'Boicy, Almada, Monjolo, O'Coí, Cuê e Guabirola. Afluentes do Rio Iguaçu: rios Tamanduá, São João e Córrego Carimã.

Segundo o Plano Diretor (2015), o núcleo de formação de Foz do Iguaçu está localizado na confluência do rio Paraná com seu afluente, rio M Boicy, com crescimento orientado para leste na direção da Vila Maracanã, a partir dos eixos das Avenidas República Argentina e Jorge Schimmelpfeng. Estas são as saídas para Curitiba e Argentina, através do antigo Porto General Meira.

A partir de 1970, com a instalação das obras da Hidrelétrica de Itaipu Binacional, provocou um dinamismo na utilização das áreas, de modo a forçar uma expansão em todas as direções (exceto para a direção leste, onde surgiu um grande número de conjuntos habitacionais e bairros com uma população de baixa renda).

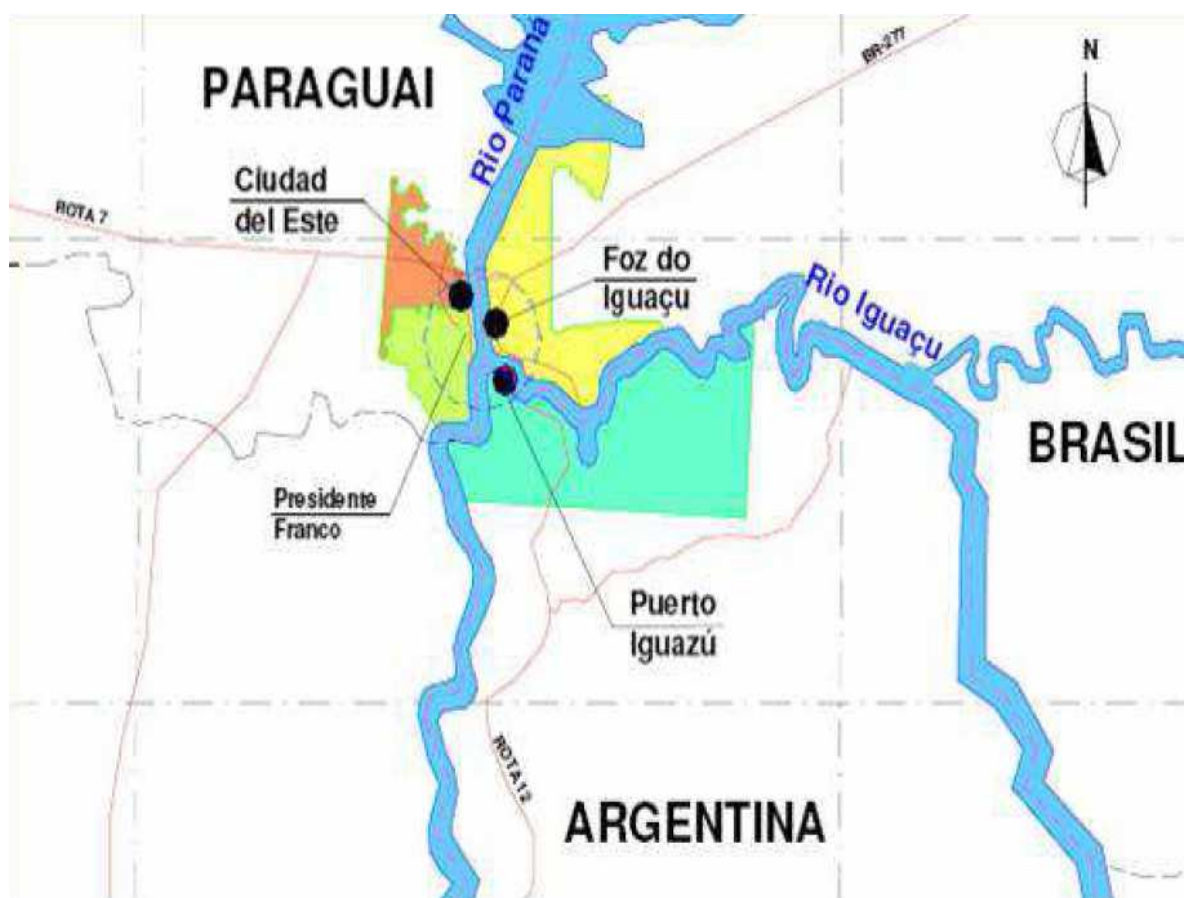
A cidade teve seu desenvolvimento no decorrer dos principais eixos viários formados na direção norte-sul e do centro para o leste, de maneira a formar uma mancha urbana em torno de grandes áreas desocupadas (vazios urbanos).

A verticalização das construções está mais evidente ao longo das Avenidas Paraná e JK, na rua Belarmino de Mendonça e no próprio centro da cidade.

No início dos anos 90 surgiram blocos uniformes e extensos de conjuntos habitacionais, com o nascimento de construções de até quatro pavimentos, destinadas às famílias de renda média capazes de se inserir em programas de financiamento.

A partir do centro da cidade, através da BR-469, chega-se às Cataratas do Iguaçu, localizada dentro do Parque Nacional do Iguaçu, e à Ponte Tancredo Neves, com passagem para a Argentina, em *Puerto Iguazu*. Pela Rota 12, em território argentino, chega-se a Posadas, capital da província de *Misiones*, e ao sul, chega-se até Buenos Aires. A figura 3 ilustra a acessibilidade do município de Foz do Iguaçu.

FIGURA 3 - ACESSIBILIDADE DO MUNICÍPIO DE FOZ DO IGUAÇU



FONTE: PMFI (2015).

Foz do Iguaçu está distante 640 km da capital do estado, Curitiba, e a 731 km do Porto de Paranaguá. Num raio de 1.600 km, situam-se cidades como São Paulo, Rio de Janeiro, Buenos Aires, Assunção e Montevidéu.

A principal via de acesso terrestre é a BR-277 ou Rodovia Panamericana, que vai do Porto de Paranaguá, passa pela capital do estado, até a Ponte Internacional da Amizade, e promove a ligação com o Paraguai, em *Ciudad del Este*. O acesso também se faz por meio da Rota Internacional 01, sentido Paraguai / Brasil, que liga Assunção a *Ciudad del Este*. Nos mapas a seguir é possível identificar a evolução urbana do município nas quatro décadas seguintes à instalação da Hidrelétrica Itaipu Binacional.

O mês de maior precipitação é maio, com média de 262,06 mm; março com 56,88 mm, e agosto com 59,25mm, são os meses menos chuvosos.

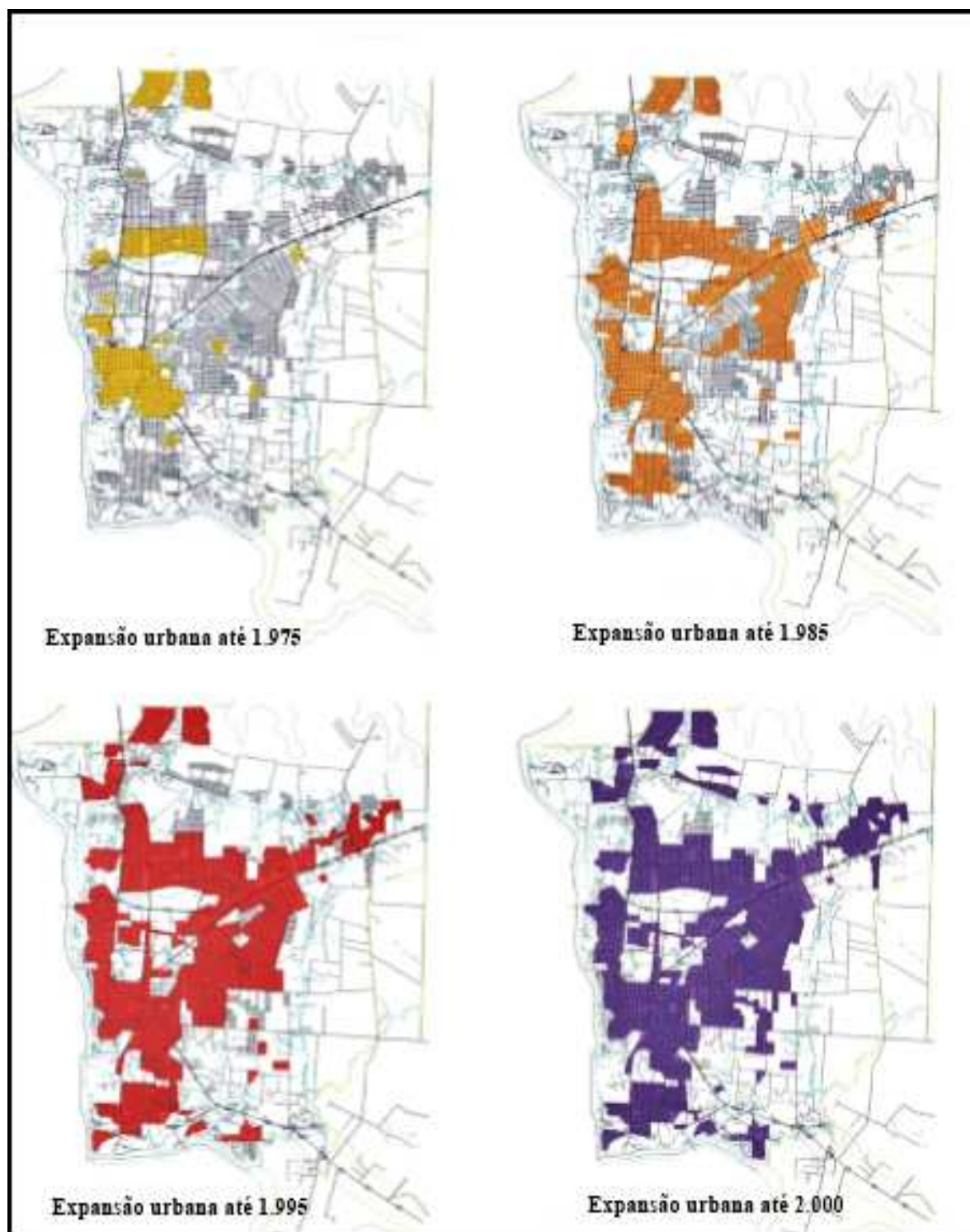
Também é importante a atuação da Baixa do Chaco, principalmente no verão, quando as temperaturas médias dos meses mais quentes ultrapassam os 22°C. Em razão de sua proximidade com a Bacia do Paraná, predomina o clima subtropical.

A média anual da umidade relativa do ar é de 73,92%, uniforme ao longo do ano. A região sofre influência dos dois grandes rios, Paraná e Iguaçu, e do lago de represamento da Hidrelétrica de Itaipu, que provoca o aumento dessa umidade.

As temperaturas máximas absolutas atingem 40°C (janeiro) e em qualquer mês podem ser superiores a 30°C. As mínimas absolutas dificilmente são inferiores a 0°C (julho). A média do mês mais quente (janeiro) é de 28,1°C e a do mês mais frio (julho) é de 14,6°C. A cidade apresenta bruscas variações de temperaturas com média anual de 27,7°C.

A figura 4 mostra o crescimento urbano do município de Foz do Iguaçu.

FIGURA 4 - CRESCIMENTO URBANO DO MUNICÍPIO DE FOZ DO IGUAÇU



FONTE: PMFI (2015).

O acentuado crescimento urbano do município aconteceu a partir da década de noventa, com a presença do processo de verticalização e aumento do número de conjuntos habitacionais.

2.2 A CLIMATOLOGIA E O URBANO NO CONTEXTO DA DENGUE

Neste capítulo são tratados aspectos concernentes à influência do clima urbano na proliferação de doenças, em específico da dengue, de forma a demonstrar as relações existentes.

2.2.1 A dengue sob a perspectiva urbana

A dengue configura-se como uma das enfermidades virais transmitidas por mosquito que mais causam infecções e uma das mais importantes para a humanidade. Nas últimas décadas, ela se tornou um problema de saúde pública e coletiva global (MINISTÉRIO DA SAÚDE, 2006).

A Organização Mundial da Saúde estima que entre 50 a 100 milhões de pessoas se infectam com a dengue anualmente em mais de 100 países.

A figura 5 demonstra o risco global de dengue, que corresponde as regiões tropicais e subtropicais que apresenta coloração em laranja, com temperaturas acima das médias térmicas mensais abaixo de 10 °C, uma vez que os países em laranja são aqueles onde a dengue foi registrada.

FIGURA 5 - RISCO GLOBAL DE DENGUE



FONTE: WHO (2016).

Por ano, cerca de 550 mil doentes no mundo carecem de hospitalização e 20 mil notificam óbito em consequência da doença. Calcula-se que 42% da população mundial vivem em áreas onde o vírus da dengue pode ser transmitido. A extensão geográfica, tanto dos vetores como do vírus, levou ao ressurgimento global das epidemias de dengue e, entre elas, a de febre hemorrágica (WHO, 2016).

Esta relação entre dengue e as temperaturas pode ser visualizada em um mapa disponibilizado pela *World Health Organization* (WHO) (2016), onde são apresentados os países compreendidos dentro das áreas de risco de transmissão da dengue.

A dengue faz parte das doenças causadas por vírus, transmitida no Brasil por um artrópode denominado *Aedes aegypti*. A doença é classificada como uma arbovirose.

As arboviroses são viroses emergentes por natureza, já que nenhuma delas é originalmente uma doença humana, uma vez que se tornam importantes quando ocorre alguma modificação ecológica significativa que altere seu *habitat* natural, de maneira a realizar as modificações de reservatórios, vetores e até mesmo virulência. (SILVA; ANGERAMI, 2008).

A reemergência de arboviroses é considerada fenômeno natural e apontada como processo de evolução e adaptação de espécies, pois durante a infecção de diferentes organismos o vírus pode adquirir maior viabilidade de infecção no hospedeiro, o que gera estirpes mais virulentas, ou melhor, adaptadas (CASSEB et al., 2013).

Além disso, as arboviroses são consideradas reemergentes, pois sua incidência na espécie humana tem aumentado nas últimas décadas, ou por existir a ameaça de aumentar num futuro próximo (CDC, 1994).

A circulação de arboviroses pode ocorrer por diferentes mecanismos, como o repasto de sangue contaminado do hospedeiro vertebrado (transmissão horizontal), transovariana e transmissão venérea (transmissão vertical) (BEATY; STEPHEN, 2005).

O vírus da dengue pertence ao gênero *Flavivirus*, família *Flaviviridae*, e sua infecção é causada por quatro sorotipos de *Flavivirus*: DEN-1, 2, 3 e 4, que produzem imunidade sorotipo específica. (BRASIL, 2008).

A doença pode manifestar-se como a dengue clássica, ou como a dengue hemorrágica e síndrome de choque por dengue. As duas últimas constituem-se como as formas clínicas mais graves (BRASIL, 2008).

Para Almeida Filho e Rouquayrol (2003):

A dengue é uma enfermidade viral aguda que se caracteriza por início súbito com febre alta, dura de 3 a 5 dias, com cefaléia intensa, mialgias, dor retroorbitária, anorexia, alterações do aparelho gastrointestinal e exantema. É transmitida pela picada de fêmeas do mosquito *Aedes Aegypti* e tem como reservatório o conjunto homem-mosquito. O mosquito torna-se infectante de 8 a 12 dias depois de alimentar-se com sangue contaminado, e continua assim pelo resto de sua vida, podendo, inclusive, transmitir a infecção, por via transovariana a seus ovos. (ALMEIDA FILHO & ROUQUAYROL, 2003, p. 279)

O *Aedes aegypti* se desenvolve através da metamorfose. Os ovos postos na superfície da água dão nascimento às larvas que, durante esta fase, posteriormente se transformam em pupas, e finalmente em alados, de maneira a assumir a forma adulta dos mosquitos.

O tempo de evolução do ovo ao alado varia de acordo com as condições da água, da temperatura e da alimentação da larva. Em uma mesma desova a fêmea coloca os ovos em vários recipientes, o que garante a sobrevivência e a dispersão de sua prole (BRASIL, 2008).

A figura 6 mostra a fêmea do mosquito *Aedes aegypti*.

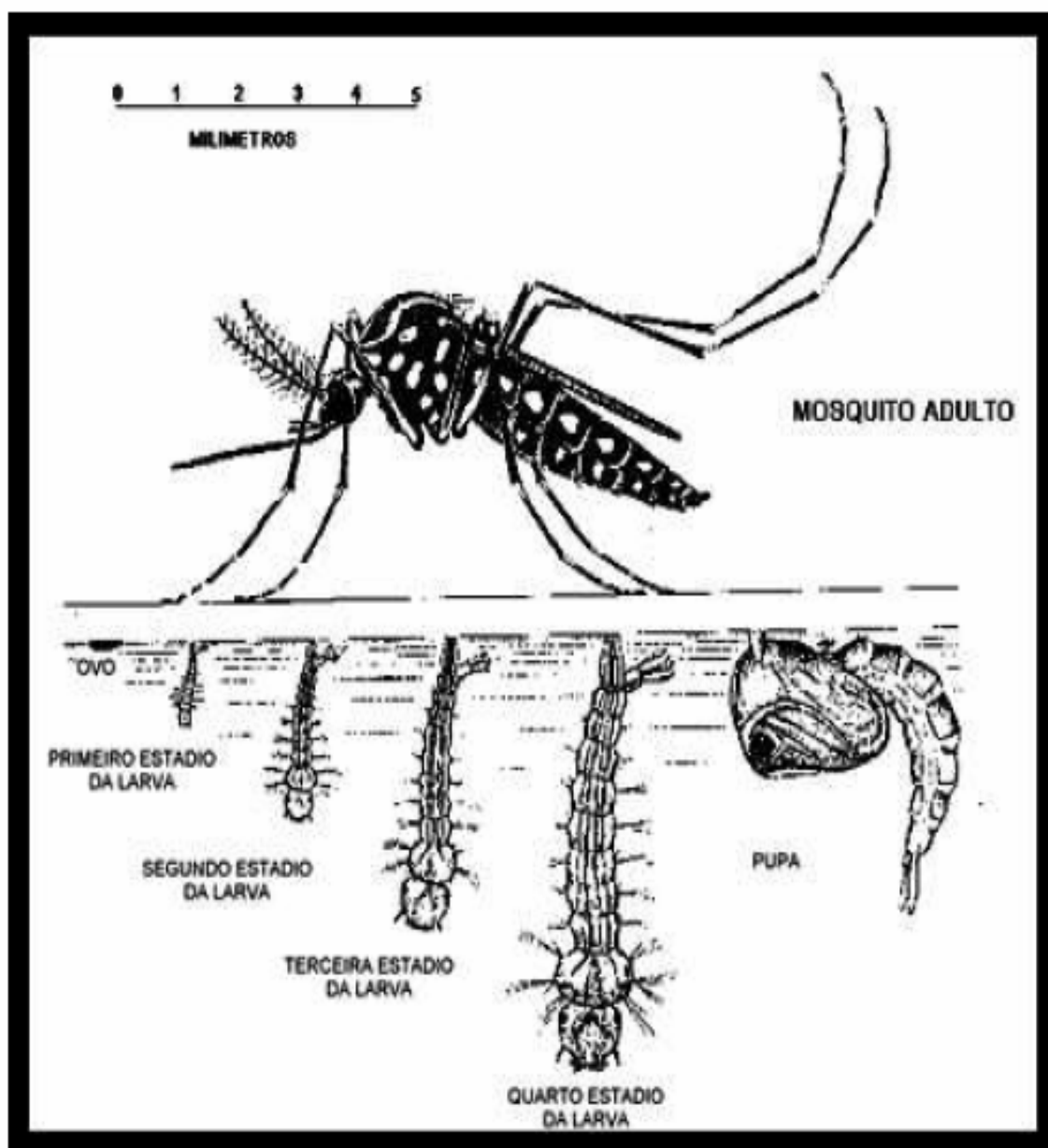
FIGURA 6 - FÊMEA DO MOSQUITO *Aedes Aegypti*



FONTE: Brasil (2008).

A figura 7 mostra as fases de desenvolvimento do *Aedes aegypti*.

FIGURA 7 - *AEDES AEGYPTI* NAS FASES DE DESENVOLVIMENTO



FONTE: Brasil (2008).

No Brasil, os programas de combate e controle da dengue se figuram como a maior campanha de saúde pública. A taxa de incidência da doença aumenta consideravelmente em anos de epidemias nacionais (MINISTÉRIO DA SAÚDE, 2006).

A taxa de incidência de dengue corresponde ao número de casos novos confirmados de dengue (clássico e febre hemorrágica de dengue) por 100 mil habitantes, na população residente em determinado espaço geográfico, no ano considerado (DATASUS, 2012).

A compreensão desta dinâmica global de circulação se tornou essencial para o entendimento dos processos de saúde das populações, principalmente porque o fluxo internacional de pessoas e mercadorias se apresenta como um importante condicionante no processo de difusão de doenças.

De acordo com Organização Panamericana de Saúde (OPAS, 2013) a incidência e as epidemias de dengue no mundo aumentaram nos últimos 35 anos. Até a década de 1950 eram notificados casos de dengue em apenas nove países; na década de 1980, 26 países passaram a notificar os casos e, a partir de 1990, foram informados casos em mais de 100 países ao redor do mundo. Somente no ano de 2002, 69 países notificaram casos positivos.

O cenário global atual se afirma com o desenvolvimento da hiperendemicidade em muitos centros urbanos de cidades localizadas entre os trópicos (WHO, 2016). Nas Américas, durante o final da década de 1970, intensas epidemias de dengue já eram registradas na Colômbia. Posteriormente, durante a década de 1980, vários países da América do Sul voltaram a registrar epidemias, entre eles, Brasil, Bolívia, Paraguai, Equador e Peru. Desde então, a doença tem afetado constantemente, com surtos e epidemias recorrentes, quase todos os países da América Central e do Sul (PIEROTE, 2009).

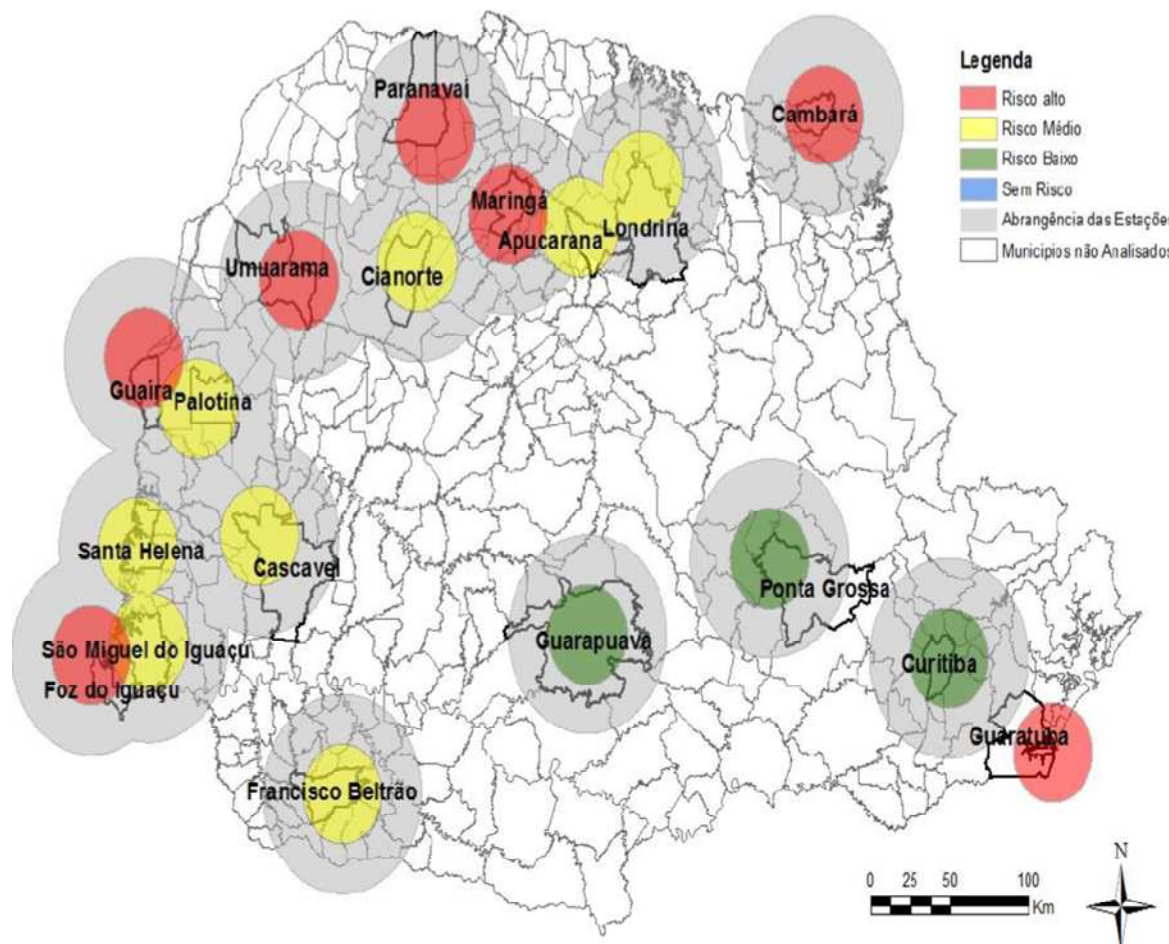
Ainda sobre a evolução dos casos de dengue, nos países latinos da América do Sul, ocorreu aumento significativo nos últimos anos, com destaques para os anos de 2002, 2007 e 2010, quando as epidemias registraram elevados índices se comparadas com os outros anos no mesmo período. Estes anos também marcaram o início dos registros de epidemias em países que até então só notificavam pequenos surtos e casos importados, neste caso destaque para a Argentina e Paraguai (MELLO, 2011).

A análise de incidência da doença no espaço urbano e a recorrência de epidemias de dengue no Brasil, para compreender a dinâmica da doença, tornam-se imperativos ao conhecimento e controle da doença (MERBITZ et al., 2012).

Dentre os inúmeros problemas que os processos de urbanização podem acarretar, evidenciam-se a expansão de arboviroses como a dengue. Ambientes urbanos se encaixam como *habitats* ideais para os vetores desta doença (BARRETO; TEIXEIRA, 2008).

A figura 8 demonstra os riscos de dengue no Estado do Paraná para o ano de 2015:

FIGURA 8 - RISCO DE DENGUE NO ESTADO DO PARANÁ



FONTE: Ministério da Saúde (2015).

Para o Estado do Paraná, no qual se localiza o município de Foz do Iguaçu, a manifestação da dengue se tornou um grande desafio após a década de 1990. A localização geográfica do Estado, que outrora parecia constituir-se num limite para o avanço da doença, foi descaracterizada.

A partir da década de noventa, os surtos e posteriormente as epidemias, se tornaram recorrentes principalmente na região norte e oeste do Paraná. Neste mesmo período, a dengue também avançou para o restante dos Estados da região sul do Brasil (DATASUS, 2012).

2.2.2 A influência da dinâmica do clima urbano na dengue

A saúde humana é fortemente influenciada pelo clima através das condições térmicas, de dispersão (ventos e poluição) e de umidade do ar, que exerce destacada influência sobre a manifestação de muitas doenças, epidemias e endemias, e cria condições favoráveis ao desenvolvimento dos transmissores de doenças contagiosas (MENDONÇA, 2000).

Nas relações entre o clima e a saúde, deve-se considerar a qualidade das condicionantes socioambientais urbanas. O clima de cada cidade apresenta certo grau de variabilidade que pode torná-lo mais ou menos regular, principalmente quando considerados os estágios variados de organização e gestão territorial das áreas urbanas (COSTA; TEIXEIRA, 1999).

Os principais fatores climáticos para dispersão de vetores e doenças são: temperatura, precipitação, umidade e velocidade do vento (CORRÊA, 2006).

O clima é um elemento importante para a saúde e a qualidade de vida. Entretanto, o impacto de determinado clima sobre o organismo não corresponde de maneira exata à percepção do prazer e conforto que ele proporciona; em alguns casos pode existir um grande hiato entre os mesmos (BESANCENOT, 2012).

O clima urbano é um sistema que abrange o comportamento das variáveis climatológicas de um determinado espaço terrestre e sua urbanização, ou seja, é o conjunto formado pelas variáveis climáticas e as características do ambiente construído (AYOADE, 1998).

As características do ambiente construído referem-se à tipologia urbana, à altura, à volumetria, aos afastamentos, à taxa de ocupação e as variáveis climáticas. Estas últimas que mais recebem influência do ambiente construído são: a temperatura, a umidade e os ventos (MENDONÇA, 2000).

As principais enfermidades que se associam diretamente ao clima nas cidades brasileiras são as de veiculação hídrica, como exemplo a dengue, que dependem da quantidade e da forma de disposição da água no sistema urbano.

Ademais, a gestão dos serviços de saúde assume papel fundamental ao permitir o acesso da população afetada com rapidez e qualidade, além de adotar com eficiência sistemas de prevenção (ALEIXO, 2011).

Quanto à incidência de certas doenças que acometem o ser humano, o fator climático é importante, porque além de afetar a resistência do corpo humano a algumas doenças, influencia o crescimento, a propagação e a difusão de alguns organismos patogênicos ou de seus hospedeiros. Algumas doenças tendem a ser predominantes em certas zonas climáticas, enquanto outras, particularmente as contagiosas, seguem um padrão sazonal na sua incidência (AYOADE, 1998).

Foi nesse sentido que Ayoade salientou:

[...] Vários fatores climáticos interferem de modo marcante no aparecimento e na manutenção de determinadas doenças infecciosas e parasitárias. A temperatura, a umidade relativa do ar, o índice pluviométrico, o grau de nebulosidade e os ventos, constituem elementos dos mais importantes no desenvolvimento de certos vetores, bem como no ciclo evolutivo de determinados protozoários em numerosos artrópodes. (AYOADE, 1998, p. 24).

A água que se acumula de forma inadequada e que permanece no sistema urbano torna-se foco de proliferação de insetos e fonte potencial de enfermidades de veiculação hídrica.

A concentração de chuvas aliadas a um ambiente com temperaturas ideais cria uma situação ótima para a reprodução e proliferação do *Aedes aegypti*. Embora possa manter uma população considerável durante as estações menos chuvosas, a custa dos criadouros semipermanentes e independentes das chuvas (caixas de água, cisternas, latões etc.), é durante a estação chuvosa que sua população realmente alcança níveis elevados e de importância para fins de transmissão de patógenos (CONSOLI, 1994).

A temperatura e a precipitação podem influenciar na transmissão da dengue, de modo a impactar a população do vetor. A abundância do vetor predominante (*Aedes aegypti*) é regulada pela precipitação.

Os vetores estimulam o desenvolvimento dos ovos. Já a temperatura influencia na habilidade do mosquito para sobreviver e determina seu desenvolvimento e taxas reprodutivas (BARRETO; TEIXEIRA, 2008).

A reprodução e a evolução do mosquito encontram condições impróprias para a manutenção de sua sobrevivência, nas temperaturas acima dos 40°C e abaixo dos 10° C (BECERRA et al., 2013).

Becerra et al. evidenciou que para algumas regiões do Brasil:

A amplitude de temperatura favorável ao ciclo de vida das populações de *Aedes Aegypti* em condições de laboratório encontra-se entre 22°C e 30°C, e que os extremos de temperatura de 18°C e 34°C apresentaram efeitos negativos sobre o desenvolvimento e a fecundidade do inseto. (BECERRA et al., 2013, p. 8).

As variações climáticas afetam todos os continentes. No Brasil não é diferente. O Estado do Paraná tem apresentado, em diversos aspectos, mudanças ligadas ao ciclo hidrológico e à temperatura. Algumas cidades registraram aumento de suas temperaturas mínimas desde o início da década de 1970 (SILVA; GUETTER, 2003).

O Estado do Paraná passou a ter invernos mais brandos e de menor duração. Já a primavera e o outono, com o aumento das temperaturas, ampliaram o período mais quente do ano (NOGAROLLI, 2005).

A expansão da malha urbana associada à verticalização e adensamento das estruturas citadinas condicionam o clima urbano e, conseqüentemente, impactam nas condições de conforto térmico da população e na qualidade de vida. Os cenários decorrentes dos processos de densificação das cidades e do aumento da impermeabilização do solo contribuem para a ampliação da formação dos microclimas e aumento de temperaturas no nível do transeunte.

Na busca pela construção de um urbanismo sustentável, que preze pela qualidade de vida da população, é fundamental o estudo das questões que impactam na produção dos climas urbanos.

Ressalta-se que a qualidade do espaço urbano está diretamente relacionada à qualidade de vida na cidade, da qual é parte integrante os conceitos de conforto térmico e climas urbanos. Assim, o clima urbano deve ser considerado como componente de qualidade do ambiente e, portanto, de contribuição para a qualidade de vida no meio urbano (ALMEIDA FILHO; ROUQUAUYROL, 2003).

As modificações das configurações urbanas interferem nas variáveis climáticas, altera suas grandezas e forma um mosaico de microclimas, do qual o clima urbano é composto. O clima urbano e seus microclimas só podem ser compreendidos pela intermediação da arquitetura, que forma os recintos urbanos, e cujo sentido está em seu entorno (AYOADE, 1998).

Para a análise do clima urbano é necessário entender o espaço em sua tridimensionalidade, ao considerar seus elementos constituintes na busca do estabelecimento efetivo dos fatores de interferência.

Os estudos climáticos no meio urbano têm sido realizados por pesquisadores em várias localidades do mundo, uma vez que as cidades formam aglomerados urbanos crescentes e esses têm exercido grande influência na mudança de temperatura e na ventilação natural (DIEGUES, 2004).

O fenômeno das ilhas de calor urbano pode ser explicado como o resultado das modificações dos parâmetros da superfície e da atmosfera pela urbanização (LEFF, 2002).

Na segunda metade do século XXI estudos acerca da temática multiplicaram-se em função da expansão das redes de observação meteorológica (BESANCENOT, 2012).

O desdobramento dos estudos da área de saúde entendeu a necessidade do conhecimento das variáveis climáticas e a sua dinâmica de atuação nas configurações urbanas para minorar o desconforto climático (CONFALONIERI, 2003).

A análise do clima das cidades em função de sua urbanização mostra-se de suma relevância, especialmente no século XXI onde a maioria da população mundial se encontra nos centros urbanos.

2.3 A DIMENSÃO DA URBANIZAÇÃO COM O CLIMA

As elevadas temperaturas e os índices de umidade relativamente altos têm sido atribuídos às aceleradas taxas de urbanização e aos altos índices de densidade de construção, de verticalização e industrialização.

Alterações climáticas graves são provocadas ainda pelo desmatamento. A derrubada e a queimada de florestas aumentam a temperatura do ar e deixam a superfície devastada, sem condições de reter a energia do sol nem de gerar fluxos ascendentes de ar.

As aceleradas taxas de urbanização da sociedade atual levaram à concentração da população, à ocupação desordenada das cidades, junto a insetos adaptados ao ambiente modificado (CORRÊA, 2006).

Com o crescimento acelerado e concentrado das cidades essas alterações climáticas ocasionam agravos a doenças tipicamente sazonais, além de favorecerem a ocorrência de doenças endêmicas e epidêmicas (KENNETH, 2008).

Segundo Nedel (2010), a abrangência do climático e do urbano conduz a noção de espaço que inclui o espaço concreto e tridimensional (planos horizontal e vertical) no qual age a atmosfera e os espaços relativos necessários à compreensão do fenômeno urbano.

2.3.1 Os fatores climáticos e ambientais para análise urbana

Nos espaços abertos há uma maior oscilação das variáveis climáticas quando comparado com os ambientes internos, o que torna mais complexa a sua avaliação.

Destaca-se que no ambiente urbano as variáveis climáticas que recebem maior influência das configurações urbanas são a temperatura do ar, a umidade do ar, a velocidade do vento, a precipitação, a radiação solar e a pressão atmosférica.

O conforto térmico urbano depende das interações entre as variáveis climáticas e as características urbanas. É de fundamental interesse o entendimento da interrelação entre as variáveis climáticas em diferentes ambientes com a dengue.

De acordo com a ISO 7726 (1998), a temperatura do ar representa a temperatura da massa de ar existente no entorno do corpo humano e influencia na sensação de conforto térmico, à medida que quanto menor essa for em relação à temperatura da pele, maior será a remoção de calor por convecção e condução (SAFATLE, 2007).

A umidade do ar é descrita pela quantidade de vapor de água contido na atmosfera, uma vez que a umidade relativa do ar é inversamente proporcional à temperatura (MINAYO, 2002).

A radiação é expressa nas medições microclimáticas por meio da temperatura radiante e a ISO 7726 (1998) a define como a temperatura uniforme de um ambiente imaginário, no qual a transferência de calor por radiação do corpo humano é igual à transferência de calor em um ambiente real não uniforme. Isso pressupõe que os efeitos do ambiente real no homem, que geralmente é heterogêneo, e o ambiente virtual que é definido como homogêneo, são idênticos (MELLO, 2011).

O conceito de temperatura radiante média permite o estudo das trocas radiativas entre o homem e o seu meio ambiente. Em cada ambiente ocorre uma troca contínua de energia radiante que é refletida, absorvida e transmitida.

O conforto térmico recebe influência tanto das radiações de ondas curtas, solares, quanto das radiações de ondas longas, terrestres. De acordo com Oke (1987), a radiação solar é percebida na forma de luz visível e de radiações infravermelhas, e a radiação terrestre é emitida em função da temperatura e emissividade dos objetos.

As trocas de energia entre a superfície do solo e a atmosfera são efetuadas tanto por condução quanto por radiação, uma vez que podem ocorrer oscilações de temperatura em ciclos diários ou anuais. Essas oscilações dependem da cobertura vegetal presente na superfície do solo, e contribuem sensivelmente na amplitude térmica diária do solo (COSTA; TEIXEIRA, 1999).

A atmosfera é um sistema dinâmico e por isso a sua pressão varia continuamente no espaço e no tempo, em decorrência do deslocamento de massas de ar tanto na vertical quanto na horizontal.

Na análise do clima é necessário conhecer a pressão atmosférica, ou seja, o peso que o ar exerce por unidade de área sobre a superfície terrestre. As temperaturas são fatores que interferem sobre os níveis de intensidade da pressão atmosférica, pois em regiões mais frias as moléculas de ar unem-se de modo a proporcionar um aumento na pressão e em regiões mais quentes as moléculas de ar afastam-se e deixam o ar menos denso, de maneira a diminuir pressão (NEDEL, 2010).

As condições ambientais artificialmente criadas pelo homem como modificações ou destruição da paisagem natural, emissão de poluentes ambientais, tipos de habitação e a organização do espaço urbano influenciam nas condições do processo de desenvolvimento do vetor da dengue (PIEROTE, 2009).

A destruição da paisagem natural pode modificar o clima através do adensamento urbano, da construção de edifícios, da retirada da cobertura vegetal, que alteram não somente os regimes pluviométricos, como os outros parâmetros climáticos globais e regionais (MENDONÇA; LEITÃO, 2008).

A organização do espaço urbano resultante da segregação da população carente faz com que os pobres urbanos construam habitações em morros ou na margem de rios, ou seja, locais sem infraestrutura (SANTOS, 1994). A emissão de poluentes ambientais altera o clima regional e contribui para a maior incidência das doenças, já que essa modificação dá origem a fenômenos meteorológicos extremos (CONFALONIERI, 2003).

2.4 A MINERAÇÃO DE DADOS NA GESTÃO URBANA

A evolução da tecnologia da informação possibilitou que grandes volumes de dados possam ser armazenados e processados mais rapidamente com uma maior precisão. Em décadas passadas, o problema residia na capacidade física de armazenamento e processamento.

Atualmente a questão está focada na competência de análise desses dados, de forma a extrair informações úteis.

Os dados no formato natural em que estão armazenados, podem não apresentar conhecimento, mas se ajustados e manipulados por um processo de mineração, revelam informações que não seriam possíveis de se obter por meio de técnicas estatísticas normais ou ferramentas analíticas.

2.4.1 *Análise Espacial e o Sistema de Informação Geográfica (SIG)*

O processamento digital pode ser entendido como o conjunto de procedimentos relativos à manipulação e análise de dados através do computador.

Pode-se dizer que o principal objetivo do processamento digital é melhorar o aspecto dos dados para a análise humana e fornecer outros fatores que serão indispensáveis à sua interpretação.

A taxonomia mais utilizada segundo Câmara et al. (2001) para caracterizar os problemas de análise espacial considera três tipos de dados:

Eventos ou padrões pontuais - são fenômenos expressos através de ocorrências identificadas como pontos localizados no espaço, denominados processos pontuais. Exemplos: localização de crimes, ocorrência de doenças e localização de espécies vegetais.

Superfícies contínuas - estimadas a partir de um conjunto de amostras de campo, que podem estar regular ou irregularmente distribuídas. Usualmente, este tipo de dado é resultante do levantamento de recursos naturais.

Áreas com contagens e taxas agregadas - trata-se de dados associados a levantamentos populacionais, como censos e estatísticas de saúde, e que originalmente se referem aos indivíduos localizados em pontos específicos do espaço.

Por razões de confidencialidade estes dados são agregados em unidades de análise, usualmente delimitados por polígonos fechados (setores censitários, zonas de endereçamento postal, municípios), bairros.

O termo *estatística espacial* engloba um conjunto de técnicas de análise geográfica que utiliza técnicas quantitativas e qualitativas para caracterizar o fenômeno em estudo, as quais procuram descrever a variação espacial do fenômeno, a partir de amostras disponíveis (CÂMARA et al., 2001).

Dessa forma, os fenômenos podem ser expressos através de ocorrências pontuais ou em áreas, cujas localizações são expressas em um SIG por meio de pontos, linhas e polígonos.

O termo geoprocessamento refere-se a um conjunto de conhecimentos que utilizam técnicas matemáticas e computacionais para o tratamento da informação geográfica. Esta tecnologia influencia de maneira crescente as áreas da cartografia, recursos naturais, transportes, comunicações e planejamento urbano e regional, dentre outros. Já o termo SIG é aplicado para sistemas que realizam o tratamento computacional de dados geográficos.

Um sistema de informações georreferenciadas armazena os atributos dos dados que estão georreferenciados, isto é, localizados na superfície terrestre segundo uma projeção cartográfica. O processo de armazenar a geometria dos objetos geográficos e todos os seus atributos insere-se como necessidade básica para um SIG. Para cada objeto geográfico, o SIG necessita armazenar seus atributos e as várias representações gráficas associadas. Devido à sua variedade de utilizações que, de acordo com Medeiros (1986), inclui temas como geofísica, agricultura, floresta, cartografia, cadastro urbano e redes de concessionárias (água, energia e telefonia), há pelo menos três grandes maneiras de utilizar um SIG:

- Como ferramenta para produção de mapas;
- Como suporte para análise espacial de fenômenos;
- Como banco de dados geográficos, com funções de armazenamento e recuperação de informação espacial.

Ainda, de acordo com Medeiros (1986), estas três maneiras de utilização do SIG são mais convergentes que conflitantes, e refletem a importância relativa do tratamento da informação geográfica, no desenvolvimento de projetos ou dentro de uma instituição.

O SIG pode ajudar na análise integrada de dados de ambiente e saúde graças à sua capacidade de relacionar dados ambientais e epidemiológicos que estão armazenados em camadas diferentes. Os macrodeterminantes das doenças sejam ambientais, sociais ou econômicos, ocorrem muitas vezes “fora” das pessoas. Portanto, para se relacionar os problemas de saúde com seus determinantes devem-se unir dados de saúde, referidos à população, a dados ambientais, referidos a algo “externo” à população. Cada um desses dados é oriundo de um sistema de informação diferente (CARVALHO et al., 2000).

Uma das principais aplicações dos mapas na epidemiologia é facilitar a identificação de áreas geográficas e grupos da população que apresentam maior risco de adoecer ou morrer prematuramente e que, portanto, precisam de maior atenção, seja preventiva, curativa ou de promoção da saúde.

A epidemiologia espacial também permite reconhecer que a frequência, a distribuição e a importância dos diversos fatores que influenciam no aumento de determinados riscos para a saúde não são, necessariamente, os mesmos em todos os grupos populacionais.

O uso do SIG pode identificar grupos que compartilham determinantes de risco similares. O reconhecimento desses grupos facilita a identificação de intervenções sociais e de saúde para diminuir ou eliminar os determinantes específicos de risco para a saúde. Essa aplicação da epidemiologia implica uma reorganização dos serviços de saúde que responda não apenas às demandas de atenção, mas também, fundamentalmente, às necessidades de saúde não atendidas (OPAS, 2007).

A análise de dados distribuídos pelo espaço geográfico se apresenta cada vez mais valorizada na gestão de saúde, por apontar novos subsídios para o planejamento e a avaliação das ações baseadas na análise da distribuição espacial das doenças, a localização dos serviços de saúde e dos riscos ambientais (BARCELLOS; BASTOS, 1996).

A produção de mapas, que permitam visualizar situações de risco à saúde resultante da interseção e da complementaridade de eventos, é coerente com um conceito de vigilância em saúde de base territorial (SOUZA; CARVALHO, 2005).

As análises realizadas com base em dados espaciais possibilitam ao técnico avaliar não só quantitativamente os dados, como também relacionar as informações de saúde com dados ambientais, socioeconômicos e com a posição que o evento ocupa na superfície terrestre, a fim de acompanhar as permanentes mudanças do espaço geográfico e detectar áreas e população sujeita aos agravos de saúde.

O SIG é utilizado cada vez mais na área da saúde, uma vez que otimiza a análise da situação de saúde e das condições de vida da população e do ambiente, possibilita trabalhar com informações de diferentes origens e formatos. O crescente acesso aos computadores e aumento na disponibilidade de dados são fatores que impulsionam a sua incorporação na Saúde Pública (MINISTÉRIO DA SAÚDE, 2006).

A disponibilização de banco de dados em saúde com a informação de endereço e a estruturação de SIG em diversos municípios brasileiros tem possibilitado o uso crescente da análise de padrões espaciais para diferenciar áreas dentro dos municípios (SANTOS et al., 2001).

O uso do SIG na saúde mostra a possibilidade de associar informações de diversas bases de dados, contribui com o avanço das análises espaciais para os sistemas de vigilância à saúde. A produção de mapas permite visualizar situações de risco à saúde resultante da intersecção e complementaridade dos eventos, coerente com um conceito epidemiológico de vigilância do espaço (CÂMARA et al., 2001).

A análise espacial de padrões epidemiológicos tem valor na análise das relações entre saúde e ambiente e no seu controle, pois o conhecimento da estrutura e dinâmica espacial possibilita a caracterização da situação em que ocorrem eventos de saúde, permite ações de controle, alocação de recursos e preparação de ações de emergência (HYGEIA, 2008).

A recente disponibilidade de bases cartográficas para algumas cidades representa, sem dúvida, um avanço para o planejamento integrado de diferentes setores.

Nesse sentido, o interesse das SMS no uso do SIG para a área urbana tem sido notável, embora muitas vezes, o geoprocessamento em saúde é utilizado incorporado, como ferramenta de visualização de eventos de saúde em mapas, sem definição de uma proposta epidemiológica conceitual de vigilância do espaço (BARCELLOS; BASTOS, 1996).

2.4.2 A descoberta do conhecimento em base de dados

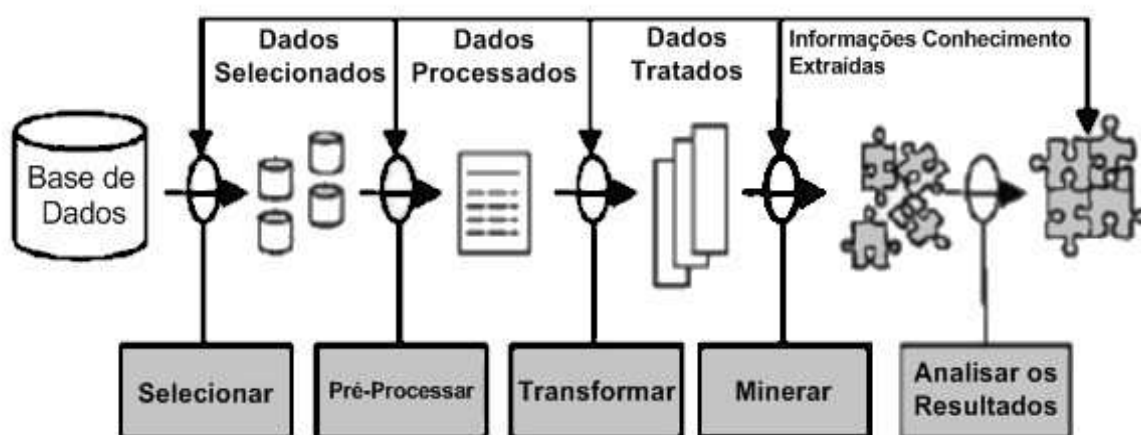
A descoberta de conhecimento é um processo que tem a finalidade de desvendar padrões, a partir de modelos válidos que servirão para a criação de novos arquétipos. Estes terão a utilidade de melhorar a compreensão de um problema para uma tomada de decisão (FAYYAD et al., 1996).

Introduzida em 1982 por Zdzislaw Pawlak, a teoria é conhecida por possuir propriedades que permitem eliminar variáveis ou atributos que não tem relevância para um determinado propósito, onde são divididos os conjuntos de atributos e os mesmos têm a capacidade de manter as propriedades. A eliminação de atributos irrelevantes é o procedimento que caracteriza essa teoria (CARVALHO, 2001).

A mineração de dados faz parte de um processo maior de tratamento de informações chamado *Knowledge Discovery in Database (KDD* - descoberta de conhecimento na base de dados), que consiste primordialmente em estruturar da melhor forma um banco de dados. (REZENDE et al., 2003).

O processo de *KDD* possui cinco etapas distintas, ilustradas na figura 9, descritas por: seleção, preparação ou pré-processamento dos dados, transformação, a mineração dos dados e a interpretação ou análise de dados (CABENA, 1998).

FIGURA 9 - ETAPAS DO PROCESSO DE DESCOBERTA DO CONHECIMENTO



FONTE: Cabena (1998).

As etapas do processo *KDD* são descritas:

- **Seleção / Escolha dos Dados (*Data Selection*):** etapa na qual ocorre a identificação e onde as informações serão escolhidas e depois inseridas no processo de *KDD*. Nesta fase define-se quais dados devem entrar, assim como são realizadas as tarefas de limpeza e filtragem dos dados, e corrige-se possíveis valores nulos, duplicados ou inconsistentes.
- **Preparação dos Dados (*Data Preparation*):** etapa em que os dados passam por uma análise, ou seja, os dados são preparados. Neste momento os dados são selecionados, há correção de dados ruidosos (etapa do processo de limpeza, responsável pela remoção de discrepâncias e impurezas contidas nos dados que serão analisados).

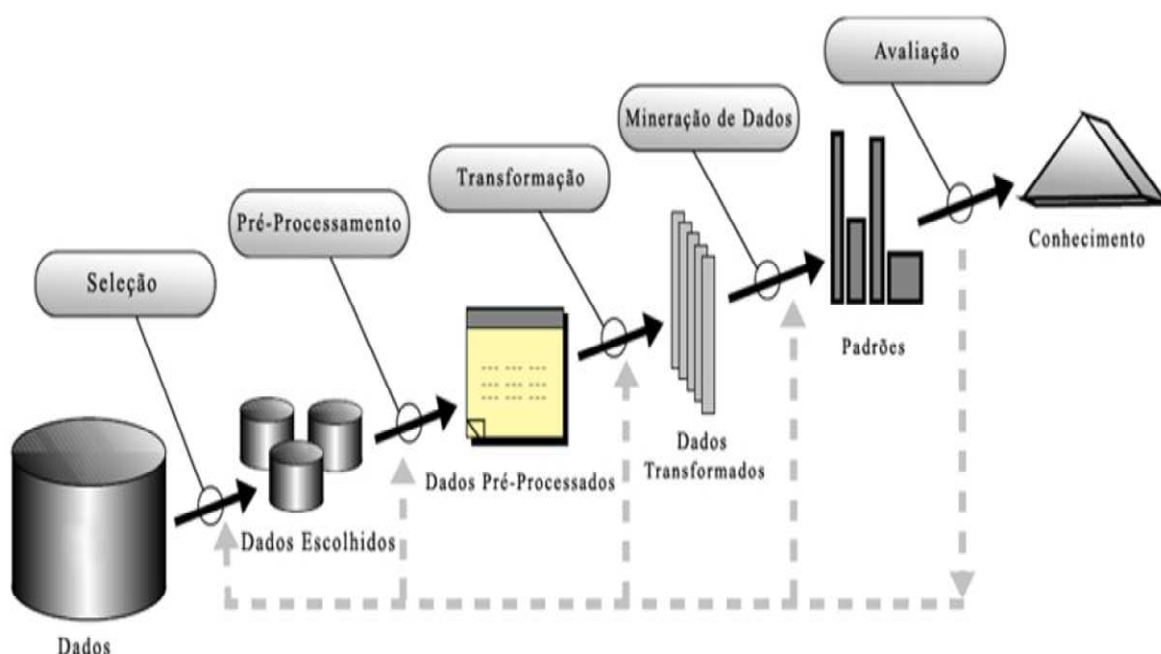
Etapa também conhecida como *Data Cleaning* (Limpeza dos Dados), onde são tratados os valores ausentes e, caso necessário, também a remoção de campos irrelevantes para o processo.

- **Transformação dos Dados (*Data Transformation*):** esta etapa tem como objetivo facilitar a análise através da transformação dos dados. Deve ser feito um modelo de dados analíticos que represente a consolidação, integração e reestruturação dos dados selecionados e processados pelas etapas anteriores. Após a definição do modelo, os dados serão submetidos a um refinamento para que então sejam utilizados pelos algoritmos de mineração de dados.
- **Mineração dos Dados (*Data Mining*):** como salientado anteriormente, esta é considerada a principal etapa do processo *KDD*. É onde os padrões de comportamento são descobertos. Primeiramente deve-se definir qual o objetivo da mineração de dados e quais algoritmos são mais apropriados para a tarefa. Ex: sumarização, classificação, regressão, associação, agrupamento.
- **Interpretação / Análise dos Dados (*Data Interpretation*):** etapa onde os resultados (conhecimentos) são analisados. Engloba a interpretação dos padrões descobertos, de modo que possa ocorrer um possível retorno de alguma etapa anterior do processo.

A mineração de dados tem como objetivo a extração de conhecimento em grandes volumes de dados. Apresenta uma característica multidisciplinar, dada pelas relações existentes com as mais diversas áreas como banco de dados, aprendizado de máquina (*machine learning*), estatística, recuperação de informação, computação paralela e distribuída (CARVALHO, 2001).

A figura 10 mostra a relação entre a mineração de dados e *KDD*

FIGURA 10 - RELAÇÃO ENTRE KDD E MINERAÇÃO DE DADOS



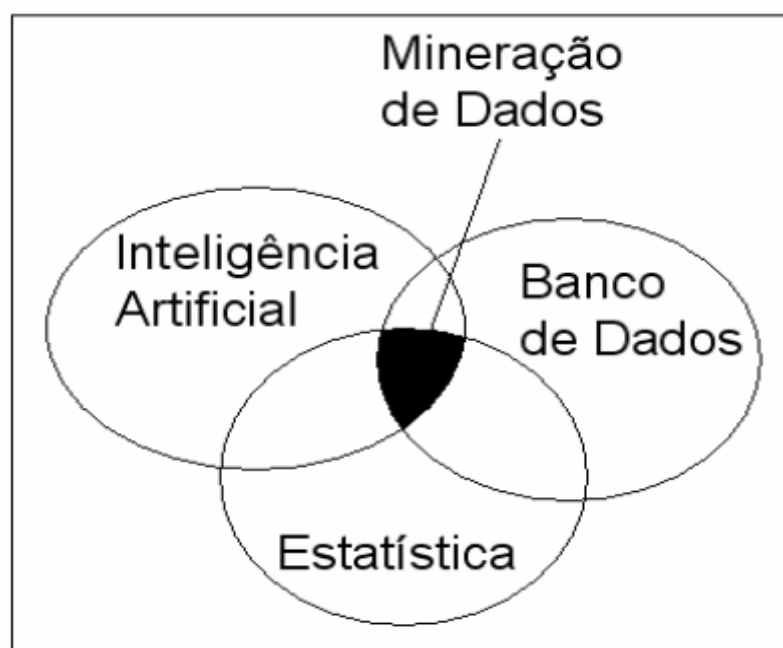
FONTE: Carvalho (2001).

2.4.3 A mineração de dados

A mineração de dados surgiu nos anos de 1980 e era uma ferramenta de análise de padrões nos dados independente das demais etapas do processo de *KDD* (FAYYAD, et al., 1996). É responsável por localizar padrões (sejam eles de comportamento, consumo, faturamento etc.) de um banco de dados. Também é o processo responsável pela seleção de métodos que visam os melhores resultados da tarefa em questão, uma vez que corresponde ao núcleo ou a principal etapa do processo de *KDD* (CARVALHO, 2001).

O processo de mineração de dados consiste na aplicação de várias áreas, técnicas e conceitos, assim como diferentes tecnologias ou ciências, tais como estatística, inteligência artificial, aprendizado de máquina, conforme mostra a figura 11. Tal procedimento sempre focado no objetivo, que é o de encontrar padrões e tendências para apoio à tomada de decisão (CABENA, 1998).

FIGURA 11 - MINERAÇÃO DE DADOS COM UTILIZAÇÃO DE DIFERENTES ÁREAS



FONTE: Cabena (1998).

Segundo Wiley (2006) a mineração de dados corresponderá a uma das dez tecnologias mais emergentes, de maneira a impactar no desenvolvimento de soluções para o mundo.

2.4.3.1 Estatística

A mineração de dados descende de três linhas fundamentais, das quais a mais antiga é a estatística clássica, que foi a primeira responsável pela existência da mineração de dados. A estatística envolve vários conceitos adotados nesse trabalho, como médias, análise de conjuntos, intervalos de dados e suas correlações de relacionamentos (LOPES, 2001).

2.4.3.2 Inteligência artificial

A Inteligência Artificial (IA) é considerada a segunda linha ou linhagem da mineração de dados. A IA veio em oposição à estatística, pois foi construída a partir da heurística, ou seja, aproximação progressiva na resolução de um dado problema.

A inteligência artificial surgiu da idéia de imitar o pensamento humano na resolução de problemas do seu cotidiano, e reproduz a resolução de problemas estatísticos (LUGER, 2004).

Segundo Luger (2004), a inteligência artificial requer máquinas mais robustas para o processamento computacional. Este processamento era inviável até os anos de 1980. Naquela época teve início a fabricação de computadores com melhor processamento e mais baratos, fato que popularizou cada vez nas décadas seguintes. A inteligência artificial faz uso de algoritmos genéricos para predição, classificação, entre outras tarefas.

2.4.3.3 *Aprendizado de máquina*

O aprendizado de máquina (*machine learning*) é a terceira e última linha da mineração de dados. Trata da junção das duas linhas anteriores à estatística e à inteligência artificial (LUGER, 2004). Segundo Carvalho (2001), a mineração de dados adapta o aprendizado de máquina e suas técnicas para aplicações científicas e de negócios. A MD utiliza-se dessas técnicas que são usadas em conjunto para estudar, analisar os dados em um banco e achar padrões e tendências.

A utilização do aprendizado de máquina facilita a localização das tendências que apontam para ocorrência dos padrões a serem encontrados. Leva-se em conta a quantidade considerável de registros armazenados no banco de dados. Na mineração de dados não há uma única técnica que resolva todos os problemas, pois existem diferentes métodos que podem ser utilizados para muitos objetivos. Cada método possui seus algoritmos específicos, com suas peculiaridades, vantagens e desvantagens. De acordo com o objetivo a ser alcançado, a escolha da técnica pode variar entre predição, regressão, classificação e segmentação (*Clustering*), associação e padrões sequenciais (GOEBEL; GRUENWALD, 1999).

A classificação é a tarefa que consiste em prover ordenação de um grande volume de dados em pequenos grupos, subgrupos ou classes mais compreensíveis. Esta técnica tem como objetivo melhorar o entendimento ou o controle de uma situação (KLÖSGEN; ZYTKOW, 2002).

A predição é a tarefa que tem função de fazer a avaliação do valor futuro ou desconhecido de um atributo. Tem como base os dados conhecidos que descrevem o comportamento passado deste atributo (REZENDE et al., 2003).

A tarefa da regressão consiste na análise da interdependência que existe entre os valores dos atributos de uma mesma tabela, e produz automaticamente um modelo ou função que, a partir de valores existentes, pode determinar os valores para novos atributos (GOEBEL; GRUENWALD, 1999).

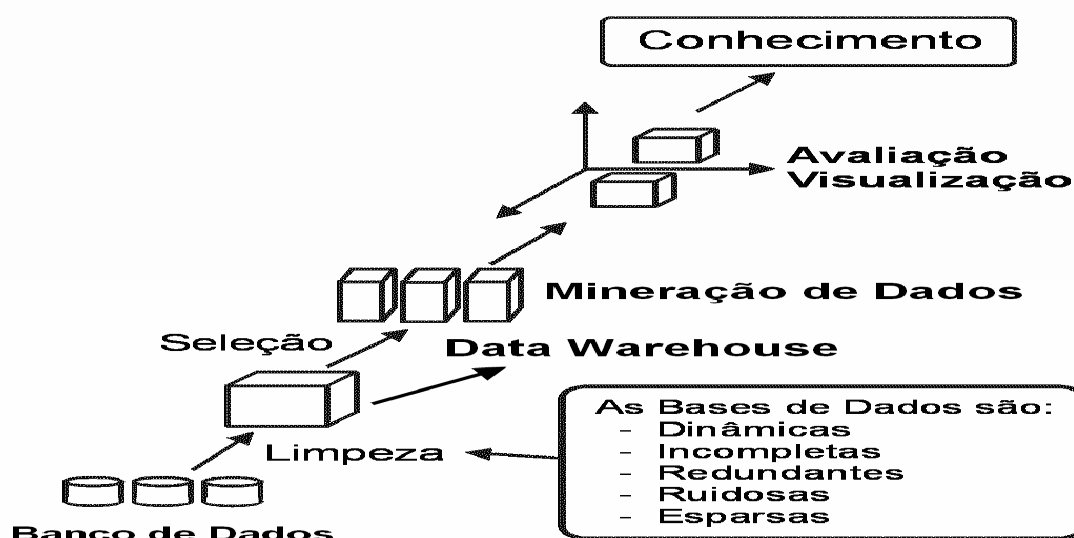
A segmentação ou também chamado de agrupamento de dados é o processo que tem como função dividir uma população ou amostra heterogênea de dados em vários subgrupos ou *clusters* mais homogêneos (CABENA, 1998).

A associação é o método que determina quais ocorrências de dados implicam na existência de outras ou em uma mesma transação (CARVALHO, 2001).

Os padrões sequenciais são responsáveis por detectar padrões de sequência entre transações; como um conjunto de itens encontra-se frequentemente acompanhado durante um período de tempo por outro conjunto de itens (CABENA, 1998). A Mineração de Dados parte geralmente de diversas fontes de dados, passa por uma consistência e logo após se constrói um banco de dados com uma estrutura mais adequada para aplicação das técnicas de MD (CARVALHO, 2001).

A seguir se tem a avaliação, visualização dos resultados e a extração de conhecimento, para que possam ser utilizados nas tomadas de decisões. Vale destacar que o processo de mineração tem início desde a coleta de dados. Na Figura 12 são apresentados os passos fundamentais para a obtenção do conhecimento através da utilização das técnicas de mineração de dados.

FIGURA 12 - INTERAÇÃO DA MINERAÇÃO DE DADOS COM O BANCO DE DADOS



FONTE: Carvalho (2001).

3 METODOLOGIA

A metodologia utilizada nesta pesquisa foi composta por uma análise da bibliografia pertinente ao tema, seguida de coleta de dados junto aos especialistas no domínio, Secretaria Municipal de Saúde, Departamento de Vigilância Epidemiológica, estação meteorológica de altitude e superfície. Na sequência, foram iniciadas as análises dos resultados de maneira cíclica e evolutiva.

Este é um estudo de abordagem descritiva, com natureza quantitativa de análise espacial e temporal da ocorrência dos casos de dengue no espaço urbano de Foz do Iguaçu.

O presente trabalho extraiu conhecimento sobre a gestão urbana no contexto de dengue associada às variáveis climatológicas e urbanas de diferentes fontes de informação, sejam de material bibliográfico, dados reais e históricos sobre a temática e também de especialistas das áreas envolvidas (gestão urbana, saúde pública e gestão ambiental).

Para o desenvolvimento da metodologia quantitativa foram necessárias quatro principais etapas: revisão bibliográfica, coleta dos dados, preparação dos dados e modelagem.

3.1 REVISÃO BIBLIOGRÁFICA

Primeiramente foram buscados na literatura artigos científicos relevantes sobre a temática de gestão urbana e dengue associada à saúde pública. Por exemplo, tais buscas foram realizadas através de palavras-chave, *como urbano, urbanismo, cidades, dengue, climatologia*, entre outras. Analogamente foram procurados artigos científicos publicados em inglês em periódicos internacionais.

Foram priorizados na busca periódicos nacionais e internacionais relacionados à temática pesquisada.

Além dos artigos científicos, também foram procurados documentos que relatassem internações hospitalares induzidas pela incidência de dengue ou relatórios técnicos.

Em toda a documentação foram procuradas relações qualitativas e quantitativas das variáveis associadas ao presente estudo.

3.2 COLETA DOS DADOS

A coleta de uma base de dados espaciais e temporais implica em uma diversidade de parâmetros obtidos de escalas numéricas não lineares. Para definir em que escalas de tempo e espaço seria viável trabalhar foi necessário primeiramente minerar esta quantidade de informações. A definição das escalas foi realizada através do confronto de diversos parâmetros medidos em relação ao espaço e ao tempo, mas de forma simultânea.

3.2.1 *Dados de variáveis climatológicas*

Os dados climatológicos de pluviometria, temperatura, umidade relativa do ar, pressão atmosférica e velocidade dos ventos foram obtidos das estações meteorológicas de superfície do INMET e ICEA localizadas na cidade de Foz do Iguaçu. Dessa forma, os dados foram obtidos e selecionados para o tratamento de preparação para compor a matriz de dados temporais.

Os dados pluviométricos foram trabalhados com a precipitação mensal atual, precipitação mensal anterior, precipitação acumulada de três meses, precipitação máxima em uma hora, precipitação acumulada do mês anterior e mês atual.

Os dados de temperatura foram trabalhados com as temperaturas máxima e mínima mensal, do ponto de orvalho e média mensal.

Os dados de umidade relativa do ar, pressão atmosférica e velocidade dos ventos foram considerados os valores médios mensais das variáveis.

3.1.2 *Dados de variáveis urbanas*

Os dados de parâmetros urbanos compõem a matriz de dados espaciais, uma vez que o município foi dividido em 13 zonas regionais, cada uma com sua peculiaridade. Com isso foram obtidas as características e as variáveis urbanas de saneamento em relação a cada zona em estudo.

O município apresenta quatro estações de tratamento de esgoto, duas estações de tratamento de água, um aterro sanitário e um antigo lixão, que era utilizado como aterro sanitário até o ano de 2011.

3.3 PREPARAÇÃO DOS DADOS

Nos problemas de mineração de dados, se faz necessária uma etapa de preparação de dados, de forma adequá-los para análise que se pretende efetuar. Com tantas fontes de dados é importante a padronização dos dados na maneira como os dados são coletados e armazenados.

A preparação dos dados para a modelagem é fundamental e consome a maior parte do tempo em projetos de Mineração de Dados. O desafio é preparar os dados de forma que a informação contida neles seja exposta da melhor maneira para as ferramentas de mineração (PYLE, 1999).

Um aspecto importante da preparação é procurar incorporar conhecimento prévio do domínio da área de aplicação nos dados preparados. Outro aspecto importante, para o tipo de problema apresentado, são as séries temporais, isto é, atributos que são medidos ao longo do tempo em intervalos fixos, como a temperatura do ar, a umidade relativa e a precipitação pluvial.

No processo de preparação dos dados são necessárias transformações nos dados e derivação de novos atributos para que a dimensão temporal seja incorporada no forma de dados usuais reconhecidos pelos algoritmos tradicionais de mineração (WEISS; INDURKHYA, 1998).

3.3.1 *Regionalização dos dados de incidência de dengue*

O município de Foz do Iguaçu é dividido em 13 zonas com dados da incidência de dengue catalogadas pela Secretaria Municipal de Saúde.

Portanto, para cada bairro é quantificado o número de casos de incidência de dengue.

3.3.2 *Discretização das variáveis para extração das regras*

A tarefa de discretização é aplicada quando se pretende realizar classificação ou associação, constitui em transformar valores numéricos contínuos em n intervalos que serão as n categorias que se deseja obter. Basicamente, todo o esforço está em decidir quantas categorias serão necessárias e onde ficarão os limites dos intervalos.

A discretização permite que novos atributos sejam derivados a partir dos atributos originais com o intuito de facilitar a extração de determinada informação de forma eficaz. O novo conjunto de atributos pode substituir ou ser agregado aos atributos originais.

As variáveis dos dados espaciais e temporais foram discretizadas em três intervalos de classe, de acordo com os valores dos tercís de modo a compor as matrizes para a realização do processo da modelagem.

3.4 MODELAGEM

A etapa de modelagem pode ser subdividida em duas fases:

- A) Análise multivariada;
- B) Construção de modelos quantitativos.

A) *Análise multivariada*

A análise multivariada corresponde a um grande número de métodos e técnicas que utilizam, simultaneamente, todas as variáveis na interpretação teórica do conjunto de dados obtidos (MANLY, 1986).

A aplicação da análise multivariada consiste em identificar padrões qualitativos entre as variáveis envolvidas no estudo. Os resultados destas análises permitem uma identificação clara das relações entre as variáveis envolvidas na construção dos modelos quantitativos (SOUZA; EBECKEN, 2012).

B) *Construção dos modelos quantitativos*

Nesta etapa são geradas regras de associação entre as variáveis envolvidas. Uma regra de associação é um modelo bastante simples e de fácil interpretação que relaciona causa e efeito pela designação Se/Então (AGRAWAL et al., 1993).

As regras de associação também podem ter caráter preditivo e então são chamadas de regras de classificação (SOUZA, 2014).

3.4.1 Montagem da matriz dos dados espaciais da incidência de dengue

As ferramentas baseadas em mineração de dados requerem para o estudo de um fenômeno uma estrutura composta por entrada (variáveis envolvidas no fenômeno) e saída (variáveis do próprio fenômeno). Neste estudo a estrutura montada foi uma matriz, composta por linhas e colunas.

As linhas são formadas pelas 13 zonas do município. As colunas são as variáveis que quantificam o total anual de pacientes por dengue em cada bairro além das características urbanas de cada zona, tais como topografia, saneamento, socioeconomia, hidrografia, entre outras.

Contudo, novas variáveis foram criadas por meio de cálculo da distância do centróide de cada zona até as estruturas físicas da cidade tais como da distância das estações de tratamento de água, estações de tratamento de esgoto, aterros sanitários e corpos hídricos.

3.4.2 Montagem da matriz dos dados temporais da incidência de dengue

Na montagem da matriz de dados temporais as linhas são formadas pelos 52 registros correspondentes aos meses do período de 2010 a 2015. As colunas são as variáveis que quantificam a incidência de dengue em cada mês.

Para cada mês também foram inseridas as variáveis climatológicas tais como umidade relativa, pressão atmosférica, velocidade do vento, temperaturas, níveis de precipitação e calculados os índices acumulados de chuva.

3.4.3 Geração das regras

Os modelos com regras foram construídos com as técnicas de mineração de dados. As regras de associação, como o próprio nome sugere, são associações de causa e efeito entre os parâmetros de uma base de dados indicados principalmente por dois valores, suporte e confiança.

Alguns algoritmos de mineração de dados, especialmente os de classificação requerem que os dados estejam de forma de atributos categorizados. Assim é necessário transformar um atributo contínuo em categórico.

Foram extraídas regras de associação entre as variáveis totais anuais de incidência de dengue e as variáveis urbanas.

Foram construídas também modelos para previsão de dengue nos próximos meses com as regras de classificação.

4 RESULTADOS

O presente capítulo contém a aplicação das técnicas de mineração de dados e sistemas de informações geográficas para a elaboração do modelo de predição de incidência de dengue.

Primeiramente foi feita a aplicação da ferramenta de georreferenciamento para fazer a distribuição das variáveis urbanas e espacialização dos casos de incidência de dengue na cidade através dos bancos de dados coletados; e em seguida, foi realizada a aplicação das técnicas de mineração de dados.

A seguir são apresentados os resultados obtidos durante o desenvolvimento desta dissertação.

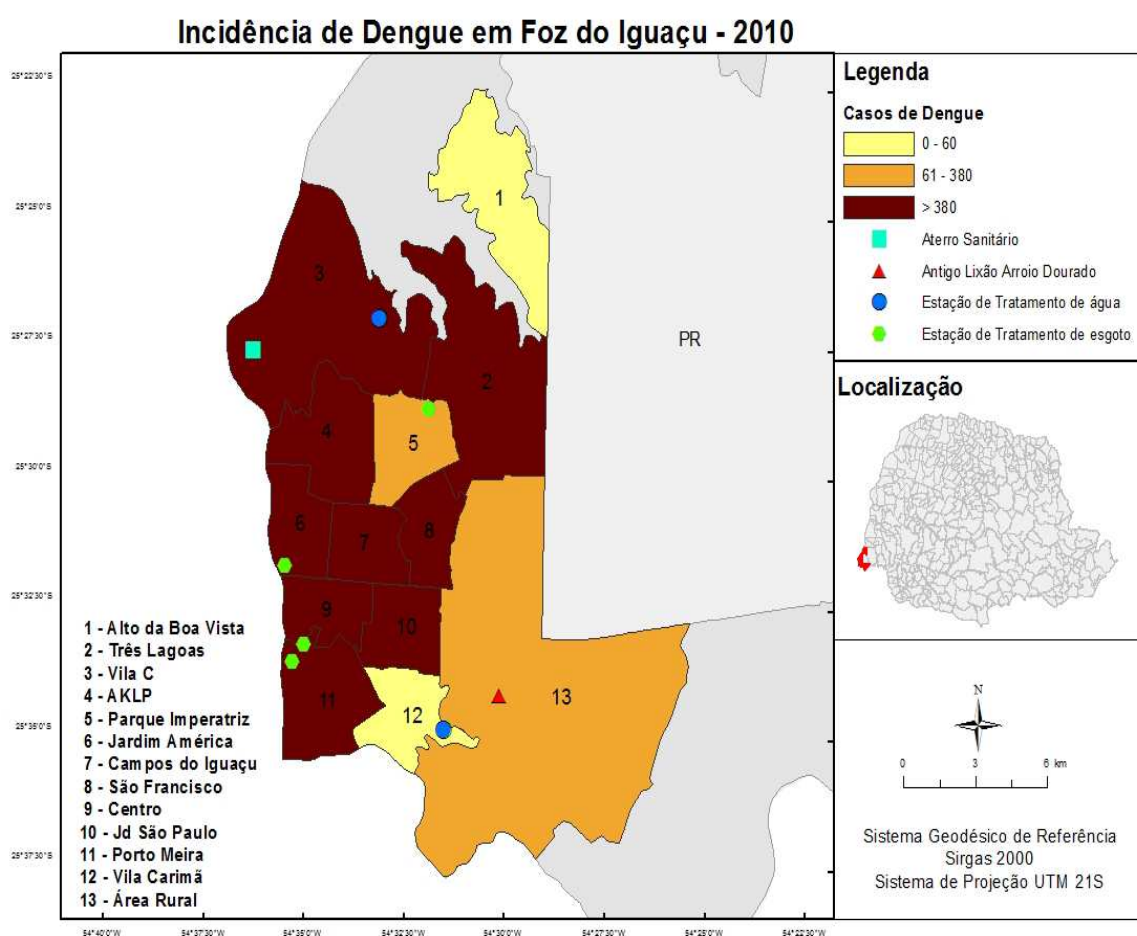
4.1 ESPACIALIZAÇÃO DA INCIDÊNCIA DE DENGUE EM FOZ DO IGUAÇU

O município de Foz do Iguaçu está dividido em treze zonas. As figuras de 13 até 18 mostram o total anual de incidência de dengue para cada zona no período de 2010 a 2015.

Na análise da distribuição de ocorrência de dengue no ano de 2010, observa-se na figura 13 que onze entre as treze zonas apresentam alta incidência de dengue (terceiro tercil), ou seja, 84,6% das zonas.

As zonas que apresentaram maiores números de casos da doença são: Vila C, Três Lagoas, AKLP, Parque Imperatriz, Jardim América, Campos do Iguaçu, São Francisco, Centro, Jardim São Paulo e Porto Meira.

FIGURA 13 - ESPACIALIZAÇÃO DA INCIDÊNCIA DE DENGUE EM FOZ DO IGUAÇU EM 2010

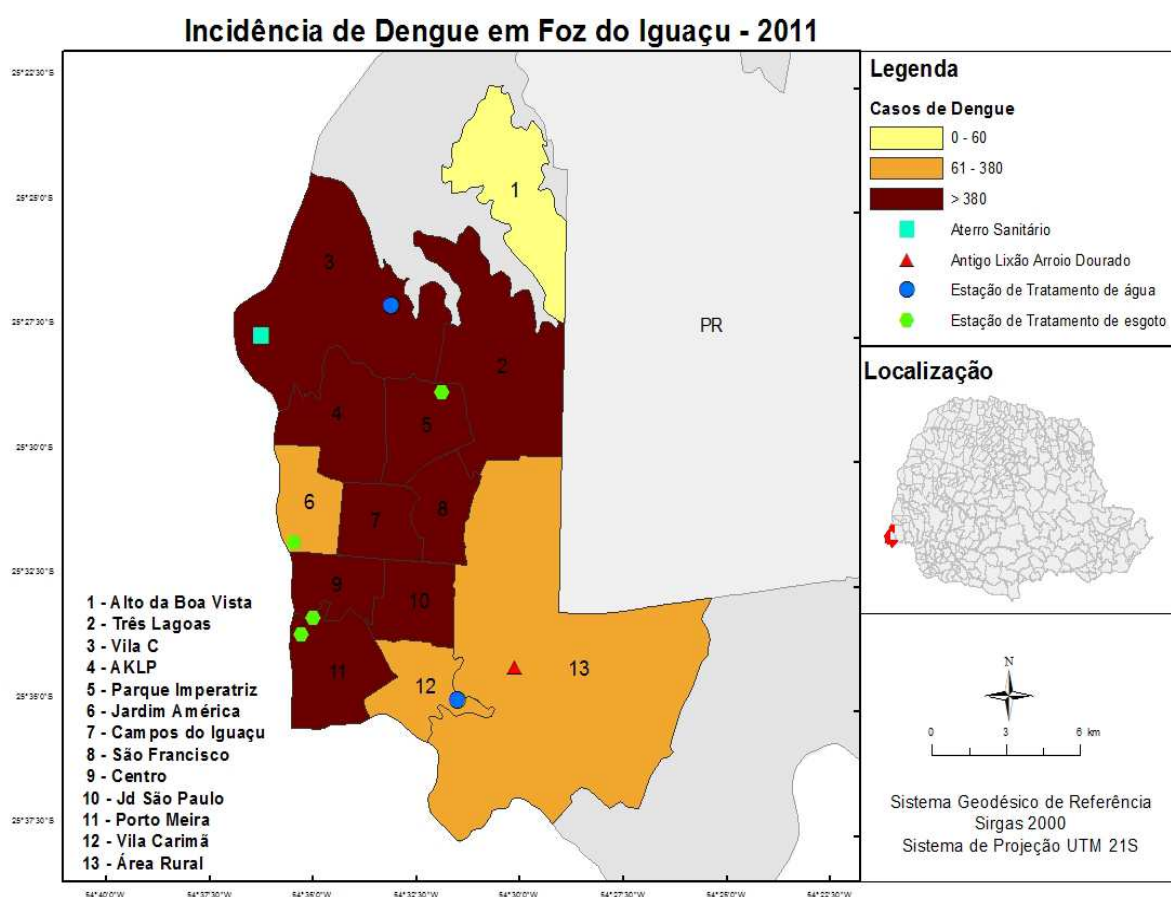


FONTE: O Autor (2017).

Na análise da distribuição de ocorrência de dengue no ano de 2011, observa-se na figura 14 que doze entre as treze zonas apresentam alta incidência de dengue, ou seja, 92,3% das zonas.

As zonas que apresentaram maiores números de casos da doença são: Vila C, Três Lagoas, AKLP, Parque Imperatriz, Jardim América, Campos do Iguaçu, São Francisco, Centro, Jardim São Paulo, Porto Meira, Vila Carimã e Área Rural.

FIGURA 14 - ESPACIALIZAÇÃO DA INCIDÊNCIA DE DENGUE EM FOZ DO IGUAÇU EM 2011

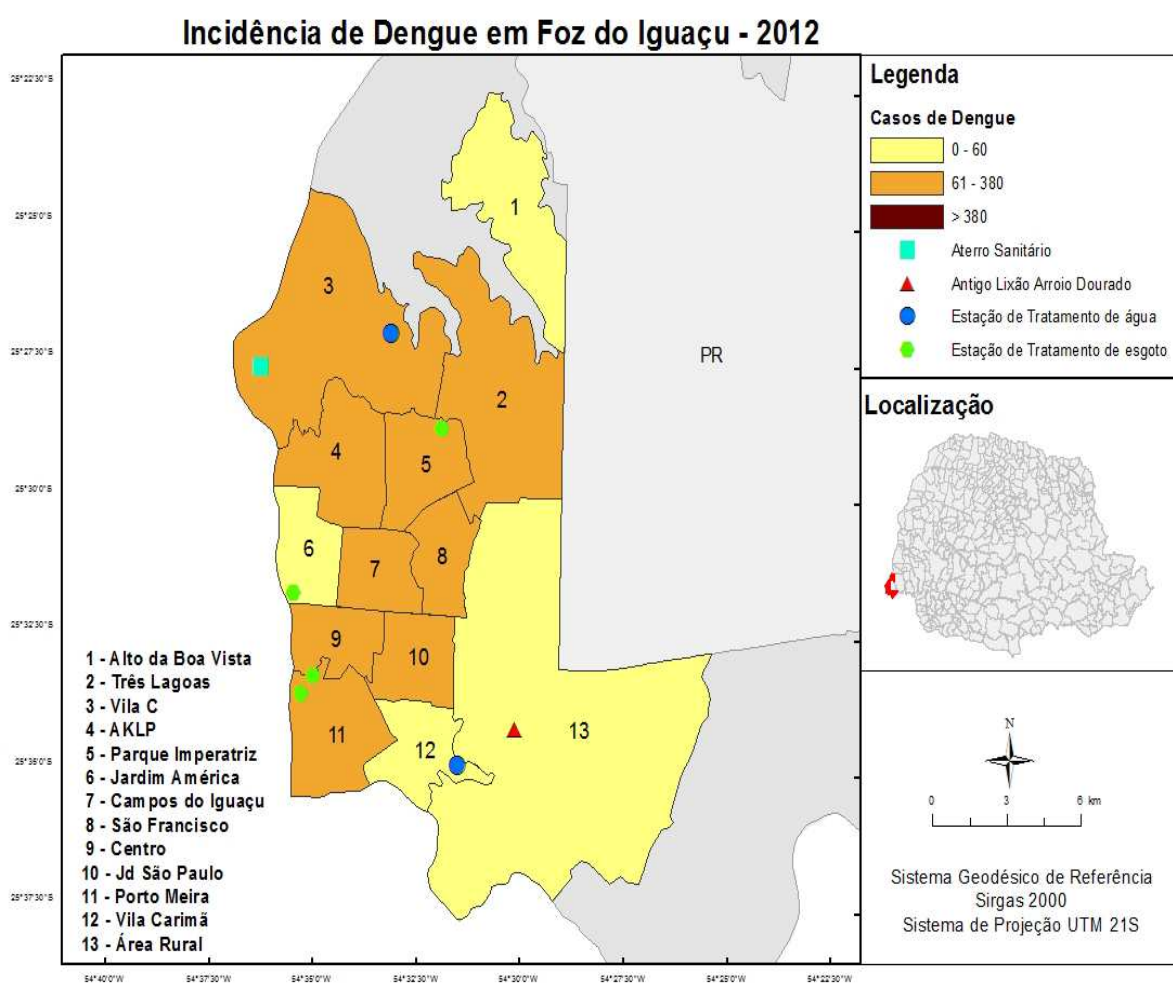


FONTE: O autor (2017).

Na análise da distribuição de ocorrência de dengue no ano de 2012, observa-se na figura 15 que nove entre as treze zonas apresentam alta incidência de dengue, ou seja, 69,2% das zonas.

As zonas que apresentaram maiores números de casos da doença são: : Vila C, Três Lagoas, AKLP, Parque Imperatriz, Campos do Iguaçu, São Francisco, Centro, Jardim São Paulo e Porto Meira.

FIGURA 15 - ESPACIALIZAÇÃO DA INCIDÊNCIA DE DENGUE EM FOZ DO IGUAÇU EM 2012

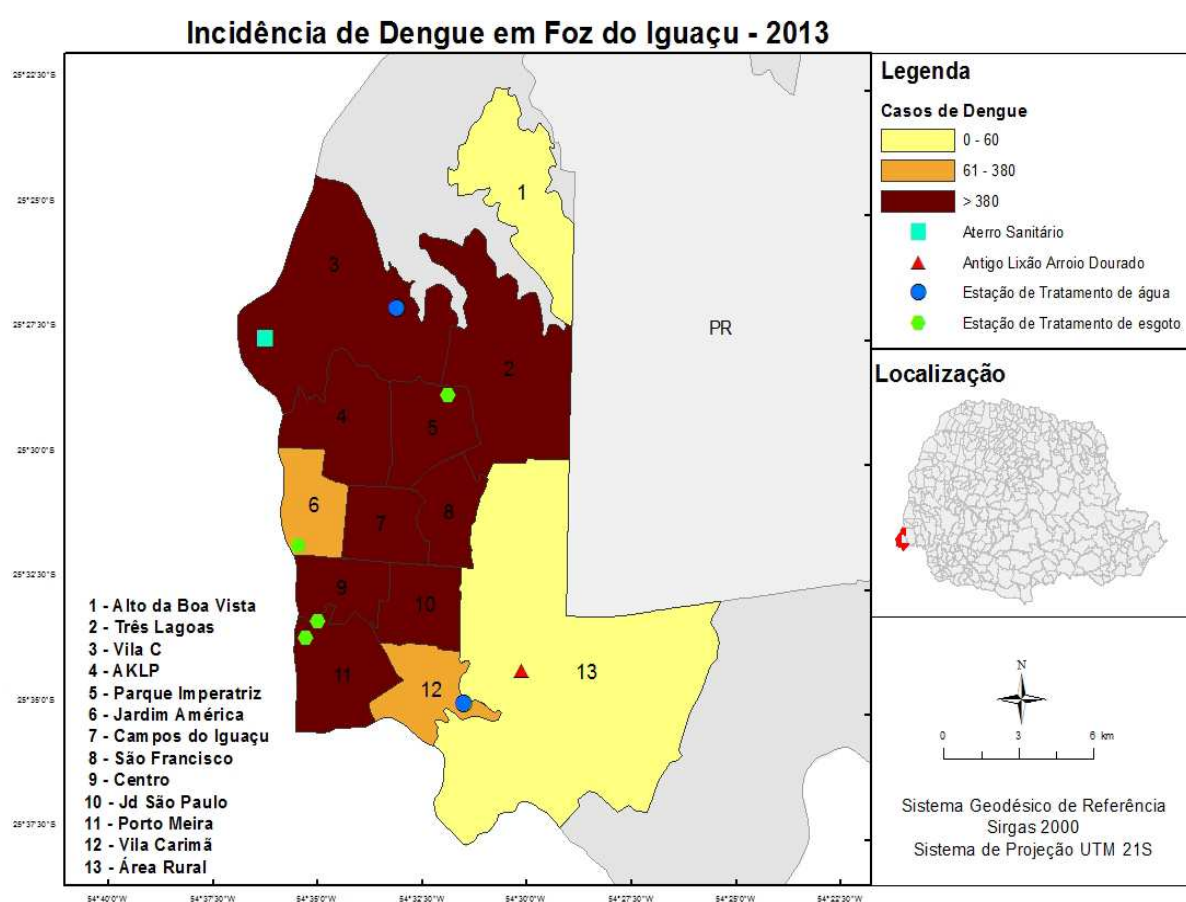


FONTE: O autor (2017).

Na análise da distribuição de ocorrência de dengue no ano de 2013, observa-se na figura 16 que onze entre as treze zonas apresentam alta incidência de dengue, ou seja, 84,6% das zonas.

As zonas que apresentaram maiores números de casos da doença são: : Vila C, Três Lagoas, AKLP, Parque Imperatriz, Jardim América, Campos do Iguaçu, São Francisco, Centro, Jardim São Paulo, Porto Meira e Vila Carimã.

FIGURA 16 - ESPACIALIZAÇÃO DA INCIDÊNCIA DE DENGUE EM FOZ DO IGUAÇU EM 2013

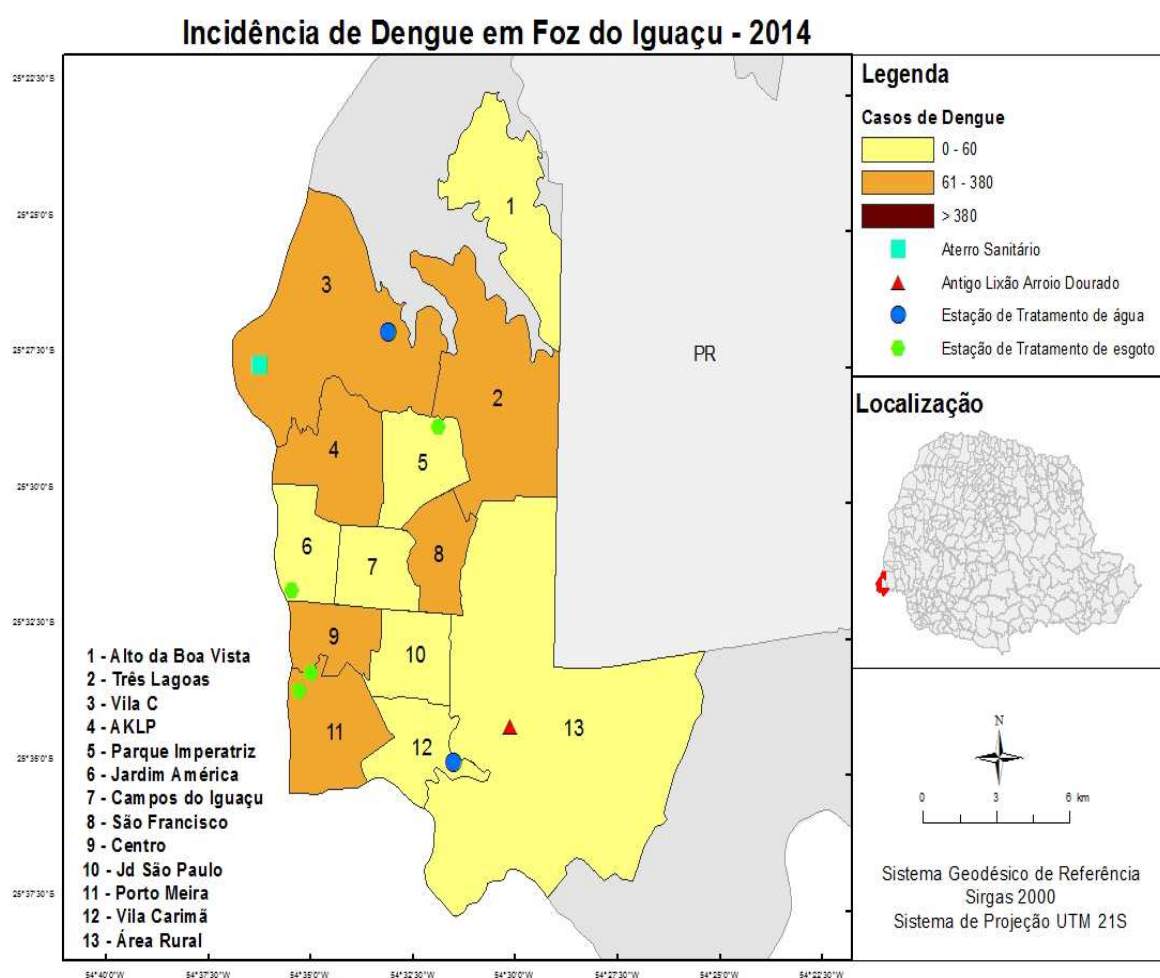


FONTE: O autor (2017).

Na análise da distribuição de ocorrência de dengue em 2014, observa-se na figura 17 que seis entre as treze zonas apresentam alta incidência de dengue, ou seja, 46,1% das zonas.

As zonas que apresentaram maiores números de casos da doença são: Vila C, Três Lagoas, AKLP, São Francisco, Centro e Porto Meira.

FIGURA 17 - ESPACIALIZAÇÃO DA INCIDÊNCIA DE DENGUE EM FOZ DO IGUAÇU EM 2014

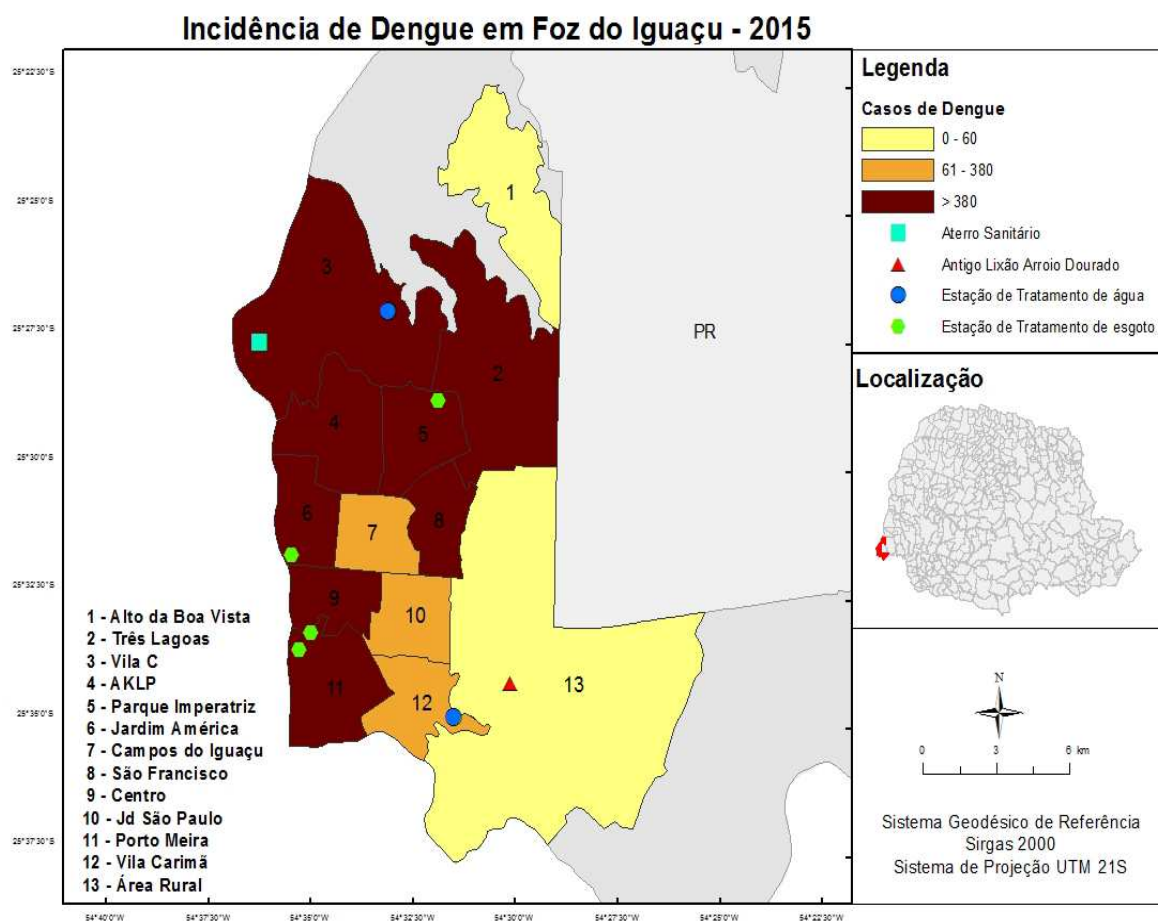


FONTE: O autor (2017).

Na análise da distribuição de ocorrência de dengue em 2015, observa-se na figura 18 que onze entre as treze zonas apresentam alta incidência de dengue, ou seja, 84,6% das zonas.

As zonas que apresentaram maiores números de casos da doença são: Vila C, Três Lagoas, AKLP, Parque Imperatriz, Jardim América, São Francisco, Centro e Porto Meira.

FIGURA 18 - ESPACIALIZAÇÃO DA INCIDÊNCIA DE DENGUE EM FOZ DO IGUAÇU EM 2015



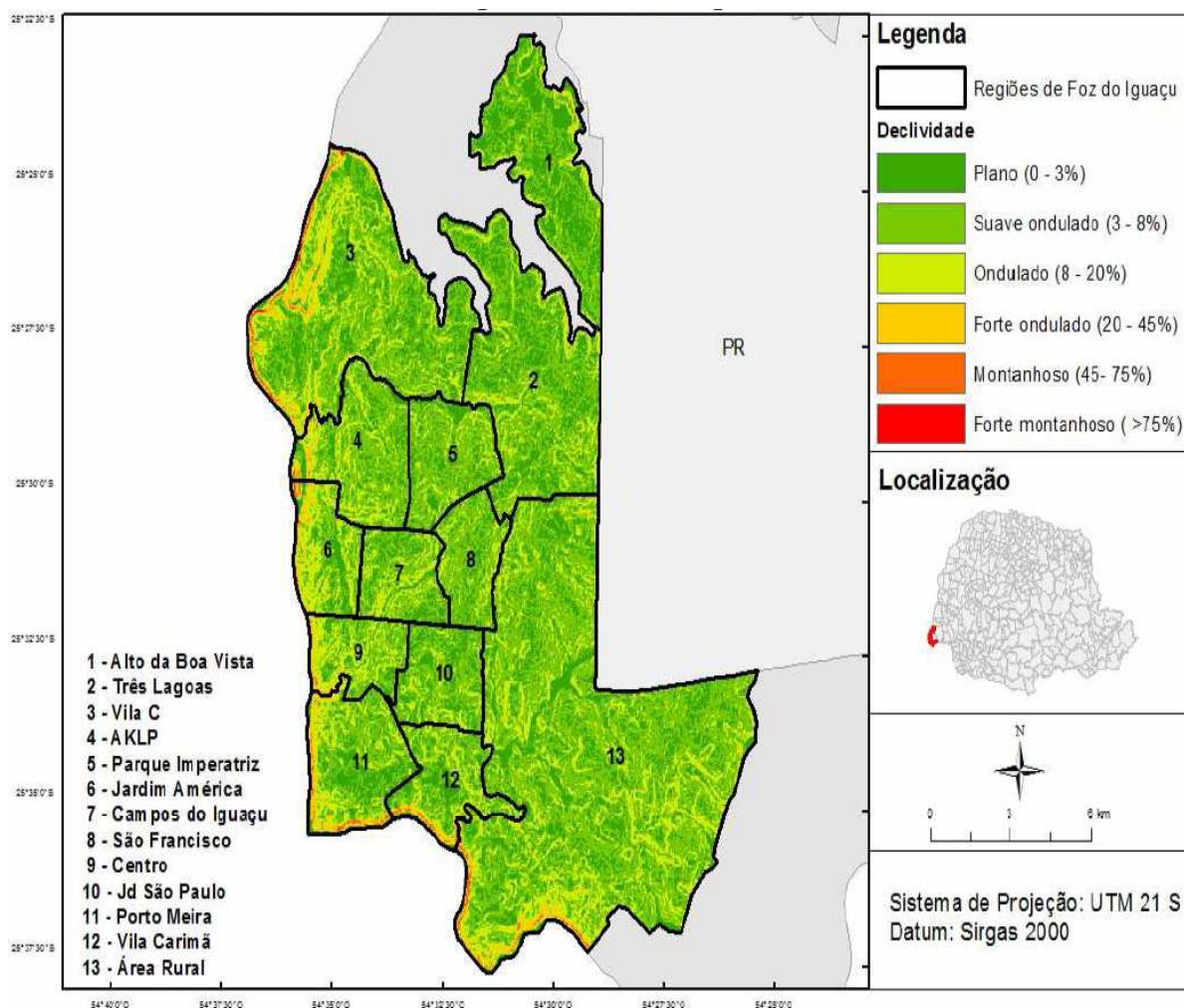
FONTE: O autor (2017).

Na análise da distribuição de incidência de dengue no período de 2010 a 2015, observa-se que nos anos 2010, 2011, 2103 e 2015 ocorreram os maiores números de casos, quando mais de 80% dos bairros tiveram registros no terceiro tercil.

4.2 GERAÇÃO DOS MAPAS DE PARÂMETROS ESPACIAIS

Na construção da matriz de dados espaciais, foram levantados os dados das declividades dos terrenos das 13 zonas em estudo, percebe-se que as regiões 6 (Jardim América) e 12 (Vila Carimã) apresentam declividades acentuadas e as regiões 5 (Parque Imperatriz) e 8 (São Francisco) são as que apresentam menores declividades. A figura 19 mostra as declividades das regiões de Foz do Iguaçu.

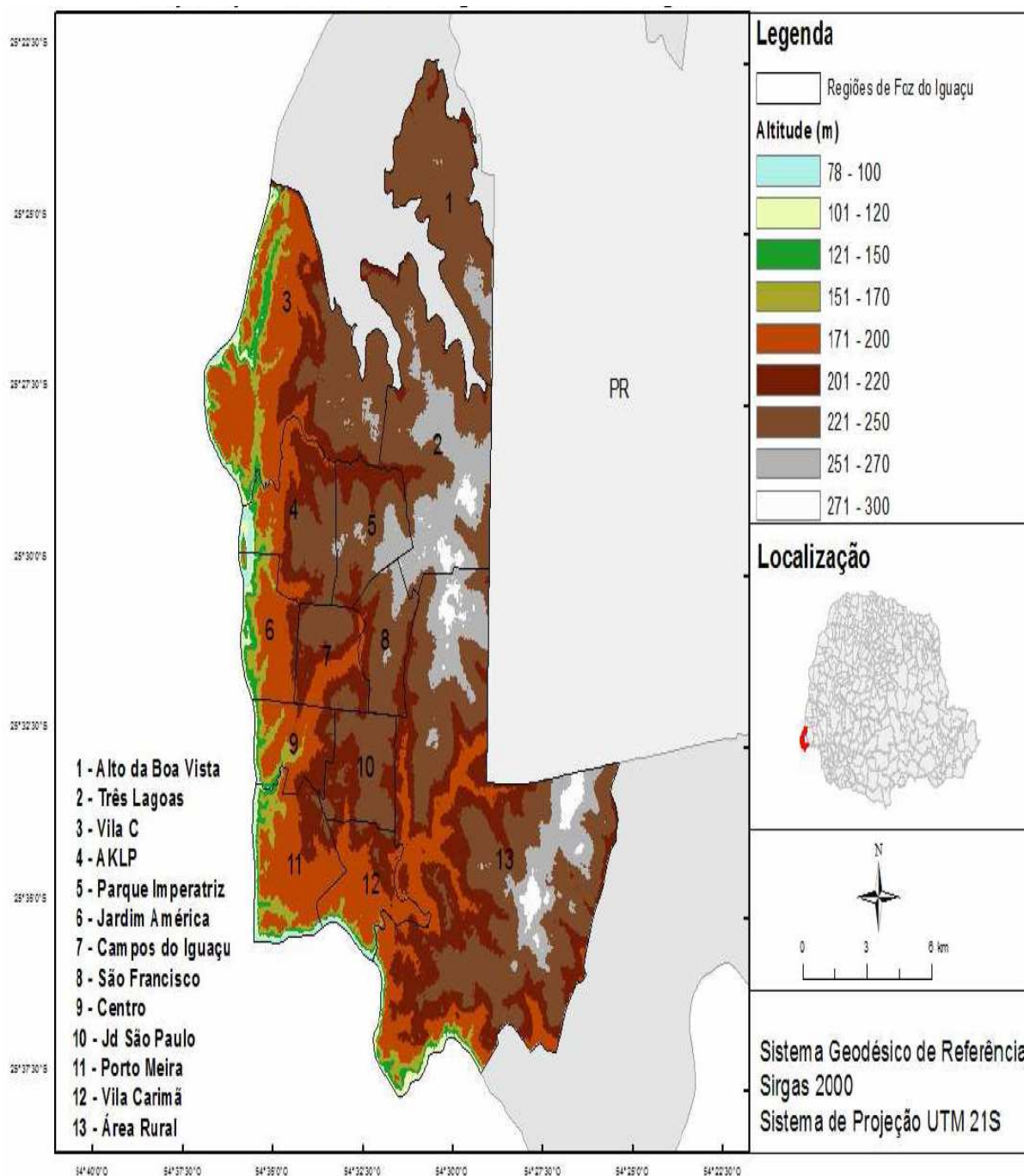
FIGURA 19 - DECLIVIDADE DAS REGIÕES DE FOZ DO IGUAÇU



FONTE: O autor (2017).

No processo em estudo foi elaborado o mapa hipsométrico das regiões em análise, conforme demonstrada na figura 20.

FIGURA 20 - MAPA HIPSOMÉTRICO DAS REGIÕES DE FOZ DO IGUAÇU

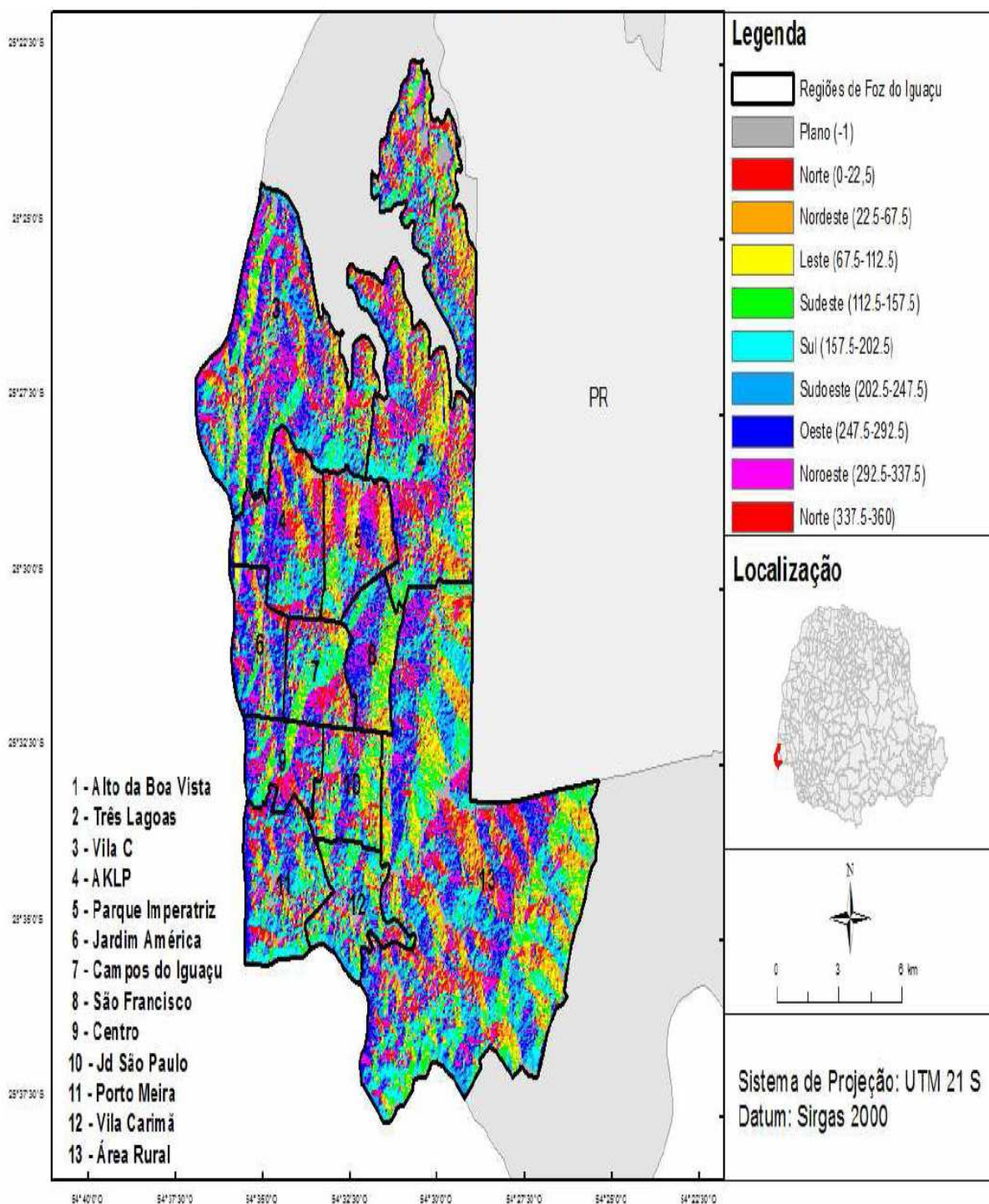


FONTE: O autor (2017).

Percebe-se na figura 20 que a região 13 (Área Rural) apresenta maior altitude e que a região 6 (Jardim América) possui menor altitude no município.

A figura 21 mostra a orientação geográfica das regiões do município de Foz do Iguaçu que foram gerados pela ferramenta *Aspect*.

FIGURA 21 - ORIENTAÇÃO DAS VERTENTES DAS REGIÕES DE FOZ DO IGUAÇU



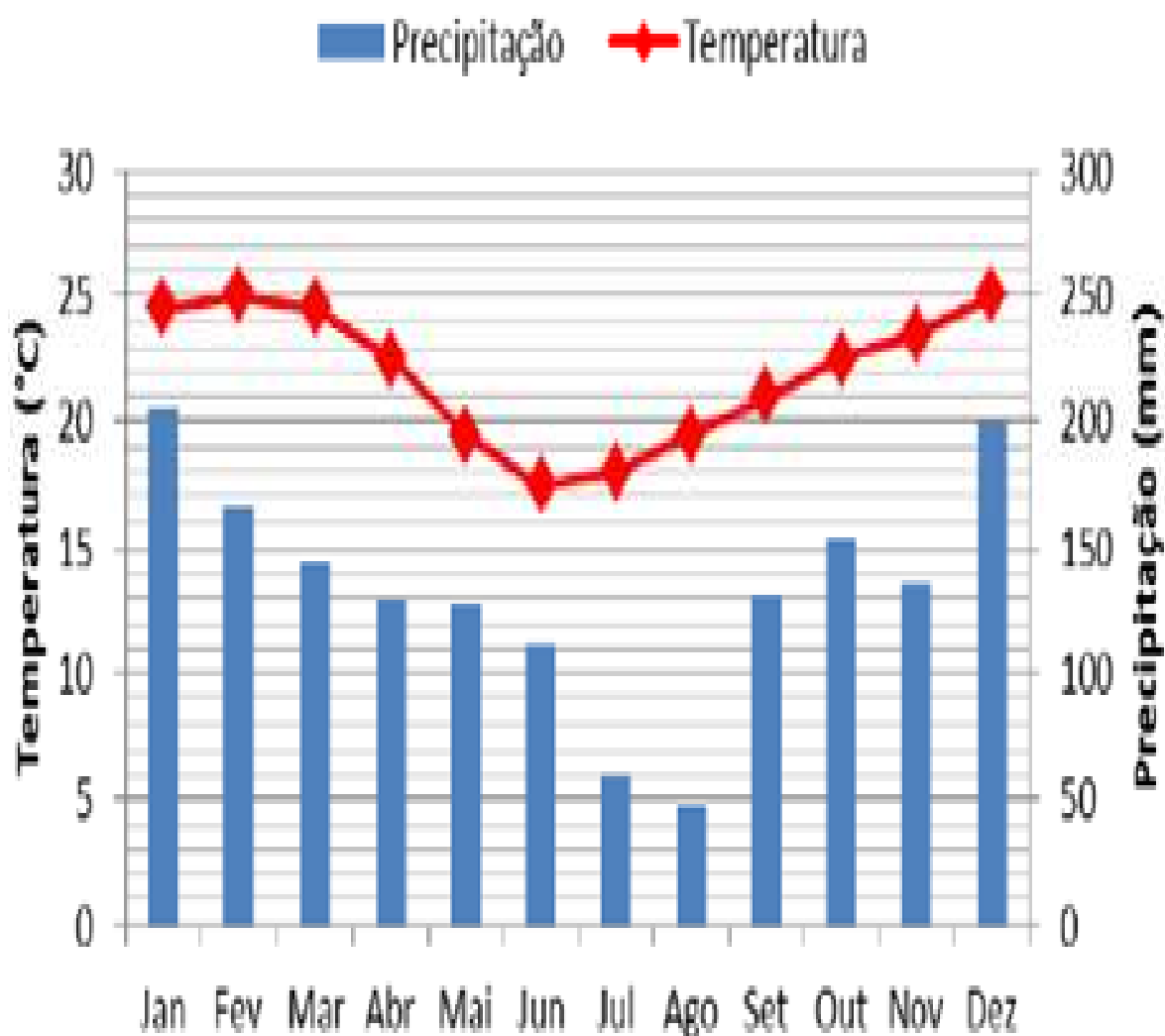
FONTE: O autor (2017).

Percebe-se na figura 21 o mapeamento da orientação das vertentes dos terrenos, ou seja, a direção da variação de declividade das zonas do município.

4.3 CORRELAÇÕES CLIMATOLÓGICAS E URBANAS COM A DENGUE

Conforme analisado neste estudo, a influência da temperatura no início da epidemia, com temperaturas médias de janeiro e fevereiro de 25°C, foi essencial na proliferação da doença. No término da epidemia, no final do mês de junho e nos meses de julho a agosto, as temperaturas mínimas chegaram a abaixo de 10°C, o que foi fundamental na limitação do desenvolvimento do mosquito. O gráfico 1 mostra o climatograma do município de Foz do Iguaçu.

GRÁFICO 1 - CLIMATOGRAMA DE FOZ DO IGUAÇU



FONTE: INMET (2015).

Em relação à influência dos eventos extremos como as enchentes, a partir das análises realizadas, foi possível evidenciar que uma enchente pode criar condições favoráveis para o ciclo viral de dengue, principalmente se os resíduos deixados pela inundação passam a servir como criadouros.

Com base nos resultados e discussões obtidos com o estudo de caso sobre a diferenciação das temperaturas em diferentes espaços da cidade, foi notável a concentração das maiores temperaturas nas áreas mais urbanizadas da cidade.

Os dados de temperatura captados durante o estudo na cidade ofereceram condições plenamente favoráveis à reprodução do *Aedes aegypti*; todavia, os riscos de proliferação da dengue foram reduzidos em alguns momentos do decorrer da pesquisa devido à carência de chuvas intermitentes, o qual demonstra a potencialidade dos fatores ambientais climáticos.

Sobre a distribuição da doença, foi possível identificar a repetição dos padrões de concentração dos casos observados, uma vez que a falta de planejamento principalmente no saneamento e a alta densidade demográfica aliada ao modo de vida da população (caracterizada pelo descarte de resíduos considerados ótimos para o desenvolvimento de criadouros), foram importantes condicionantes de risco e vulnerabilidade da dengue.

Outro condicionante que contribui para a situação de risco e vulnerabilidade à dengue é a condição de localização da cidade de Foz do Iguaçu na tríplice fronteira, pois esta resulta na intensa circulação de pessoas e mercadorias o que, consequentemente, facilita a circulação de sorotipos.

Neste estudo foi aplicada a análise de agrupamentos, em sua aplicação, constitui uma metodologia numérica multivariada e engloba uma variedade de técnicas e algoritmos, onde o objetivo é encontrar e separar objetos em grupos similares.

O agrupamento de dados ou *Clustering* consiste em um processo de agrupamento de objetos dentro de classes ou *clusters*. Cada classe é composta por objetos ou dados que apresentam alta similaridade quando comparados entre eles e baixa similaridade quando comparados com objetos de outras classes (CARVALHO, 2001).

Entre os métodos de agrupamento encontram-se os métodos hierárquicos e os métodos de repartição. No primeiro grupo, se destaca a árvore de agrupamento ou *dendrograma* que utiliza as dissimilaridades ou distâncias entre objetos para formar as classes. Quando os dados contêm uma clara estrutura em termos de sua similaridade, essa estrutura é refletida em uma árvore hierárquica que agrupa as classes em ramificações distintas. A maneira mais popular e adequada para calcular as distâncias entre objetos em um espaço multidimensional é por meio do cálculo das distâncias Euclidianas (HAN; KAMBER; PEI, 2006).

Como uma técnica de mineração de dados, a análise dos *clusters* pode resultar em maior conhecimento sobre padrões de distribuição e correlações interessantes entre dados, ou também pode servir para observar as principais características de cada classe quando é planejado gerar um modelo preditivo. Uma das principais vantagens desta técnica é que pode trabalhar com diferentes tipos de atributos como dados binários, nominais ou categóricos, dados ordinários ou mistura entre esses tipos. No entanto, as unidades de medida das variáveis podem afetar seu desempenho, por isso é recomendável padronizar as medidas para que todas as variáveis tenham o mesmo peso.

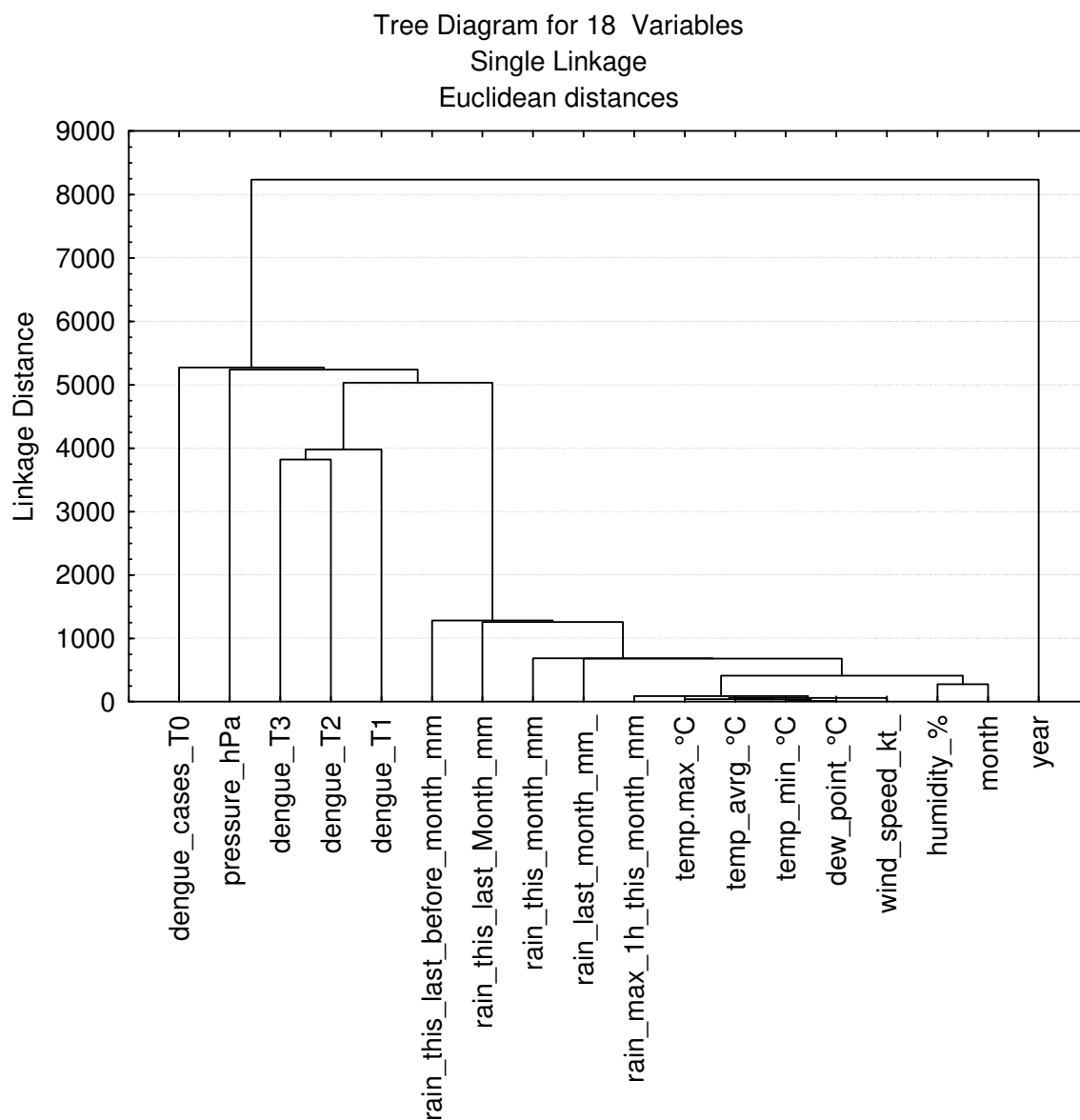
Segundo Wiley (2006), embora a distância euclidiana seja uma medida de dissimilaridade, às vezes ela é referida como uma medida de semelhança, pois quanto maior seu valor, menos parecidos são os indivíduos ou unidades amostrais.

Entre os métodos de repartição, o mais conhecido é a técnica baseada no centróide: o método *k-means*. Esse método divide um grupo de n objetos em k classes de forma tal que a similaridade intra-classes seja alta e a similaridade inter-classes seja baixa.

O método *k-means* é um processo iterativo que, primeiro, agrupa aleatoriamente os objetos em k classes. A partir deste ponto, calcula o valor médio para cada classe e reorganiza os objetos segundo sua distância com esse valor, e procura sempre a classe mais similar.

O gráfico 2 mostra o *dendrograma* das variáveis temporais relacionadas com a incidência de dengue no município de Foz do Iguaçu.

GRÁFICO 2 - DENDROGRAMA DAS VARIÁVEIS TEMPORAIS



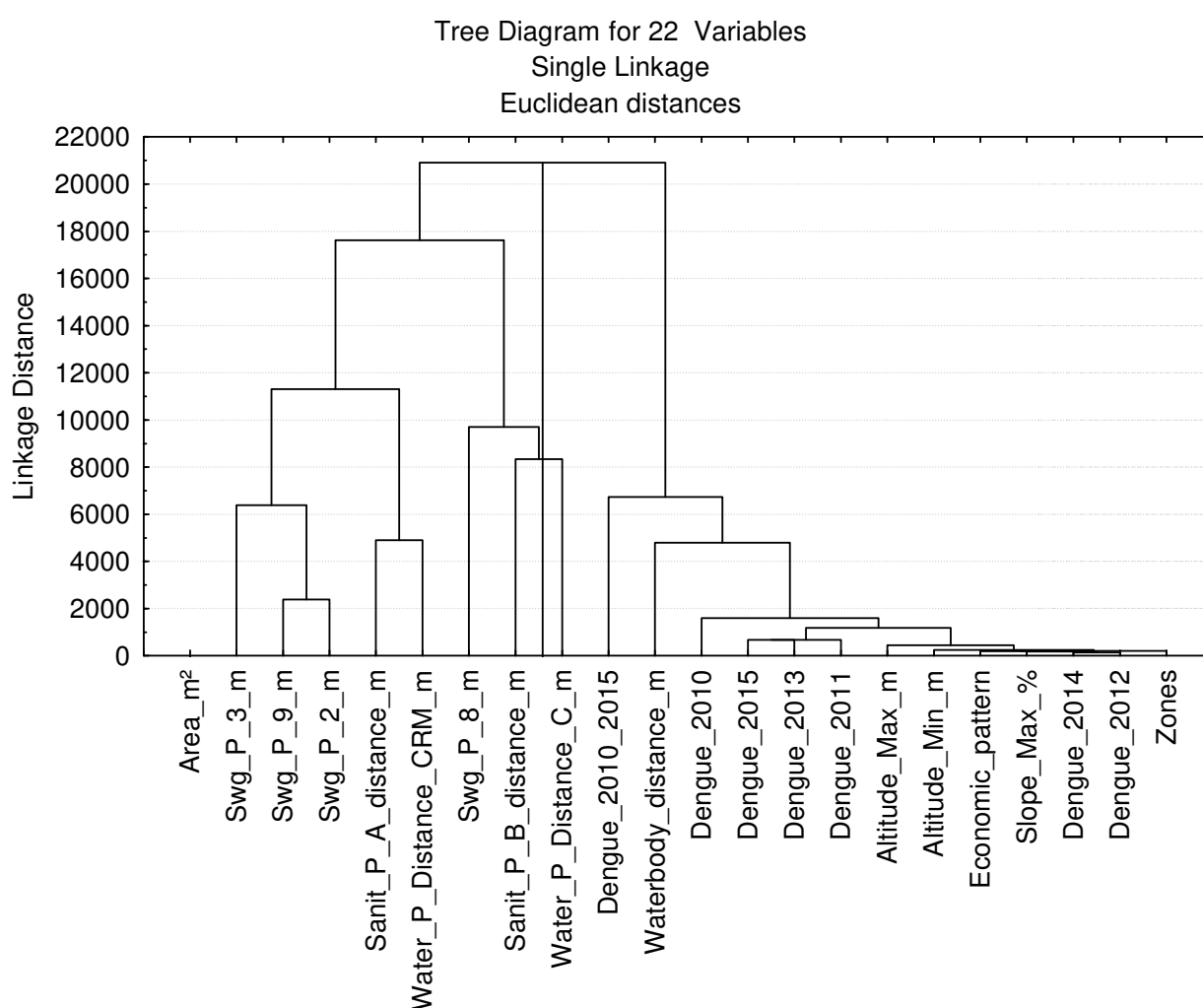
FONTE: O autor (2017).

Verifica-se no gráfico 2 que as variáveis casos de dengue (T0) e pressão atmosférica estão no mesmo galho das variáveis de predição e são as que possuem maior semelhança, no *dendrograma*, por possuírem menor distância Euclidiana, uma vez que essas formam um grupo.

O gráfico ilustra, ainda que de forma qualitativa uma relação da incidência de dengue com a pressão atmosférica, uma vez que tal relação pode ser explicada fisicamente, pois a instabilidade do tempo e conseqüentemente a precipitação é regulada pelos padrões da pressão atmosférica.

O gráfico 3 mostra o *dendograma* das variáveis espaciais relacionadas com a incidência de dengue no município de Foz do Iguaçu.

GRÁFICO 3 - DENDROGRAMA DAS VARIÁVEIS ESPACIAIS

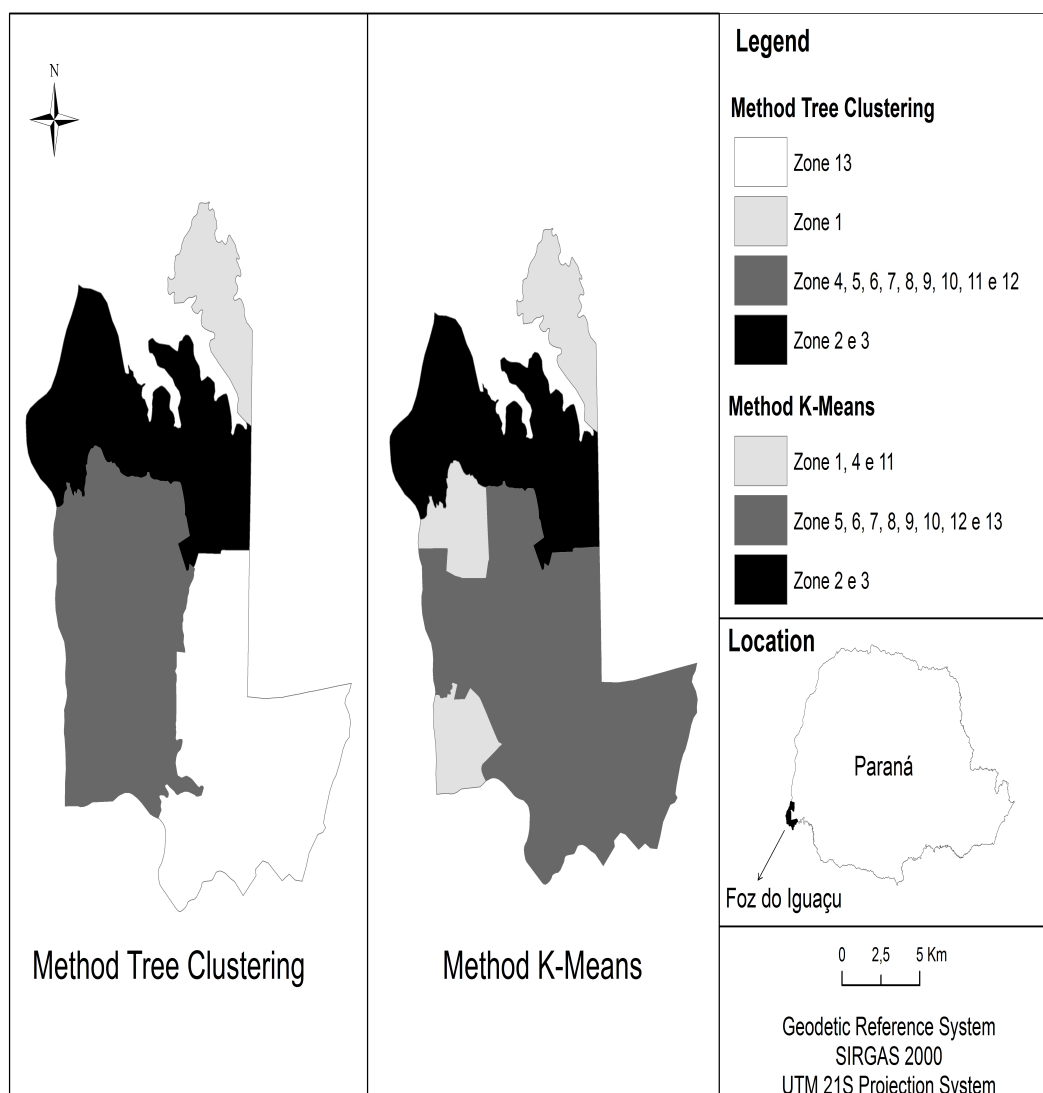


FONTE: O autor (2017).

Percebe-se que existe um padrão qualitativo de relação entre os casos de incidência de dengue com a distância entre os corpos hídricos, padrão econômico e topografia, pois estão no mesmo galho das variáveis de predição e são as que possuem maior semelhança. Espera-se que quanto menor for a declividade da zona, maior a probabilidade de água acumular.

A figura 22 mostra a comparação dos métodos de agrupamento de dados *Tree Clustering* e *K-means* com as respectivas zonas do município de Foz do Iguaçu para os casos de incidência de dengue no ano de 2015 e demonstram semelhanças na obtenção dos resultados encontrados.

FIGURA 22 - COMPARAÇÃO DOS MÉTODOS *TREE CLUSTERING* E *K-MEANS*

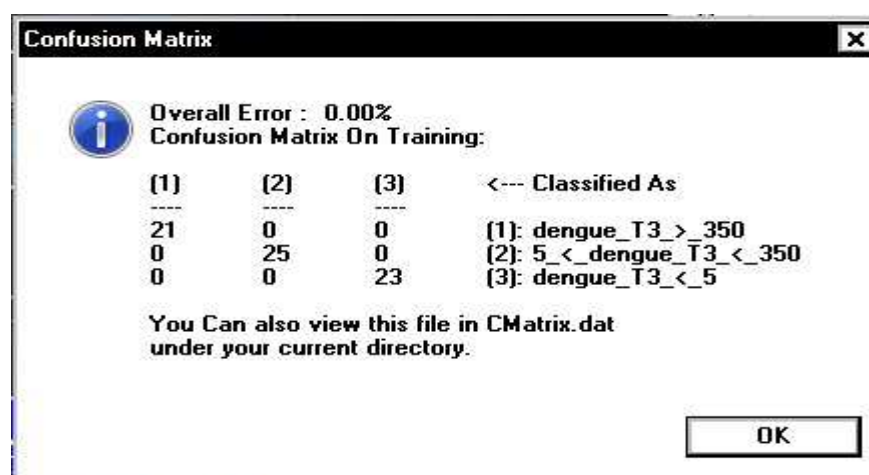


FONTE: O autor (2017).

As figuras 23, 24 e 25 mostram as matrizes confusão dos casos de incidência de dengue por variáveis climatológicas no terceiro, segundo e primeiro mês à frente respectivamente dos modelos de previsão.

Quando um conjunto de dados não é balanceado, ou seja, o número de elementos de cada classe é diferente, o resultado é melhor analisado por meio de uma matriz de confusão, uma vez que mostra com mais detalhes a operação do classificador. A matriz de confusão é uma matriz quadrada onde as colunas representam as n classes reais de entrada e as n linhas das n classes de saída do classificador. A figura 23 mostra a matriz de confusão para previsão de incidência de dengue no terceiro mês à frente.

FIGURA 23 - MATRIZ CONFUSÃO PARA VARIÁVEL DE PREDIÇÃO T3

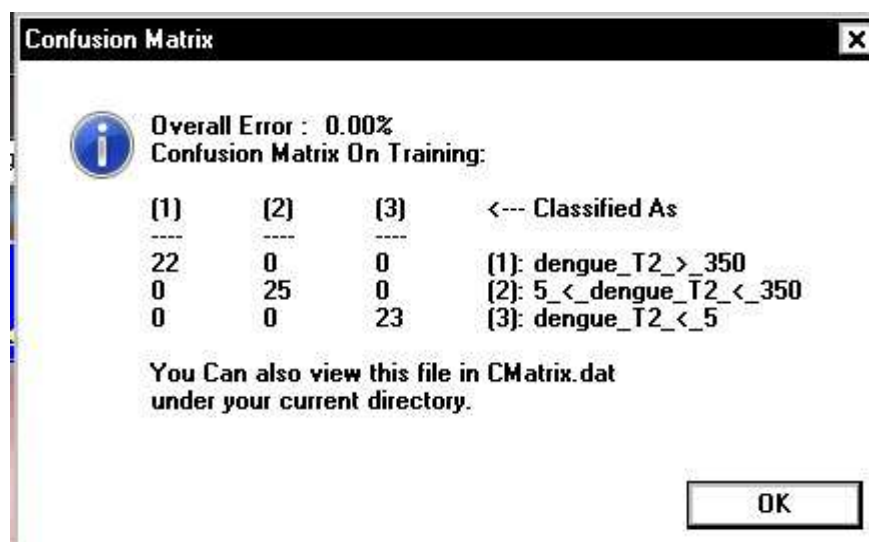


FONTE: O autor (2017).

Observa-se na figura 23 que todos valores pertencentes a diagonal principal da matriz de confusão representam casos de incidência de dengue, uma vez que todos os elementos representados por valores que estão fora da diagonal principal irão compor a matriz de erros. Na faixa do número de pacientes de 5 a 350, o erro observado no modelo é de 0,00% e a acurácia corresponde a 100%.

A figura 24 mostra a previsão de incidência de dengue para dois meses à frente.

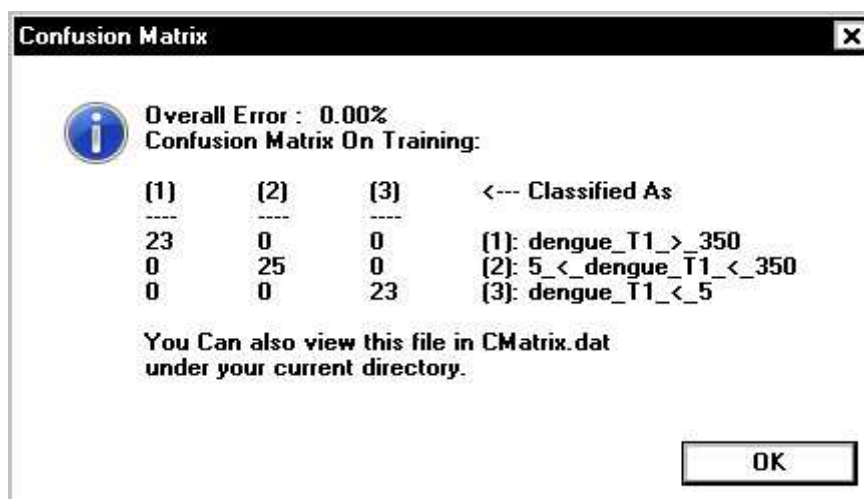
FIGURA 24 - MATRIZ CONFUSÃO PARA VARIÁVEL DE PREDIÇÃO T2



FONTE: O autor (2017).

A figura 25 mostra a previsão de incidência de dengue para um mês à frente.

FIGURA 25 - MATRIZ CONFUSÃO PARA VARIÁVEL DE PREDIÇÃO T1



FONTE: O autor (2017).

A modelagem das regras foi realizada em duas fases, onde a primeira corresponde a discretização dos dados e a divisão em percentil, e, a segunda, a aplicação dos dados no software CBA.

As regras são compostas por: antecedente (A), o consequente (C), o suporte da regra (s%), a confiança da regra (c%). O valor suporte da regra expresso em número de vezes que ocorre A ou C (s), e o número de vezes em que os dois itens A e C ocorrem simultaneamente (s").

O suporte e a confiança das regras são indicadores da utilidade e da certeza das regras descobertas, respectivamente. O valor de s% significa a porcentagem de registros de um determinado valor de A, nos quais também ocorreu C. Por outro lado, c% significa a porcentagem dos registros com valor C que ocorreram pela presença de um valor determinado de A. Neste sentido, as regras são consideradas interessantes se elas satisfazem um limiar de suporte mínimo e um limiar de confiança mínima.

O conjunto 1 é composto por três regras, cada uma relativa aos casos de incidência de dengue para um mês futuro. No conjunto 1, a regra referente foi a seguinte:

```
SE:      dew_point_øC_<_15
         rain_this_month_mm_<_100
ENTÃO  dengue_T1_<_5
        (5.634% 100.000% 4 4 5.634%)
```

Os índices da regra proposta representam:

5.634% - é o valor suporte da regra obtido para as variáveis (dew_point_øC_<_15 e rain_this_month_mm_<_100), ou seja, do total de registros da base de dados, 4 registros foram computados e resulta em um dengue_T1_<_5, número baixo de incidência da doença para um mês à frente.

100.00% - é a confiança da regra, ou seja, dado que o primeiro item aconteça (dew_point_øC_<_15 e rain_this_month_mm_<_100), existe uma probabilidade de 100.00% que aconteça o segundo item (dengue_T1_<_5).

6 - é o valor suporte da regra expresso em número de vezes que ocorre os itens (dew_point_øC_<_15 e rain_this_month_mm_<_100) ou dengue_T1_<_5.

6 - é o número de vezes em que os dois itens ocorrem simultaneamente.

Ainda no conjunto 1, foi encontrada a seguinte regra:

```
SE      15_<_dew_point_øC_<_19
        humidity_%_<_72
        340_<_rain_this_last_before_month_mm_<_450
ENTÃO  5_<_dengue_T1_<_350
        (7.042% 100.000% 5 5 7.042%)
```

Os índices da regra proposta representam:

7.042% - é o valor suporte da regra obtido para as variáveis (15_<_dew_point_øC_<_19; humidity_%_<_72 e 340_<_rain_this_last_before_month_mm_<_450), ou seja, do total de registros da base de dados, 5 registros foram computados e resulta em um 5_<_dengue_T1_<_350, número médio de incidência da doença para um mês à frente.

100.00% - é a confiança da regra, ou seja, dado que o primeiro item aconteça (15_<_dew_point_øC_<_19; humidity_%_<_72 e 340_<_rain_this_last_before_month_mm_<_450), existe uma probabilidade de 100.00% que aconteça o segundo item (5_<_dengue_T1_<_350).

5 - é o valor suporte da regra expresso em número de vezes que ocorre os itens (15_<_dew_point_øC_<_19; humidity_%_<_72 e 340_<_rain_this_last_before_month_mm_<_450) ou 5_<_dengue_T1_<_350.

5 - é o número de vezes em que os dois itens ocorrem simultaneamente.

A última regra do conjunto 1 foi a seguinte:

```
SE      dew_point_øC_>_19
        dengue_cases_T0_>_350
ENTÃO  dengue_T1_>_350
        (11.268% 100.000% 8 8 11.268%)
```

Os índices da regra proposta representam:

11.268% - é o valor suporte da regra obtido para as variáveis (dew_point_øC_>_19 e dengue_cases_T0_>_350), ou seja, do total de registros da base de dados, 8 registros foram computados e resulta em um dengue_T1_>_350, número alto de incidência da doença para um mês à frente.

100.00% - é a confiança da regra, ou seja, dado que o primeiro item aconteça (dew_point_øC_>_19 e dengue_cases_T0_>_350), existe uma probabilidade de 100.00% que aconteça o segundo item (dengue_T1_>_350).

10 - é o valor suporte da regra expresso em número de vezes que ocorre os itens (dew_point_øC_>_19 e dengue_cases_T0_>_350) ou dengue_T1_>_350.

10 - é o número de vezes em que os dois itens ocorrem simultaneamente.

O quadro 2 apresenta o resumo do conjunto 1 de regras.

QUADRO 2 - CONJUNTO 1 DE REGRAS DE CLASSIFICAÇÃO

SE:	dew_point_øC_<_15 rain_this_month_mm_<_100
ENTÃO	dengue_T1_<_5 (5.634% 100.000% 4 4 5.634%)
SE	15_<_dew_point_øC_<_19 humidity_%_<_72 340_<_rain_this_last_before_month_mm_<_450
ENTÃO	5_<_dengue_T1_<_350 (7.042% 100.000% 5 5 7.042%)
SE	dew_point_øC_>_19 dengue_cases_T0_>_350
ENTÃO	dengue_T1_>_350 (11.268% 100.000% 8 8 11.268%)

FONTE: O autor (2017).

Observa-se no conjunto 1 das regras de classificação que a soma do suporte apresentou resultado de 23,94%, ou seja, a temperatura do ponto de orvalho, ponto de condensação de água na atmosfera, explica o número da morbidade de dengue para a previsão de um mês à frente.

O conjunto 2 é composto por três regras, cada uma relativa aos casos de incidência de dengue para dois meses à frente.

O quadro 3 apresenta o resumo do conjunto 2 de regras.

QUADRO 3 - CONJUNTO 2 DE REGRAS DE CLASSIFICAÇÃO

SE	temp.max_øC_<_25 rain_this_last_Month_mm_<_220 5_<_dengue_cases_T0_<_350
ENTÃO	dengue_T2_<_5 (7.143% 100.000% 5 5 7.143%)
SE	25_<_temp.max_øC_<_29 12_<_rain_max_1h_this_month_mm_<_23 1012_<_pressure_hPa_<_1016
ENTÃO	5_<_dengue_T2_<_350 (5.714% 100.000% 4 4 5.714%)
SE	temp.max_øC_>_29 dew_point_øC_>_19
ENTÃO	dengue_T2_>_350 (20.000% 100.000% 14 14 20.000%)

FONTE: O autor (2017).

Observa-se no conjunto 2 das regras de classificação que a soma do suporte apresentou resultado de 32,86%, ou seja, a temperatura máxima explica um terço da morbidade de dengue para a previsão de dois meses à frente.

O conjunto 3 é composto por três regras, cada uma relativa aos casos de incidência de dengue para três meses à frente.

O quadro 4 apresenta o resumo do conjunto 3 de regras.

QUADRO 4 - CONJUNTO 3 DE REGRAS DE CLASSIFICAÇÃO

SE	rain_max_1h_this_month_mm_<_12 20_<_wind_speed_kt__<_25	
ENTÃO	dengue_T3_<_5 (7.246% 100.000% 5 5 7.246%)	
SE	12_<_rain_max_1h_this_month_mm_<_23 15_<_dew_point_øC_<_19 wind_speed_kt__<_20 humidity_%_>_80	
ENTÃO	5_<_dengue_T3_<_350 (5.797% 100.000% 4 4 5.797%)	
SE	rain_max_1h_this_month_mm_>_23 temp_avrg_øC_>_24 = Y	
ENTÃO	dengue_T3_>_350 (14.493% 100.000% 10 10 14.493%)	

FONTE: O autor (2017).

Observa-se no conjunto 3 das regras de classificação que a soma do suporte apresentou resultado de 27,54%, ou seja, o máximo de chuva em uma hora no mês explica a morbidade de dengue para a previsão de três meses à frente.

O conjunto 4 é composto por três regras, referente à análise espacial, cada uma relativa aos casos de incidência de dengue no período de 2010 a 2015 e relacionados com as variáveis urbanas.

O quadro 5 apresenta o resumo do conjunto 4 de regras.

QUADRO 5 - CONJUNTO 4 DE REGRAS DE CLASSIFICAÇÃO

SE:	Sanit_P_B_distance_m_<_5000 Dengue_2010_<_500
ENTÃO:	Dengue_2010_2015_<_1500 (7.143% 100.000% 9 9 7.143%)
SE:	5000_<_Sanit_P_B_distance_m_<_9000 Swg_P_9_m_<_5000
ENTÃO:	1500_<_Dengue_2010_2015_<_2700 (23.077% 100.000% 3 3 23.077%)
SE:	9000_<_Sanit_P_B_distance_m_ 1000_<_Dengue_2012_
ENTÃO:	2700_<_Dengue_2010_2015_ (8.043% 100.000% 9 9 8.043%)

FONTE: O autor (2017).

Observa-se nas regras do conjunto 4 que o suporte apresentou resultado de 38,26%, ou seja, a distância do aterro sanitário em relação aos pontos afetados pela doença explica mais de um terço dos casos de incidência de dengue no período de 2010 a 2015.

5 CONCLUSÃO

Os elementos geográficos presentes no meio urbano foram essenciais na dispersão da doença. Dentre os principais elementos que promoveram a incidência da doença ficou evidenciado que o acúmulo de água é que propicia a formação dos criadouros e infestação do vetor.

O estudo realizou o georreferenciamento dos casos notificados de incidência de dengue no município no período de 2010 a 2015, onde percebeu-se que os anos de 2010, 2011, 2013 e 2015 ocorreram os maiores números de casos de incidência da doença.

Foi elaborado um referencial teórico e revisão bibliográfica sobre a temática dengue associada aos dados climatológicas e variáveis urbanas, uma vez que ficou evidenciada que a mineração de dados é uma atividade multidisciplinar e engloba o desenvolvimento de modelos aplicáveis a fenômenos em engenharia nas diversas escalas: espacial, temporal e multivariada.

Ao realizar análises espaciais dos pontos localizados, com a integração de dados através do sistema de informação geográfica, os SIG's provaram ser bastante úteis para a análise do problema de incidência de dengue e permitiram uma visualização clara entre as variáveis do tempo e espaço, como também visualizar os casos de proliferação e a densidade de aglomeração dos pontos de vulnerabilidades.

Observa-se as correlações climatológicas e ambientais com as ocorrências de dengue no espaço urbano. Os resultados obtidos na pesquisa demonstraram que as condições climáticas do município, as incidências de chuvas, setores populosos, condições geográficas favoreceram a transmissão e evolução da doença nos períodos estudados.

Os resultados apresentados demonstraram que as correlações climatológicas das variáveis pressão atmosférica, velocidade do vento, temperatura, pluviosidade e umidade do ar com a incidência de dengue, bem como as variáveis urbanas analisadas favoreceram a transmissão e evolução da doença.

Portanto, esse estudo elaborou o modelo da predição de incidência de dengue para três, dois e um mês futuro, baseado por correlações de dados climatológicos e apresentou acurácia de 100%.

REFERÊNCIAS BIBLIOGRÁFICAS

AGRAWAL, R.; IMIELIŃSKI, T.; SWAMI, A. **Mining association rules between sets of items in large databases**. In: ACM SIGMOD Record. ACM, 1993. p.207-216.

ALEIXO, Natacha C. R. Percepção e riscos: abordagem socioambiental do processo saúde-doença. In: **Mercator**, Fortaleza, v. 10, n. 22, p. 191-208, mai./ago. 2011.

ALMEIDA, P. G. Mosquitoes (Diptera, Culicidae) and their medical importance for Portugal: challenges for the 21st century. **Acta Médica Portuguesa**, v. 24, n. 6, p. 961–74, 2013.

ALMEIDA FILHO, N; ROUQUAUYROL, M. Z. **Epidemiologia e Saúde**. 06 Ed. Rio de Janeiro, 2003. 728p.

AYOADE, J. O. **Introdução à Climatologia dos Trópicos**. 5ed. São Paulo: Diefel, 1998.

BARCELLOS, C.; BASTOS, F.I. Geoprocessamento ambiente e saúde: uma união possível? **Cadernos de Saúde Pública**. n 12. 1996.

BARRETO, M. L.; TEIXEIRA, M. G. Dengue no Brasil: situação epidemiológica e contribuição para uma agenda de pesquisa. **Estudos avançados**. n 64, v 22. São Paulo, 2008.

BECERRA, J. M. et al. Transport and health: a look at three Latin American cities. **Cadernos de Saúde Pública**, v. 29, n. 4, p. 654-666, 2013.

BESANCENOT, J. **Climat et santé**. Paris: PUF, 2012.

BEATY, B.; STEPHEN, H. Natural Cycles of Vector-Borne Pathogens. In: **Biology of Disease Vectors**. p. 168–185, 2005.

BRASIL. Ministério da Saúde. **Doenças infecciosas e parasitárias: guia de bolso**. Brasília: Departamento de Vigilância Epidemiológica. Secretaria de Vigilância em Saúde. Ministério da Saúde, 2008.

CABENA, P. et al., **Discovering Data Mining: From concept to implementation**. Upper Saddle River – NJ: Prentice Hall, 1998.

CÂMARA, G.; MONTEIRO, A. M. V; FUCKS, S. D; CARVALHO, M. S. **Análise espacial e geoprocessamento**. INPE, 2001.

CARVALHO, L. A. V., **Data Mining: A mineração de Dados no Marketing, Medicina, Economia, Engenharia e Administração**. São Paulo: Editora Érica, 2001.

CARVALHO, M. S.; PINA, M. F.; SANTOS, S. M. **Conceitos Básicos de Sistemas de Informação Geográfico e Cartografia Aplicados à Saúde**. Brasília: Organização Pan-Americana da Saúde/Ministério da Saúde, 2000.

CASSEB, A. D. R. et al. Arbovírus: Importante Zoonose Na Amazônia Brasileira. **Veterinária e Zootecnia**, v. 20, n. 3, p. 391–403, 2013.

CDC. **Addressing emerging infectious disease threats**: a prevention strategy for the United States. Centres for Disease Control and Prevention, 1994.

CONFALONIERI, U. E. C. Variabilidade climática, vulnerabilidade social e saúde no Brasil. In: **Mudanças Climáticas: Repercussões Globais e Locais**. Terra Livre, São Paulo, Ano 19, n. 20, v. 1, p. 51- 63. 2003.

CEOHATS-Committee of the Environmental and Occupational Health of the American Thoracic Society. Health effects of outdoor pollution. **Am J Respir Crit Care Med**; n. 153, p. 3-50, 1996.

CONSOLI, R., OLIVEIRA, R. L. **Principais mosquitos de importância sanitária no Brasil**. FIOCRUZ, 1994.

CORRÊA, R. L. **Estudos Sobre a Rede Urbana**. Rio de Janeiro: Bertrand Brasil, 2006.

COSTA, M. da C. N.; TEIXEIRA, M. da G. L. C. A concepção de “espaço” na investigação epidemiológica. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 15, n. 2, p. 271-279, 1999.

DATASUS. Departamento de Informática do SUS (DATASUS). **Guia de Epidemiologia e Saúde Pública**. 4ed. Brasília, 2012.

DIEGUES, A.C.S.A. **O mito moderno da natureza intocada**. 5ª ed. – São Paulo: Hucitec; Núcleo de Apoio à Pesquisa sobre Populações Humanas e Áreas Úmidas Brasileiras, USP, 2004.

FAYYAD, U. M.; PIATESKY-SHAPIRO, G.; SMYTH, P. **From Data Mining to Knowledge Discovery: An Overview**. In *Advances in Knowledge Discovery and Data Mining*, AAAI Press, 1996.

GOEBEL, M.; GRUENWALD, L., **A Survey of Data Mining and Knowledge Discovery Software Tools**. New Zealand: SIGKDD Explorations, 1999.

HAN, J.; KAMBER, M.; PEI, J. **Data mining: concepts and techniques**. Second ed. San Francisco: Morgan Kaufmann Publishers, 2006.

HOGAN, D. J.; MANDAROLA JR. E. **População e mudanças climáticas: dimensões humanas das mudanças ambientais globais**. Campinas: Núcleo de Estudos de População-Nepo/ Unicamp; Brasília: UNFPA, 2009.

HYGEIA, **Revista Brasileira de Geografia Médica e da Saúde**. v.3, n. 6, p.163-175, 2008.

IPCC. Intergovernmental Panel on Climate Change. **Relatório do Intergovernmental Panel on Climate**. Washington, IPCC, Publicação Científica e Técnica N.65. D.C.: IPCC, 2v., 2015.

ISO 7726/1998. **Ergonomics of the thermal environment-Instruments and methods for measuring physical quantities**, 1998.

KENNETH, L. **Climate and vectorborne diseases**. Published by Elsevier Inc. on behalf of American Journal of Preventive Medicine, 2008.

KLÖSGEN, W.; ZYTKOW, J. M. **Handbook of Data Mining and Knowledge Discovery**. New York: Oxford University Press, 2002.

LEFF, E. **Saber Ambiental: Sustentabilidade, racionalidade, complexidade, poder**. Petrópolis, RJ: Vozes, 2002.

LOPES, P. A., **Probabilidades e Estatística**. Rio de Janeiro: Reichmann & Affonso Editores, 2001.

LUGER, G. F., **Inteligência Artificial – Estruturas e Estratégias para a resolução de problemas complexos**. Porto Alegre: Bookmann, 2004.

MANLY, B. F. J. **Multivariate statistical methods: a primer**. London: Chapman and Hall, 1986.

MARTINS, L. C. et al. Poluição atmosférica e atendimentos por pneumonia e gripe em São Paulo, Brasil. **Revista de Saúde Pública**, v. 36, n. 1, p. 88-94, 2012.

MEDEIROS, J. S. de. **Desenvolvimento metodológico para a detecção de alterações da cobertura vegetal através da análise digital de dados MSS Landsat**. 1986. 81f. Dissertação (Mestrado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 1986.

MELLO, T. N. A. Política (e ação) pública, território e o papel da geografia. In: **Revista da ANPEGE**, v. 7, n. 1, número especial, p. 11-19, out., 2011.

MENDONÇA, F. A. Aspectos da Interação Clima-Ambiente-Saúde Humana: Da relação sociedade-natureza à sustentabilidade ambiental. **Revista RAEGA**, n4, p. 85 - 99. Curitiba, UFPR, 2000.

MENDONÇA, F. A.; LEITÃO, S. A. M. Riscos e Vulnerabilidade socioambiental urbana: uma perspectiva a partir dos recursos hídricos. In: **GeoTextos**, vol. 04, n. 01 e 02, p.145-163, 2008.

MERBITZ, H. et al. GIS-based identification of spatial variable enhancing heat and poor air quality in urban áreas. **Applied Geography**. n. 33, p. 94-106, 2012.

MINAYO, M. C. S. **Saúde e ambiente sustentável: estreitando nós**. Rio de Janeiro: Fiocruz, 2002.

MINISTÉRIO DA SAÚDE. **Série: Capacitação e atualização em Geoprocessamento em saúde. Abordagem Espaciais na Saúde Pública. Série B. Textos Básicos de Saúde**, Brasília – DF, 2006.

NEDEL, A. Sistemas Atmosféricos e Desastres Naturais. In: CEOS WGEDU Workshop, 2010, Santa Maria. **Anais...** Santa Maria: Instituto Nacional de Pesquisas Espaciais, 2010.

NOGAROLLI, M. **Tendências climáticas do Paraná**. Dissertação (Mestrado em Geografia) - Universidade Federal do Paraná, Curitiba, 2005.

OKE, T. R. **Boundary layer climates. Second Ed.** London: Methuen, p. 435, 1987.

OPAS. **Organização Panamericana da Saúde nas Américas: 2007**. Washington, Publicação Científica e Técnica n. 622. D.C.: OPAS, 2 v., 2007.

OPAS. **Organização Panamericana da Saúde Saúde nas Américas: 2013**. Washington, OPAS, Publicação Científica e Técnica n. 531. D.C.: OPAS, 1 v., 2013.

PIEROTE, A. B. **Políticas Públicas de Controle e Prevenção da Dengue em Londrina na Percepção da População e dos Agentes de Saúde**. Dissertação (Mestrado em Geografia) – Universidade Estadual de Londrina, Londrina, 2009.

PYLE, Dorian. **Data preparation for data mining**. Morgan Kaufmann, 1999.

PMFI. Prefeitura Municipal de Foz do Iguaçu. **Plano Diretor de Desenvolvimento Integrado**. Foz do Iguaçu, 2015.

REITER P. et al. Texas Lifestyle Limits Transmission of Dengue Virus. f. 86-89. **Emerging Infectious Diseases**, v. 9, n. 1, Boston/USA, January, 2013.

REZENDE, S. O.; PUGLIESI, J. B.; MELANDA, E. A.; PAULA, M. F., Mineração de Dados. In: REZENDE, S. O. (Org.). **Sistemas Inteligentes - Fundamentos e Aplicações**. Barueri, SP: Manole Ltda, 2003.

RYDIN, Y. et al. Shaping cities for health: complexity and the planning of urban environments in the 21st century. **Lancet**, v. 379, n. 9831, p. 2079, 2012.

SAFATLE, A. **Para onde vamos? Especial Cidade**. São Paulo, n. 9, p. 18-25, jun. 2007.

SANTOS, Milton. **A urbanização brasileira**. São Paulo: Hucitec, 1994.

SANTOS, S. M.; BARCELLOS, C.; CARVALHO, M. S; FLORES, R. Detecção de aglomerados espaciais de óbitos por causas violentas em Porto Alegre, Rio Grande do Sul, Brasil, 1996. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 17, n. 5, p. 1141-1151, 2001.

SILVA, L. J.; ANGERAMI, R. N. **Viroses emergentes no Brasil**. [s.l.] SciELO. Editora FIOCRUZ, 2008.

SILVA, M. E. S.; GUETTER, A. K. Mudanças climáticas regionais observadas no Estado do Paraná. In: **Mudanças Climáticas: Repercussões Globais e Locais**. **Terra Livre**, São Paulo, Ano 19, v.I, n.20. p51 a 63. 2003.

SILVA, J. S.; MARIANO, Z. F.; SCOPEL, I. Dengue no Brasil e as Políticas de Combate ao *Aedes aegypti*: da tentativa de erradicação às políticas de controle. In: HYGEIA, **Revista Brasileira de Geografia Médica e da Saúde**. v.3, n. 6, p.163-175, 2008.

SOUZA, F. T. **A data-based model to locate mass movements triggered by seismic events in Sichuan, China**. Environmental monitoring and assessment, v. 186, n. 1, p. 575-587, 2014.

SOUZA, F. T.; EBECKEN, N. F. F. A data based model to predict landslide induced by Rainfall in Rio de Janeiro City. **Geotechnical and Geological Engineering**, v. 30, n. 1, p. 85-94, 2012.

SOUZA, R. S.; CARVALHO, M. S., **Análise de dados espaciais em saúde pública: métodos, problemas, perspectivas**. Escola Nacional de Saúde Pública. Fundação Oswaldo Cruz: Rio de Janeiro, Brasil. 2005.

ULTRAMARI, C.; FIRMINO, R. J. Urban Beings or City Dwellers? The complementary concepts of urban and city. **City & Time**, v. 4, p. 1-12, 2010.

UN. United Nations. Dept. of Economic and Social Affairs. **Population Division. World urbanization prospects: the 2011 revision: data tables and highlights**. 2011 rev. ed. New York: United Nations; 2012.

WILEY, John. **Data Mining Methods and Models**. New Jersey, 2006.

WEISS, S. M.; INDURKHIA, N. **Predictive data mining: a practical guide**. San Francisco: Morgan Kaufmann, 1998. 228 p.

World Health Organization. **Global alert and response. Dengue/dengue haemorrhagic fever**. Disponível em <<http://www.who.int/csr/disease/dengue/en/>> Acesso em: 17 nov. 2016.