
UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA

**AVALIAÇÃO DE MÉTODOS PARA DETECÇÃO DE
CONGLOMERADOS ESPACIAS PARA DADOS DE DENGUE EM JOÃO
PESSOA**

ANA HERMÍNIA ANDRADE E SILVA

Orientador:

Prof. Dr. Ronei Marcos de Moraes

João Pessoa

2010

ANA HERMÍNIA ANDRADE E SILVA

**AVALIAÇÃO DE MÉTODOS PARA DETECÇÃO DE
CONGLOMERADOS ESPACIAIS PARA DADOS DE DENGUE EM JOÃO
PESSOA**

Monografia apresentada ao departamento
de Estatística da Universidade Federal
da Paraíba, para obtenção do título de
Bacharel em Estatística.

Orientador: Prof. Dr. Ronei Marcos de
Moraes

Monografia sob o título *Avaliação de Métodos para Detecção de Conglomerados Espaciais para Dados de Dengue em João Pessoa*, defendida por Ana Hermínia Andrade e Silva e aprovada em 08 de julho de 2010, para obtenção do título de Bacharel em Estatística.

Banca Examinadora

Prof. Dr. Ronei Marcos de Moraes - UFPB

Prof. Dra. Andréa Vanessa Rocha - UFPB

Prof. MsC. Marcelo Rodrigo Portela Ferreira - UFPB

*Dedico carinhosamente aos meus pais,
João e Graça.*

Agradecimentos

Primeiramente à Deus, por todas as graças a mim concedidas.

Aos meus pais, João e Graça, por serem meus exemplos de perseverança e dedicação.

Aos meus irmãos, Anna Karlla e Christiano.

*Ao meu namorado Tiago, por aguentar todos os meus estresses durante a graduação e
pelo apoio de sempre!*

À Sadraque, pelo companheirismo em todos os momentos da graduação.

*Aos meus colegas de laboratório, Renata, Danielly, Andréa e Ana Carla, além dos
colegas que já seguiram adiante, por todos os momentos de alegria e pelas experiências
compartilhadas. Em particular um agradecimento especial a Andréa e Danielly que
contribuíram com esta monografia.*

Aos colegas de curso.

À todos os meus amigos, em particular, Vanessa, Riquinho, Carol e Andréa.

Aos amigos da Crisma Tambauí.

À todos os professores do departamento de Estatística da UFPB.

Ao meu orientador desde o início da graduação, Ronei, por todos os ensinamentos.

*Aos membros da Banca Examinadora, pela atenção e dedicação no exame desta
monografia.*

À todos que contribuíram de alguma forma com esta monografia.

”Sob tortura, os números confessam tudo o que quisermos”

Zuernir Ventura

Resumo

O dengue é uma doença que vem causando inquietação por parte de técnicos, pesquisadores e da sociedade em geral. Em João Pessoa são observados valores de incidência baixos ou nulos em grande parte dos bairros, entretanto há uma concentração de altos valores em bairros específicos da cidade. Neste trabalho foi analisada a incidência de dengue no município de João Pessoa, no período de 2001 a 2009, onde foram aplicados os métodos da Estatística M (Poisson, Padrão e Rossi) e o Método de Tango. Comparando com os mapas de incidência gerados, a Estatística M utilizando o método de Rossi foi o que melhor retratou a situação da incidência de dengue no município. Com este método detectou-se principalmente conglomerados de valores baixos ou nulos nas regiões oeste e sudoeste, como também conglomerados de valores altos nas regiões norte e sudeste.

Palavras-chave: Dengue, análise espacial, detecção de conglomerados

Sumário

Agradecimentos	p. iv
Resumo	p. vi
Lista de Figuras	p. ix
1 Introdução	p. 1
Objetivos	p. 3
2 Metodologia	p. 4
2.1 Materiais	p. 4
2.2 Métodos	p. 6
2.2.1 Incidência	p. 6
2.2.2 Projeção Geométrica	p. 6
2.2.3 Conceitos de Geoprocessamento	p. 7
2.2.3.1 Dados Cadastrais	p. 8
2.2.3.2 Mapa Cadastral	p. 8
2.2.4 Conceitos Preliminares	p. 9
2.2.4.1 Matriz de Pesos	p. 9
2.2.4.2 Largura de Banda	p. 9
2.2.5 Detecção de Conglomerados Espaciais	p. 9

2.2.5.1	Método de Tango	p. 11
2.2.5.2	Estatística M	p. 12
2.2.6	Aspectos do <i>software</i> Geosurveillance	p. 14
2.2.7	Comparação entre Métodos de Detecção de Conglomerados Es- paciais	p. 15
3	Resultados	p. 17
3.1	Análise Descritiva	p. 17
3.2	Detecção de conglomerados espaciais	p. 22
3.2.1	Método de Tango	p. 23
3.2.2	Estatística M	p. 26
3.2.2.1	Aplicação do método de Poisson	p. 26
3.2.2.2	Aplicação do método Padrão	p. 30
3.2.2.3	Aplicação do método de Rossi	p. 33
3.3	Escolha do melhor método	p. 36
4	Conclusões	p. 38
	Referências	p. 39

Lista de Figuras

1	Legenda para testes retrospectivos do Geosurveillance	p. 14
2	Valores máximos e médios da incidência de dengue	p. 19
3	Histograma e qq-plot da incidência de dengue nos anos de 2001 a 2003	p. 20
4	Incidência de dengue no período de 2001 a 2005	p. 21
5	Incidência de dengue no período de 2006 a 2009	p. 22
6	Aplicação do método de Tango no período de 2001 a 2005	p. 24
7	Aplicação do método de Tango no período de 2006 a 2009	p. 25
8	Aplicação da Estatística M (método Poisson) no período de 2001 a 2005	p. 29
9	Aplicação da Estatística M (método Poisson) no período de 2006 a 2009	p. 30
10	Aplicação da Estatística M (método Padrão) no período de 2001 a 2005	p. 32
11	Aplicação da Estatística M (método Padrão) no período de 2006 a 2009	p. 33
12	Aplicação da Estatística M (método Rossi) no período de 2001 a 2005 .	p. 35
13	Aplicação da Estatística M (método Rossi) no período de 2006 a 2009 .	p. 36
14	Histograma e qq-plot da incidência de dengue nos anos de 2004 a 2006	p. 41
15	Histograma e qq-plot da incidência de dengue nos anos de 2007 a 2009	p. 42

1 Introdução

O dengue é uma doença causada por arbovírus que afeta 56 países. Seu nível endêmico elevado desafia a vigilância, estratégias de controle e o conhecimento atual de prevenção de doenças. Estimativas internacionais indicam que 50 milhões de pessoas são infectadas por ano no mundo, dentre estas os casos de dengue hemorrágico alcançam aproximadamente 500.000, com pelo menos 12.000 mortes (TEIXEIRA et al., 2005).

Estima-se que 2,5 bilhões de pessoas ao redor do mundo residam em áreas com risco potencial de transmissão do vírus (DUARTE; FRANÇA, 2006). Segundo Tauil (2002), diversos fatores de risco são associados ao surgimento do mosquito transmissor. Dentre eles estão a localização em áreas tropicais e subtropicais, devido ao fato de que a temperatura e a umidade nestas áreas favorecerem a proliferação de seu transmissor. Outros fatores de risco que também têm grande influência são: o crescimento populacional, a urbanização inadequada e sem planejamento, migrações, viagens aéreas, aquecimento global, bem como a deficiência dos sistemas de saúde pública e, por consequência, das campanhas de controle do vetor transmissor do dengue.

A ocorrência de sucessivas epidemias de dengue em todo o mundo e no Brasil em particular, vem sendo objeto de inquietação por parte de técnicos, pesquisadores e da sociedade em geral, diante das potenciais consequências, como o grau de transmissão generalizado da doença, por todos os meses do ano, com aumento da morbidade e letalidade (FERREIRA; VERAS; SILVA, 2009). Na Paraíba observa-se que o padrão epidemiológico do

dengue é endêmico, caracterizado por meses de baixa endemicidade e alta endemicidade. Mais especificamente no município de João Pessoa são observados valores de incidência baixos ou nulos em grande parte dos bairros, entretanto há uma concentração de altos valores em bairros específicos da cidade (SOUZA; VIANNA; MORAES, 2007). Um trabalho feito por Lucena e Moraes (2009) indica que as regiões de maior risco relativo da doença nos anos 2004 e 2005 foram as regiões norte e sudeste do município.

A detecção de padrões espaciais no alastramento da endemia e de aglomerados espaciais é relevante para o combate mais efetivo às causas da doença. Segundo Mondini et al. (2005), são importantes as pesquisas relacionadas ao comportamento das incidências de endemias e da duração da transmissão nas áreas que compõem a região de estudo. A identificação de áreas que apresentam risco distinto das demais áreas de risco constante torna-se interesse dos pesquisadores (COULSTON; RITTERS, 2003). Verificar se a distribuição espacial de um evento adverso à saúde é aleatória ou não se torna primordial para o seu combate (SANKOH et al., 2001). Em particular, a escolha do melhor índice de aglomeração espacial, juntamente com uma análise descritiva baseada em mapas cadastrais, visa uma combinação de técnicas que levará a criação de um modelo seguro de monitoramento dos casos, sugerindo áreas preferenciais de combate a doenças (SANCHES, 2000). De posse dessas informações, torna-se viável estabelecer medidas eficazes de combate, tendo como resultado uma imediata melhoria das condições de saúde da população (MARÇAL JÚNIOR; SANTOS, 2004).

Objetivos

O objetivo desse trabalho foi analisar espacialmente a incidência de dengue no município de João Pessoa, no período de 2001 a 2009. Os objetivos específicos foram:

- Verificar a existência de conglomerados espaciais da incidência de dengue;
- Verificar a adequação dos métodos de detecção de conglomerados espaciais da Estatística M e Tango.

2 *Metodologia*

Neste capítulo são apresentados os materiais utilizados para a análise, como base de dados, arquivos de mapas e os *softwares* Geosurveillance e R. Logo após são expostos os métodos necessários para proceder a análise, como o cálculo da incidência de dengue, que por sua vez necessita da projeção populacional para os anos do estudo. Além disso, neste capítulo, encontram-se os métodos de detecção de conglomerados espaciais utilizados, assim como, uma forma adequada de comparação de tais métodos.

2.1 Materiais

Os dados utilizados nesse trabalho são referentes às ocorrências de casos de dengue no município de João Pessoa e foram obtidos junto à Secretaria Municipal. As informações foram subdivididas entre os bairros do município, no período de 2001 a 2009.

As informações populacionais foram obtidas no site do Instituto Brasileiro de Geografia e Estatística (IBGE) e correspondem ao Censo 2000 e a Contagem Populacional de 2007. Para efeito de estudo, foram consideradas as populações por bairro de moradia dos indivíduos. As informações estão disponíveis em: <<http://www.ibge.gov.br>>.

Os arquivos de mapas que foram utilizados são do formato *shapefile*, sendo referentes aos mapas digitais do município de João Pessoa e foram obtidos juntamente ao Departamento de Geoprocessamento da Prefeitura Municipal de João Pessoa. Justifica-se a

utilização de dois mapas devido à ocorrência de uma mudança na divisão dos bairros no ano de 2006. O primeiro mapa está dividido em 66 bairros e foi utilizado como base para o estudo no período de 2001 a 2005. Para o período de 2006 a 2009, foi utilizado o mapa que contém 64 bairros. Além da mudança na junção de alguns bairros, houve também uma diminuição na área da reserva florestal do município.

O *software* utilizado para verificar a presença de conglomerados espaciais da incidência de dengue foi o Geosurveillance, um *software* de georeferenciamento que combina rotinas de estatística espacial com funções básicas do Sistema de Informação Georeferenciada (SIG) para executar testes prospectivos e retrospectivos para detecção de conglomerados espaciais. Testes retrospectivos são aqueles que testam a hipótese nula de que os dados seguem uma distribuição aleatória no espaço geográfico versus a hipótese de que a distribuição dos dados é não aleatória, com dados já conhecidos, o que inclui detecção de conglomerados espaciais. Já os testes prospectivos envolvem a inclusão de novos dados no processo. Neste caso foram utilizados testes retrospectivos (LEE; YAMADA; ROGERSON, 2007).

Com o Geosurveillance, pode-se aplicar dois tipos de métodos retrospectivos para detecção de conglomerados: O método da Estatística M e o método de Tango. A versão utilizada foi a versão 1.1 disponível em <<http://www.acsu.buffalo.edu/rogerson/geosurv.htm>>.

O *software* R consiste em um sistema de análises estatísticas composto por um amplo e variado conjunto de métodos estatísticos tradicionais e modernos, um *software* livre e obtido gratuitamente. A sua utilização foi necessária para gerar mapas coropléticos da incidência de dengue no período de estudo, assim como verificação da normalidade dos dados.

A versão utilizada neste estudo foi a 2.8.1 (R Development Core Team, 2008) e para a geração dos resultados foram utilizadas as seguintes bibliotecas: *shapefiles* - utilizada para leitura e importação dos mapas; *DCluster* - para a geração dos resultados com saídas em forma de mapa e *Nortest* - que é usada para proceder testes de normalidade. O *software* R está disponível em: <<http://www.r-project.org/>>.

2.2 Métodos

2.2.1 Incidência

A utilização da incidência de cada subregião permite a observação da ocorrência do fenômeno em relação à população da subregião. Este cálculo permite uma visão relativa deste fenômeno, o que não é possível ao utilizarmos o número absoluto de casos (MEDRONHO, 2004). O cálculo da incidência é feito da seguinte forma:

$$inc_j = \frac{obs_j}{pop_j}, \quad (2.1)$$

onde obs_j é o número de ocorrências do fenômeno em estudo na subregião j ; pop_j é a população na subregião j .

2.2.2 Projeção Geométrica

Para a análise da incidência de dengue foi necessária a projeção populacional para todos os anos do período estudado, exceto os anos 2000 e 2007 que correspondem às informações obtidas juntamente com o IBGE. Para tanto, foi utilizado o método geométrico (SPIEGELMAN, 1968), iniciando, então, pela taxa de crescimento populacional (r) dada pela equação 2.2:

$$r(\Delta t) = \left(\frac{P(t + \Delta t)}{P(t)} \right)^{1/\Delta t} - 1, \quad (2.2)$$

onde $P(t)$ é a população no tempo t ; $P(t + \Delta t)$ é a população no tempo $t + \Delta t$; t é a data base; Δt é o intervalo entre a data base e a data a ser estimada.

A partir da taxa de crescimento obtida, a população é projetada segundo a equação (2.3):

$$P(t + \xi) = P(t)(1 + r)^\xi, \quad (2.3)$$

onde $P(t + \xi)$ corresponde a população projetada para o tempo $t + \xi$.

Observações:

- Para projeção da população de cada bairro o crescimento populacional foi considerado constante, visto que não estavam disponíveis dados específicos para cada bairro para modelar o crescimento individual no período de estudo;
- A projeção obtida só é válida para o tempo t utilizado como base, neste caso o meio de cada ano, ou seja, o mês de julho.

2.2.3 Conceitos de Geoprocessamento

O geoprocessamento é o processamento informatizado de dados georreferenciados. Para entender melhor esse conceito uma definição importante é a de espaço geográfico:

Definição 2.2.1 *Espaço geográfico é uma coleção de localizações na superfície da Terra, onde ocorrem fenômenos geográficos, sendo definido em funções de suas coordenadas, sua altitude e sua posição relativa (SOUZA; MORAES, 2003).*

Por sua vez, o espaço geográfico pode ser modelado seguindo duas visões complementares (SOUZA; MORAES, 2003):

geocampo: representa a distribuição espacial de uma variável, que possui valores em todos os pontos pertencentes a uma região do espaço geográfico, em um dado tempo t .

geo-objetos: são entidades distintas e localizáveis que compõem uma determinada região do espaço geográfico.

Neste trabalho, o espaço geográfico foi modelado utilizando os geo-objetos, que são os bairros do município de João Pessoa. As informações no espaço geográfico são denominadas de informações espaciais, que se subdividem em duas partes, a parte gráfica e a parte não-gráfica. A parte gráfica representa simbolicamente o geocampo ou os geo-objetos, enquanto a parte não-gráfica armazena as características quantitativas do geo-campo ou dos geo-objetos. A parte gráfica pode ser mapa temático, mapa cadastral, redes, imagens e modelo numérico de terreno, enquanto que a parte não gráfica pode ser dados de área ou dados de superfície (SOUZA; MORAES, 2003). Nesse trabalho os dados utilizados foram os dados de área representados pelo mapa cadastral. Estes conceitos estão explicados detalhadamente a seguir.

2.2.3.1 Dados Cadastrais

O tipo de informação utilizada neste trabalho se enquadra na categoria de dado cadastral, que pode ser definido como:

Definição 2.2.2 *Dados Cadastrais são dados referentes à indivíduos localizados em pontos específicos do espaço que, por motivo de confidencialidade ou de tratamento estatístico, são agregados em unidades de análise, usualmente delimitadas por polígonos fechados (SOUZA; MORAES, 2004).*

2.2.3.2 Mapa Cadastral

A representação gráfica dos dados cadastrais é denominada de **Mapa Cadastral**. Este mapa permite associar duas componentes dos dados de área: a parte gráfica, que representa simbolicamente uma área no mapa, e a parte não gráfica, que armazena as características qualitativas e quantitativas desta área.

2.2.4 Conceitos Preliminares

2.2.4.1 Matriz de Pesos

Para a construção da Matriz de Pesos (ROGERSON, 2001) se utiliza a função de peso gaussiana, da seguinte forma:

$$w_{ij} = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{d_{ij}^2}{2\sigma^2(n/A)}\right); \quad i, j = 1, 2, \dots, n, \quad (2.4)$$

onde σ é a largura de banda (determina o alisamento); d_{ij} é a distância da localização i para a localização j ; A é a área da região; n é o número de subregiões.

Os pesos definidos desta maneira são dimensionados de modo que $\sigma = 1$ implica que a maioria dos pesos é dada a subregiões de um raio aproximadamente igual à distância média entre centróides. À medida que σ aumenta, ocorre um maior alisamento e os pesos atribuídos às subregiões distantes começam a aumentar (ROGERSON, 2001).

2.2.4.2 Largura de Banda

A largura de banda, denotada por σ , é de escolha arbitrária e determina o nível de suavização espacial. Existem dois tipos de largura de banda, a largura de banda simples e a série de larguras de banda. A largura de banda simples consiste na utilização de um único valor, enquanto que ao utilizar uma série de valores como largura de banda pode-se fixar um valor mínimo e um valor máximo variando de acordo com um valor também fixado (LEE; YAMADA; ROGERSON, 2007).

2.2.5 Detecção de Conglomerados Espaciais

Os estudos de detecção de conglomerados são importantes para a identificação de regiões geográficas que apresentam risco elevado em relação a ocorrência de um determinado evento, seja uma doença ou uma epidemia, em suas diversas formas. Em especial, é de interesse tentar verificar a hipótese de que a epidemia apresenta uma distribuição espacial aleatória ou se possui conglomerados de alta incidência.

Sob este aspecto, se faz necessário entender o conceito de distribuição aleatória. As áreas que apresentam um risco significativamente elevado são denominadas conglomerados espaciais. Uma definição para este conceito é:

Definição 2.2.3 *Conglomerado Espacial é um conjunto de áreas que apresentam um risco significativamente elevado quando considerada a hipótese nula (H_0) de que os eventos são gerados aleatoriamente sobre a região de estudo. Um conglomerado pode ser caracterizado como temporal, espacial ou espaço-temporal, dependendo da variável (espaço e/ou tempo) sobre a qual é realizada a análise de contagem dos eventos (COSTA; SCHRERRER; ASSUNÇÃO, 2006).*

Os conglomerados temporais podem ser caracterizados por áreas de caráter temporal, quando a distribuição espacial dos casos é ignorada; espacial, quando especifica-se um determinado intervalo de tempo observando-se os casos sem que seja necessário conhecer o período exato no tempo de ocorrência; ou espaço-temporal quando deve ser analisado não somente a existência de aglomerações no espaço, mas também no tempo. De modo geral, o objetivo é encontrar um grupo de eventos limitado em relação ao tamanho e concentração tal que seja improvável a sua ocorrência por mero acaso.

Existem vários métodos de detecção de conglomerados espaciais, como Varredura Scan e o método de Besag e Newell (LUCENA; MORAES, 2009). Dentre as várias formas de implementação existem a estatística ajustada e a não-ajustada. A forma de cálculo da estatística ajustada dá origem ao método de detecção utilizando a Estatística M, enquanto a não ajustada dá origem ao método de detecção de Tango (LEE; YAMADA; ROGERSON, 2007). Observe as formas de cálculo respectivamente a seguir:

$$U^2 = \sum_{i=1}^n \left(\sum_{j=1}^n \left(\frac{w_{ij}}{\sqrt{esp_j}} \right) (obs_j - esp_j) \right)^2, \quad (2.5)$$

$$U^2 = \sum_{i=1}^n \left(\sum_{j=1}^n w_{ij} (obs_j - esp_j) \right)^2, \quad (2.6)$$

Observação: Tanto a estatística ajustada quanto a não ajustada podem ser aproximadas utilizando a distribuição normal ou a qui-quadrado. A distribuição normal é utilizada para valores altos de σ , enquanto a distribuição qui-quadrado é utilizada para valores baixos de σ .

2.2.5.1 Método de Tango

O método implementado inicialmente por Tango (1995) considerava apenas o número de observações de cada subregião, não levando em consideração a discrepância entre o número de observações e o número esperado de observações. Lawson et al. (1999) propuseram uma nova estatística, baseando-se no conceito de Whittermore et al. (1989). A estatística de Tango é dada então por:

$$U^2 = \sum_{i=1}^n \left(\sum_{j=1}^n w_{ij} (obs_i - esp_j) \right)^2, \quad (2.7)$$

Seguindo Rogerson (1999), assim como Tango (1995), temos:

$$V_p = \Delta(p) - pp', \quad (2.8)$$

onde $\Delta(p)$ é a matriz diagonal das proporções esperadas na região i , (p_i) ; p é o vetor com os p'_i s.

A esperança e a variância de U^2 são dadas respectivamente pelas equações (2.9) e (2.10):

$$E[U^2] = \frac{1}{O^+} Tr(AV_p), \quad (2.9)$$

$$V[U^2] = \frac{2}{(O^+)^2} Tr(AV_p)^2, \quad (2.10)$$

onde A é uma matriz de dimensões $n \times n$; $a_{ii} = \sum_{i=k}^n w_{ki}^2$; $a_{ij} = \sum_{i=k}^n w_{ki}w_{kj}$; O^+ é o número de casos; $Tr(AV_p)$ é o traço da matriz AV_p .

Logo, a estatística de Tango dada pela aproximação distribuição pela normal, que tem distribuição normal padrão é dada pela equação (2.11):

$$T_G = \frac{\left(\frac{U^2}{(O^+)^2}\right) - E(U^2)}{\sqrt{V[U^2]}}. \quad (2.11)$$

Para a aproximação pela distribuição qui-quadrado, definindo ν como:

$$\nu = \left[\frac{[Tr(AV_p)^2]^{1.5}}{[Tr(AV_p)]^3} \right]^2, \quad (2.12)$$

Temos que, a estatística de Tango dada pela aproximação pela distribuição qui-quadrado é dada pela equação (2.13):

$$T'_G = \nu + adj.T_g\sqrt{2\nu}, \quad (2.13)$$

onde ν é o número de graus de liberdade definido pela equação (2.12)

Observação: A distribuição normal é utilizada para valores altos de σ , enquanto a distribuição qui-quadrado é utilizada para valores baixos de σ .

2.2.5.2 Estatística M

O método de detecção de conglomerados espaciais Estatística M nos permite aplicar tanto a estatística global quanto a local. Segundo Rogerson (2001) a Estatística M é dada por:

$$M = max_i \left(M_i = \frac{\sum_{j=1}^n w_{ij}z_j}{\sqrt{\sum_{j=1}^n w_{ij}^2}} \right) \quad (2.14)$$

onde w_{ij} é o elemento da matriz gaussiana de pesos entre as regiões i e j ; z_j é a variável normal na região j ; M_i é a estatística local de cada subregião e é espacialmente ponderada, utilizando os z -scores (z_j), em uma localidades i ;

Para a obtenção do z_j podem ser utilizadas três diferentes formas de transformação, baseados nos valores observados (obs_j) e esperados (esp_j) de cada subregião j . Os métodos foram denominados como Método de Poisson (equação 2.15), Método Padrão (equação 2.16) e Método de Rossi (equação 2.17):

$$z_j = \frac{(obs_j - esp_j)}{\sqrt{esp_j}}, \quad (2.15)$$

$$z_j = \sqrt{obs_j} + \sqrt{obs_j + 1} - \sqrt{4esp_j + 1} \quad (2.16)$$

e

$$z_j = \frac{obs_j - 3esp_j + 2\sqrt{obs_j esp_j}}{2\sqrt{esp_j + j}}. \quad (2.17)$$

O valor esperado de cada região j , denotado por esp_j é calculado segundo a equação 2.18:

$$esp_j = pop_j \frac{\sum_{j=1}^n obs_j}{\sum_{j=1}^n pop_j}, \quad (2.18)$$

onde pop_j é a população da subregião j .

Para detectar a existência do conglomerado espacial a Estatística M tem um valor crítico (M^*), dado pela equação 2.19:

$$M^* = \sqrt{\sqrt{\pi} \ln \left(\frac{4\alpha(1 + 0.81\sigma^2)}{n} \right)}, \quad (2.19)$$

onde α é o nível de significância; σ é a largura de banda; n é o número de subregiões.

Observação: A hipótese nula, de que não há conglomerado espacial, se dá quando a Estatística M é menor ou igual que o valor crítico (M^*).

2.2.6 Aspectos do *software* Geosurveillance

O *software* Geosurveillance foi utilizado para a aplicação de métodos de detecção de conglomerados espaciais. Para a aplicação de tais técnicas os dados a serem utilizados devem ser indexados à parte não-gráfica do mapa, que por sua vez deve ser do formato *shapefile*, ou seja, deve ser composto de três arquivos, ".shp", ".shx" e ".dbf". O *software* não aceita inserções de populações zeradas na base de dados.

No que diz respeito aos testes retrospectivos pode-se gerar mapas coropléticos com seis diferentes cores. Cada cor corresponde a um intervalo de valores da Estatística M, de acordo com o valor crítico M^* (no caso do *software* Geosurveillance a estatística de Tango é considerada como um caso particular da Estatística M sem nenhum dos ajustes).



Fonte: Adaptado de Lee, Yamada e Rogerson (2007)

Figura 1: Legenda para testes retrospectivos do Geosurveillance

Pode-se observar na Figura 1 que, em alguns casos elas são construídas utilizando os valores mínimos e máximos da Estatística M ou podem ser construídas utilizando o valor crítico M^* . Observe que a cor *azul escuro* indica a presença de conglomerados espaciais de valores baixos ou nulos, podendo ocorrer quanto a estatística máxima local é $-2/3$ do

valor mínimo ou $-M^*$. A cor *vermelha* indica a presença de conglomerados espaciais de valores altos, podendo ocorrer quando a estatística máxima local é $2/3$ do valor máximo ou M^* .

Com a aplicação dos testes retrospectivos obtêm-se como saída, além do mapa, um arquivo do tipo *texto* contendo as informações do teste. As informações contidas neste arquivo, para o caso da Estatística M (método de Poisson) e Tango, são:

- O valor máximo da Estatística M;
- O valor crítico M^* (varia apenas com a troca do mapa);
- O valor da estatística U^2 ;
- A estatística do teste e, no caso da utilização do qui-quadrado, também os graus de liberdade;
- p-valor do teste;
- Os valores observados obs_j das subregiões;
- Os valores esperados esp_j das subregiões;
- Os valores de M das subregiões.

No caso dos outros dois métodos da Estatística M esse arquivo contém o valor máximo de M e os valores observados e esperados de cada subregião.

2.2.7 Comparação entre Métodos de Detecção de Conglomerados Espaciais

Para a escolha do método mais adequado para verificar a existência de tendências espaciais não se pode simplesmente compará-los entre si, devido ao fato de que cada um tem sua forma de implementação. Em se tratando de um problema de saúde pública, pela

ótica epidemiológica, para comparar métodos de detecção de conglomerados, compara-se os resultados de cada método com os mapas de incidência da doença, visto que as regiões de alta incidência da doença são as regiões de possíveis aglomerações espaciais, que também são as áreas de prioridade do combate à doença. Além de localizar as regiões de alta incidência, os métodos têm que representar a sua variabilidade (ROTHMAN; GREENLAND, 1998).

3 *Resultados*

Neste capítulo são expostos os resultados obtidos com a aplicação da metodologia descrita anteriormente, como os métodos de detecção de conglomerados espaciais, assim como a análise descritiva da doença. Para possibilitar a utilização dos dados de incidência de dengue, para os dois mapas utilizados, os dados foram indexados utilizando o *software* ArcGis, de acordo com o anexo. Para facilitar a interpretação da incidência, esta foi multiplicada por 100.000, logo, deve ser interpretada para cada 100.000 habitantes. Como subregião utilizou-se os bairros do município de João Pessoa.

3.1 *Análise Descritiva*

Os dados utilizados neste trabalho são referentes aos casos notificados de dengue em João Pessoa no período de 2001 a 2009. Foi feita uma triagem nos dados para retirar informações duplicadas e indivíduos que não informaram o bairro de residência ou nenhuma variável que permitisse a sua localização de moradia, totalizando em 7890 casos.

Para possibilitar o cálculo da incidência de dengue a população de cada bairro do município foi projetada para todos os anos do período, exceto para os anos 2000 e 2007. Para tanto se utilizou o método geométrico, considerando que o crescimento populacional de cada ano permaneceu constante. Calculou-se então a incidência de dengue para cada bairro do município, para os nove anos de estudo.

Tabela 1: Análise exploratória da incidência de dengue

Ano	Mínimo	Máximo	Média	Mediana	Desvio Padrão
2001	0	379,75	42,57	8,01	74,760
2002	0	740,68	105,51	49,63	140,91
2003	0	192,26	21,84	1,52	37,55
2004	0	456,77	46,08	22,85	76,85
2005	0	383,34	89,16	66,46	91,30
2006	0	370,7	83,51	47,21	98,98
2007	0	31974,63	438,77	334,87	429,39
2008	0	967,74	126,64	64,46	170,50
2009	0	633,85	45,75	18,39	87,29

Como pode-se observar na Tabela 1, em todos os anos de estudo existem bairros de incidência nula. O ano em que a incidência de dengue atingiu o máximo do período foi o ano de 2007, onde a maior incidência foi de 31974,63 para cada cem mil habitantes (Figura 2). O ano de 2007 foi o ano em que ocorreu a epidemia da doença no município. Observe também que em média, no ano de 2007 obteve-se o maior valor no período de estudo, igual a 438,77 para cada cem mil habitantes. O ano em que observou-se os valores máximos e de média mais baixos em relação ao restante do período foi o ano de 2003, que foram respectivamente 192,26 e 21,84 para cada cem mil habitantes (Figura 2).

Comparando visualmente a média com a mediana observa-se que, para todos os anos de estudo, a distribuição é assimétrica. Para confirmar essa suspeita aplicou-se o teste de normalidade de Lilliefors no *software* R com nível de significância adotado igual a 0,05. Este teste é uma derivação do teste de Kolmogorov-Smirnov, com o diferencial da ponderação dos valores centrais com as caldas da distribuição (SIEGEL, 1975).

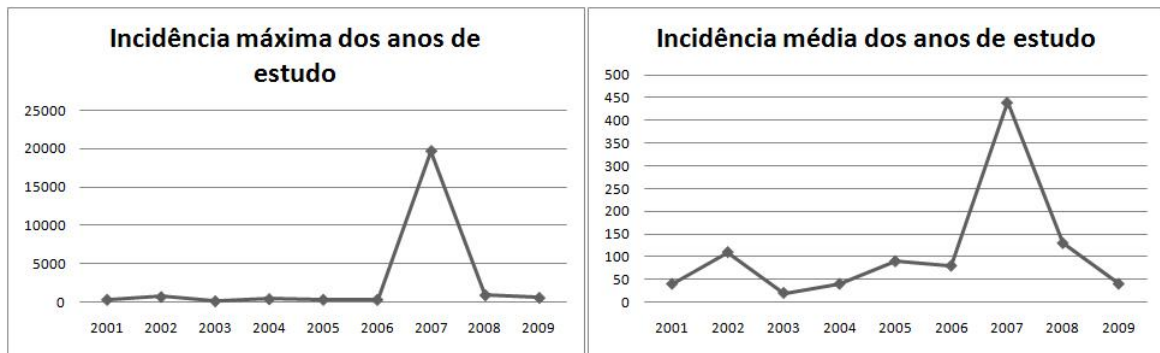


Figura 2: Valores máximos e médios da incidência de dengue

Foi obtido um p-valor menor que $2 * 10^{-16}$ para todos os anos do estudo, podendo-se concluir então que os dados não apresentam uma distribuição normal. Para melhor observar a assimetria dos dados e confirmar o teste de normalidade de Lilliefors foi gerado um histograma e um qq-plot para cada ano de estudo, observe a seguir na Figura 3 os gráficos dos anos 2001 a 2003. Nos outros anos obtiveram-se resultados semelhantes, em anexo.

Para melhor visualizar a distribuição espacial da incidência de dengue em João Pessoa foram gerados mapas coropléticos para cada ano de estudo. Para todo o período houve bairros de média ou alta incidência, em relação aos outros bairros, próximo a reserva florestal, na parte sudeste do município, como se pode ver na Figuras 4 e 5, apenas no ano de 2007, que foi o ano de maior incidência do período (1975 casos/ 100 mil habitantes), todos os bairros em torno da reserva florestal tiveram média ou alta incidência em relação ao restante dos bairros.

Em se tratando do comportamento geral da incidência de dengue, nos anos 2001 e 2002 as regiões de maior incidência foram as regiões norte e sudeste. Já os anos que tiveram apenas o norte como região de maior incidência foram os anos 2003, 2004 e 2009 (Figuras 4 e 5).

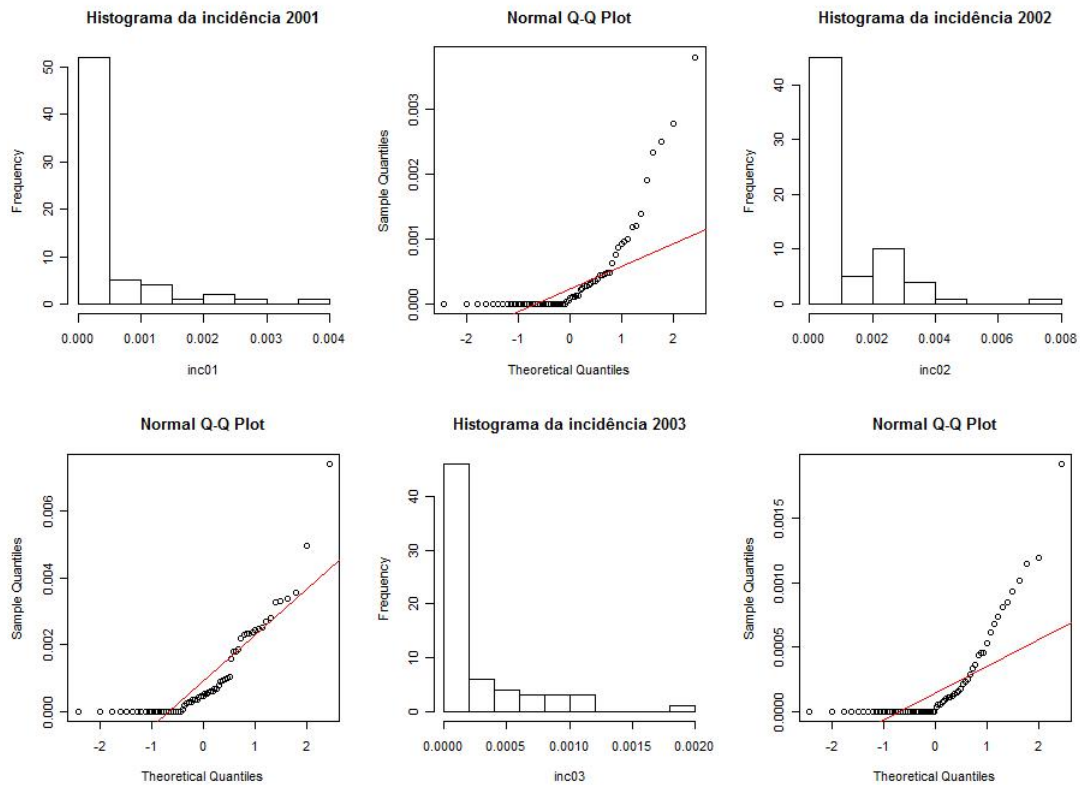


Figura 3: Histograma e qq-plot da incidência de dengue nos anos de 2001 a 2003

Nos anos 2005 e 2006 as regiões de maior incidência foram norte, sudeste e sul. Em particular, no ano de 2007 só não observou-se alta incidência na região sudoeste, e no ano de 2008 a alta incidência foi observada nas regiões leste e sudeste (Figuras 4 e 5). Observe que na região norte houve bairros de alta incidência em todos os anos do período de estudo, apesar de que no ano de 2009 observou-se apenas um bairro de alta incidência (Figuras 4 e 5), e na região sudeste observou-se bairros de alta incidência em seis anos do estudo (2001, 2002, 2005, 2006, 2007 e 2008) (Figuras 4 e 5).

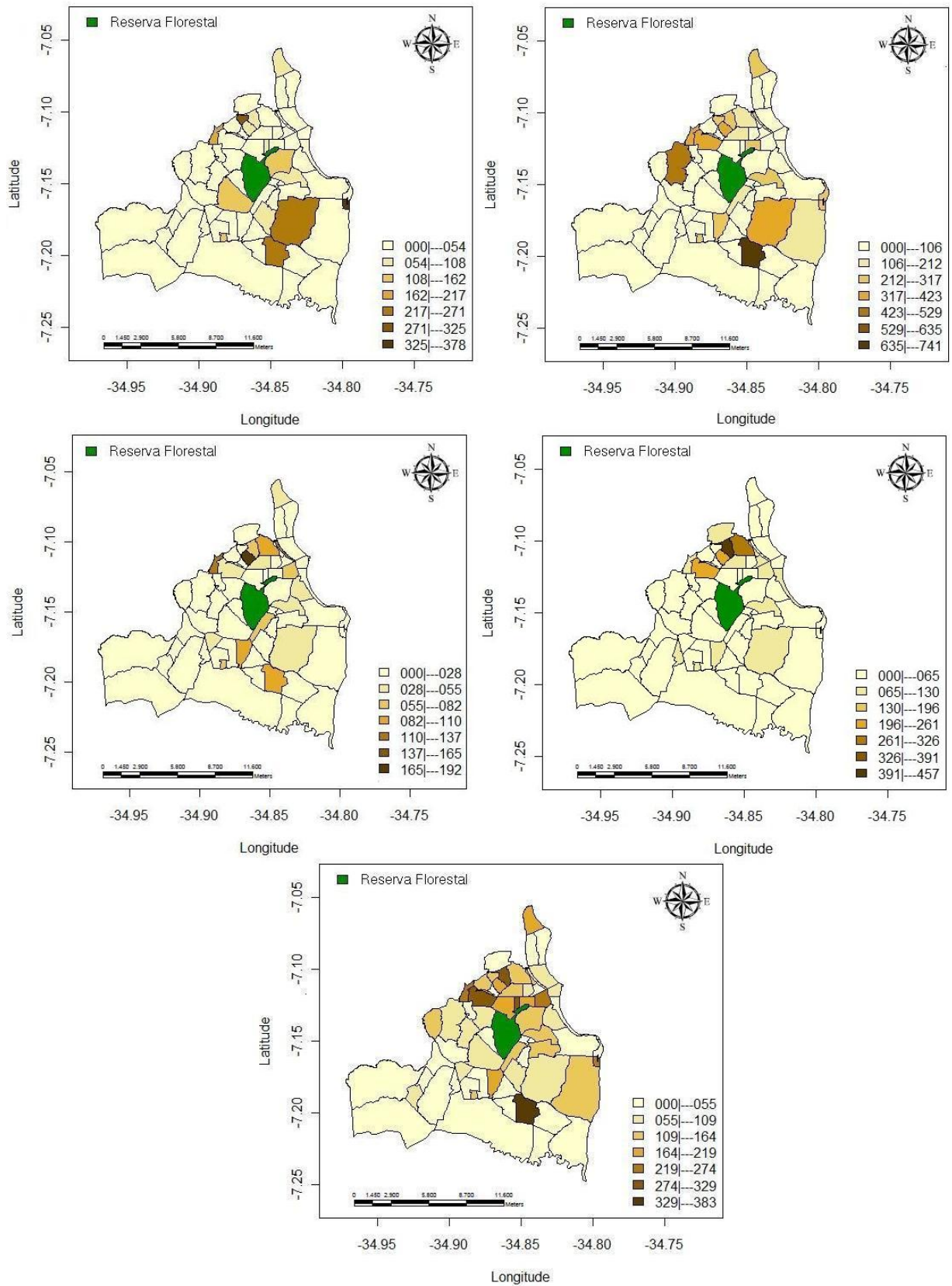


Figura 4: Incidência de dengue no período de 2001 a 2005

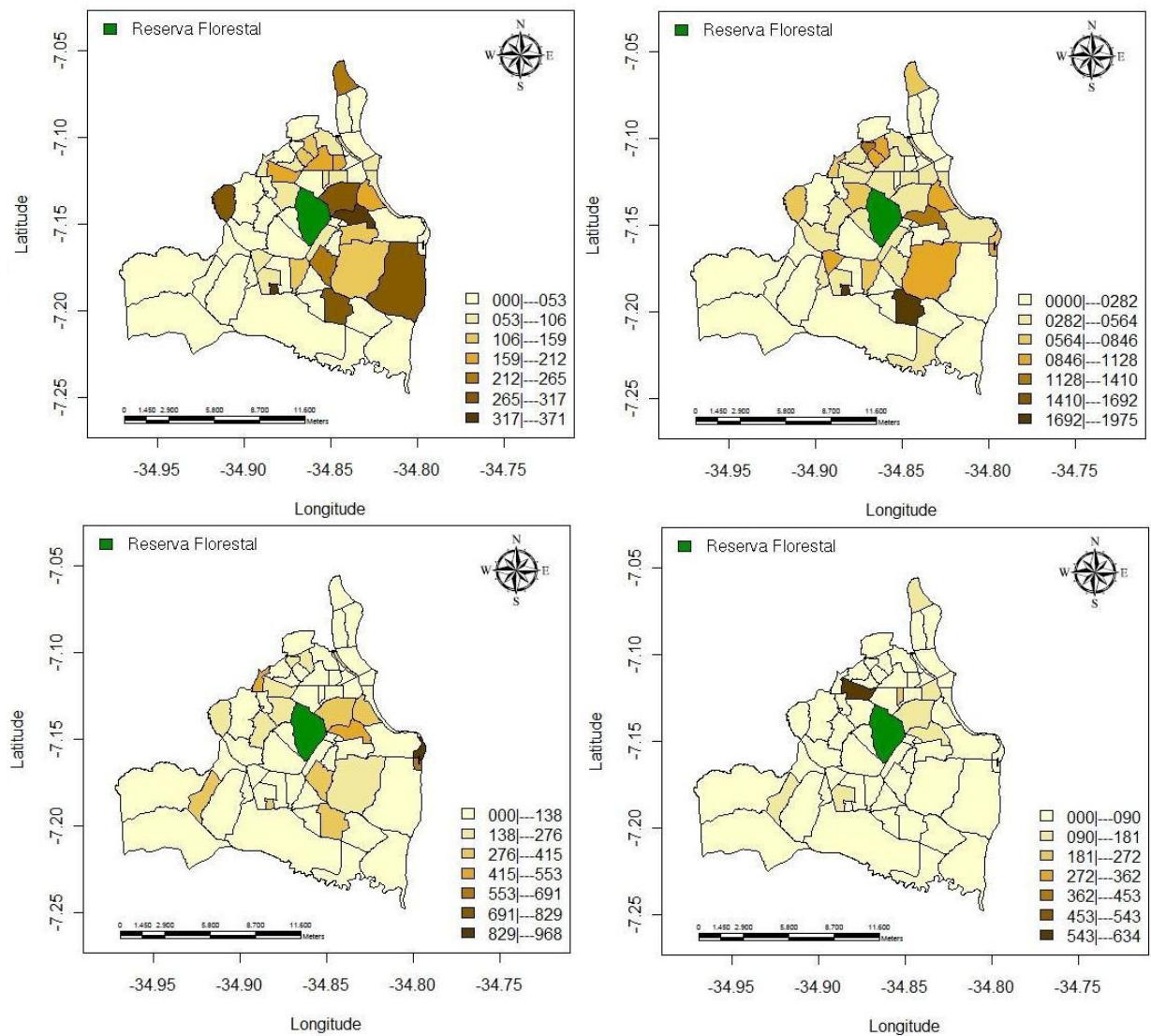


Figura 5: Incidência de dengue no período de 2006 a 2009

3.2 Detecção de conglomerados espaciais

Para a detecção de conglomerados espaciais de dengue foram utilizados dois diferentes métodos de detecção de conglomerados, o método de Tango e a Estatística M, com nível de significância igual a 0,05. Adotou-se também o valor da largura de banda $\sigma = 1$, pois assumiu-se que o desvio padrão de σ se estende desde o centro da região a uma distância igual à distância média entre os centróides dos bairros. Para assumir este valor para σ partiu-se do fato de que existem bairros de alta incidência próximos de bairros com incidência nula.

3.2.1 Método de Tango

Para a detecção de conglomerados espaciais da incidência de dengue foi utilizado o método de Tango. O ajuste utilizado foi pela aproximação da distribuição qui-quadrado, devido ao valor de $\sigma = 1$. Com a aplicação deste método obteve-se as seguintes estatísticas globais:

Tabela 2: Aplicação do método de Tango

Ano	U^2 (não-ajustado)	T máximo	χ^2	g.l.	p-valor
2001	18839,832	86,08	331,928	3,903	<0,0001
2002	53165,084	113,77	113,77	3,903	<0,0001
2003	986,572	9,01	39,963	3,903	<0,0001
2004	9104,999	34,21	176,356	3,903	<0,0001
2005	3,903	39,13	127,162	3,903	<0,0001
2006	13078,458	32,81	143,024	3,799	<0,0001
2007	222254,484	249,13	474,945	4,039	<0,0001
2008	9776,595	44,64	79,398	4,040	<0,0001
2009	2565,329	12,91	62,061	4,085	<0,0001

Pode-se verificar na Tabela 2 que em todos os anos de estudo o método de Tango rejeitou a hipótese nula de que a distribuição espacial da incidência de dengue é aleatória (p-valor < 0,05). Observe ainda que o ano em que ocorreu o maior valor de U^2 foi o ano de 2007, seguido pelo ano de 2002, que foram os anos de maior incidência do período. Para visualizar a detecção de conglomerados foram gerados mapas coropléticos com os valores de T.

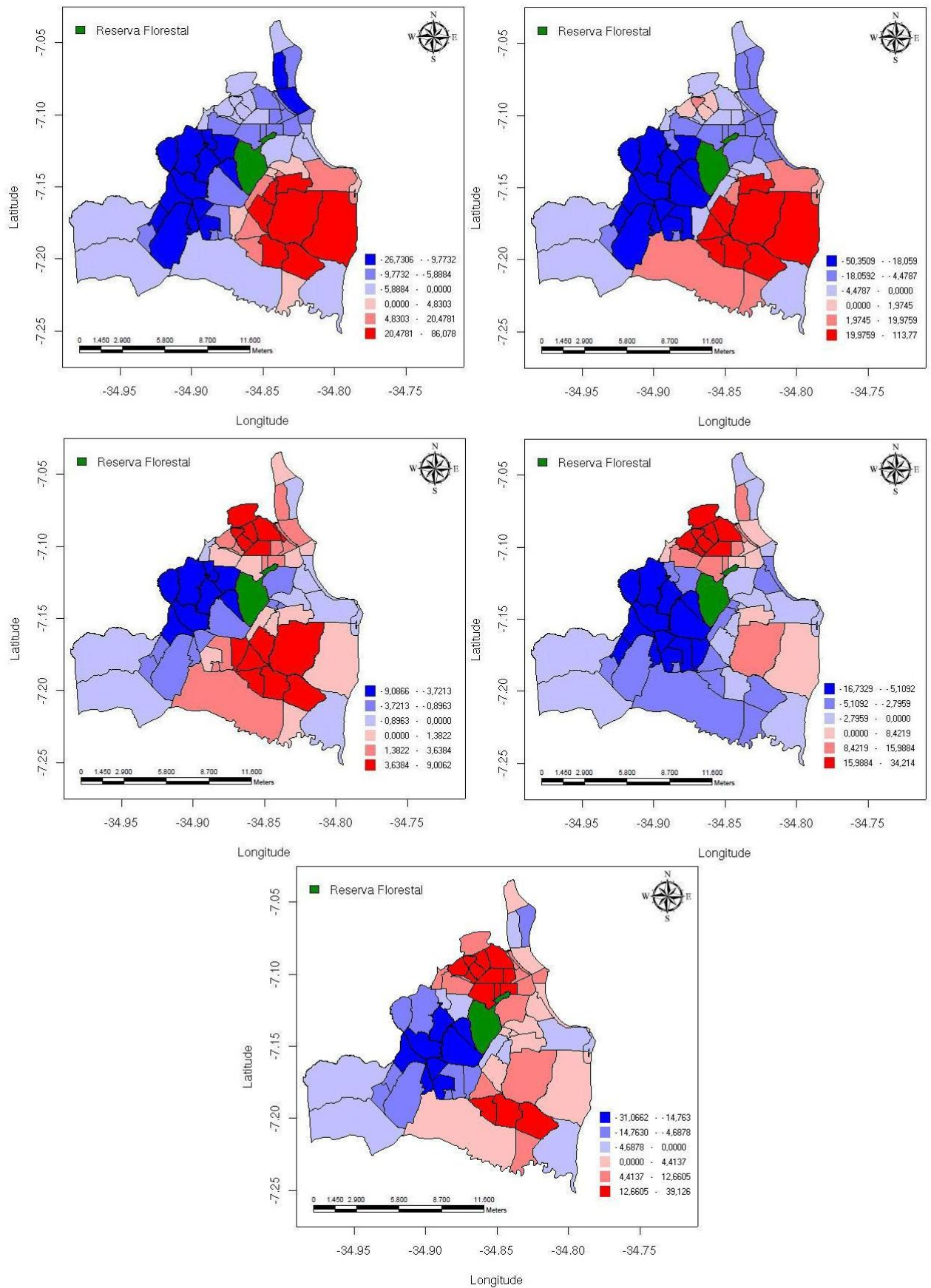


Figura 6: Aplicação do método de Tango no período de 2001 a 2005

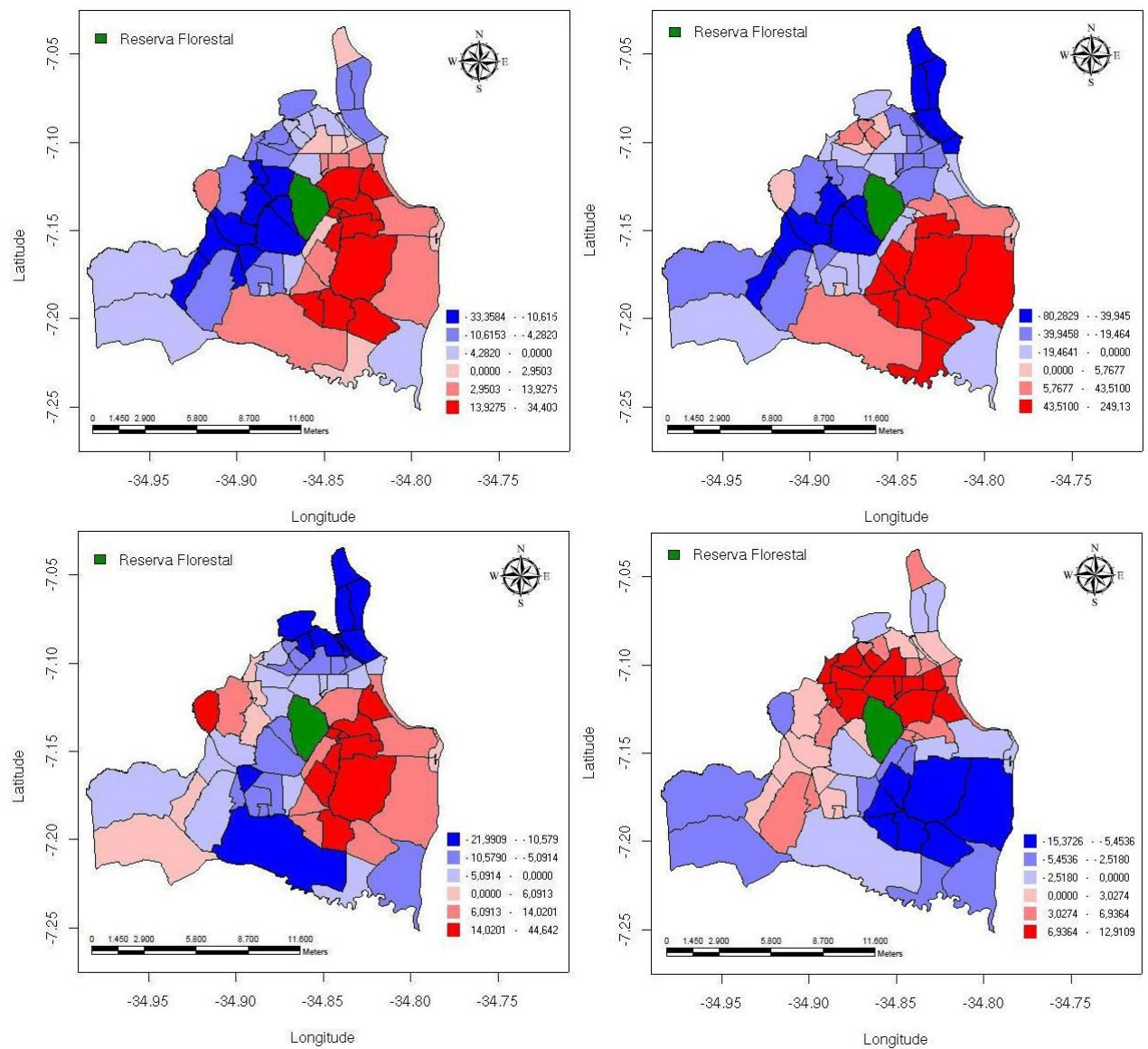


Figura 7: Aplicação do método de Tango no período de 2006 a 2009

Com o método de Tango detectou-se conglomerados de valores baixos ou nulos e de valores altos para todos os anos do estudo. Em relação aos conglomerados de valores baixos ou nulos foram detectados conglomerados na região oeste para todos os anos do estudo, exceto em 2008 e 2009 (Figuras 6 e 7). Já na região sudoeste detectou-se conglomerados nos anos 2001, 2002 e 2007 (Figuras 6 e 7). Outros conglomerados desse tipo detectados com o método foram na região norte, nos anos 2001, 2007 e 2008, e na região sudeste, no ano de 2009 (Figuras 6 e 7).

Em relação aos conglomerados de valores altos foram detectados conglomerados na região sudeste em todos os anos de estudo exceto nos anos 2004 e 2009 (Figuras 6 e 7). Na região norte, por sua vez, detectou-se conglomerados espaciais nos anos 2003, 2004, 2005 e 2009. Já na região leste detectou-se conglomerados espaciais nos anos 2001, 2002, 2006 e 2008 (Figuras 6 e 7). Outros conglomerados desse tipo detectados pelo método foram na região sul em 2007, e na região oeste no ano de 2008 (Figuras 6 e 7).

3.2.2 Estatística M

O segundo método utilizado para a detecção de conglomerados espaciais no município de João Pessoa foi o Estatística M. Por sua vez existem três tipos de métodos de ajuste para esta estatística: Poisson, Padrão e Método de Rossi. Com o objetivo de optar pelo melhor ajuste foram aplicados os três métodos para todos os anos do período de estudo utilizando a aproximação pela distribuição qui-quadrado, devido ao valor de $\sigma = 1$.

3.2.2.1 Aplicação do método de Poisson

Segundo Rogerson (2001) a aplicação da Estatística M utilizando o método de Poisson fornece, na maioria das vezes, os melhores resultados. Com a aplicação desse método obteve-se os seguintes resultados globais:

Tabela 3: Aplicação do método de Poisson

Ano	$U^2(\text{ajustado})$	M máximo	χ^2	g.l.	p-valor
2001	952,967	11,89	104,513	4,104	<0,0001
2002	1228,276	135,324	14,56	4,104	<0,0001
2003	215,767	22,011	3,91	4,104	<0,0001
2004	1252,161	137,997	12,65	4,104	<0,0001
2005	821,013	89,746	6,05	4,104	<0,0001
2006	565,339	62,578	5,89	4,004	<0,0001
2007	1429,706	160,483	15,48	3,992	<0,0001
2008	466,373	50,938	4,83	3,992	<0,0001
2009	705,240	78,483	6,69	4,037	<0,0001

Observe na Tabela 3 que o teste global de detecção de conglomerados espaciais rejeitou a hipótese nula de aleatoriedade na distribuição espacial da incidência (p-valor < 0,05). Verifique também que os anos em que se obtiveram os maiores valores da estatística U^2 foram os anos 2007 e 2002 em que apresentou-se os maiores valores de incidência do período.

Para verificar a existência de conglomerados espaciais foram gerados mapas coropléticos com os valores da Estatística M. Observou-se que em todos os anos de estudo foram detectados conglomerados espaciais de valores baixos ou nulos na região sudoeste do município. Outros conglomerados deste tipo foram detectados na região sul nos anos 2003 e 2008 e na região norte nos anos 2007 e 2009 (Figuras 8 e 9).

Já os conglomerados de valores altos foram detectados na região norte nos anos 2002, 2003, 2004, 2005, 2007 e 2009 (Figuras 8 e 9). Foram detectados conglomerados na região sudeste nos anos 2001, 2002, 2003, 2006, 2007 e 2008 (Figuras 8 e 9) e ainda na região sul nos anos 2003, 2005, 2006 e 2007 (Figuras 8 e 9).

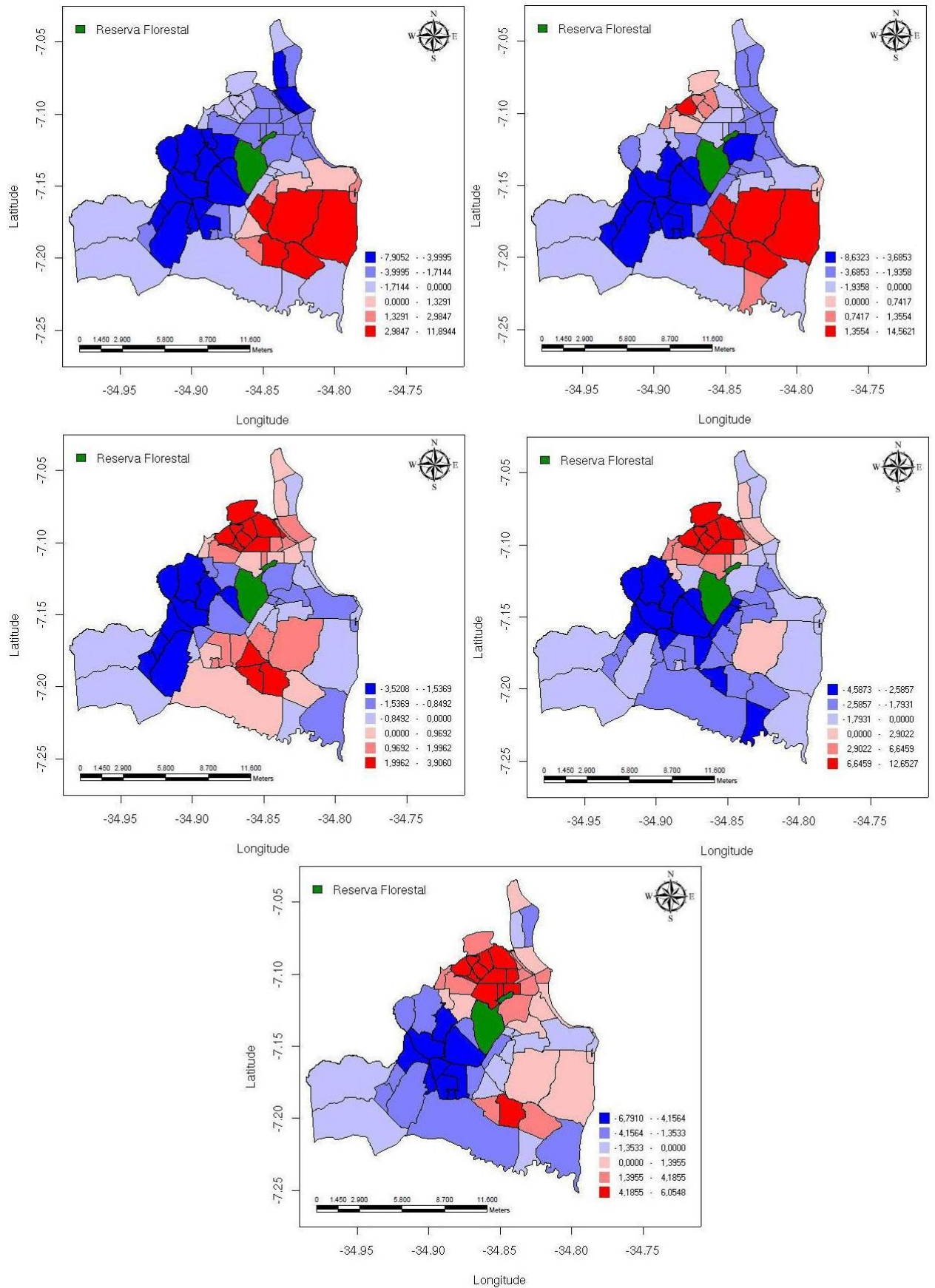


Figura 8: Aplicação da Estatística M (método Poisson) no período de 2001 a 2005

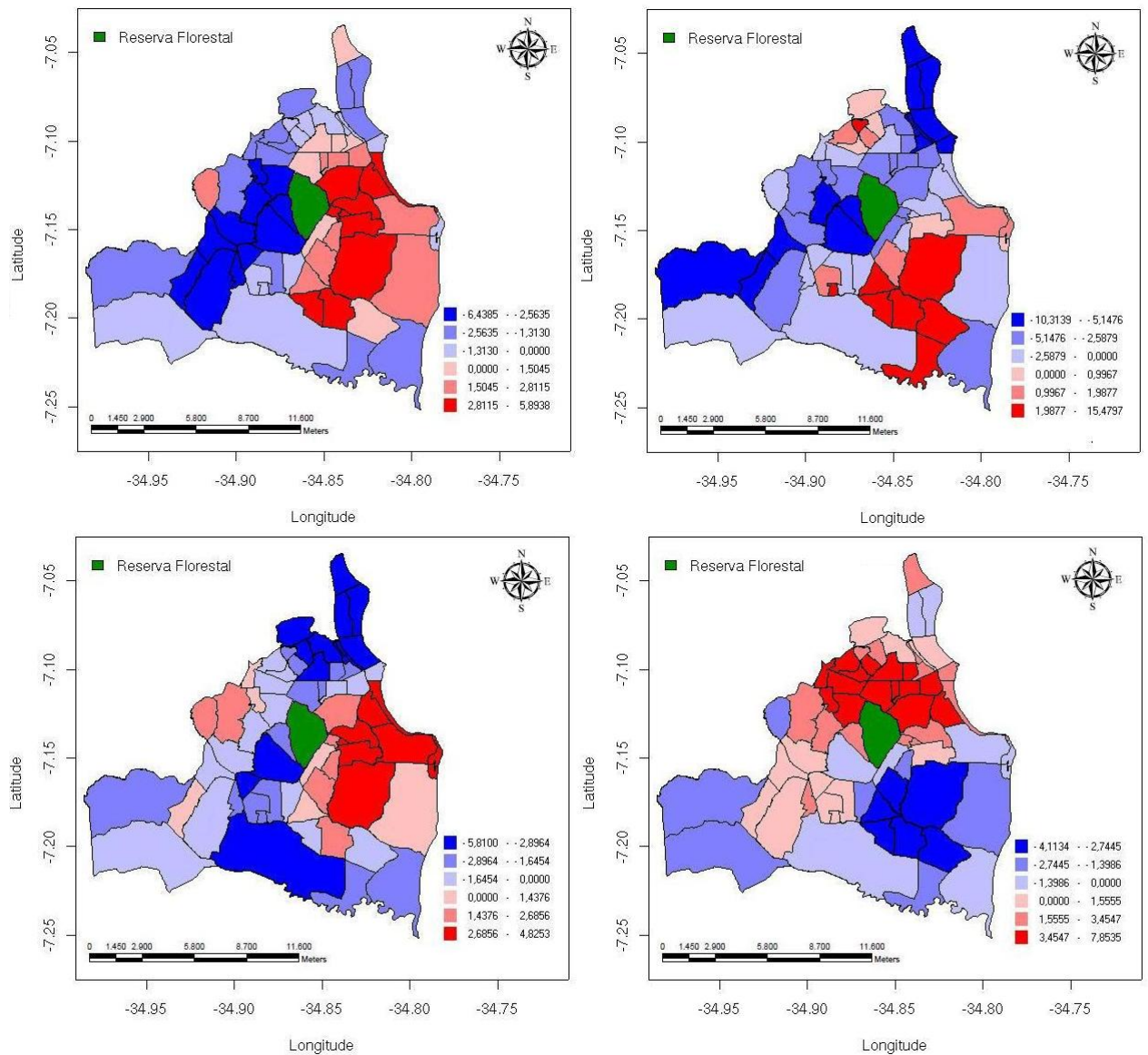


Figura 9: Aplicação da Estatística M (método Poisson) no período de 2006 a 2009

3.2.2.2 Aplicação do método Padrão

O segundo método aplicado para detecção de conglomerados espaciais utilizando a Estatística M foi o método padrão. Em relação à estatística global, para todos os anos de estudo o p -valor $< 0,0001$, ou seja, a distribuição espacial da incidência de dengue é não aleatória. Além disso, o M^* foi igual a 3,038 para o período de 2001 a 2005 e de 3,029 para o período de 2006 a 2009. Para os anos de estudo os valores máximos estão na tabela a seguir.

Tabela 4: Valores máximos de M para o método Padrão

Ano	2001	2002	2003	2004	2005	2006	2007	2008	2009
M max	8,524	9,107	3,038	9,360	5,058	3,914	8,134	3,029	2,809

Para observar o resultado local da detecção de conglomerados espaciais foram gerados mapas coropléticos com os valores da Estatística M. Em relação a detecção de conglomerados de valores nulos ou baixos, detectou-se conglomerados desse tipo em todas as regiões do município nos anos 2001, 2002 e 2007 (Figuras 10 e 11). Em relação à região oeste, o único ano em que não se detectou conglomerados espaciais de valores baixos ou nulos foi 2008. Além desses, detectou-se conglomerados espaciais de valores baixos ou nulos na região sul nos anos 2004, 2005, 2006, 2008 e 2009 (Figuras 10 e 11). Outros conglomerados espaciais de valores baixos foram detectados na região norte, nos anos 2004 e 2005 e na região sudeste em 2009 (Figuras 10 e 11).

Em se tratando da detecção de conglomerados de valores altos, não foi detectado nenhum conglomerado nos anos 2003, 2007, 2008 e 2009 (Figuras 10 e 11). Apesar do ano de 2007 ter sido o de maior incidência do período, não detectou-se nenhum conglomerado, uma vez que se baseia no valor máximo da incidência, porém a média da incidência também foi alta (Tabela 1). Nos anos 2001 e 2002 (Figura 10) foram detectados conglomerados na região sudeste. Já na região norte houve conglomerados no ano de 2004. Outro conglomerado detectado pelo método foi em 2006 na região leste do município (Figura 11).

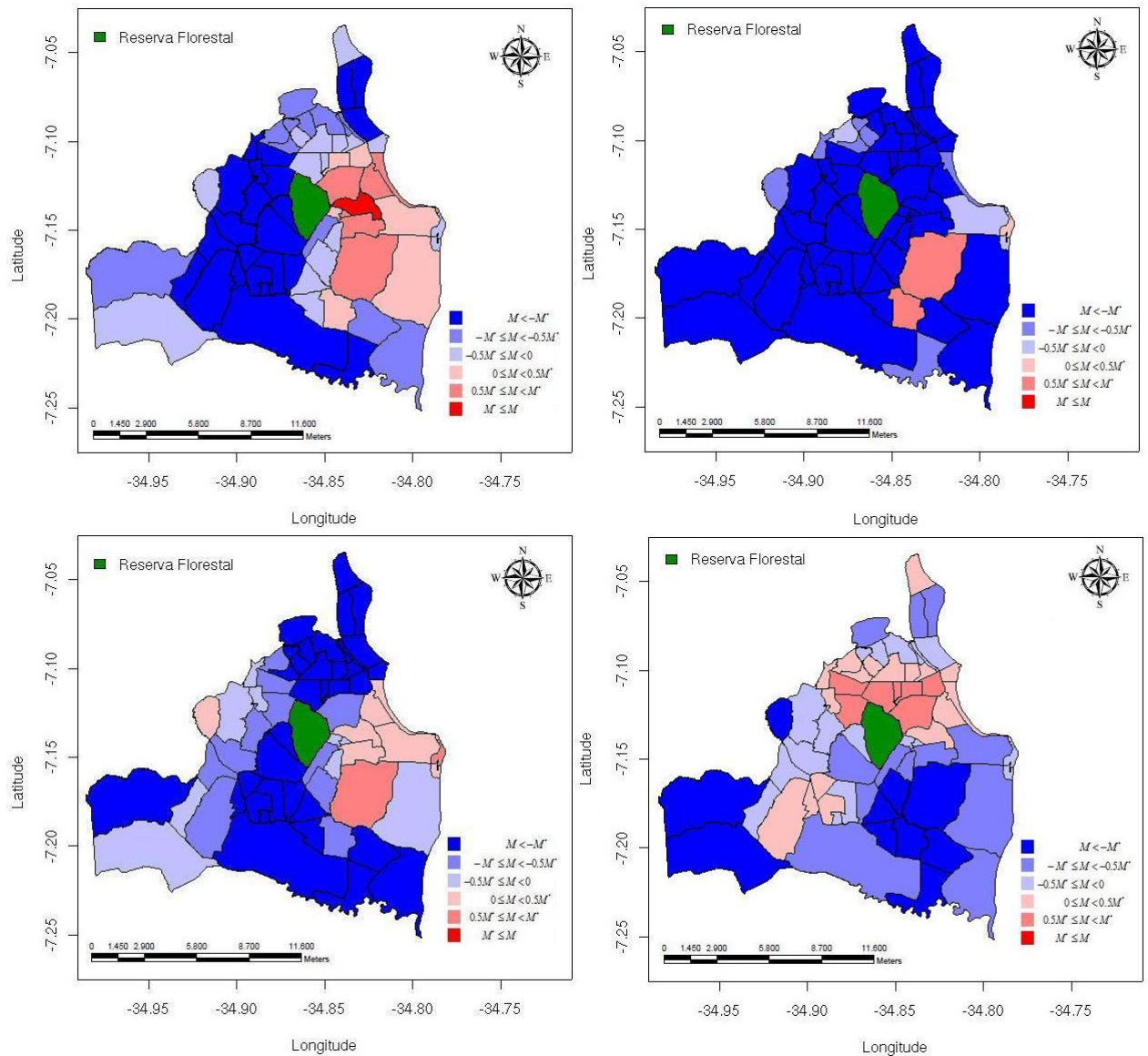


Figura 11: Aplicação da Estatística M (método Padrão) no período de 2006 a 2009

3.2.2.3 Aplicação do método de Rossi

O terceiro método aplicado para detecção de conglomerados espaciais da incidência de dengue foi o método de Rossi. Em relação à estatística global, para todos os anos de estudo o $p\text{-valor} < 0,0001$, ou seja, a distribuição espacial da incidência de dengue é não aleatória. Além disso, o M^* foi igual a 3,038 para o período de 2001 a 2005 e de 3,029 para o período de 2006 a 2009. Para os anos de estudo os valores máximos estão na tabela seguinte.

Tabela 5: Valores máximos de M para o método Rossi

Ano	2001	2002	2003	2004	2005	2006	2007	2008	2009
M max	8,524	9,107	1,560	9,360	5,058	3,914	8,134	1,560	4,517

Em relação à detecção de conglomerados de valores baixos ou nulos foram detectados conglomerados espaciais na região oeste e sudoeste em todos os anos do estudo exceto em 2006 e na região sul detectou-se conglomerados em todos os anos a não ser em 2003 (Figuras 12 e 13). Outros conglomerados detectados foram detectados nos anos 2001, 2002 e 2003 na região norte (Figuras 12 e 13).

Com respeito aos conglomerados de valores altos não detectou-se conglomerados para os anos 2003 e 2008 (Figuras 12 e 13). Na região norte detectou-se conglomerados nos anos 2004, 2005 e 2009 e na região sudeste detectou-se nos anos 2001, 2002, 2006 e 2007 (Figuras 12 e 13). Outro conglomerado deste tipo foi detectado em 2006 na região oeste.

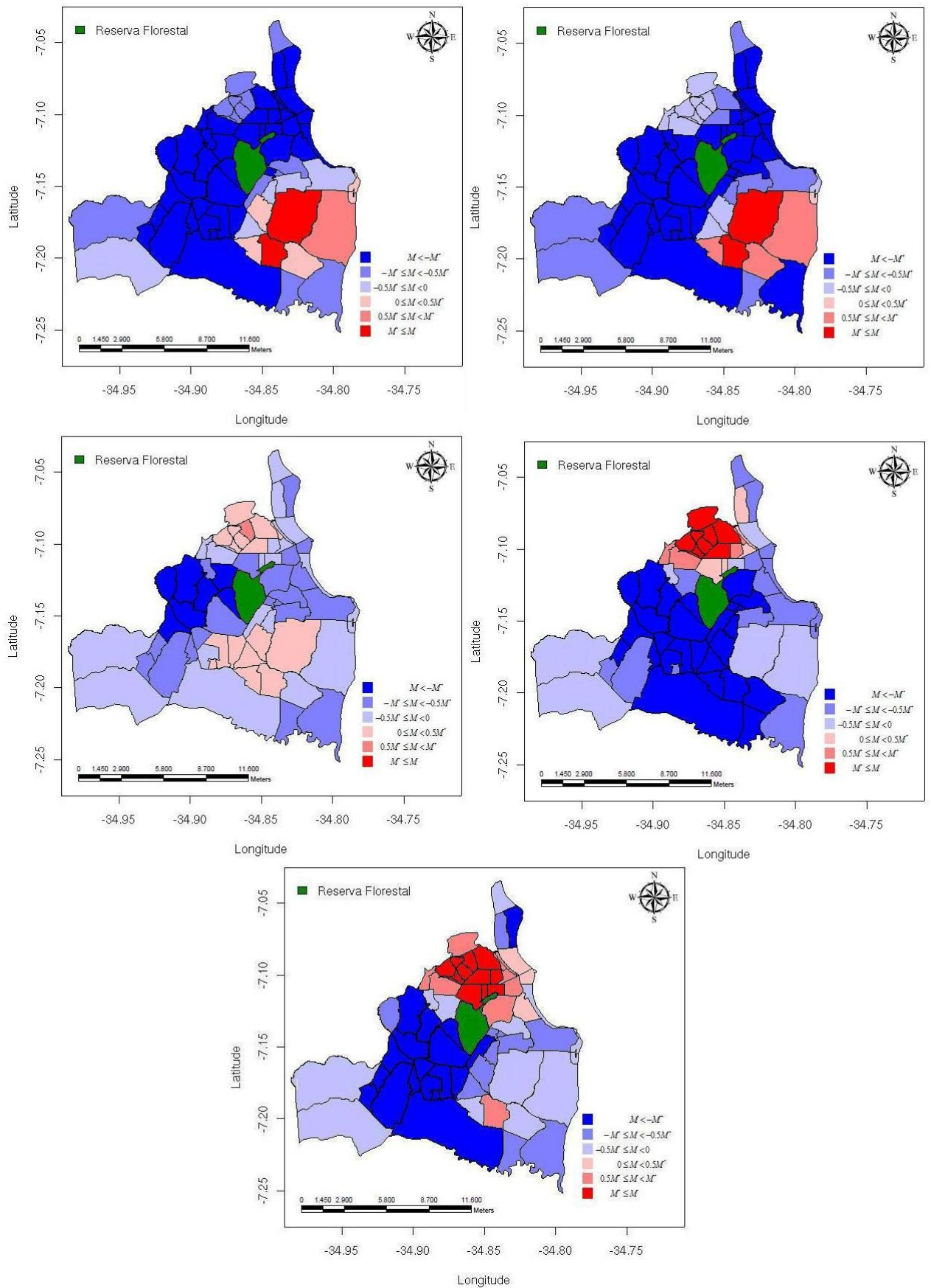


Figura 12: Aplicação da Estatística M (método Rossi) no período de 2001 a 2005

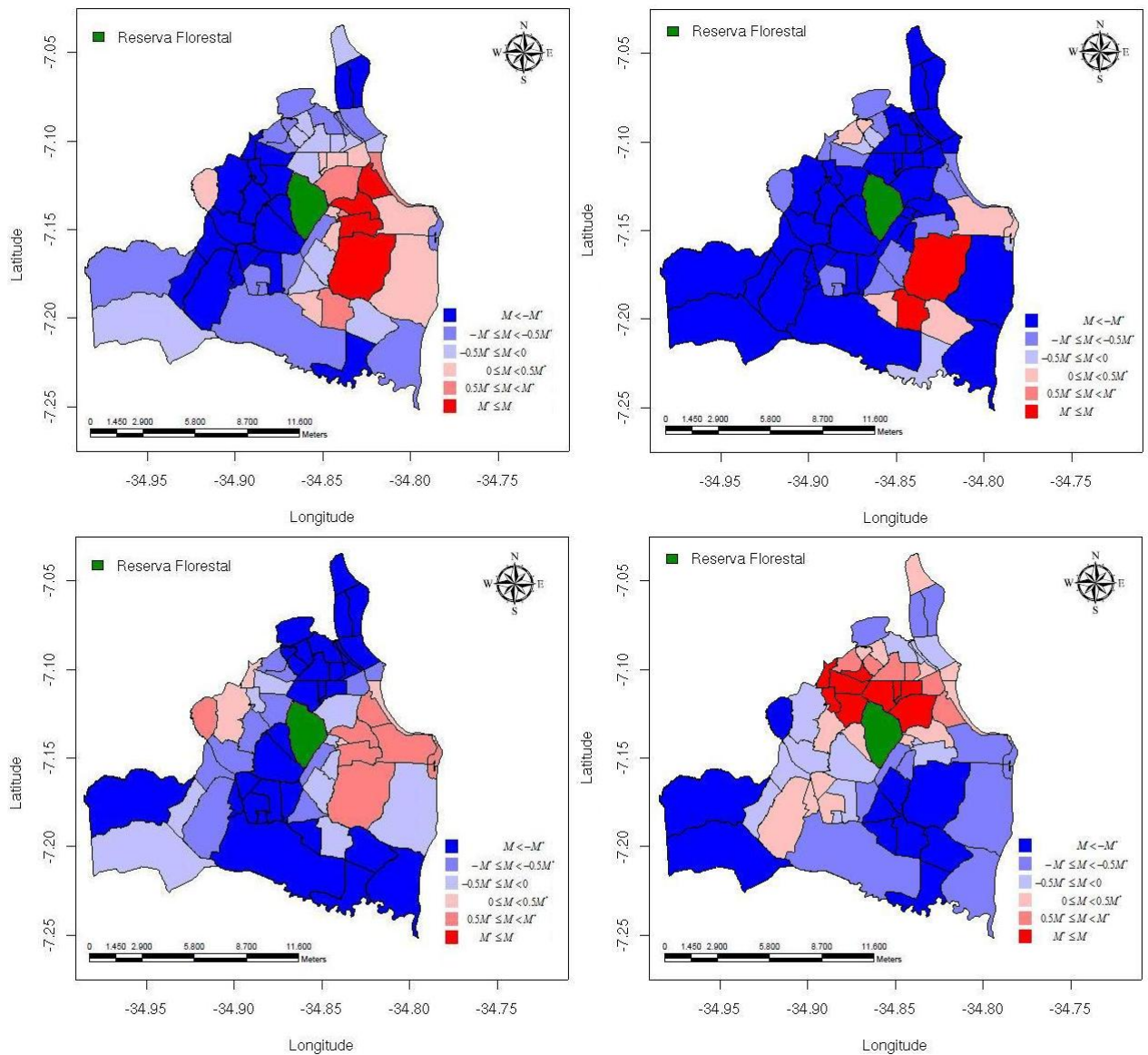


Figura 13: Aplicação da Estatística M (método Rossi) no período de 2006 a 2009

3.3 Escolha do melhor método

Para a análise da incidência de dengue no município de João Pessoa foram aplicados métodos de detecção de conglomerados espaciais, o Método de Tango e a Estatística M com três variações (Poisson, Padrão e Rossi). Para escolher o método mais adequado para esses dados não se pode simplesmente compará-los entre si, devido ao fato de que cada um tem sua forma de implementação. Visto isso, para a escolha do melhor método pode-se comparar cada um com os mapas de incidência, pois estes mostram o que se

espera encontrar em termos de conglomerados espaciais.

Na comparação dos métodos de detecção com os mapas de incidência, levando-se em consideração os conglomerados de valores altos, verificou-se que os métodos de Poisson e Tango foram mais eficientes na detecção de conglomerados espaciais maiores, ou seja, com um maior número de bairros, assim como os métodos Padrão e Rossi foram mais eficientes na detecção de conglomerados menores, com menos bairros. Mais especificamente, observou-se que com o método de Tango, mesmo para conglomerados grandes, captou-se muitos bairros com incidência baixa ou zerada. Já para os conglomerados pequenos, com o método de Rossi captou-se mais bairros com incidência alta do que com o método Padrão.

De uma forma geral, a tendência dos dados de incidência de dengue, neste período, foi de se localizar em poucos bairros próximos, devido à discrepância de seus valores. Os casos em que ocorreram conglomerados maiores foram os anos de menor incidência, como 2005 e 2006, onde com o método de Poisson obteve-se resultado mais próximo à realidade do que com Rossi, porém, nos outros anos com o método de Rossi obteve-se resultado de melhor ou igual forma a realidade da distribuição espacial de incidência de dengue.

4 *Conclusões*

Para a análise da incidência de dengue foram utilizadas informações obtidas juntamente à Secretaria Municipal de Saúde, no período de 2001 a 2009. Para tanto, além da utilização da estatística descritiva, incluindo os mapas cadastrais, foram aplicados dois métodos de detecção de conglomerados espaciais: o método de Tango e a Estatística M, em suas três variações (Poisson, Padrão e Rossi). Tais métodos foram aplicados utilizando a aproximação pela distribuição qui-quadrado, devido ao valor de $\sigma = 1$.

Verificou-se que os anos com maior incidência de dengue no período de estudo foram os anos 2007 e 2002, enquanto que o ano 2003 apresentou a menor incidência do período estudado. Com a aplicação dos métodos de detecção de conglomerados espaciais concluiu-se que, para conglomerados menores, a aplicação da Estatística M utilizando o método de Rossi possibilitou a obtenção de melhores resultados, de acordo com o esperado pelo mapa de incidência, enquanto que, para conglomerados maiores, a aplicação da Estatística M utilizando o método de Poisson possibilitou a obtenção dos melhores resultados. Para a escolha de apenas um método que represente todo o período, optou-se pelo método de Rossi, uma vez que este foi mais adequado para detecção de conglomerados espaciais da incidência de dengue, na maioria dos anos de estudo. De acordo com o método, em todo o período houve a presença de conglomerados espaciais de valores baixos ou nulos na região oeste e em quase todo o período na região leste. Detectou-se conglomerados espaciais de valores altos, principalmente nas regiões sudeste e norte do município.

Referências

- COSTA, M. A.; SCHRERRER, L. R.; ASSUNÇÃO, R. M. Detecção de conglomerados espaciais com geometria arbitrária. *Informática Pública*, v. 8, p. 23–34, 2006.
- COULSTON, J. W.; RITTERS, K. H. Geographic analysis of forest health indicators using spatial scan statistics. *Environmental Management*, v. 31, p. 764–773, 2003.
- DUARTE, H. H. P.; FRANÇA, E. B. Qualidade dos dados da vigilância epidemiológica da dengue em belo horizonte, mg. *Cadernos de Saúde Pública*, v. 40, p. 134–142, 2006.
- FERREIRA, I. T. R. N.; VERAS, M. A. S. M.; SILVA, R. A. Participação da população no controle da dengue: uma análise da sensibilidade dos planos de saúde de municípios do estado de são paulo, brasil. *Cadernos de Saúde Pública*, v. 26, 2009.
- LAWSON, A. et al. Comparison of general tests for spatial clustering. *Disease Mapping and Risk Assessment for Public Health*, v. 8, p. 111–117, 1999.
- LEE, G.; YAMADA, I.; ROGERSON, P. A. *GeoSurveillance: Software for monitoring change in geographic patterns*. State University os New York, Buffalo, 2007. Disponível em: <<http://www.acsu.buffalo.edu/~rogerson/geosurv.htm>>.
- LUCENA, S. E. F.; MORAES, R. M. Análise do desempenho dos métodos scan e besag e newell para identificação de conglomerados espaciais do dengue no município de joão pessoa entre os meses de janeiro de 2004 e dezembro de 2005. *Boletim de Ciências Geodésicas*, v. 15, p. 544–561, 2009.
- MARÇAL JÚNIOR, O.; SANTOS, A. Infestação por aedes aegypti (diptera: Culicidae) e incidência do dengue no espaço urbano: um estudo de caso. *Caminhos da Geografia*, v. 15, p. 241–251, 2004.
- MEDRONHO, R. A. *Epidemiologia*. 2. ed. [S.l.]: Ateneul, 2004.
- MONDINI, A. et al. Análise espacial da transmissão de dengue em cidade de porte médio do interior paulista. *Revista de Saúde Pública*, v. 39, p. 444–451, 2005.
- R Development Core Team. *R: A language and environment for statistical computing*. Vienna, Austria, 2008. ISBN 3-900051-07-0. Disponível em: <<http://www.R-project.org>>.
- ROGERSON, P. A. The detection of clusters using a spatial version of the chi-square goodness-of-fit statistic. *Geographical Analysis*, v. 31, 1999.
- ROGERSON, P. A. Spatial monitoring of geographic patterns: an application to crime analysis. *Computes, Environment and Urban Systems*, v. 25, p. 539–556, 2001.

- ROTHMAN, K. J.; GREENLAND, S. *Modern Epidemiology*. 2. ed. USA: [s.n.], 1998.
- SANCHES, O. Princípios básicos de procedimentos estatísticos aplicados na análise de dados de vigilância em saúde pública: uma revisão. *Cadernos de Saúde Pública*, v. 16, p. 317–333, 2000.
- SANKOH, O. A. et al. Clustering of childhood mortality in rural burkina faso. *International Journal of Epidemiology*, v. 30, p. 485–492, 2001.
- SIEGEL, S. *Estatística não paramétrica para ciências do comportamento*. São Paulo: McGraw Hill do Brasil, 1975.
- SOUZA, I.; MORAES, R. M. *Utilização de sistemas de informação geográfica na análise espacial de dados de saúde pública na Paraíba entre os anos de 1998 e 2001*. João Pessoa, Paraíba, 2003. Disponível em: <www.de.ufpb/ ronei>.
- SOUZA, I.; MORAES, R. M. *Análise espacial da relação entre pluviometria e incidência dos casos notificados de dengue da Paraíba em 2004*. João Pessoa, Paraíba, 2004. Disponível em: <www.de.ufpb/ ronei>.
- SOUZA, I. C. A.; VIANNA, R. P. T.; MORAES, R. M. Modelagem da incidência do dengue na paraíba por modelos de defasagem distribuída. *Cadernos de Saúde Pública*, v. 23, p. 2623–2630, 2007.
- SPIEGELMAN, M. *Introduction to Demography*. London: Havard University Press, 1968.
- TANGO, T. A class of tests for detecting ‘general’ and ‘focused’ clustering of rare diseases. *Statistics in Medicine*, v. 14, p. 2323–2334, 1995.
- TAUIL, P. L. Aspectos críticos do controle do dengue no brasil. *Cadernos de Saúde Pública*, v. 18, p. 867–871, 2002.
- TEIXEIRA, M. G. et al. Dengue and dengue hemorrhagic fever epidemics in brazil: what research is needed based on trends, surveillance, and control experiences. *Cadernos de Saúde Pública*, v. 21, p. 1307–1315, 2005.
- WHITTERMORE, A. S. et al. A test to detect clusters of disease. *Biometrika*, v. 74, p. 631–635, 1989.

Anexos

Histogramas e QQ-plot's

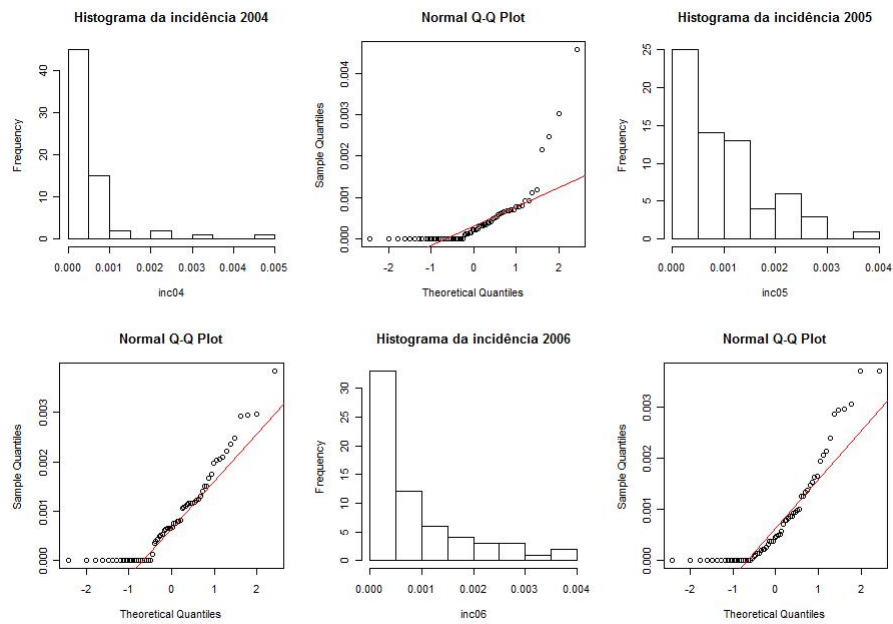


Figura 14: Histograma e qq-plot da incidência de dengue nos anos de 2004 a 2006

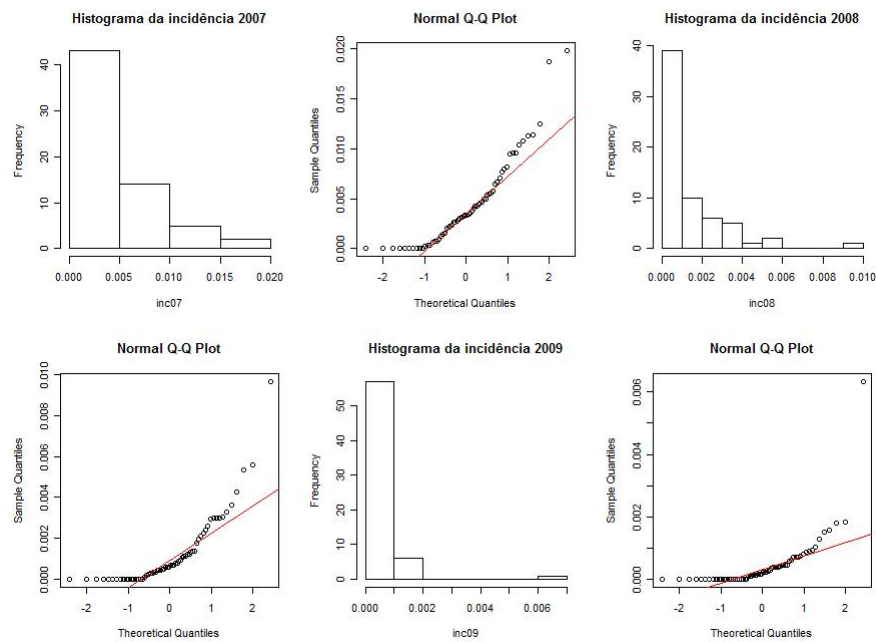


Figura 15: Histograma e qq-plot da incidência de dengue nos anos de 2007 a 2009

Comandos utilizados

• Mapas de incidência no *software* R

Exemplo para o ano de 2006:

```
library(DCluster)
library(shapefiles)

mapa= read.shape("C:/Ana Herminia/ferramentas geral/shape_jp_2007correto/jp_bairros_2_region.pol")
dados = read.csv2("C:/Ana Herminia/Monografia/tabela para mapas de incidencia.csv")
attach(dados)

#mapa cadastral
mapa = Map2poly(mapa, as.character(dados$Cod))

#2006
classes1 = c(-Inf,0.000530,0.001059,0.001589,0.002118,0.002648,0.003175,Inf)
legenda1 = c("000|---053","053|---106","106|---159","159|---212","212|---265","265|---317","317|---371")
var = inc06

cores = c("#FFFFCE", "#F1EA9E", "#EAC859", "#E6A925", "#B07B09", "#865B00", "#593D00")

plot(mapa, xlab="Longitude", ylab="Latitude", col=cores[findInterval(var, classes1, all.inside=TRUE)], forcefill=F)
legend(x="bottomright", legend=legenda1, fill=cores, bty="n")
```

• Indexação dos dados com o mapa no ArcGis

- Criar uma coluna no banco de dados do shape igual a coluna do banco de dados que se deseja indexar;
- Importar o shape no ArcGis;
- Clicar com o botão direito em "join and relates" e depois clicar em "join";
- Na janela que abrir após o passo anterior selecionar o banco de dados que se deseja indexar ao mapa e as colunas iguais no shape e na tabela;
- Clicar novamente em "join";
- Para salvar esta alteração deve criar um novo shape clicando com o botão direito nas propriedades do shape e clicando em "export data".