

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 2005¹

Performance review of scan and besag and newell methods to identify spatial clusters of dengue in João Pessoa city between January 2004 and December 2005

**SADRAQUE ENÉAS DE FIGUEIRÊDO LUCENA
RONEI MARCOS DE MORAES**

Universidade Federal da Paraíba – UFPB
Centro de Ciências Exatas e da Natureza – CCEN
Departamento de Estatística – DE
Laboratório de Estatística Aplicada ao Processamento de Imagens e
Geoprocessamento – LEAPIG
Cidade Universitária s/n
João Pessoa - PB - Brasil 58051-900
sadraquelucena@gmail.com, ronei@de.ufpb.br

RESUMO

A determinação do melhor índice de aglomeração espacial é primordial quando se trata de estudos epidemiológicos, pois a partir dos seus resultados é possível delinear políticas de saúde pública preventivas e eficazes no combate a doenças como o dengue. Neste artigo realizou-se uma análise do comportamento dos métodos Scan circular e o proposto por Besag e Newell utilizando dados de notificações de dengue em João Pessoa, PB. Para efeito de estudo, os dados mensais foram agregados por bairro para os anos de 2004 e 2005. Após aplicação, verificou-se que ambos os métodos detectaram conglomerados principalmente nas regiões norte e sudeste do município, embora a região norte incluísse um maior número de bairros durante quase todo período. Para comparar os métodos foi utilizada uma

¹ Uma versão prévia deste artigo foi apresentada no II Simpósio Brasileiro de Ciências Geodésicas e Tecnologias da Informação (II SIMGEO) – setembro/2008 - Recife-PE –Brasil.

modificação particular no método de Besag e Newell. O método Scan detectou conglomerados de alto e baixo risco, enquanto que com o método modificado de Besag e Newell foram identificados apenas conglomerados de alto risco, porém este último o fez com maior precisão.

Palavras-chave: dengue, identificação de conglomerados espaciais, método Scan, método de Besag e Newell.

ABSTRACT

The identification of the best spatial cluster methodology in epidemiological studies because from those results is possible to define prevention and effective policies in public health to combat diseases, as dengue. In this paper it was made a comparative analysis of circular Scan method and that one proposed by Besag and Newell using data of dengue notifications from João Pessoa, Brazil. To perform this study, monthly data was aggregated by neighborhood for the years 2004 and 2005. After application, it was verified for both methods detected clusters mainly in north and southeast regions of the city, but north region included the biggest number of neighborhoods during all period. To compare both methods it was used a particular modification on Besag and Newell's method. Scan method has detected high and low risk clusters, but modified Besag and Newell's method has identified high risk clusters only and performed that with more accuracy.

Keywords: dengue, spatial cluster identification, Scan method, Besag and Newell's method.

1. INTRODUÇÃO

Os Sistemas de Informação Geográfica (SIG) têm sido amplamente usados na análise da distribuição espacial de doenças (HIGGS e GOULD, 2001). Em saúde pública, a busca pela determinação de regiões cuja taxa de ocorrência de agravos é acima do esperado torna-se importante para uma intervenção eficaz (ANDRADE *et al.*, 2004). Nestes casos é de interesse dos pesquisadores identificarem dentre as áreas que apresentam risco aproximadamente constante, aquelas cujo risco observado é distinto das demais (COULSTON e RITTERS, 2003). Para tanto, verificar se a distribuição espacial de um evento adverso à saúde é aleatória ou não se torna primordial para prevenção de epidemias (SANKOH *et al.*, 2001).

A escolha do melhor índice de aglomeração espacial, aliado ao estudo descritivo baseado em mapas cadastrais, visa uma combinação de técnicas que levará a criação de um modelo mais robusto de monitoramento dos casos, sugerindo áreas preferenciais de combate a doenças (SANCHES, 2000). Com esses conhecimentos em mãos e medidas eficazes de combate, tem-se por consequência uma imediata melhoria das condições de saúde da população (MARÇAL JUNIOR e SANTOS, 2004).

O dengue *Aedes aegypti* (Diptera: Culicidae) emergiu como um problema de saúde mundial, sendo evidenciado pela série de epidemias nas regiões tropicais e

subtropicais do mundo (ENDY *et al.*, 2002). No estado da Paraíba verifica-se um padrão epidemiológico endêmico, caracterizado por meses de alta e baixa endemicidade cuja maior frequência ocorre durante os meses de abril a junho (SOUZA, 2007). No município de João Pessoa a incidência é mais elevada no primeiro semestre. 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, 2007).baixos

Dada a deficiência nos instrumentos de monitoramento de regiões de alta incidência, a combinação de estudos descritivos e de modelagem por geoprocessamento possibilitam uma política de saúde pública preventiva e eficaz no combate a doenças como o dengue (MONDINI *et al.*, 2005). No caso deste trabalho, o interesse recai sobre a possibilidade de aglomerações espaciais da incidência do dengue, o que pode indicar presença de zonas com necessidade de intervenção (KREMPI, 2004). Tendo em vista o grande número de técnicas existentes, a definição do melhor método é essencial para obtenção de resultados concretos e objetivos, propiciando um monitoramento eficaz e conciso.

Neste artigo é realizada uma análise do comportamento dos métodos Scan circular proposto por Kulldorff e Nagarwalla (1995) e o método apresentado por Besag e Newell (1991) utilizando dados de notificações de dengue no município de João Pessoa, PB. Para efeito de estudo serão analisados os casos de dengue coletados mensalmente durante o período de janeiro de 2004 a dezembro de 2005. Assim, o objetivo deste trabalho consiste em comparar as metodologias abordadas e, além disso, analisar o desempenho de ambos a partir do ponto de vista epidemiológico, visto que os métodos se propõem a identificar áreas de risco elevado na ocorrência de um evento. Desse modo, serão confrontados os métodos ao passo em que se realizará um estudo epidemiológico do tipo ecológico (MORGENSTERN, 1982) (MEDRONHO, 2003), avaliando o comportamento dos casos de dengue pelo município. Os dados foram agregados por bairro de moradia do indivíduo a partir das bases de dados disponíveis pelo SINAN e IBGE, complementadas pelas informações disponíveis na Secretaria Municipal de Saúde de João Pessoa.

2. METODOLOGIA

Define-se um conglomerado como um conjunto de áreas que apresentam um risco significativamente alto em relação a um determinado evento, embora não seja esperada tal observação nestas áreas (KNOX, 1988). Tais conglomerados podem ser caracterizados segundo três classificações: espacial, temporal e espaço-temporal. A primeira classificação se trata da observação de valores elevados em determinada(s) área(s) sem a interferência do período de tempo considerado. O segundo diz respeito à alta ocorrência de casos em um intervalo de tempo, ignorando-se, de certa forma, a distribuição espacial (COSTA *et al.*, 2006). Por fim, a última classificação é caracterizada pela busca de aglomerações não somente no

espaço, mas também no tempo (LIAN, 2007). Para efeito deste estudo será considerada apenas a detecção de conglomerados espaciais.

No que diz respeito às técnicas de detecção de conglomerados, estas são caracterizadas em dois grupos: os testes focados de conglomerados e os testes genéricos de conglomerados (KNOX, 1988). O primeiro grupo refere-se aos métodos nos quais se procura evidências de que há um risco elevado ao redor de uma fonte suspeita sob a ocorrência de um evento. O segundo grupo busca a existência de que as regiões não identificadas possuem um risco significativamente maior que o valor observado em toda área de estudo. Estes tipos de abordagens são adequados a dados de contagem agregados por área, de modo que são conceitos associados a alguns métodos de análise espacial (BEATO FILHO *et al.*, 2001). Neste artigo são abordados os testes genéricos de conglomerados Scan circular e o método proposto por Besag e Newell.

Na aplicação dos testes, considera-se a hipótese nula de homogeneidade do padrão espacial da variável em estudo, ou seja, há ausência de conglomerados. Com isto, é assumido que a incidência sobre uma variável é constante em todas as áreas, o que implica em um número de casos esperado proporcional ao número de residentes em uma dada área (COSTA e ASSUNÇÃO, 2005). Em termos matemáticos, temos a definição a seguir.

Definição 1. Seja uma região de interesse particionada em n sub-regiões indexadas por $i = 1, \dots, n$, na qual se deseja estudar um determinado evento. Denota-se por c_i a variável aleatória que representa o número de casos observados em cada sub-região e E_i seus respectivos números de casos esperados. A hipótese nula H_0 de não existência de conglomerados é dada por (KNOX, 1988):

$$H_0: c_i \sim \text{Poisson}(E_i = \lambda P_i) \quad (1)$$

onde λ é o número de observações sobre toda região de interesse e P_i corresponde ao tamanho da população na sub-região i . Para especificar as posições geográficas das sub-regiões, podem ser usadas as respectivas coordenadas dos centróides (TANGO, 2005). Uma estimativa para λ é obtida por:

$$\hat{\lambda} = \frac{\sum_{i=1}^n c_i}{\sum_{i=1}^n P_i} \quad (2)$$

2.1 Método de Varredura Scan

O objetivo da estatística Scan, proposto por Kulldorff e Nagarwalla (1995) é identificar conglomerados em unidades cuja ocorrência de um evento é significativamente mais provável dentro de uma área do que fora dela. Para tanto, o método trabalha com o conjunto de todos os possíveis candidatos a conglomerados, sobre os quais são sobrepostos os círculos de raio arbitrário centrados nos

centróides de cada sub-região de estudo (COULSTON e RITTERS, 2003). Para cada círculo há o interesse na probabilidade de haver casos dentro (p) e for a (q) dos círculos (GÓMEZ-RUBIO *et al.*, 2005). Dois modelos probabilísticos podem ser empregados na utilização deste método: o modelo Binomial e o Poisson (KULLDORFF e NAGARWALLA, 1995). Tendo em vista que para a análise deste trabalho foi utilizado apenas o modelo de Poisson, serão apresentadas apenas as formulações para este.

Para o modelo de Poisson é computado o raio cujos valores de $p(z)$ e $q(z)$ maximizam a função de verossimilhança condicionada ao total de casos observados. Admitindo-se z como um candidato a conglomerado, considera-se a estatística (GÓMEZ-RUBIO *et al.*, 2005)

$$KN = \max_{z \in Z} \frac{L(z, \hat{p}(z), \hat{q}(z))}{L_0} \quad (3)$$

onde Z é o conjunto de todos os possíveis candidatos a conglomerados, $\hat{p}(z)$ e $\hat{q}(z)$ são as estimativas de $p(z)$ e $q(z)$. L_0 é definido como:

$$L_0 = \frac{C^C (M - C)^{M-C}}{M^M} \quad (4)$$

onde C é o total de casos observados na região considerada e M a população total sob toda região de estudo. $L(z)$ é definido como:

$$L(z) = \frac{\exp[-p.n_z - q(M - n_z)]}{C!} p^{C_z} q^{C-c_z} \prod_i c_i \quad (5)$$

onde \exp representa a função exponencial; c_z e c_i são, respectivamente o número de casos no círculo z e no círculo i e n_z é o número de indivíduos em risco no círculo z .

Tendo em vista as formulações apresentadas, o método Scan trabalha como descrito a seguir. Para cada centróide da região de estudo é calculada sua distância em relação a todos os outros centróides. Em seguida é criado um círculo posicionado sobre o centróide cujo raio é aumentado continuamente, de modo a envolver um novo centróide. Para cada ponto inserido no círculo z , calculam-se o número de casos c_z e a população n_z dentro do círculo. Obtém-se o valor de KN para cada par (c_z, n_z) , dado pela equação (3). Ao final é registrado o maior valor de KN e calculadas as simulações de Monte Carlo para avaliar a significância do teste (MOURA, 2006).

Na simulação de Monte Carlo são gerados diversos conjuntos de dados independentes como mesmo número de casos que o conjunto original e são distribuídos ao acaso entre as sub-regiões de acordo com a hipótese nula. Para cada conjunto calcula-se a estatística do teste KN e ordenam-se os valores. Se a

estatística obtida para o conjunto original estiver entre os $100\alpha\%$ valores mais altos das estatísticas dos conjuntos de dados independentes, rejeita-se a hipótese nula ao nível de significância α (MOURA, 2006).

2.2 Método Besag e Newell

Assim como o método Scan apresentado, Besag e Newell propuseram uma saída visual de conglomerados verossímeis de formato circular que tomam por base círculos centrados sobre os centróides de cada sub-região (OPENSHAW et al., 1988). Denotado por BN , o método determina o raio necessário para que o círculo contenha pelo menos k de casos em seu interior. O procedimento é iniciado com um círculo de raio igual a zero. Se este contém k ou mais casos, o processo é interrompido; caso contrário, o raio é aumentado até incluir o centróide mais próximo. Logo, são adicionados os respectivos casos e a população exposta ao risco. Deste modo, os raios são definidos de tal forma a incluir novos centróides quando necessário, sendo executado este procedimento até que haja pelo menos k casos inseridos no círculo (COSTA e ASSUNÇÃO, 2005).

Seja C o número total de casos em toda região de estudo e Y a população total exposta ao risco na região. Sejam ainda $C_{j(i)}$ e $M_{j(i)}$ o número de casos e a população acumulada das j áreas mais próximas ao centróide i . A estatística do teste baseia-se na variável aleatória L , definida como o mínimo de áreas próximas necessárias para que se obtenha um número k de casos mais próximos do centróide i . Formalmente, temos (COSTA e ASSUNÇÃO, 2005):

$$L = \min \{j : C_{j(i)} \geq k\} \quad (6)$$

Para cada centróide é necessário verificar a existência de aglomeração espacial. Portanto, a partir do valor l observado para L , o nível de significância do teste é definido por $P(L \leq l)$, que busca testar a hipótese nula (ausência de aglomeração espacial). A significância, denotada por $p_k(i)$, é calculada através da seguinte expressão (COSTA e ASSUNÇÃO, 2005):

$$p_k(i) = P(L_i \leq l_i) = 1 - \sum_{j=1}^{k-1} \frac{(M_{j(i)}C/M)^j}{j!} \exp(-M_{j(i)}C/M) \quad (7)$$

onde M_j é a população observada em j áreas.

Um aglomerado é dito significativo se o p-valor obtido é menor que a significância adotada. Assim, após a obtenção de todos os círculos que englobam um número k de casos por meio do método de Besag e Newell, são desenhados no mapa apenas os círculos significativos (p-valor $< \alpha$). Geralmente a escolha de α é feita para permitir os muitos testes simultâneos e são considerados valores menores que a significância usual 0,05 ou 0,01 (COSTA e ASSUNÇÃO, 2005).

Tendo em vista que o presente artigo tem como objetivo comparar os métodos descritos, COSTA e ASSUNÇÃO (2005) sugeriram uma modificação no método de Besag e Newell, de modo a permitir uma comparação adequada em relação ao método Scan. A justificativa para esta modificação é devido ao método de Besag e Newell precisar da determinação do número de casos e do nível de significância para busca de conglomerados, enquanto o método Scan não tem esta dependência. Isso ocorre por que o método Scan busca todos os candidatos a conglomerado e calcula a estatística k e o nível de significância por simulação de Monte Carlo sob a hipótese nula de aleatoriedade, ao invés de determiná-los *a priori*. Além disso, a sobreposição de círculos do método de Besag e Newell pode gerar conglomerados não circulares. Assim a modificação – que foi denominada de CABN (COSTA e ASSUNÇÃO, 2005) – consiste em eliminar a dependência de k e da significância no método de Besag e Newell, utilizando a seguinte estatística:

$$T = \min_k \{ \min_{z_k \in Z_k} [p_k(i)] \} \quad (8)$$

onde Z_k é o conjunto de todas as regiões circulares z_k centradas nos centróides e contendo k casos próximos e $p_k(i)$ é dado pela equação 2. Neste artigo a variação do parâmetro k foi entre 2 casos e o máximo observado em um único bairro naquele mês.

Para desenvolvimento do estudo foi utilizado o *software* R (R DEVELOPMENT CORE TEAM, 2007). A aplicação dos métodos se deu pelo uso do módulo *DCluster*, que é composto por um conjunto de funções para detecção de aglomerações espaciais.

3. RESULTADOS E DISCUSSÃO

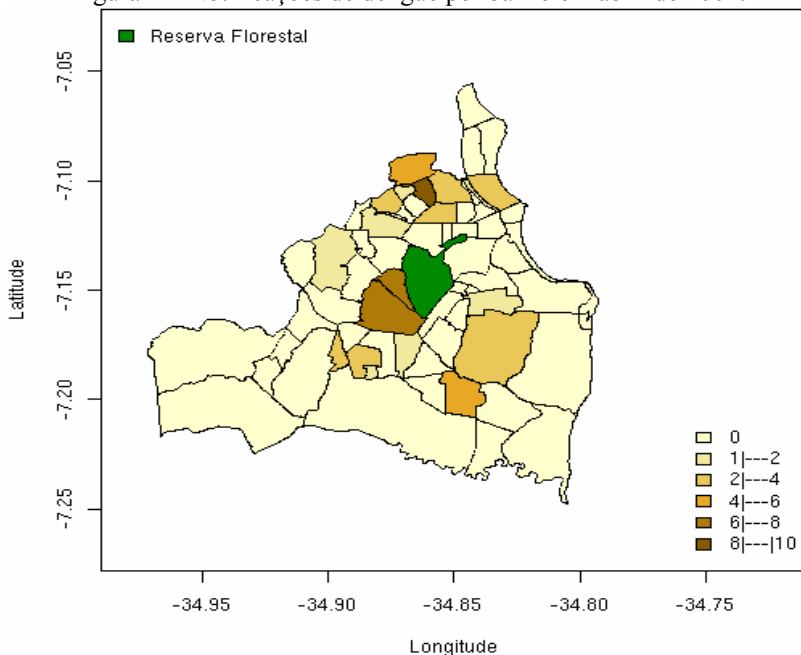
A aplicação de ambos os métodos se deu sobre as notificações de dengue por bairro considerando cada um dos meses dos anos de 2004 e 2005. Para comparação entre os métodos Besag e Newell e varredura Scan foi utilizada a distribuição espacial dos casos de dengue no município e esta Seção fornece uma breve descrição dos mapas cadastrais obtidos, bem como de cada método em separado. Em seguida é realizada uma análise comparativa entre os conglomerados significativos obtidos por ambas as técnicas. Os conglomerados identificados são apresentados por mapas nos quais os pontos em vermelho representam os centróides significativos. Cada círculo em vermelho representa o raio necessário para identificação de um conglomerado, de tal forma que dentro de cada um se pode verificar o número de centróides necessários para a existência de conglomeração. A significância adotada foi $\alpha = 0,05$, para possibilitar a comparação entre eles.

3.1 Análise dos mapas cadastrais

Para os meses de janeiro a junho do ano de 2004 foram observadas as maiores frequências de notificações de casos de dengue por bairro. Ao longo dos cinco

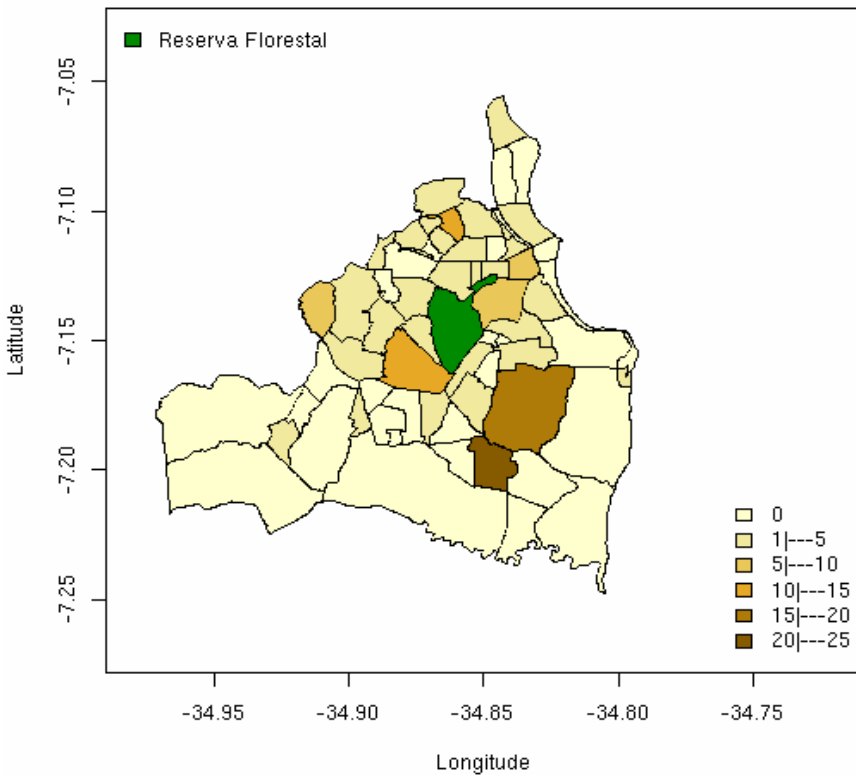
primeiros meses desse ano, foi verificado que a região norte apresentou sempre os bairros com os maiores números de notificações de dengue. Dois bairros do oeste próximos à Reserva Florestal, assim como do sudeste, apresentaram valores relativamente mais elevados que os demais. Nos meses de junho, julho e agosto um bairro do sudeste (Mangabeira) apresentou maior número de registros. Nos demais meses houve uma gradativa queda na frequência de notificações. Assim, foram identificados valores distintos dos demais na região norte, oeste próximo à Reserva Florestal e sudeste do município, como mostrado na Figura 1 para o mês de abril.

Figura 1 – Notificações de dengue por bairro em abril de 2004.



É perceptível a presença de frequência mais elevada entre os meses de maio e agosto do ano de 2005. No decorrer dos meses foram verificados valores elevados em relação aos demais bairros no norte do município. A região sudeste também apresentou o mesmo comportamento de frequências elevadas, como apresentado na Figura 2 para o mês de julho. No oeste os maiores valores observados foram em bairros próximos à Reserva Florestal, assim como no ano anterior. O que se observa ainda neste ano, é que a quantidade de bairros com número de notificações diferente de zero foi maior que no ano anterior.

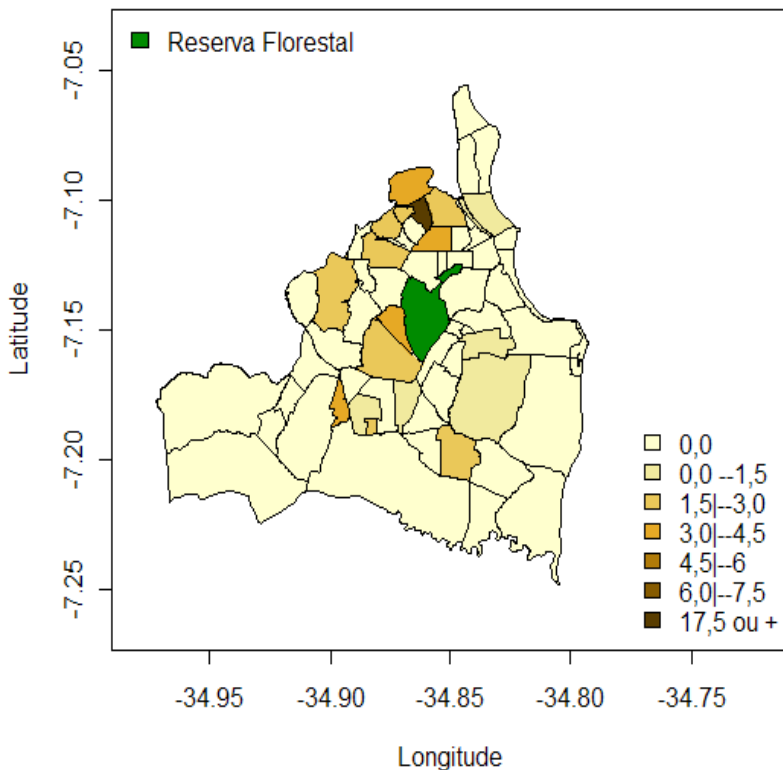
Figura 2 – Notificações de dengue por bairro em julho de 2005.



3.2 Análise dos mapas de risco

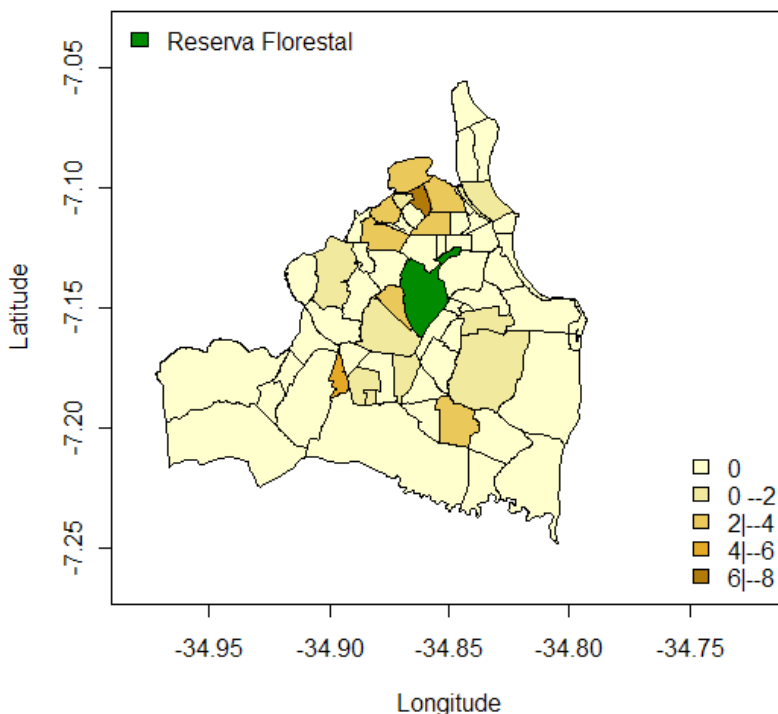
Foram obtidos os mapas de risco para os bairros do município de modo a permitir a visualização das áreas cujo risco de notificação de dengue é maior que os demais bairros. Entre os meses de janeiro e julho de 2004 os maiores riscos observados ocorreram na região norte do município, cujos riscos eram no mínimo dez vezes superiores ao risco total da população. A Figura 3 apresenta o mapa de riscos no mês de abril de 2004, confirmando a afirmação apresentada. Diferentemente do padrão observado nos outros meses, em agosto o maior risco observado foi em um bairro da região sul e em outubro o maior risco foi na região oeste do município. Nos demais meses os maiores riscos ocorreram na região norte.

Figura 3 – Mapa de riscos para abril de 2004.



No decorrer dos seis primeiros meses de 2005 os riscos mais elevados predominaram na região norte do município, havendo riscos um pouco menores, embora ainda elevados, na região sudeste (bairro Valentina). Como se pode ver na Figura 4, que representa o fenômeno ocorrido mês de junho de 2005, os maiores riscos ocorreram no norte do município e no sudeste. Em julho e agosto houve riscos elevados no norte, sudeste e oeste. Em setembro os riscos mais elevados estavam no norte e nos meses posteriores ocorreram também no oeste e sudeste do município.

Figura 4 – Mapa de riscos para junho de 2005.



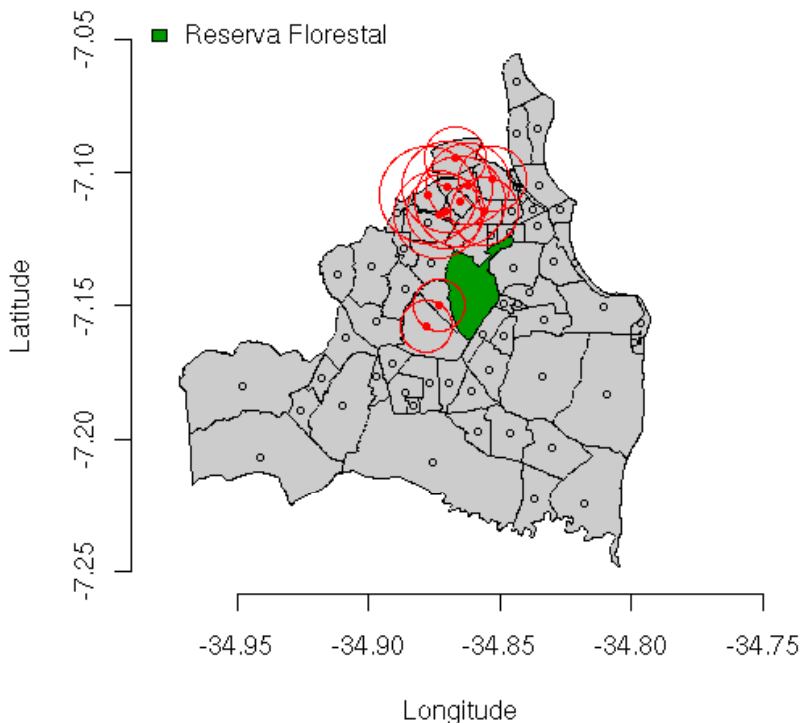
3.3 Método de varredura Scan

Para o método Scan foi admitida a significância $\alpha = 0,05$ e definido que os candidatos a conglomerados não poderiam superar 10% da população do município. A escolha deste percentual se deu mediante a comparação entre o mapa de incidência e o mapa gerado pelo método Scan, de modo que os conglomerados identificados não fossem maiores que o conglomerado real.

A partir da aplicação do método Scan para os meses do ano de 2004 foi verificado que ao longo de todo o período houve a presença de persistentes conglomerados na região norte. Os bairros que estiveram presentes em grande parte dos meses foram Alto do Céu, Bairro dos Ipês, Bairro dos Estados, Mandacaru, Treze de Maio, Padre Zé, Roger, Jardim das Acácias e Tambiá. Na porção oeste do município foi verificada a presença de conglomerados apenas no mês de abril, formados por dois bairros próximos à Reserva Florestal (Cristo e Varjão). No sudeste foram observados conglomerados de bairros próximos apenas em junho e julho. Houve também um conjunto de conglomerados de bairros vizinhos das regiões sudeste e sul no mês de março. Nos demais meses os conglomerados

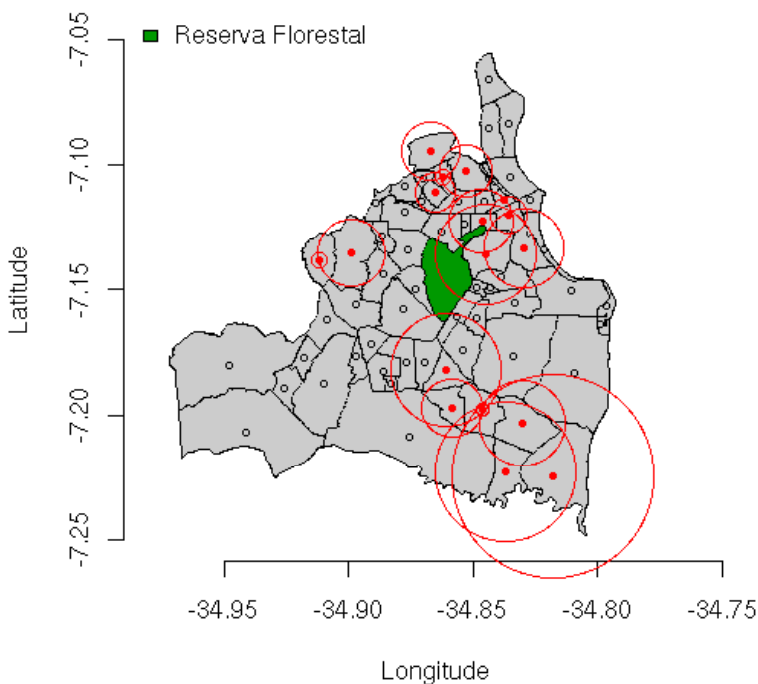
identificados estavam distribuídos esparsamente pelas regiões. A Figura 3 apresenta o comportamento do método Scan para o mês de junho de 2004.

Figura 5 – Método Scan para o número de notificações de dengue em abril de 2004.



Assim como em 2004, no ano de 2005 houve a presença de conglomerados na região norte em todos os meses, entretanto não foi observada preferência sobre determinados bairros. Os conglomerados do mês posterior não foram formados sempre por muitos dos bairros significativos no mês prévio. Diferentemente do observado em 2004, houve conglomerações no sudeste em muitos dos meses de 2005. Como pode ser visto na Figura 4, nesta região os bairros que apresentaram conglomerados sempre foram Cuiá, Planalto Boa Esperança, Valentina, Paratibe, Muçumagro e Barra de Gramame. Na região oeste do município houve conglomerados em três meses do ano, sendo os bairros identificados próximos à fronteira do município ou à Reserva Florestal.

Figura 6 – Método Scan para o número de notificações de dengue em julho de 2005.



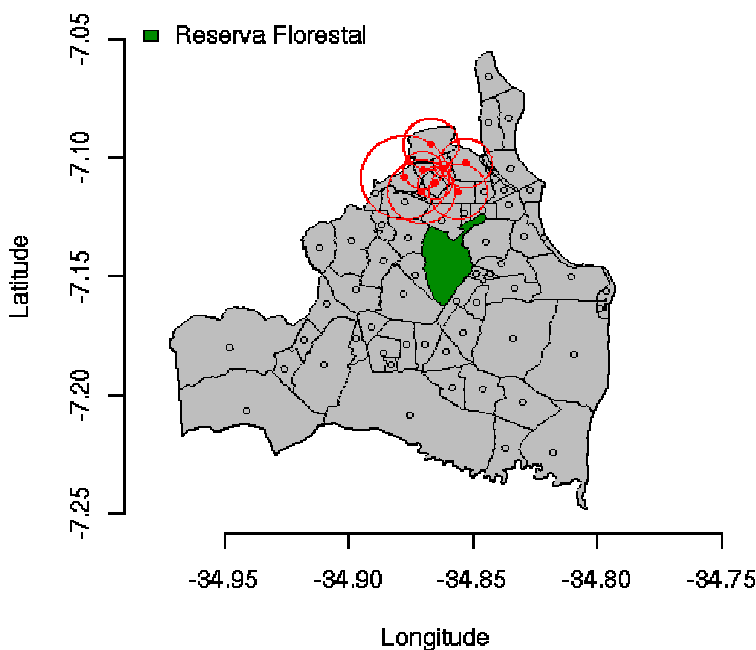
3.4 Método Besag e Newell modificado

Basicamente, a evolução mensal dos anos estudados apresentou padrões distintos no tocante aos conglomerados espaciais observados. No primeiro semestre de 2004 o método identificou uma predominância de conglomerados na região norte do município, enquanto que em 2005 os conglomerados não persistiram por muitos meses nos mesmos bairros. Ainda sobre 2005, o método observou uma alternância na presença de conglomerados entre as regiões norte e sudeste ao longo do ano.

No ano de 2004 foram identificados conglomerados persistentes na região norte ao longo de todo primeiro semestre. A Figura 5 mostra essa persistência no mês de abril de 2004, quando da aplicação do CABN. Ainda levando em consideração os primeiros seis meses do ano, apenas no mês de março houve presença de conglomerados em outra região além do norte do município: esta ocorrência foi na região sul, envolvendo três bairros apenas. No mês de julho foi verificada a presença de conglomerados apenas na região sudeste e nos dois meses seguintes foram detectados conglomerados novamente apenas na região norte. No

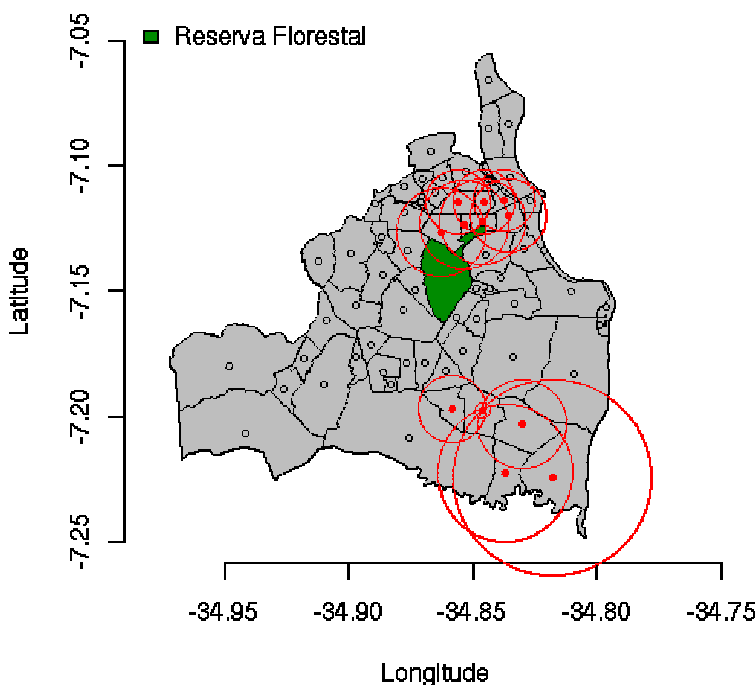
último trimestre do ano não foi identificada a presença de conglomerados de notificações de dengue no município.

Figura 7 – Método CABN para o número de notificações de dengue em abril de 2004.



Em 2005 os conglomerados identificados nas regiões não persistiram por muitos meses em uma mesma região, diferentemente do ocorrido no ano anterior. No mês de janeiro não foram identificados conglomerados significativos no município, ocorrendo apenas na região sudeste no mês seguinte. Em março os mesmos conglomerados verificados na região sudeste foram observados e além disso houve também em bairros ao norte do município. Entre abril e junho continuaram sendo identificados conglomerados na região sudeste e neste último mês houve também na região sudeste. A partir do mês de julho, no qual se observaram conglomerados apenas no sudeste, como se pode ver na Figura 6, nos meses que se seguiram os conglomerados foram se alternando entre as regiões norte e sudeste no decorrer dos meses.

Figura 8 – Método CABN para o número de notificações de dengue em julho de 2005.



3.5 Comparação entre os métodos

Basicamente ambas as metodologias indicaram presença de conglomerados nas regiões norte e sudeste do município. De modo geral, para cada mês o método CABN tendeu a identificar conglomerados apenas em uma única região, ao contrário do método de varredura Scan. Além disso, houve também uma diferença interessante: CABN não indicou presença de conglomerados no último trimestre de 2004 e em janeiro de 2005. Isto pode ser devido ao pequeno número de registros de notificações de dengue nesse período.

O confronto entre os mapas cadastrais e os mapas gerados pelos métodos Scan e CABN mostrou resultados diferentes em se tratando das regiões. Para efeito comparativo, foram analisados os conglomerados identificados por ambas as metodologias e o esperado com base nos mapas de risco. O motivo desta abordagem

diz respeito ao ponto de vista epidemiológico, o qual espera que a ocorrência de conglomerados de uma doença não seja muito distinta da região em que ela foi notificada.

Com base no parágrafo anterior, o método de Besag e Newell tendeu a identificar conglomerados apenas na região em que se observou o maior número de notificações, dificilmente apresentando mais de uma área ao mesmo tempo. Os conglomerados identificados pelo CABN correspondiam aos bairros com risco mais elevado de ocorrência de dengue, enquanto o Scan apontou conglomerados de altos e de baixos valores. Esse pode ser o motivo pelo qual o método Scan identificou conglomerados em regiões distintas do mapa para um mesmo mês e o CABN não.

Na ocorrência de poucos bairros com alto risco, o CABN identificou conglomerados envolvendo menos bairros. Em contra partida, o método Scan tendeu a apresentar conglomerados em outras áreas que também tinham riscos mais baixos, chegando a envolver também outros bairros um pouco mais distantes, por vezes com riscos menores.

Na região sudeste do município, o método CABN quando não apresentou o mesmo conglomerado que o Scan, em todos os demais casos envolveu um número menor de bairros nos conglomerados. Dado que na região sudeste há bairros muito populosos vizinhos de bairros menos populosos, a restrição do percentual de população utilizado no método Scan tenha causado esse leve aumento no número de bairros dentro dos conglomerados. Por outro lado, nas demais regiões, o CABN não identificou conglomerados em áreas com apenas um bairro de alto risco isolado e com vizinhos sem qualquer registro de notificações.

Do ponto de vista epidemiológico, a identificação de conglomerados de bairros cujo risco de ocorrência de uma doença é alto, tem prioridade maior frente a conglomerados de baixos valores. Sob este ponto de vista, neste estudo, o método CABN revelou vantagens sobre o Scan, pois identificou esses conglomerados e incorporou aos mesmos menos bairros com baixo risco.

4. CONCLUSÕES

Foi realizado um estudo mensal do comportamento do dengue no município de João Pessoa, PB, entre janeiro de 2004 e dezembro 2005 com base nos registros de notificação e no risco de ocorrência da doença. Além disso, foi feito um estudo comparativo entre os métodos de varredura Scan e o método proposto por Besag e Newell, levando em consideração uma modificação particular, buscando avaliar seus desempenhos em relação à endemia. A comparação realizada não foi apenas sobre as metodologias, mas também a partir do ponto de vista epidemiológico.

A utilização de ambos os métodos levou à determinação de conglomerados na região norte e sudeste do município. Observou-se ainda que os conglomerados na região norte foram mais abrangentes no tocante ao número de bairros identificados.

Ao serem confrontados os métodos, verificou-se que o CABN tendeu a identificar conglomerados apenas em uma região do município na maioria dos

meses. Os conglomerados identificados por esse método eram formados por bairros com riscos elevados. Por outro lado o Scan identificou conglomerados de alto risco (em geral na região norte do município) e de baixo risco (em geral na região sudeste). Entretanto, nas regiões em que ambos os métodos identificaram conglomerados, o CABN continha no máximo a mesma quantidade de bairros que o Scan, embora geralmente o número menor que com o Scan. Este último, por vezes incorporou no conglomerado, bairros com baixo risco nos conglomerados de alto risco e vice-versa. Do ponto de vista epidemiológico, neste estudo, o método CABN revelou-se superior ao Scan, pois identificou conglomerados de bairros cujo risco de ocorrência de uma doença são altos com maior precisão.

5. AGRADECIMENTOS

Este trabalho tem suporte do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), por meio de bolsas de pesquisa PIBIC.

REFERÊNCIAS BIBLIOGRÁFICAS

- ANDRADE, A.L.S.S. *et al.* Population-based surveillance of pediatric pneumonia: use of spatial analysis in an urban area of Central Brazil. *Cad. Saúde Pública*, v. 20, n. 2, p. 411-421, 2004.
- BEATO FILHO, C.C.; ASSUNÇÃO, R.; SILVA, B.F.S.; MARINHO, F.C.; REIS, I.A.; ALMEIDA, M.C.M. Conglomerados de homicídios e o tráfico de drogas em Belo Horizonte, Minas Gerais, Brasil, de 1995 a 1999. *Cad. Saúde Pública*, Rio de Janeiro, v. 17, n. 5, p. 1163-1171, 2001.
- BESAG, J.; NEWELL, J. The detection of clusters in rare diseases. *Journal of the Royal Statistical Society*, v. 154, p. 143-155, 1991.
- COSTA, M.A.; SCHERRER, L.R.; ASSUNÇÃO, R.M. Detecção de Conglomerados Espaciais com Geometria Arbitrária. *Informática Pública*, v. 8, n. 1, p. 23-34, 2006.
- COSTA, M.A.; ASSUNÇÃO, R.M. A fair comparison between the spatial scan and the Besag-Newell disease clustering tests. *Environmental and Ecological Statistics*, Springer, v.12, p. 301-319, 2005.
- COULSTON, J.W.; RITTERS, K.H. Geographic analysis of forest health indicators using spatial scan statistics. *Environmental Management*, v. 31, n. 6, p. 764-773, 2003.
- ENDY, T.P. *et al.* Spatial and temporal circulation of dengue virus serotypes: a prospective study of primary school children in Kamphaeng Phet, Thailand. *American Journal of Epidemiology*, v. 156, p. 52-59, 2002.
- GÓMEZ-RUBIO, V.; FERRÁNDIZ-FERRAGUD, J.; LÓPEZ, A. Detecting clusters of disease with R. *Journal of Geographical Systems*, v. 7, n. 2, p. 189-206, 2005.
- HIGGS, G.; GOULD, M. Is there a role for GIS in the 'new NHS'? *Health Place*, v. 7, n. 3, p. 247-259, 2001.

- KREMPI, A.P. Explorando recursos de estatística espacial para análise da acessibilidade da cidade de Bauru. 2004. 94 f. Dissertação (Mestrado em Engenharia de Transportes) – Universidade de São Paulo, São Carlos.
- KNOX, E.G. Detection of clusters. In: Methodology of enquiries into disease clustering. ELLIOTT, P. (editor). London, Small Area Health Statistics Unit, 1989.
- KULLDORFF, M.; NAGARWALLA, N. Spatial disease clusters: detection and inference. *Statistics in Medicine*, v. 14, p. 799-810, 1995.
- LIAN, M.; WARNER, R.D.; ALEXANDER, J.L.; DIXON, K.R. Using geographic information systems and spatial and space-time scan statistics for a population based risk analysis of the 2002 equine West Nile epidemic in six contiguous regions of Texas. *International Journal of Health Geographics*, v. 6, n. 7, set. 2007.
- MARÇAL JUNIOR, 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. Estudos Ecológicos. In: Medronho, R.A. et al. Epidemiologia. São Paulo: Editora Ateneu, 2003, p 191-198.
- MONDINI, A. ; CHIARAVALLOTI NETO, F. ; SANCHES, M. G. Y. ; LOPES, J. C. C. . Análise espacial da transmissão de dengue em cidade de porte médio do interior paulista. *Revista de Saúde Pública*, v. 39, n. 3, p. 444-451, 2005.
- MORGENSTERN, H.; Uses of ecological Analysis in Epidemiologic Research. *Am. J. Public Health*, v. 72, n. 12, p.1336-1344, 1982.
- MOURA, F.R. Detecção de clusters espaciais via algoritmo scan multi-objetivo. 2006. Dissertação (Mestrado em Estatística) – Universidade Federal de Minas Gerais, Belo Horizonte.
- OPENSHAW, S.; CRAFT, A.W.; CHARLTON, M.; BIRCH, J.M. Investigation of leukaemia clusters by use of a geographical machine. *Lancet*, p. 272-273, 1988.
- R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Disponível em: <http://www.R-project.org/>. Acesso: 5 abril 2007.
- SANCHES, O. Princípios básicos de procedimentos estatísticos aplicados na análise de dados de vigilância em saúde pública: uma revisão. *Cad. Saúde Pública*, v. 16, n. 2, p. 317-333, 2000.
- SANKOH, O.A.; YÉ, Y.; SAUERBORN, R.; MÜLLER, O.; BECHER, H. Clustering of childhood mortality in rural Burkina Faso. *International Journal of Epidemiology*, v. 30, p. 485-492, 2001.
- 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, n.11, p. 2623-2630, 2007.

TANGO, T; TAKAHASHI, K. A flexibly shaped spatial scan statistic for detecting clusters. *International Journal of Health Geographics*, v. 4, n. 11, 2005.

(Recebido em abril de 2009. Aceito em julho de 2009)