

## **Análise de conglomerados de homicídios por arma branca e arma de fogo na cidade de João Pessoa entre 2002 e 2005.**

**Daniel Matos de Carvalho, Ronei Marcos de Moraes**

Depto de Estatística, UFPB.

58.051-900, João Pessoa, PB

danielmatosc@hotmail.com, ronei@de.ufpb.br

**Resumo:** *Sabe-se que a violência é um sério problema de saúde pública, afetando a qualidade de vida de todos. A análise espacial é uma ferramenta poderosa que pode auxiliar as políticas de segurança e saúde pública na melhoria da qualidade de vida das pessoas de nossa sociedade. Este trabalho tem como objetivo descrever o comportamento dos homicídios em João Pessoa e realizar uma análise comparativa de algumas das técnicas de detecção de conglomerados e dependência espacial. Dentre as várias técnicas existentes, foram utilizados: Índice de autocorrelação espacial Getis e Ord, Método Besag e Newell, Método de Varredura Scan proposto por Kulldorff e GAM-K. Os conglomerados de homicídios encontram-se localizados na região Noroeste, com os bairros do Centro, Jaguaribe, Trincheiras e Tambaí sendo estes os principais focos de homicídios. Outros conglomerados foram identificados de forma isolada, o principal localiza-se na região Sudoeste de João Pessoa com o foco no Bairro das Indústrias. De forma geral, pode-se dizer que os métodos apresentaram um certa semelhança no padrão da localização espacial de conglomerados de homicídios. Os métodos identificaram a região Noroeste como a principal área de ocorrência de homicídios, ocorrendo pequenas variações na localização e tamanho dos conglomerados.*

**Palavras-chave:** Dependência espacial, Autocorrelação Espacial, Conglomerados, Métodos Scan, Homicídios

### **Introdução**

A análise espacial é o estudo quantitativo de fenômenos que são localizados no espaço, permitindo a identificação de padrões espaciais na distribuição dos mesmos [21]. Compreender a distribuição espacial de dados oriundos de fenômenos ocorridos no espaço, constitui hoje um grande desafio para a elucidação de questões centrais em diversas áreas do conhecimento humano, por exemplo, em saúde, geologia, agronomia e entre outras. Quando se estuda dados agregados por unidade político-administrativa (estados, municípios, bairros, setor censitário, etc.) deseja-se estudar dentre vários fatores espaciais, o fenômeno denominado dependência espacial. Segundo Cressie (1993) “a dependência espacial está presente em todas as

direções e fica mais fraca a medida que aumenta a dispersão na localização dos dados”.

Nos estudos epidemiológicos busca-se identificar conglomerados de doenças, que são definidos como regiões com risco significativamente elevado de incidência de uma enfermidade que não seja esperado no período considerado [22][18]. Estes estudos são relevantes para a identificação de regiões geográficas que apresentam o risco elevado de ocorrência de uma determinada doença ou uma epidemia [3][5][8].

Devido ao grande problema que acarreta à saúde pública e à qualidade de vida social, inúmeros trabalhos sobre violência e mortalidade por homicídios vêm sendo desenvolvidos em diversas áreas do conhecimento humano [11][16][20]. No Brasil da década 1990 os homicídios foram responsáveis por 33% dos óbitos por causas externas [19]. Os homicídios lideravam a mortalidade no Brasil no ano de 2000 com 38,3% do total, que corresponde a 124 pessoas assassinadas por dia no país [12]. As armas de fogo contribuíram com mais de 50% dos casos em 1991 e com cerca de 70% no ano de 2000 [19]. Assim, a análise espacial pode auxiliar as políticas de segurança e saúde pública no estudo da violência descrevendo e analisando seu comportamento espacial.

Visando compreender o comportamento das diversas técnicas de detecção de conglomerados e de análise de autocorrelação espacial, foram estudadas e aplicadas as seguintes técnicas: Índice de Autocorrelação Getis e Ord, Método Besag e Newell, Varredura Scan e GAM-K. Foram utilizados os dados de homicídios por arma de fogo e arma branca da cidade de João Pessoa no período de 2002 a 2005 fornecidos pela Secretaria de Segurança e da Defesa Social do Estado da Paraíba.

### **Metodologia**

#### *Análise de Dados de Área*

Segundo Assunção (2001) dados de área são dados referentes a um mapa particionado em áreas contíguas e disjuntas. Em cada área pode-se medir uma ou mais variáveis aleatórias e possivelmente covariáveis de interesse que supostamente afetam a distribuição de probabilidade. Estes dados ocorrem com muita frequência quando lidamos com eventos agregados por

municípios, bairros ou setores censitários, onde não se dispõem da localização exata dos eventos, mas de um valor por área [6]. Análises espaciais são baseadas tanto em métodos gráficos e visuais como também em técnicas numéricas, normalmente estatísticas. Quando são utilizadas técnicas de estatística espacial com dados de área, ocorre a produção de índices, gráficos de espalhamento, que medem a associação espacial e mapas. Estes índices auxiliam na identificação de agrupamentos, áreas de transição e casos atípicos [14].

### Risco Relativo

Considerando um estudo da incidência de uma certa doença, numa região geográfica D, formada por n áreas contíguas e disjuntas, denotadas por  $A_1, A_2, \dots, A_n$ . Seja X uma variável aleatória, definida como o número de ocorrências de uma determinada doença, então, as observações das n áreas serão definidas por  $x_1, x_2, \dots, x_n$  [6]. Um indicador que descreve a intensidade de ocorrência de uma determinada doença é a taxa de incidência (ou risco) de X, denotado por  $\theta$ . Seja  $\theta_i$  a taxa de incidência na região  $A_i$  e  $\Theta$  a taxa de incidência global da região D, para  $i = 1, 2, \dots, n$ . Podemos estimar estes riscos por (1) e (2), onde  $y_i$  é o número de pessoas em risco na área  $A_i$ .

$$\theta_i = \frac{x_i}{y_i} \quad (1)$$

$$\Theta = \frac{\sum_{i=1}^n x_i}{\sum_{i=1}^n y_i} \quad (2)$$

A comparação de taxas de incidência entre duas ou mais áreas só faz sentido quando estas apresentam populações similares [3]. Como a suposição de populações idênticas é muito restritiva, costuma-se utilizar as taxas de incidência padronizadas ou riscos relativos denotadas por  $\psi_i$ , para  $i = 1, 2, \dots, n$ . Este risco padronizado pode ser estimado da seguinte forma:

$$\psi_i = \frac{\theta_i}{\Theta} \quad (3)$$

### Índice de Autocorrelação Espacial: Getis e Ord

A análise da autocorrelação espacial, também chamada de dependência espacial, visa identificar a estrutura de correlação espacial dos dados medindo a autocorrelação entre as áreas [2]. Os índices de

autocorrelação global fornecem uma medida de associação para toda a área de estudo, o que é útil para a caracterização da área como um todo. Entretanto, quando a região de interesse apresenta um grande número de sub-regiões é provável que ocorram diferentes regimes de associação espacial em diferentes sub-regiões. Para evidenciar tais regimes é necessário utilizar os indicadores locais de autocorrelação espacial. Estes indicadores fornecem uma medida para cada área permitindo um estudo mais detalhado da associação espacial dos dados. Formalmente, a estatística Getis e Ord Global é dada por [2]:

$$G(d) = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij}(d) x_i * x_j}{\sum_i \sum_j x_i * x_j} \quad (4)$$

$$w_{ij}(d) = \begin{cases} 1, & \text{se } d_{ij} \leq d \\ 0, & \text{se } d_{ij} > d \end{cases} \quad (5)$$

onde:

$x_i$  e  $x_j$  são os valores observados nas áreas  $A_i$  e  $A_j$  respectivamente, para  $i = 1, 2, \dots, n$  e  $j = 1, 2, \dots, n$ .

$w_{ij}$  - são os pesos da matriz de proximidade relacionadas com a distância crítica d, em relação a distância entre as áreas  $A_i$  e  $A_j$

Os indicadores locais produzem estatísticas específicas para cada polígono em estudo, permitindo assim, a identificação de agrupamentos, objetos anômalos e de mais de uma regime espacial [2]. O índice local de Getis e Ord é dado por (6):

$$G_i(d) = \frac{\sum_{j=1}^n w_{ij}(d) x_j}{\sum_{j=1}^n x_j}, \quad j \neq i, \quad (6)$$

A interpretação da estatística de Getis e Ord é realizada com a utilização da padronização da estatística obtida. Desta forma, valores positivos e significativos indica aglomeração de valores elevados. Valores negativos e significativos indicam um aglomeração espacial de valores pequenos.

### Teste Genéricos de Conglomerados

Knox (1988) [15] define um conglomerado como um grupo de ocorrências geograficamente limitado de tamanho e concentração tais que seja improvável de

ocorrer por mero acaso. Conglomerado pode ser definido também como áreas de risco significativamente elevado mas não esperado. Seja uma região geográfica onde se deseja estudar um determinado evento. Suponha ainda que esta região seja dividida em  $n$  sub-regiões onde em cada região esteja associado um número de observações de casos  $c_i$  e o valor esperado  $E_i$  para  $i = 1, 2, \dots, n$ . Seja ainda  $M_i$  o número de pessoas em risco na sub-região  $i$  e  $\lambda$  a taxa global de ocorrência de casos, para  $i = 1, 2, \dots, n$ . O modelo de aleatoriedade ou ausência de aglomeração espacial é definido como:

$$H_0 : c_i \sim \text{Poisson}(E_i = \lambda M_i) \quad (7)$$

A estimativa da taxa global pode ser calculada segundo a equação (8)

$$\hat{\lambda} = \frac{\sum_{i=1}^n c_i}{\sum_{i=1}^n M_i} \quad (8)$$

### Método Besag e Newell

O método proposto por Besag e Newell, geralmente denominado BN, é um método visual que procura identificar conglomerados verossímeis de formato circular [13]. Este baseia-se nas coordenadas dos centróides das possíveis zonas de risco supondo círculos centrados nos centróides de cada sub-área. Em geral, o método calcula o raio necessário para obter um círculo que reúna um número de centróides que acumulem pelo menos  $k$  casos. O algoritmo do método aumenta o raio de busca de cada centróide sucessivamente de forma a agrupar os centróides mais próximos, absorvendo o número de casos e população. Este algoritmo é finalizado quando o raio de busca agrupar um número de caso igual ou superior a um  $k$  pré-fixado. Após a aplicação deste algoritmo uma estatística de teste, para cada centróide, é calculada testando a hipótese de aleatoriedade ou ausência de aglomeração espacial dado pela equação (7). Sob a hipótese de aleatoriedade ou ausência de aglomeração espacial o valor- $p$ , dado pela equação (9) do centróide  $i$  é calculado da seguinte forma [9]:

$$\text{valor} - p = 1 - \frac{\sum_{j=1}^{k-1} (M_j * C / M)^j}{j!} * \exp\left(\frac{-M_j * C}{M}\right) \quad (9)$$

onde  $C$  é o número de casos em toda região de estudo e  $M$  é a população total em risco na região. O método

Besag e Newell identifica os centróides com aglomerados significativos, ou seja, a área de risco será a região com centróides com conglomerados significativos. A escolha do parâmetro  $k$  (número de casos que devem ser buscados) é crucial. Um valor de  $k$  muito grande levaria a um grande número de falsos positivos, isto é, muitos conglomerados significativos que não correspondem à realidade, enquanto que um  $k$  muito pequeno leva a poucos conglomerados que também não correspondem ao real.

### Método de Varredura Scan

A estatística de Varredura *Scan* Kulldorf é um teste de conglomerado que vem sendo muito utilizado nos últimos anos, devido a sua eficiência em detectar tais regiões [18]. Considerando  $n$  sub-regiões candidatas a formarem conglomerados com seus respectivos centróides. O algoritmo do método *Scan* Kulldorf inicialmente possui um círculo que englobe apenas o centróide de cada sub-região. Em termos gerais, o algoritmo aumenta o raio do círculo englobando um novo centróide até que o círculo reúna todos os centróides da região de estudo [4]. Em outras palavras, tem-se um conjunto de possíveis conglomerados com os raios variando desde a situação onde somente o centróide em questão esteja na região circular, até um círculo que acumule todos os centróides da região. Este conjunto de possíveis conglomerados pode ser reduzido se for definido um parâmetro de condição para o raio de busca, de modo que nenhum candidato possua mas que uma certa porcentagem da população. Sejam os parâmetros  $(z, p, r)$  onde  $z \in Z$  representa os círculo na região de interesse,  $p$  a probabilidade de um indivíduo qualquer dentro de  $z$  seja um caso e  $r$  a probabilidade de um indivíduo fora de  $z$  seja um caso [9]. O método de Varredura *Scan* é baseado na função de verossimilhança, podendo ser utilizadas as distribuições de Bernoulli ou Poisson (dado pela equação 7). Assim, a hipótese de aleatoriedade completa para o modelo de Bernoulli é dado por:

$H_0 : p = r$  (Cada Indivíduo é igualmente provável de se tornar um caso)

$H_1 : p > r$

Os parâmetros da função verossimilhança podem ser estimados da seguinte forma:

$$\hat{p} = \frac{c_z}{n_z} \quad (10)$$

$$\hat{r} = \frac{(C - c_z)}{(M - n_z)} \quad (11)$$

onde,  $c_z$  é o número de casos no círculo  $z$ ,  $n_z$  é o número de indivíduos em risco no círculo  $z$ ,  $C$  é o total

de casos na região de estudo e por fim  $M$  é a total de indivíduos na região de estudo.

Para encontrar o aglomerado mais verossímil fixa-se e calcula-se  $p(z)$  e  $r(z)$  que maximiza a função de máxima verossimilhança. De modo geral, é realizado uma varredura sobre todos os possíveis aglomerados definidos em  $Z$ . Após a definição do aglomerado mais verossímil é associado uma estatística de teste da razão de verossimilhança dada pela equação (12)

$$k = \frac{L(\hat{z}, p(\hat{z}), r(\hat{z}))}{L_0} \quad (12)$$

onde  $L_0$  é a função de máxima verossimilhança sob a hipótese nula. As funções de verossimilhança para a distribuição de Poisson e Bernoulli são dadas respectivamente pelas equações (13) e (14):

$$L_0 = \frac{C^C (M-C)^{(M-C)}}{M^M} \quad (13)$$

$$L_0 = \exp \left( \frac{-C}{C!} \right) \left( \frac{C}{M} \right)^C \prod_i c_i \quad (14)$$

Existe uma dificuldade em determinar a distribuição de  $k$ , pois esta estatística depende da distribuição da população. Assim, a significância do teste é avaliada através da distribuição exata condicionada ao número total de casos observados, obtido através do método de simulação de Monte Carlo.

### GAM-K

Geographical Analysis Machine - GAM, é um método visual desenvolvido por Openshaw et al [17], que procura avaliar a hipótese de aleatoriedade ou ausência de conglomerados (dada pela equação 7). O método funciona da seguinte forma [3]:

- I. A informação de cada área é associada a um ponto de referência, geralmente os centróides denotados aqui por  $i$ .
- II. Constrói-se uma grade sobre a região geográfica
- III. Estabelece um raio de busca, denotado por  $r_i$ ;
- IV. Em cada nó  $s$  da grade fixa-se um círculo  $z_i$  de raio  $r_i$ ;
- V. Calcula-se o número de casos  $C = \sum c_i$  e a população  $M = \sum m_i$  das áreas cujos os pontos  $i$  caem dentro do círculo  $z_i$ .

VI. Calcula-se a probabilidade de  $H_0$  ser verdade  $p_i = P(X \leq C)$ .

VII. Desenharm-se no mapa os círculos significativos  $z_i$ , ou seja, aqueles em que  $C \geq k$  e  $p_i \leq \alpha$ .

VIII. Refazer os passos de III a VI para cada valor de  $r_i$  selecionado no passo III.

O resultado final é a identificação de áreas de risco por emaranhados de círculos sobrepostos, todos significativos e com pelo menos dois eventos. Este método é considerado simples e fácil aplicação. Existe ainda uma variação do método GAM denominado GAM-K, no qual são realizados os procedimentos do GAM para  $n$  raios distintos  $r_1, r_2, \dots, r_n$  obtendo-se para cada raio os nós da grade que foram significativos para o GAM, denotados pelo vetor  $u(r_i)$ . Após a definição do vetor de raios os nós significativos são reunidos em um único vetor  $u = [u(r_1), u(r_2), \dots, u(r_n)]$ . Com as localizações dos nós significativos aplica-se o estimador de intensidade Kernel, determinado a intensidade com que os pontos da grade foram significativos para o GAM-K [1].

## Resultados

### Getis e Ord Global e Local

O estudo da dependência espacial da distribuição de um fenômeno permiti verificar padrões e tendências espaciais semelhantes ou divergentes em uma determinada área. Com a verificação da existência da dependência espacial se faz necessário identificar mas precisamente as regiões onde se encontram estes padrões. Esta identificação pode ser feita através do índice de autocorrelação local  $G_i(d)$ . O índice global  $G(d)$  indicará aglomeração espacial de valores altos quando o valor da estatística  $z$  ( $G$  padronizada) for positivo. Quando o valor da estatística  $z$  for negativo o índice indicará uma aglomeração espacial de valores baixos. A estatística deve ser levada em consideração quando a hipótese nula de inexistência de autocorrelação for rejeitada. Portanto, o índice global  $G(d)$  deve ser significativo para que o índice seja válido. Se houver razões para crer que a autocorrelação é nula pode-se concluir que não existe dependência espacial. Neste caso, não é necessário calcular os índices locais  $G_i(d)$ . Como não existe um critério para a definição de um parâmetro adequado (distância  $d$ ), foi utilizado a distância média dos centróides de cada bairro (6 quilômetros).

Para a distância média dos centróides é notado que o índice  $G$  (Tabela 1) foi significativamente diferente de zero, ao nível de significância de 5%, apenas para a variável arma categoria fogo no ano de 2003. Neste casos, só faz sentido calcular os índices  $G$  locais para o ano de 2003 para a variável arma categoria fogo. O mapa dos índices Getis e Ord Local está dividido em em dois grupos negativos e positivos. Estas classes representam as significâncias dos índices a partir dos

seguintes símbolos: \* significativo a 5% e \*\* significativo a 1%. No ano de 2003 (Figura 1), o índice local G identificou duas áreas com aglomeração espacial o primeiro, localizado na região Noroeste, com os seguintes bairros como um aglomerado: Varjão, Jaguaribe, Trincadeiras, Varadouro, Centro, Tambiá, Roger, Jardim das Acácias, Torre, Treze de Maio, Padre Zé e Alto Céu. Este aglomerado caracteriza-se por ter altos valores de risco relativo ou estarem vizinhos a bairros com alto risco de homicídios por arma de fogo. A segunda região de aglomeração é localizada no litoral formado pelos bairros de Mangabeira, Costa do Sol, Penha e Ponta dos Seixas. Esta aglomeração se caracterizou por ter baixo valores de risco relativo ou estarem vizinhos a bairros com baixo risco de homicídios por arma de fogo.

Tabela 1 - Índice Global Getis e Ord (distância de 6km) para a variável homicídio segundo tipo de arma.

Arma	Ano	Índice G	Valor-p
Branca	2002	-0,7058	0,4803
	2003	0,4940	0,6427
	2004	1,2629	0,2066
	2005	0,5500	0,5823
Fogo	2002	0,6703	0,5027
	2003	2,8960	0,0038
	2004	-1,4678	0,1421
	2005	-0,3704	0,7111

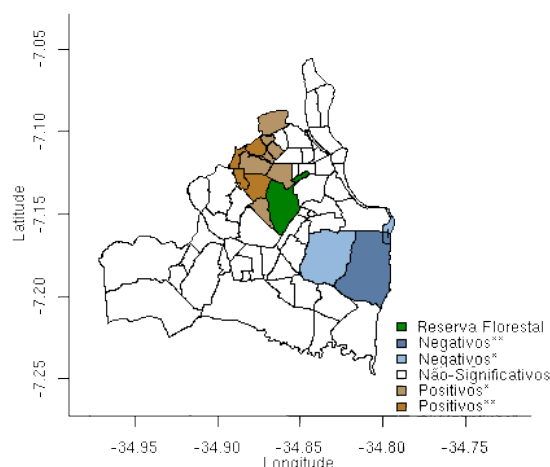


Figura 1 – Mapa Getis e Ord Local Homicídios por Arma de Fogo para distância de 6 quilômetros.

### Besag e Newell

A escolha dos parâmetros do teste (números de casos para busca - k e o nível de significância -  $\alpha$ ) foi definido da seguinte forma: para o parâmetro k foi contabilizado como sendo a maior ocorrência de homicídios de cada ano e definido um nível de significância de 0,002 para reduzir os erros de

detecção de aglomerados. A escolha destes parâmetros carece de um referencial teórico, sendo então totalmente arbitrários. As Figuras 2 e 3 mostram os conglomerados detectados pelo método Besag e Newell para os anos de 2002 e homicídios por arma branca e fogo, utilizando o número de homicídios e a população de risco de cada área. A área de risco é identificada como o emaranhado de círculos significativos. Portanto, os bairros que representarão o conglomerado espacial serão os bairros que possuírem os centróides contidos nestes emaranhados. A Figura 2 mostra o comportamento da aglomeração espacial para os homicídios por arma branca em 2002. Esta aglomeração espacial localizada na região Noroeste possui 10 bairros representados por: Ilha do Bispo, Trincadeiras, Jaguaribe, Varadouro, Centro, Tambiá, Jardim das Acácias, Roger, Padre Zé e Treze de Maio. Estes bairros possuem altos valores de risco relativo, pois para todos eles o risco relativo foi superior a 1, ou seja, o risco de homicídios é superior ao risco global de João Pessoa. Na Figura 3 pode ser visualizado o método Besag e Newell para os homicídios por arma de fogo em 2002. O emaranhado de círculos significativos representa a área de risco de ocorrência de homicídios por arma de fogo. A área de risco é formada por 9 bairros: Trincadeiras, Varadouro, Centro, Tambiá e Jardim das Acácias, Roger, Padre Zé e Treze de Maio. Para estes bairros apenas Tambiá e Jardim das Acácias tiveram risco relativo abaixo de 1, caracterizando bairros de transição, ou seja, são bairros que possuem risco relativo alto ou estão cercados por bairros que possuem alto risco relativo.

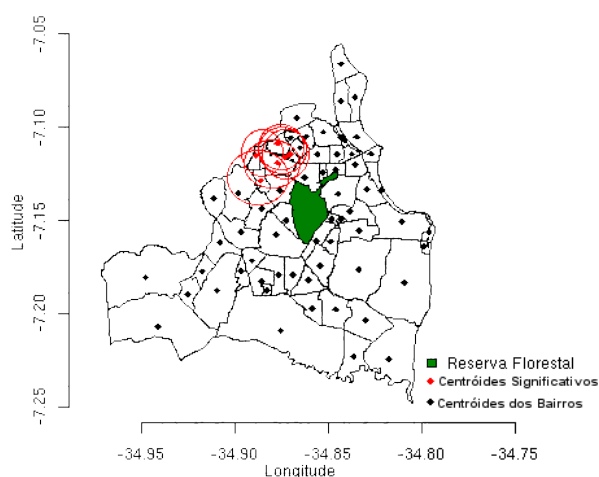


Figura 2 - Método Besag e Newell para homicídios por arma branca em 2002 com k = 6.

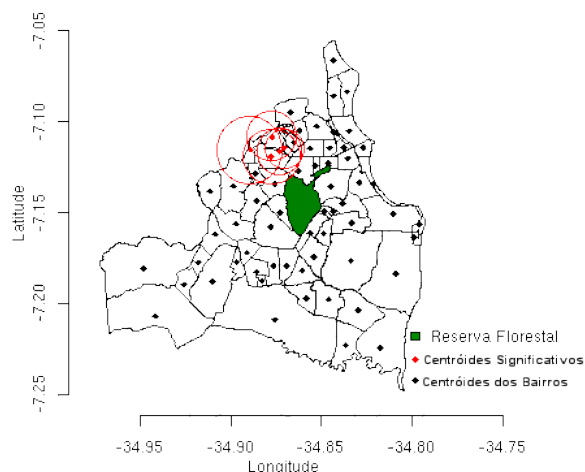


Figura 3 - Método Besag e Newell para homicídios por arma de fogo em 2002 com  $k = 5$ .

### Varredura Scan

O método de Varredura proposto por Kulldorff e Nagarwalla permite identificar conglomerados, envolvendo um número de parâmetros não cruciais definidos pelo usuário. Como mencionado anteriormente, o método de Varredura *Scan* depende de uma proporção. Aqui será utilizado os candidatos a aglomerado que não possuírem mais que 20% da população total de João Pessoa.

A Figura 4 mostra o comportamento do método de Varredura para o número de homicídios por arma branca em 2002. É verificado a existência de três regiões com aglomeração espacial uma localizada no litoral Sul de João Pessoa formado por 3 bairros: Barra de Gramame, Muçumagro, Paratibe; uma segundo aglomerado espacial localizado na região Centro-Oeste formado por 5 bairros: Cidade dos Colibris, Jardim Cidade Universitária, Anatólia, Jardim São Paulo e Bancários, e por fim um terceiro aglomerado espacial localizado na região Noroeste de João Pessoa formado por 15 bairros: Cristo, Varjão, Cruz das Armas, Ilha do Bispo, Alto do Matheus, Jaguaribe Trincadeiras, Varadouro, Centro, Tambiá, Jardim das Acácias, Roger, Padre Zé, Treze de Maio e Torre.

O aglomerado de homicídios por arma de fogo em 2002, identificado na Figura 5, abrange uma área que engloba muitos bairros da região Norte de João Pessoa. Este bairros são: Varjão, Jaguaribe, Trincadeiras, Varadouro, Centro, Tambiá, Jardim das Acácias, Roger, Padre Zé, Mandacaru, Treze de Maio, Alto Céu, Bairro dos Ipês, Bairro dos Estados, Torre, Pedro Gondim, João Agripino, Bairro São Jose, Brisamar, Tambauzinho, Expedicionários, Altiplano, Castelo Branco, Tambaú, Manaíra e Aeroclube.

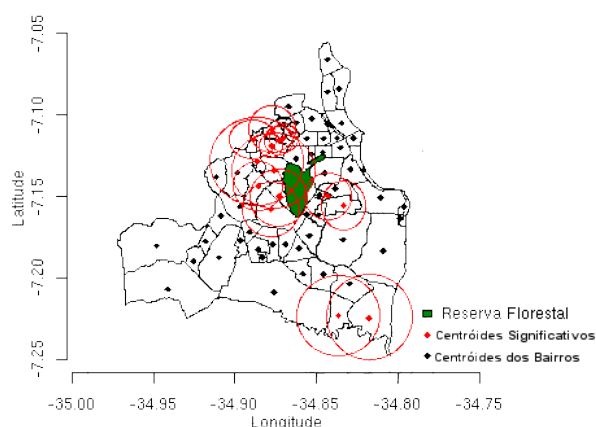


Figura 4 - Método *Scan* Kulldorff e Nargalha para homicídios por arma branca em 2002.

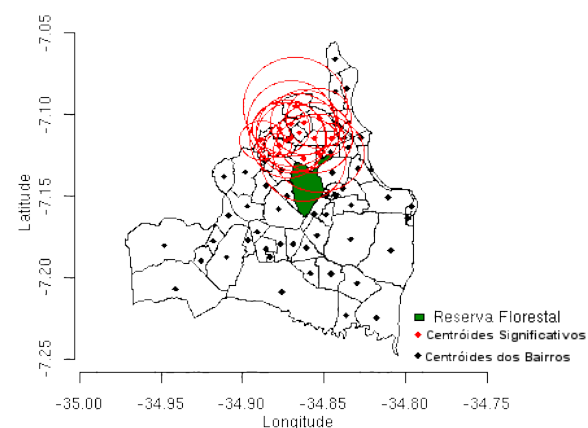


Figura 5 - Método *Scan* Kulldorff e Nargalha para homicídios por arma de fogo em 2002.

### GAM-K

Para obter o GAM-K foram utilizados cinco raios de busca na obtenção dos nós significativos do método GAM. Com a utilização do método GAM-K será possível determinar a intensidade dos pontos, ou seja, onde a intensidade for maior indicará um aglomerado de violência. O método GAM-K identificou dois focos de aglomeração espacial de homicídios por arma de fogo em 2002 (Figura 6). O primeiro foco, localizado no bairro do Jardim Veneza, apresentando baixa intensidade mas com indicio de aglomeração espacial de homicídios por arma de fogo. O segundo foco encontra-se na região Nordeste englobando os bairros do Varadouro, Centro, Tambiá, Jardim das Acácias, Roger, Padre Zé, Treze de Maio e Alto Céu. Em 2003 (Figura 7) a aglomeração espacial localizou-se no Centro-Oeste com a o foco nos bairros: Cruz das Armas, Varjão e Jaguaribe. Nestes bairros a intensidade de aglomeração é maior. Para os homicídios por arma branca de 2002 a 2005 e homicídios por arma de fogo em 2005 o método GAM-K não apresentou bom resultados, encontrando



indícios de aglomeração onde não ocorreram homicídios. Estes erros devem-se ao fato de que existe uma grande dispersão nos dados de risco relativo, fazendo com que raios grandes definidos no método GAM-K incluam casos onde não deveria, como por exemplo no bairro de Mussurê, localizado na região do extremo sudoeste, que não apresentou ocorrência de homicídios por arma branca. Entretanto, o método detectou aglomeração espacial neste bairro. Para minimizar tais problemas deve-se ser criterioso na escolha dos raios e sempre analisar os mapas GAM-K conjuntamente com os mapas de risco relativo.

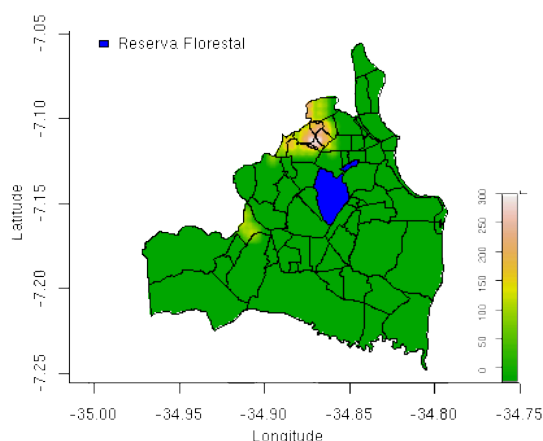


Figura 6 - Aglomerados de homicídios por arma de fogo em 2002 segundo o método GAM-K.

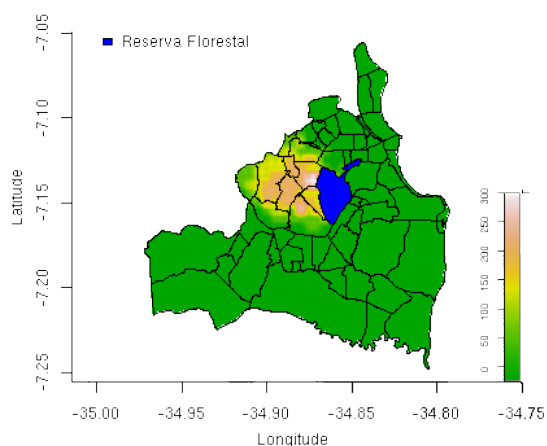


Figura 7 - Aglomerados de homicídios por arma de fogo em 2003 segundo o método GAM-K.

## Conclusão

Em termos gerais, a comparação dos resultados de cada método deve ser realizada de forma criteriosa devido a subjetividade e a falta de uma teoria para definição precisa dos parâmetros de cada método. Existem vantagens e desvantagens incorporados em cada método. Assim, a escolha da utilização de um método, deve ser minuciosa, analisando os problemas

e ajustando-os ao estudo a ser considerado. O método de Varredura *Scan* é um dos mais utilizados, entretanto, apresenta algumas desvantagens que podem ser visualizadas nos resultados aqui expostos. A principal delas é a tendência de identificar um conglomerado maior do que realmente o é. Esta deficiência pode ser vista na Figura 5 onde o método detectou a região Norte com exceções dos bairros do Bessa e Jardim Oceania (bairros situados no extremo norte do município).

Em alguns casos o método GAM-K não apresentou bom resultados, encontrando indícios de aglomeração onde não ocorreram homicídios. Estes erros devem-se ao fato de que existe uma grande dispersão nos dados de risco relativo, fazendo com que raios grandes definidos no método GAM-K incluam casos onde não deveria. Observando todos os resultados, fica claro que os conglomerados de homicídios encontram-se localizado na região Noroeste, com os bairros do Centro, Jaguaribe, Trincheiras e Tambiá sendo os principais focos de homicídios. Outros conglomerados foram identificados de forma isolada. O principal localiza-se na região Sudoeste de João Pessoa com o foco no Bairro das Industrias.

Considerando que os parâmetros definidos são os melhores possíveis, pode-se dizer que os métodos apresentaram uma certa semelhança no padrão da localização espacial de conglomerados de homicídios. Os métodos identificaram a região Noroeste como a principal área de ocorrência de homicídios tanto por arma de fogo quanto por arma branca, ocorrendo pequenas variações na localização e tamanho dos conglomerados.

## Agradecimentos

Daniel Matos de Carvalho foi aluno bolsista do Conselho Nacional de Desenvolvimento Científico Nacional – CNPq, pelo Projeto CNPq-505258/2004-8. Este trabalho tem suporte da Secretaria da Segurança e da Defesa Social do Estado da Paraíba (parceria).

## Referências

- [1] M. Ahmadi, Crime Mapping and Spatial Analysis, MSc Thesis of Science in Geoinformatics – International Institute for Geo-information Science and Earth Observation, Enschede, The Netherlands, ITC, 2003.
- [2] L. Anselin, Spatial Data Analysis with GIS: An Introduction to Application in the Social Sciences. National Center for Geographic Information and Analysis, University of California, Santa Barbara, CA 93106, Agosto 1992.
- [3] R. M. Assunção, Estatística Espacial com Aplicações em Epidemiologia, Economia e Sociologia, Associação Brasileira de Estatística, São Carlos - SP, 2001, Disponível em: <http://www.est.ufmg.br/~assuncao/art/art.html>, Acesso em: nov. 2005.
- [4] R. M. Assunção, Uma Análise de Desempenho dos Métodos Scan e Besag & Newell na Detecção de

- Conglomerados Espaciais. V Brazilian Symposium on GeoInformatics 2003
- [5] F. C. Beato, R. M. Assunção, Silva BFA, Marinho FC, Reis IA, Almeida MCM. Conglomerados de homicídios e o tráfico de drogas em Belo Horizonte, Minas Gerais, Brasil, de 1995 a 1999. *Cad Saúde Pública* 2001;17:1163-71.
- [6] G. Camara, et al. Análise espacial de áreas. In: DRUCK, S et al. Análise espacial de dados geográficos. 2. ed. São Jose dos Campos: INPE, jul. 2002. Disponível em: <<http://www.dpi.inpe.br/gilberto/livro/analise/>>
- [7] G. Camara, C. Davis., A. M. Monteiro, J. C. D'alge, ; Introdução à Ciência da Geoinformação. São Jose dos Campos: INPE, 2001, Disponível em: <<http://www.dpi.inpe.br/gilberto/livro/introd/>>
- [8] M. S. Carvalho, Souza-Santos R, "Analysis of spatial data in public health: methods, problems, and perspectives", *Cad Saúde Pública*, Vol 21, No 2., 2005, pp 361-378.
- [9] M. A. Costa, R. M. Assunção, (2005) A Fair Comparison Between the Spatial Scan and the Besag - Newell disease Clustering Tests. *Environmental and Ecological Statistics*, Vol. 12, 301-319.
- [10] N. A. C. Cressie,. *Statistics for Spatial Data*. New York: John Wiley & Sons, 1993.
- [11] L. L. Dahlberg, ; E. G. Krug, Violence a global public health problem. *Ciênc. saúde coletiva* , Rio de Janeiro, v. 11, n. 2, 2006 .
- [12] V. P. Gawryszewski, ; M. S. Koizumi,; M. H. P. Mellojorge, 2004. As causas externas no Brasil no ano 2000: comparando a mortalidade e a morbidade, *Caderno de Saúde Pública*, Rio de Janeiro, 20(4): 9951003, jul-ago, 2004, p. 1000.
- [13] B. Julian, N. James, The detection of clustes in rare diseases. *Jornal of the Royal Statistcs Society A*, (154): 143 – 155, 1991
- [14] A. P. Krempi .; Explorando Recursos de Estatística Espacial para análise da acessibilidade da cidade de Bauru. Dissertação apresentada ao Departamento de transportes, da Escola de Engenharia de São Carlos, como parte dos requisitos para obtenção do título de Mestre em Engenharia Civil, 2004.
- [15] E. G. Knox, Detection of Clusters. Em *Methodology of enquiries into disease clustering*. P. Elliott (editor). London: Small Area Health Statistics Unit, 1988
- [16] M. Kulldorff, E. J. Feuer, Miller BA, Freedman LS. Breast cancer clusters in the northeast United States: a geographic analysis. *Am J Epidemiol* 1997; 146: 161–170
- [17] S. Openshaw, et al, "Investigation of leukaemia cluster by use of a geographical analysis machine", *LANCET*, 1988, pp 272-273.
- [18] A. Ozonoff, Webster T, Vieira V, Weinberg J, Ozonoff D, Aschengrau A. Cluster detection methods applied to the Upper Cape Cod cancer data, 2005
- [19] M. F. T. Peres, P. C. SANTOS, . Trends of homicide death in Brazil in the 90s: the role of firearms. *Rev. Saúde Pública* , São Paulo, v. 39, n. 1, 2005 .
- [20] E. R. Souza, M. L. C. Lima, The panorama of urban violence in Brazil and its capitals. *Ciênc. saúde coletiva* , Rio de Janeiro, v. 11, n. 2, 2006 .
- [21] I. C. A. Souza, ; R. M. Moraes, Utilização de Sistemas de Informação Geográfica na análise espacial de dados de saúde Pública do estado da Paraíba entre os anos de 1998 a 2001. *Relatorio de Iniciação Científica PIBIC/CNPq/UFPB*. João Pessoa, 2003. Disponível em <<http://www.de.ufpb.br/onei/>>
- [22] I. C. A. Souza, R. M. Moraes, (2006). Identificação de conglomerados espaciais do dengue. *Proceedings of the Environmental and Health World Congress 2006. (EWHC'2006)*. Julho, Santos, Brasil, pg 239-243.