



*MINISTÉRIO DA CIÊNCIA E TECNOLOGIA  
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS  
CURSO DE PÓS GRADUAÇÃO EM SENSORIAMENTO REMOTO*

## **Análise exploratória de dados sócio- econômicos do município do Rio de Janeiro**

*Rosely Kimie Teruiya*

Trabalho referente a disciplina SER-301 Análise Espacial, ministrada  
pelos Profs. Drs. Gilberto Câmara e Antonio Miguel Monteiro

São José dos Campos  
1999

## *ÍNDICE*

1. INTRODUÇÃO.....	2
2. OBJETIVOS.....	2
3. MÁQUINA DE ANÁLISE GEOGRÁFICA.....	3
3.1. O algoritmo do GAM.....	3
3.2. Vantagens do GAM.....	4
3.3. Desvantagens do GAM.....	4
4. ESTIMADOR DE KERNEL.....	5
5. MATERIAIS E MÉTODOS.....	6
5.1. Banco de Dados.....	6
5.2. Programas Utilizados.....	8
6. RESULTADOS OBTIDOS.....	10
6.1. Índice de Apgar.....	11
6.2. Nascidos Vivos com Baixo Peso.....	13
6.3. Proporção de Cesáreas.....	15
6.4. Proporção de Mães Adolescentes.....	17
6.5. Proporção de Mães com Escolaridade acima de 2º grau.....	19
7. CONCLUSÕES.....	21
8. REFERÊNCIAS BIBLIOGRÁFICAS.....	22

## **1. INTRODUÇÃO**

Métodos exploratórios em análise de dados envolvem a procura de boas descrições de dados, a fim de ajudar o analista a desenvolver algumas hipóteses sobre o assunto e modelos apropriados para tais dados (Bailey e Gatrell, 1995). Em geral, os métodos exploratórios enfatizam visualizações gráficas dos dados, realçando feições características e permitindo ao analista detectar padrões, relações, valores anômalos, entre outras particularidades. Portanto, no contexto de análise espacial de dados, os resultados são apresentados em forma de mapas ou gráficos. Algumas técnicas exploratórias, quando aplicadas a um padrão de eventos pontuais, resultam em um mapa de contorno de intensidade estimada de ocorrências em toda a área de estudo. Outras técnicas exploratórias aplicadas ao mesmo conjunto de eventos podem resultar num gráfico referente ao grau de dependência espacial entre as localizações das amostras.

Como técnica exploratória foi utilizada neste trabalho a Máquina de Análise Geográfica (GAM), desenvolvida pelo geógrafo Stan Openshaw, com o objetivo de detectar clusters, porém não se preocupando em explicar o processo de aparecimento destes. Esta forma de análise traz poucas suposições a respeito dos dados, porém se enquadra perfeitamente no espírito de um método exploratório. Outra técnica utilizada para detectar padrões espaciais de distribuição das amostras foi o Estimador de Densidade de Kernel, que foi testada em dois programas, o INFO-MAP e o SPRING.

Os dados utilizados são provenientes do Sistema de Informações sobre Nascidos Vivos (SINASC) de 1994, onde foram analisadas cinco variáveis sobre o perfil de nascimentos por bairros no município do Rio de Janeiro. As variáveis utilizadas foram: Índice de Apgar, Nascidos Vivos com Baixo Peso, Proporção de Cesáreas, Mães Adolescentes e Mães com Escolaridade acima de 2º grau.

## **2. OBJETIVOS**

Os principais objetivos deste trabalho foram: testar a eficiência de métodos exploratórios para dados sócio-econômicos do município do Rio de Janeiro e, comparar os resultados com trabalhos feitos na área a partir do mesmo banco de dados.

### 3. MÁQUINA DE ANÁLISE GEOGRÁFICA

A Máquina de Análise Geográfica (GAM - Geographical Analysis Machine) é uma tentativa de análise espacial exploratória de dados, seja de pontos ou pequenas áreas, de fácil entendimento. Basicamente, o GAM necessita de uma base de dados referenciada, podendo desta forma indicar evidências de localização e importância de *clusters*. O GAM é uma ferramenta de análise exploratória, desenvolvida para ser uma máquina amigável de análise espacial onde os resultados são claros, sem a necessidade de uma estatística complexa.

Seu funcionamento consiste em examinar um grande número de círculos de vários tamanhos cobrindo completamente a região de interesse. Os círculos se sobrepõem permitindo o efeito de bordas e fornecendo um grau de análise sensitiva, sendo que os dados de cada círculo são armazenados e uma avaliação estatística é feita para saber se a taxa de incidência em cada círculo é alta.

#### 3.1. O algoritmo do GAM

O algoritmo do GAM funciona de acordo com os seguintes passos:

1. Ler os dados X,Y da população de risco e da variável de interesse.
2. Identificar o retângulo contendo os dados, o raio do círculo inicial e o grau de recobrimento.
3. Gerar um grid cobrindo este retângulo.
4. Para cada interseção do grid gerar um círculo de raio R.
5. Buscar dois valores para a população de risco e variável de interesse.
6. Aplicar um teste de significância.
7. Manter o resultado, se significativo.
8. Repetir os passos 5 a 7 até que todos os círculos sejam processados.
9. Aumentar o raio do círculo e retornar aos passos 3 a 10.
10. Criar uma superfície suavizada de excesso de incidência para os círculos significantes, usando o procedimento de suavização de kernel e agregando os resultados para todos os círculos.
11. Mapear esta superfície.

### 3.2. Vantagens do GAM

- ◆ É completamente automatizada e de fácil aplicação.
- ◆ Prévio conhecimento de hipóteses de padrões é considerado irrelevante, pois é uma técnica exploratória.
- ◆ Procura localizar *clusters* em vez de utilizar mapas estatísticos.
- ◆ A busca é geograficamente abrangente e nenhuma região é desconsiderada.
- ◆ Todas as regiões são tratadas igualmente.
- ◆ A saída é visual em vez de termos estatísticos complexos.
- ◆ É uma ferramenta de análise espacial exploratória e pode sugerir hipóteses que podem ser testadas, mais tarde, por outros métodos.

### 3.3. Desvantagens do GAM

- A tentativa deliberada de automatizar a ciência da análise estatística foi criticada por muitos especialistas.
- Necessidade de alguma teoria estatística bem desenvolvida para o método.
- A idéia de *clustering* localizado foi inicialmente considerada como sendo puramente artefato de dados devido a autocorrelação espacial.
- O processo de busca usado no GAM é compreensível e computacionalmente muito eficiente, mas não é inteligente podendo considerar regiões que não interessam como possíveis clusters.

#### 4. ESTIMADOR DE KERNEL

De acordo com Bailey e Gatrell (1995), o estimador de Kernel foi originalmente desenvolvido para obter a estimativa de densidade de probabilidade univariada ou multivariada de uma amostra observada. Estimar a intensidade de um padrão de pontos é como estimar uma densidade de probabilidade bivariada. Pode-se adaptar a estimativa bivariada de Kernel para se obter uma estimativa de intensidade do padrão de pontos.

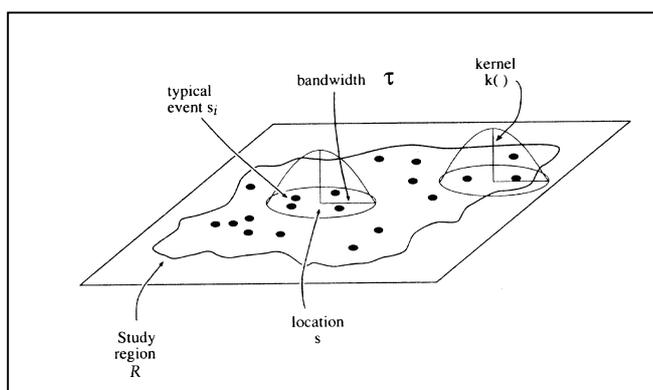


Fig. 1 – Estimador de Kernel para um padrão de pontos

De acordo com a Figura 1, suponha que  $s$  represente uma localização em uma região  $R$  e  $s_1, \dots, s_n$  são localizações de  $n$  eventos observados. Então, a densidade estimada,  $\lambda$ , na localização  $s$  é estimada segundo a equação 1 abaixo:

$$\hat{\lambda}_\tau(s) = \sum_{h_i \leq \tau} \frac{3}{\pi\tau^2} \left(1 - \frac{h_i^2}{\tau^2}\right)^2 \quad (1)$$

onde  $h_i$  é a distância entre o ponto  $s$  e a localização do evento observado  $s_i$ , e a soma só acontece para os pontos que estão a uma localização  $h_i$  que não ultrapasse  $\tau$ . A região de influência dentro da qual os eventos contribuem para o cálculo da intensidade é um círculo de raio  $\tau$  com centro em  $s$ . Observando a fórmula, verifica-se que na localização  $s$ , a uma distância de zero, o peso é  $3/\pi\tau^2$  e cai suavemente para o valor zero quando a distância é  $\tau$ .

Quando se estima o kernel sobre uma grade de localizações em  $R$  pode-se pensar em uma função tridimensional que visita cada ponto  $s$  dessa grade. Calculam-se as distâncias de cada ponto  $s$  aos eventos observados  $s_i$  dentro da região de influência limitada pela distância  $\tau$ . Essas distâncias contribuem para o cálculo da intensidade estimada no ponto  $s$ . A escolha de  $\tau$  define superfícies suaves ou não, sendo que para grandes larguras  $\tau$ , a intensidade tende a apresentar picos centrados em  $s$ .

## **5. MATERIAIS E MÉTODOS**

Basicamente foi utilizado o banco de dados do município do Rio de Janeiro para a aplicação de técnicas exploratórias como o GAM e o Estimador de Kernel nos programas INFO-MAP e SPRING.

### **5.1. Banco de Dados**

O banco de dados utilizado neste trabalho foi fornecido pela Fundação Oswaldo Cruz, contendo informações sócio-econômicas a respeito do município do Rio de Janeiro, em ambiente SPRING (Figura 2). Estes dados são provenientes do Sistema de Informações sobre Nascidos Vivos (SINASC) de 1994. O SINASC é um subsistema de informações de âmbito nacional, sob responsabilidade das Secretarias Estaduais e Municipais de Saúde, concebido para diminuir a subenumeração e melhorar a qualidade da informação sobre nascidos vivos. Este sistema é alimentado pelas declarações de nascido vivo, documento oficial emitido pelo hospital onde ocorre o nascimento (d'Orsi e Carvalho, 1998).

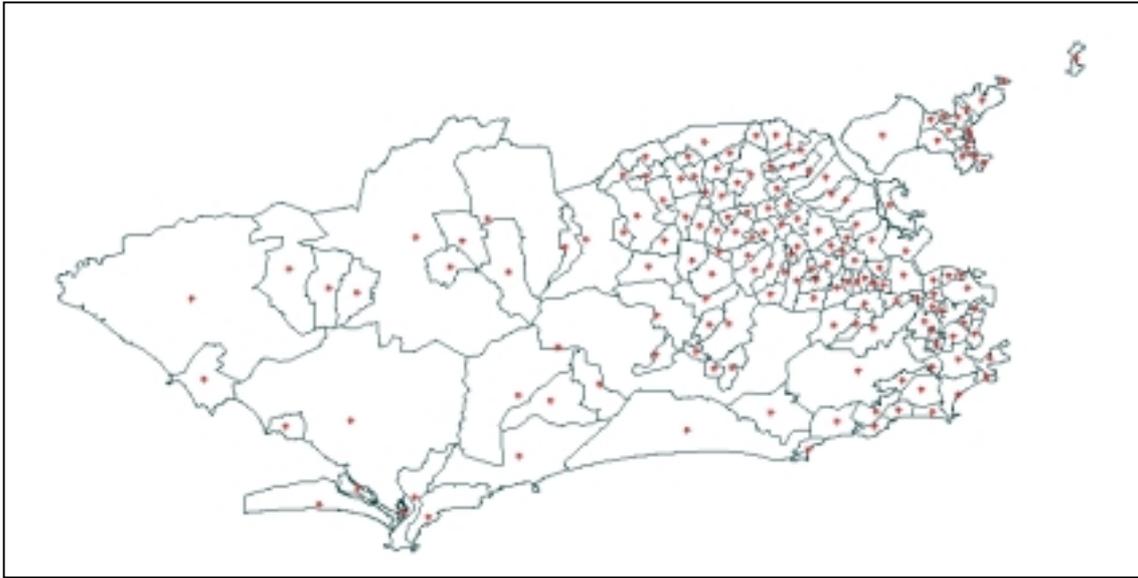


Fig. 2 – Município do Rio de Janeiro dividido por bairros, mostrando os centróides dos polígonos.

Como a quantidade de variáveis sócio-econômicas presente no banco de dados é muito grande, optou-se por ter como referência o artigo de d’Orsi e Carvalho (1998), que utilizou o mesmo banco de dados, analisando o perfil de nascimentos no município do Rio de Janeiro. Portanto, as variáveis utilizadas neste trabalho foram: Índice de Apgar, Nascidos Vivos com Baixo Peso, Proporção de Cesáreas, Mães Adolescentes e Mães com Escolaridade acima de 2º grau.

O Apgar é um índice composto que mede a vitalidade (cor da pele, respiração, batimentos cardíacos, tônus muscular e resposta a estímulos nervosos) do recém-nascido no primeiro e no quinto minuto após o nascimento. O baixo peso constitui-se um dos principais determinantes da mortalidade perinatal e o tipo de parto permite caracterizar a prática obstétrica nos hospitais. A idade e o nível de instrução são relativos às características da mãe, sendo que a escolaridade é a única variável que permite avaliar a situação sócio-econômica do recém-nascido.

## 5.2. Programas Utilizados

As cinco variáveis descritas acima, foram testadas no GAM e os resultados dos clusters comparados aos mapas gerados no SPRING e INFO-MAP através da técnica de Estimador de Kernel.

A versão do GAM utilizada foi a on-line (GAM/K), disponível na Internet no endereço eletrônico <http://www.ccg.leeds.ac.uk/smart/gam>. A entrada de dados no GAM faz-se através de dois arquivos tipo ASCII, um contendo o número identificador (ID) das amostras e suas coordenadas e outro arquivo contendo o ID, o valor da incidência e a população de risco, como mostrado na Figura 3.

	X	Y		Incidência	População
1	18.447500	54.750000	1	8	252
2	15.952500	44.850000	2	2	45
3	35.259091	92.000000	3	5	360
4	47.887500	84.550000	4	8	366
5	36.200000	79.600000	5	1	54
6	31.902857	89.200000	6	2	36
7	32.659375	73.250000	7	2	72
8	25.850000	58.300000	8	1	27
9	42.007955	69.650000	9	6	708
10	42.075000	68.050000	10	2	234
11	41.018750	65.950000	11	15	522
12	42.065909	64.550000	12	6	183
13	42.463636	63.300000	13	6	165
14	42.800000	61.500000	14	3	225
15	43.700000	59.750000	15	2	93
16	44.400000	58.150000	16	5	225
17	44.950000	55.700000	17	3	144
18	43.592424	53.400000	18	1	120
19	46.191250	51.050000	19	5	336
20	47.883333	48.600000	20	4	369

Fig. 3 – Exemplo de arquivos de entrada no GAM

Os parâmetros utilizados para as cinco variáveis foram os mesmos para o funcionamento do GAM e podem ser vistos na figura 4 abaixo:

The screenshot displays the 'GAM/K Spatial Analysis' software interface. It features several sections with adjustable parameters:

- Insert Title:** A text input field.
- Spatial Search Parameters:** Includes 'Minimum Radius' (1.0), 'Maximum Radius' (20.0), and 'Radius Step' (0.1). Below these are radio buttons for 'are the data represented as counts' (selected) and 'or reported values'.
- Statistical Things:** Includes a 'Threshold' dropdown (1.0), 'Type of Test' dropdown (Poisson Probability), and 'Measure' dropdown (Excess).
- Spatial Data Uncertainty:** Includes 'Minimum population at risk count' (1.0), 'Minimum incidence count' (1.0), 'Minimum number of datapoints' (5.0), and 'Degree of circle overlap' (0.2).
- Multiple Testing:** Includes 'Number of test runs' (0.0).
- Options you could change (if you know what you are doing):** Includes 'Samples for bootstrap' (400.0), 'Sample size for Monte Carlo testing' (100.0), and 'Parameter (HR) for Monte Carlo testing' (20.0). Below these are three radio button questions: 'Do you want a test based on a random version of your data?' (no selected), 'Do you want to find excess or deficient?' (no selected), and 'Are your input coordinates in?' (metres selected).

A 'Run GAM' button is located at the bottom of the interface.

Fig. 4 – Tela do GAM com os parâmetros utilizados para as cinco variáveis.

O programa SPRING (Sistema para Processamento de Informações Georeferenciadas) foi desenvolvido pelo Departamento de Processamento de Imagens do INPE, sendo de domínio público. A versão utilizada neste trabalho foi uma preliminar, pois a opção de Estimador de Densidade de Kernel ainda está em fase de teste. Como o banco de dados já se encontrava em ambiente SPRING, não houveram dificuldades na manipulação dos dados para testar o Estimador de Kernel.

A técnica de estimativa por Kernel também foi testada no programa INFO-MAP, que acompanha o livro *Interactive Spatial Data Analysis* (Bailey e Gatrell, 1998). Este programa tem um caráter simples e educacional, compreendendo análises estatísticas em dados espaciais. Para a execução deste trabalho, talvez a entrada dos dados no INFO-MAP tenha sido a fase que dispendeu maior parte do tempo, devido a precária importação dos dados no programa. Os bairros do Rio de Janeiro foram importados em formato PCX e devido a baixa resolução dos gráficos, os polígonos tiveram que ser editados. Como o formato do arquivo para entrada das amostras é muito complexo e quase impossível de ser editado devido a grande quantidade de polígonos existentes na área (157), optou-se por associar cada amostra ao seu devido polígono através de um identificador. Este procedimento exigiu enorme trabalho braçal, confirmando a ineficácia de entrada deste tipo de dado no INFO-MAP. Com os dados já inseridos no programa, sua manipulação é relativamente simples e foram testadas, então, as cinco variáveis para o Estimador de Densidade de Kernel.

## **6. RESULTADOS OBTIDOS**

Os resultados obtidos através do GAM e do Estimador de Densidade por Kernel foram comparados com os dados de d'Orsi e Carvalho (1998) para verificar a coerência dos resultados.

Os mapas gerados a partir dos programas SPRING e INFO-MAP para o Estimador de Kernel apresentaram resultados muito semelhantes, identificando praticamente os mesmos locais de altas evidências. Para melhor comparação entre os resultados gerados pelos dois programas, foi utilizado um raio de busca de aproximadamente 3km e os mapas foram fatiados em oito classes iguais, para ambos os programas. Já o GAM discriminou áreas muito mais restritas, com resultados diferenciados em cada uma das cinco variáveis.

A seguir serão apresentados os mapas obtidos a partir dos três métodos para as cinco variáveis (Índice de Apgar, Nascidos Vivos com Baixo Peso, Proporção de Cesáreas, Mães Adolescentes e Mães com Escolaridade acima de 2º grau).

## 6.1. Índice de Apgar

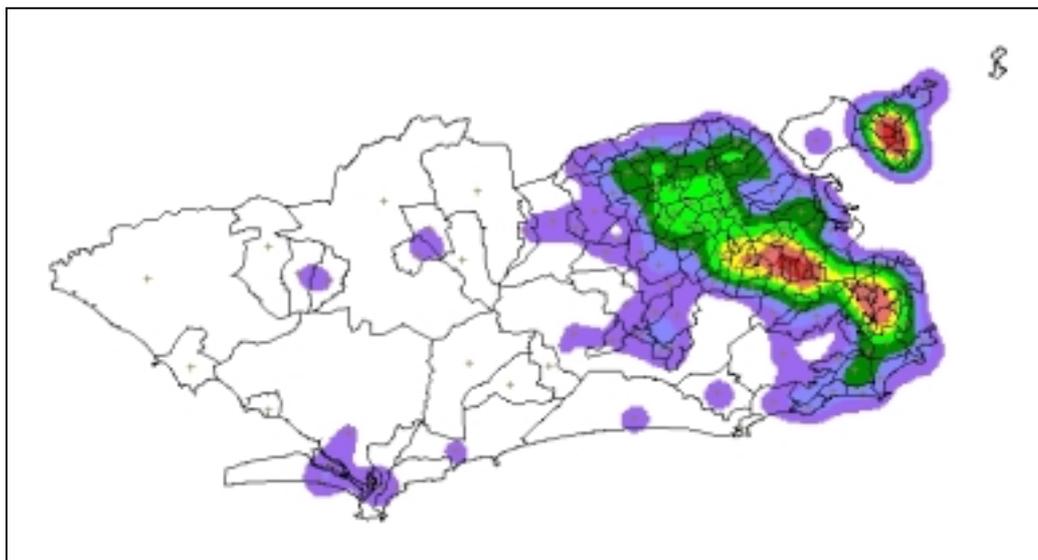


Fig. 5 – Mapa gerado pelo SPRING através do Estimador de Densidade de Kernel

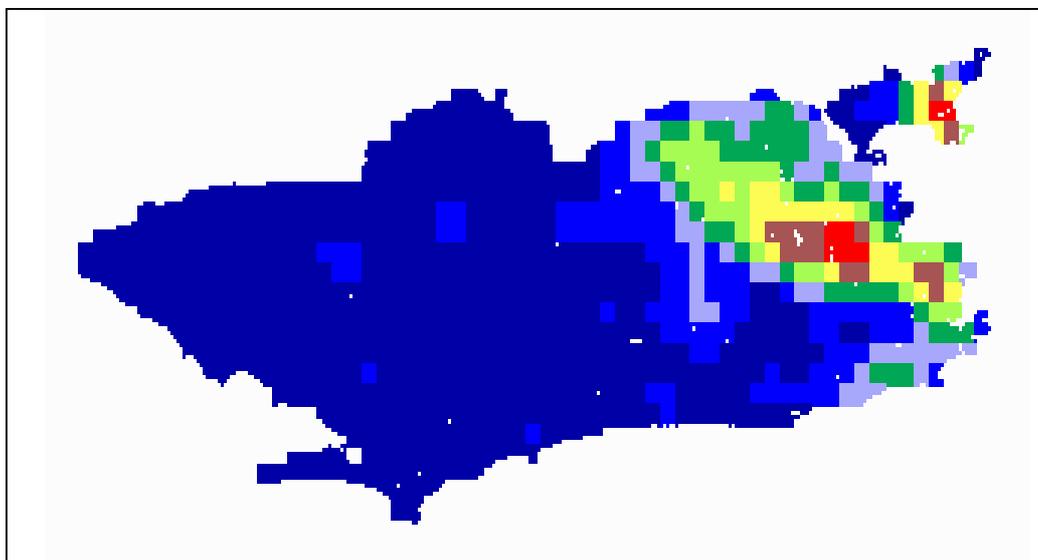


Fig. 6 – Mapa gerado pelo INFO-MAP através do Estimador de Densidade de Kernel

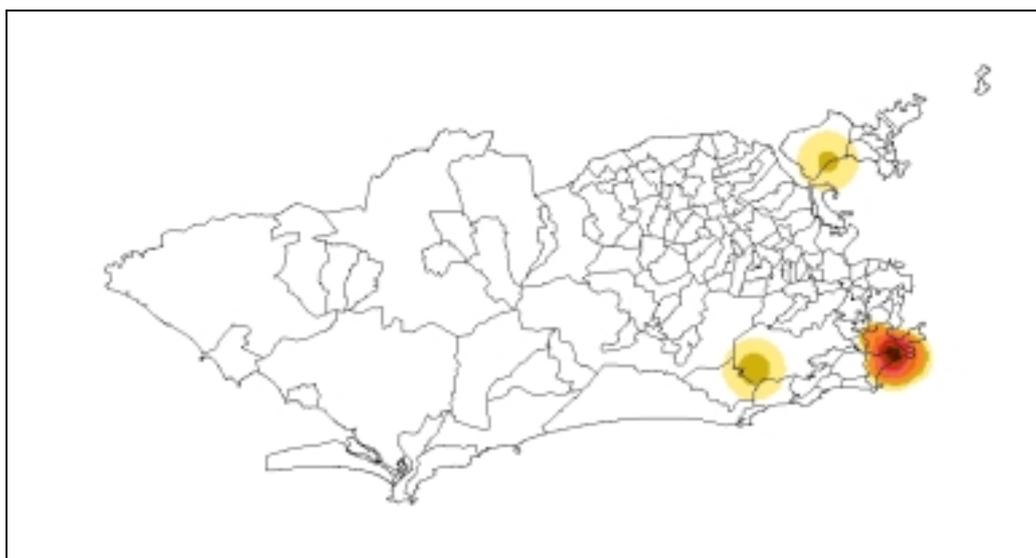


Fig. 7 – Mapa gerado pelo GAM para Apgar bom

Na comparação entre os mapas para a variável Apgar, pode-se perceber que o Estimador de Kernel gerado pelos programas SPRING e INFO-MAP apresentam resultados muito semelhantes, considerando que as cores de intensidade dos mapas são similares, aumentando gradativamente dos tons azulados até a máxima intensidade representada pela cor vermelha. Com isso, observa-se que as regiões com maiores valores de Apgar estão concentradas na porção leste do mapa, correspondendo aos bairros de parte da zona sul e norte do município, incluindo a Ilha do Governador (Figuras 5 e 6).

Como o Estimador de Kernel é um estimador de densidade de padrões, já era esperado que a técnica identificasse os clusters na porção onde a concentração de amostras é maior. Sendo assim, pode-se perceber que o estimador trabalha somente sobre as áreas onde a densidade de amostragem é grande, identificando sobre estas áreas, as maiores incidências da variável.

Já no GAM, o resultado foi bem diferente, apresentando áreas bem mais restritas. O algoritmo localizou um forte cluster na região da zona sul (Botafogo, Urca, Leme, Copacabana) e um foco menos intenso nas regiões de São Conrado, Alto da Boa Vista e Ilha do Governador. Considerando que um alto valor de Apgar associa-se a classes mais favorecidas da sociedade, estes resultados apresentam-se coerentes (Figura 7).

## 6.2. Nascidos Vivos com Baixo Peso

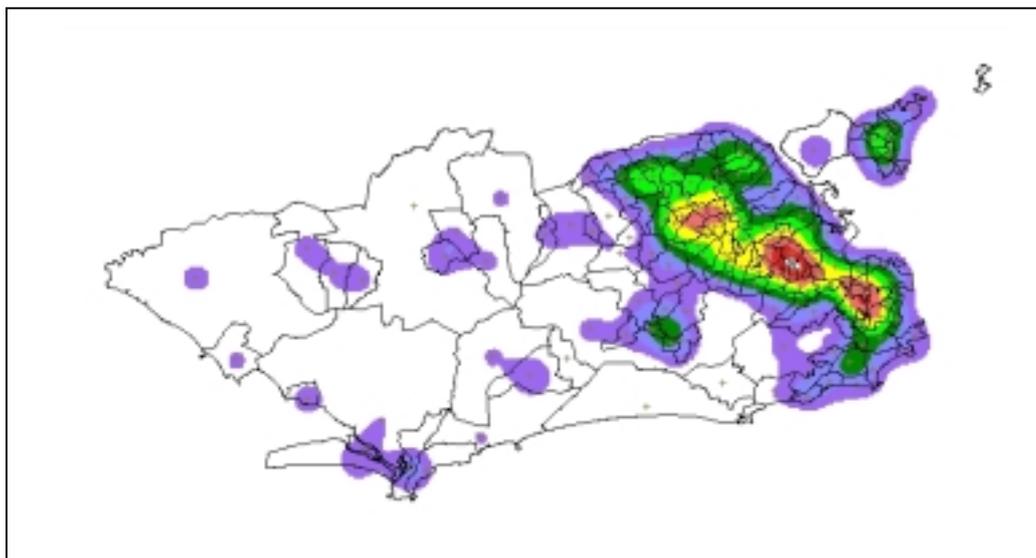


Fig. 8 – Mapa gerado pelo SPRING através do Estimador de Densidade de Kernel

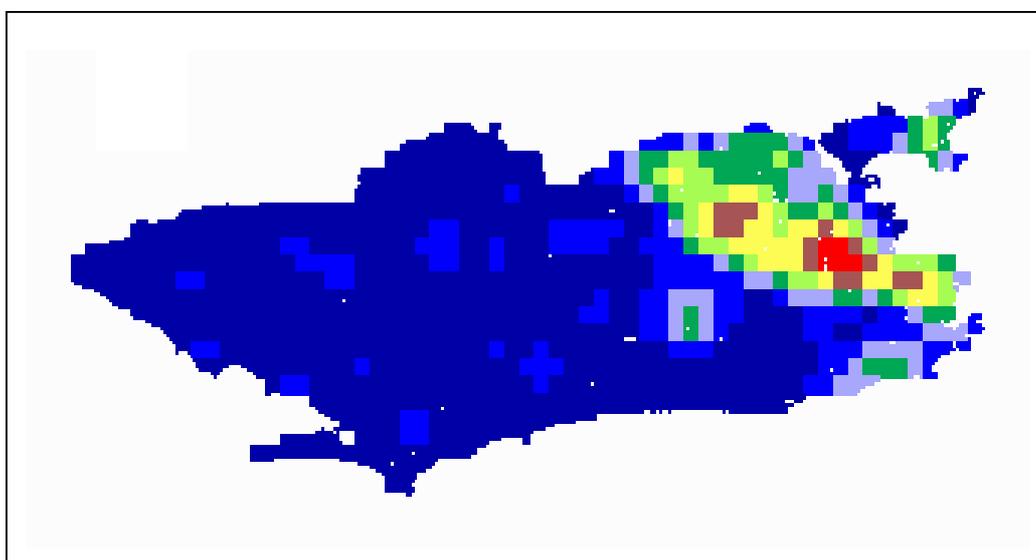


Fig. 9 – Mapa gerado pelo INFO-MAP através do Estimador de Densidade de Kernel



Fig. 10 – Mapa gerado pelo GAM para Nascidos Vivos com Baixo Peso

Na comparação dos mapas gerados a partir da técnica Estimador de Kernel para Nascidos Vivos com Baixo Peso, os resultados dos dois programas foram bem similares, como ocorrido no caso do Apgar. A escala de cores variando de tons azulados (menor incidência) ao vermelho (maior incidência) é a mesma para todas as variáveis. Como já descrito no item anterior, o estimador de Kernel identifica regiões onde a amostragem é mais densa e neste caso, aparecem três áreas de maior incidência, entre a região sul e norte do município, não incluindo a Ilha do Governador, como no caso do Apgar (Figuras 8 e 9).

No mapa gerado pelo GAM, visualiza-se somente um cluster, onde a cor vermelha apresenta maior incidência e está concentrada aproximadamente nos bairros Estácio, São Cristovão, Mangueira, Vila Isabel, entre outros. Aparentemente este resultado também apresenta coerência, devido a estes bairros serem menos favorecidos, podendo estar associado a proporção de nascidos vivos com baixo peso (Figura 10).

### 6.3. Proporção de Cesáreas

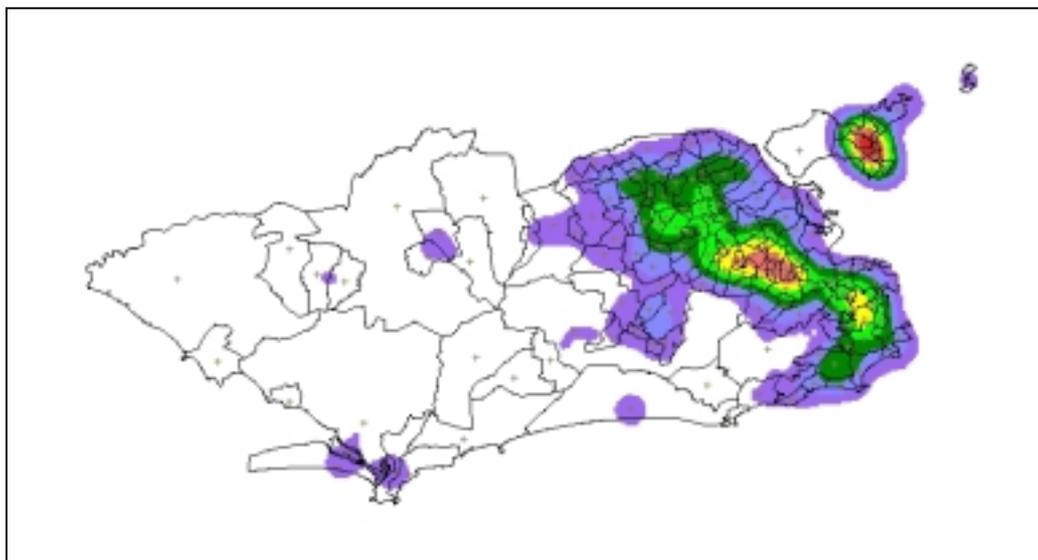


Fig. 11 – Mapa gerado pelo SPRING através do Estimador de Densidade de Kernel

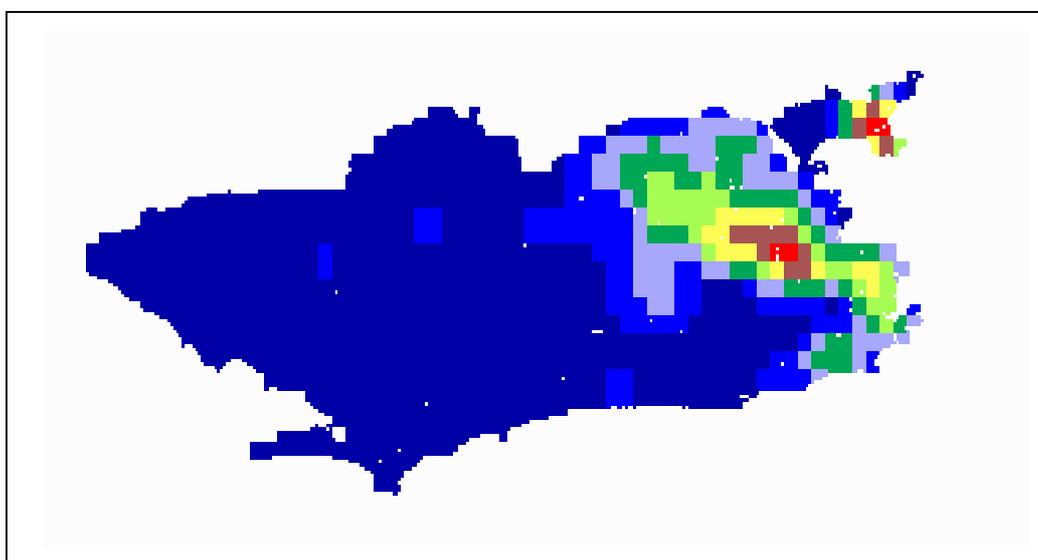


Fig. 12 – Mapa gerado pelo INFO-MAP através do Estimador de Densidade de Kernel



Fig. 13 – Mapa gerado pelo GAM para Proporção de Cesáreas

Para a variável Proporção de Cesáreas, percebe-se que na técnica Estimador de Kernel, para os dois programas, aparecem dois principais focos de incidência, representados pela cor vermelha; um na região centro-norte e outro foco na região da Ilha do Governador, diminuindo a intensidade do cluster conforme as cores gradam para o azul (Figura 11).

O GAM, para esta variável, identificou pequenos clusters espalhados pelo município, cuja cor vermelha representa as áreas de maior incidência e correspondem aproximadamente aos bairros de Botafogo, Urca, Leme, Lagoa, Barra da Tijuca na zona sul e Maracanã, Méier, Campo dos Afonsos e alguns bairros da Ilha do Governador (Figura 13). O Brasil é o país com maior número de partos por cesáreas no mundo, sendo que este tipo de parto está relacionado a classes mais favorecidas e mulheres com maior nível de instrução, refletindo uma situação de exposição ao risco desnecessário, por parte da sociedade brasileira feminina.

#### 6.4. Proporção de Mães Adolescentes

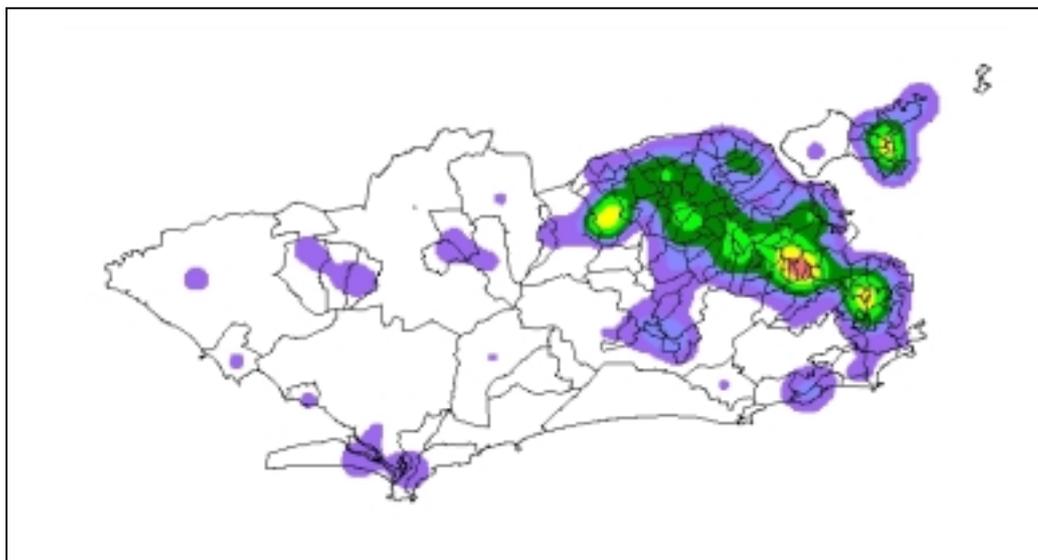


Fig. 14 – Mapa gerado pelo SPRING através do Estimador de Densidade de Kernel

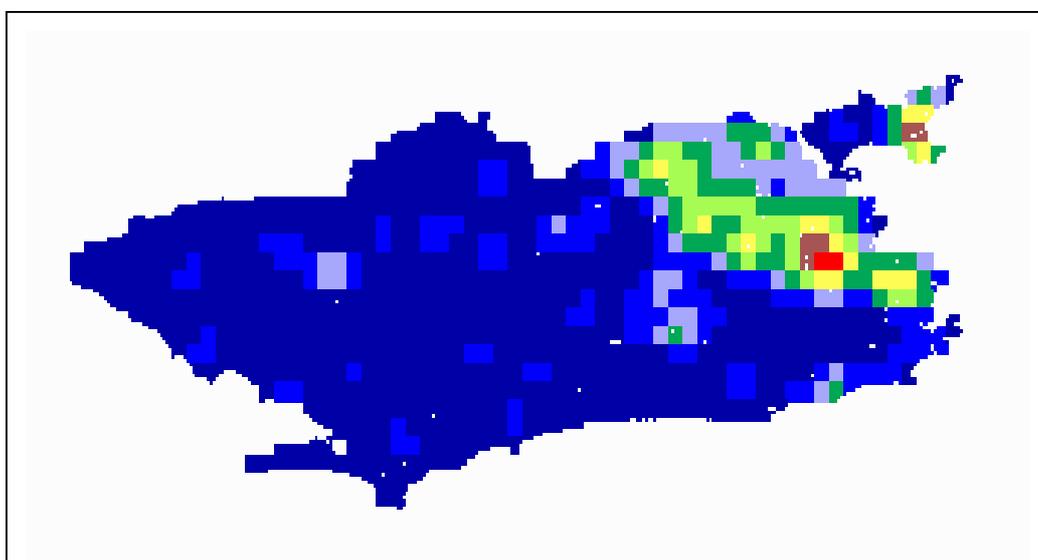


Fig. 15 – Mapa gerado pelo INFO-MAP através do Estimador de Densidade de Kernel

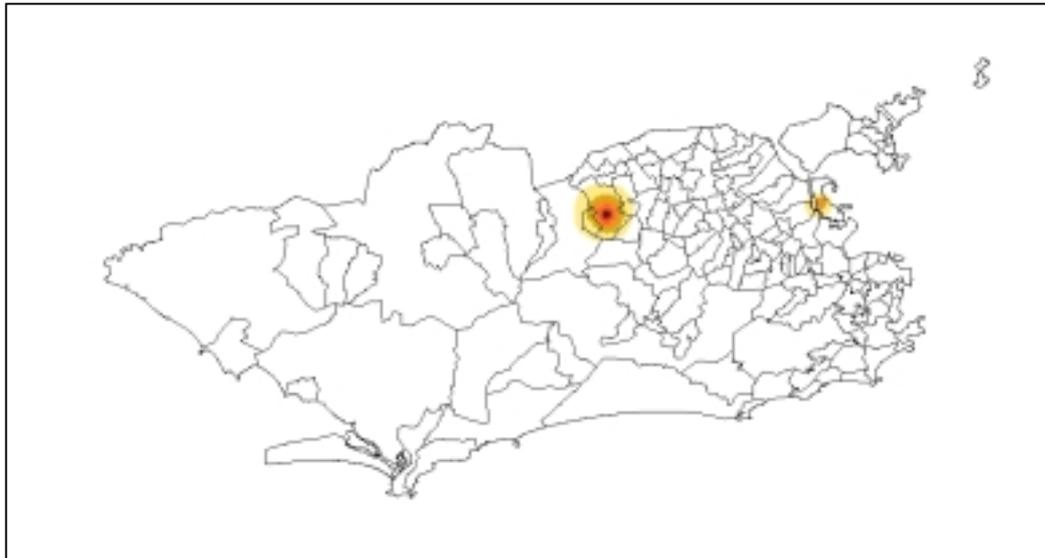


Fig. 16 – Mapa gerado pelo GAM para Proporção de Mães Adolescentes

De acordo com o já descrito anteriormente a respeito do Estimador de Kernel, para a variável Proporção de Mães Adolescentes, obteve-se o mesmo padrão nos mapas, tanto para o SPRING quanto para o INFO-MAP, aparecendo um cluster na região onde a densidade de amostragem é maior, ou seja na zona leste do mapa. Neste caso, somente dois pequenos focos de alta incidência (cor vermelha) foram localizados, concentrando-se aproximadamente nos mesmos locais gerados pela variável Proporção de Cesáreas, porém não havendo correlação entre estas variáveis (Figuras 14 e 15).

Analisando o mapa gerado pelo GAM, pode-se observar que, entre as variáveis, este foi o mapa que apresentou menor quantidade de clusters, estando o maior foco de incidência concentrado nos bairros Jardim Sulacap, Deodoro, Grumari e um pequeno cluster localizado na região do Galeão (Figura 16). Considerando que o GAM identifica os clusters somente em áreas onde a incidência de uma amostra e suas vizinhanças é alta, este resultado é aparentemente coerente, quando comparado ao trabalho de d’Orsi e Carvalho (1998), que apresentam valores altos nessas áreas ou próximas destas.

### 6.5. Proporção de Mães com Escolaridade acima de 2º grau

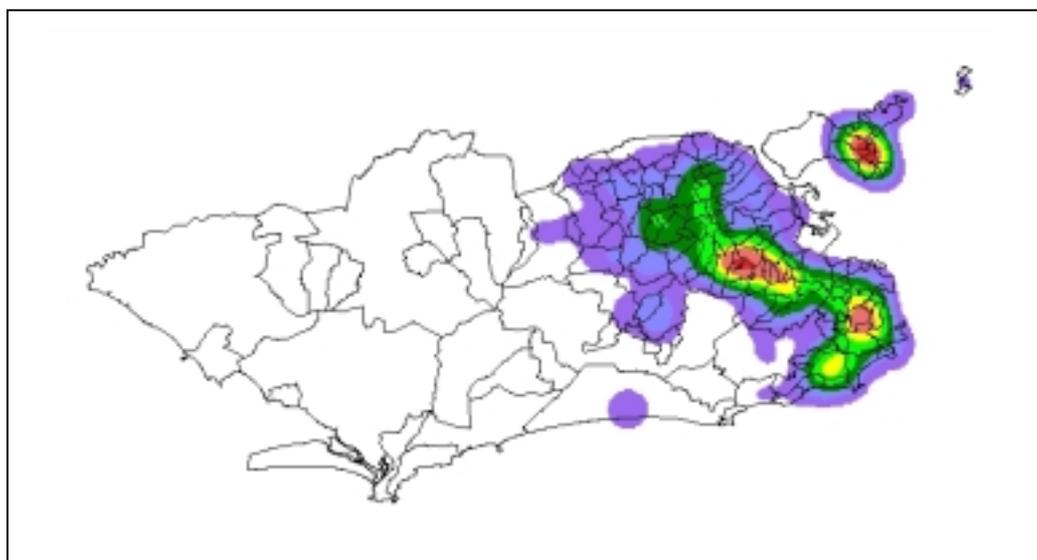


Fig. 17 – Mapa gerado pelo SPRING através do Estimador de Densidade de Kernel

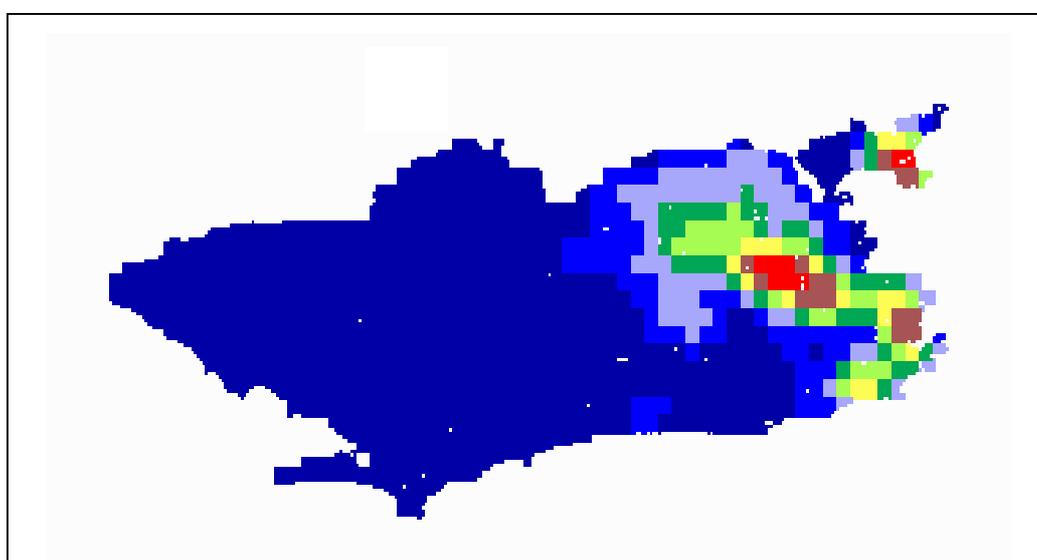


Fig. 18 – Mapa gerado pelo INFO-MAP através do Estimador de Densidade de Kernel

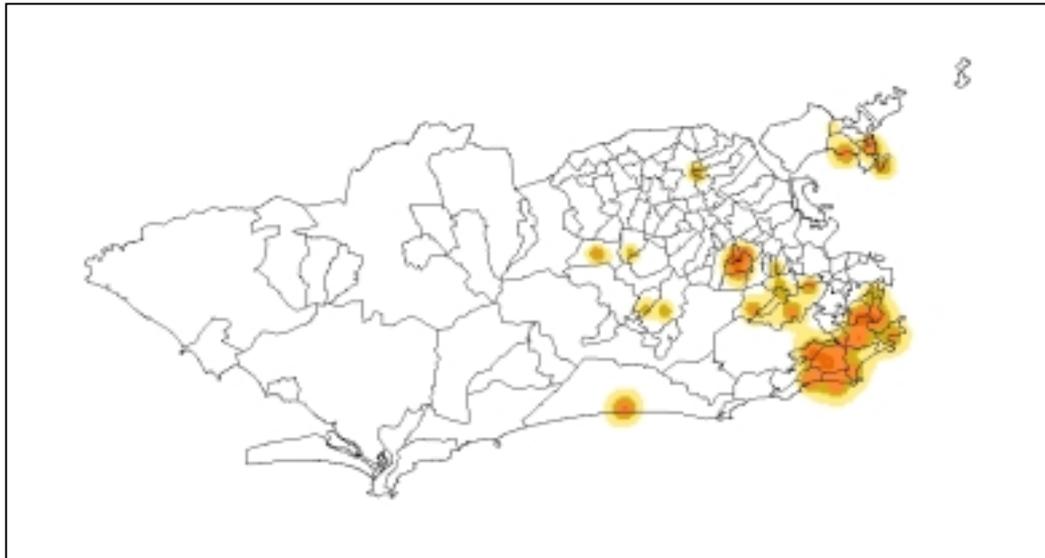


Fig. 19 – Mapa gerado pelo GAM para Proporção de Mães com Escolaridade acima de 2º grau

Nesta variável, para o Estimador de Densidade de Kernel, em ambos os mapas, foram identificadas três áreas com alta incidência (cor vermelha) nas regiões sul e norte, inclusive na Ilha do Governador (Figuras 17 e 18). Estes mapas mostraram-se muito semelhante aos demais mapas gerados por esta técnica, mesmo para diferentes variáveis, porém como já discutido, apresentam o mesmo padrão em função da densidade de amostragem.

Já o GAM localizou vários clusters para a variável Mães com Escolaridade acima de 2º grau, sendo que há um grande cluster na zona sul, ou seja, na região mais nobre do município e também outros pequenos clusters identificados nos bairros da Tijuca, Maracanã, Méier, Campo dos Afonsos, Ilha do Governador, entre outros (Figura 19). Estes resultados apresentam-se coerentes, considerando que o nível de instrução da mãe associa-se às classes sociais mais favorecidas, como pode ser concluído pelo mapa.

## 7. CONCLUSÕES

De acordo com os mapas apresentados no item anterior, gerados a partir do Estimador de Kernel e da Máquina de Análise Geográfica, pode-se concluir que para esta área de estudo os resultados obtidos através do GAM foram muito mais interessantes.

Como o estimador de Kernel trabalha com densidade de amostragem, ambos os programas identificaram clusters praticamente na mesma região (zona leste do mapa), justamente onde a quantidade de bairros e centróides associados é muito maior que em qualquer outra parte do município. Portanto, somente nestas áreas é que o estimador considera o valor da incidência, apresentando mudanças sutis, ainda que para diferentes variáveis analisadas, como pôde ser visto nos mapas apresentados.

O Estimador de Densidade de Kernel testado para os programas SPRING e INFO-MAP apresentaram resultados muito semelhantes. Considerando esta técnica uma nova opção do SPRING (ainda em fase de teste), esta mostrou-se eficiente para esse estudo, além disso, possuindo uma saída visual muito superior ao do INFO-MAP.

Os resultados obtidos através do GAM, mostraram-se coerentes à realidade sócio-econômica do município. Talvez por ser um identificador de áreas muito restritas de incidência, podem ocorrer eventuais perdas de informações importantes, em locais onde existe um alto valor de incidência, mas que o GAM não identifica como cluster pois a sua vizinhança não apresenta valores significativos. Apesar disso, os resultados foram satisfatórios, comprovando que o GAM é uma boa ferramenta para análise espacial exploratória de dados.

Por fim, pode-se concluir que este tipo de análise sócio-ecônômica é muito difícil de ser realizada, devido a dificuldade de tratamento dos dados, principalmente num município tão heterogêneo como o Rio de Janeiro, onde os contrastes sociais são enormes, confundindo a manipulação espacial dos dados.

## 8. REFERÊNCIAS BIBLIOGRÁFICAS

- BAILEY, T.; GATRELL, A. **Interactive Spatial Data Analysis**. Longman Scientific and Technical, London, 1995.
- D'ORSI, E.; CARVALHO, M.S. Perfil de nascimentos no Município do Rio de Janeiro: uma análise espacial. **Cad. Saúde Pública**, Rio de Janeiro, 14(2): 367-379, Abr-Jun, 1998.
- <<http://www.ccg.leeds.ac.uk/smart/gam>>