

Recebido em 11/2011. Aceito para publicação em 03/2012.

ANÁLISE DO PADRÃO ESPACIAL DA MORTALIDADE INFANTIL UTILIZANDO GEOESTATÍSTICA

ANALYSIS OF SPATIAL PATTERN OF INFANT MORTALITY USING GEOSTATISTICS

João Domingos Scalon¹
Simone Soares Lima Jardim²
Gérson Rodrigues Santos³
Denismar Alves Nogueira⁴

RESUMO: A mortalidade infantil é considerada um importante indicador da situação de bem-estar social de uma população. Este artigo teve como objetivo analisar a distribuição espacial da mortalidade infantil, utilizando as técnicas de geoestatística clássica e geoestatística baseada em modelos. Foram utilizados dados de contagem de mortes infantis, por setor censitário, do município de Alfenas, MG, no período de 2000 a 2004. O critério de Akaike permitiu selecionar a geoestatística baseada em modelos como a metodologia mais adequada. A metodologia selecionada permitiu construir um mapa de predição da mortalidade infantil. Os resultados demonstraram que a geoestatística baseada em modelos é uma metodologia poderosa para a análise da distribuição espacial da mortalidade infantil.

Palavras-chave: epidemiologia; saúde pública; estatística espacial.

ABSTRACT: Infant mortality is considered an important indicator of well-being of a population. The main aim of this paper was to analyze the spatial distribution of infant mortality using techniques based on both classical geostatistics and model-based geostatistics. Count data was used of infant deaths for census sectors in the city of Alfenas, MG from 2000 to 2004. The Akaike criteria, which allowed selecting the model-based geostatistics, were the best model. The selected methodology allowed building a prediction map of the infant mortality. The results showed that model-based geostatistics are a powerful methodology for analyzing the spatial distribution of infant mortality.

Keywords: epidemiology; public health; spatial statistics.

¹ Doutor em Estatística – Sheffield University e Professor Associado da Universidade Federal de Lavras - UFLA. E-mail: scalon@dex.ufla.br.

² Mestre em Estatística e Experimentação Agropecuária - UFLA. E-mail: simonesoaresjardim@yahoo.com.br.

³ Doutor em Estatística e Experimentação Agropecuária - UFLA e Professor Adjunto da Universidade Federal de Viçosa - UFV. E-mail: gerson.santos@ufv.br.

⁴ Doutor em Estatística e Experimentação Agropecuária - UFLA e Professor Adjunto da UFLA. E-mail: denismar@unifal-mg.edu.br.

1. INTRODUÇÃO

A mortalidade infantil, entendida como o número de óbitos em crianças menores de um ano, nascidas vivas, é considerada um dos mais importantes indicadores para medir a qualidade de vida de uma população (COSTA, 2008). Os dados de mortalidade infantil podem ser usados para vários tipos de análises importantes para o auxílio ao planejamento, gestão e avaliação das políticas de saúde de uma região.

Vários pesquisadores defendem que a partir do momento em que houver uma maior preocupação em melhorar a assistência à mulher e ao recém-nascido, em sala de parto e berçário, as condições socioeconômicas da população com baixa renda, o acesso à educação, ao saneamento e aos serviços básicos de saúde, os índices de mortalidade infantil passarão a diminuir (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, 2008; DUARTE *et al.*, 2002; SIMÕES, 2002). Assim, pode-se dizer que a mortalidade infantil é uma variável que pode ser controlada. Para tal, é importante que sejam desenvolvidos projetos focados nas políticas sociais que incluem ações sociais dirigidas às gestantes e aos recém-nascidos (FARAH, 2005). São necessários, também, estudos epidemiológicos que promovam o conhecimento da distribuição espacial da mortalidade infantil, para permitir que sejam conhecidas as regiões onde há uma necessidade maior de intervenção positiva por parte dos órgãos de saúde no bem-estar das populações que vivem nessas regiões (SILVA *et al.*, 2011).

Sabe-se que trabalhos que se baseiam na estrutura espacial dos dados favorecem, consideravelmente, o entendimento de eventos, dado o fato de

que, ao incorporar a dimensão espacial, não apenas trabalha-se com a delimitação geográfica do problema, mas possibilita-se, também, fazer uma análise integrada, que inclui o estudo da dependência no espaço (KEEN *et al.*, 1987; LEAL; SZWARCOWALD, 1997). Por essa razão, a estatística espacial tem se tornado cada vez mais utilizada no meio científico, já que propõe a inclusão direta das coordenadas geográficas no estudo de um fenômeno.

A classificação mais utilizada, para descrever o problema da análise e modelagem espacial, considera três tipos de dados cujos métodos não são, necessariamente, mutuamente excludentes. São eles: padrões pontuais, dados de área e geoestatística (CRESSIE, 1993). Padrões pontuais são expressos por intermédio de ocorrências identificadas como pontos localizados no espaço, como, por exemplo, os dados são as coordenadas dos casos de uma doença. Dados de área referem-se, por exemplo, a contagens de casos de uma doença em unidades de análises, usualmente delimitadas por polígonos fechados, tais como setores censitários ou municípios. A geoestatística trabalha com dados que são distribuídos continuamente em uma região (CRESSIE, 1993).

A análise espacial de dados de mortalidade infantil é, usualmente, realizada a partir de dados agregados por áreas (LEAL; SZWARCOWALD, 1997; SILVA *et al.*, 2011). Isto é devido ao fato de que a maioria dos dados de mortalidade infantil se apresenta em forma de contagens por setores censitários ou bairros. Embora os mapas resultantes possam ser facilmente interpretáveis, com a agregação pode-se perder uma estrutura espacial mais detalhada dos dados e, conseqüentemente,

as variações espaciais, em pequena escala, podem não ser detectáveis (SHIMAKURA *et al.*, 2001). Quando é possível obter as coordenadas dos casos de mortalidade infantil, a análise baseada em processos pontuais parece ser uma ferramenta mais adequada e poderosa para estimar medidas de risco de mortalidade infantil, em uma particular região (SHIMAKURA *et al.*, 2001).

Uma alternativa, para estimar os casos de mortalidade infantil de tal modo que essas estimativas possam variar continuamente na região de interesse, é utilizar a geoestatística, que pode ser abordada tanto na forma clássica como na forma baseada em modelos. A geoestatística clássica não pressupõe qualquer distribuição para o mecanismo gerador dos dados e o modo e o grau de variabilidade espacial da variável aleatória pode ser expresso por um semivariograma (CRESSIE, 1993; ISAAKS; SRIVASTAVA, 1989). A geoestatística baseada em modelos, ao contrário, incorpora, explicitamente, um mecanismo gerador dos dados, tal como a distribuição de Poisson que pode ser associada a estruturas que descrevam o padrão espacial para dados de contagens. Esse tipo de abordagem é um caso particular dos modelos lineares generalizados mistos, em que as variáveis aleatórias latentes derivam de um processo espacial (DIGGLE; RIBEIRO JR., 2007).

Outra particularidade, na análise espacial da mortalidade infantil, é que, em geral, não se trabalha com as contagens dos casos nos recortes da região, preferindo-se utilizar a taxa de mortalidade infantil, que é dada pela relação entre o número de óbitos de crianças menores de um ano e o número de nascidos vivos, em determinado local e período, calculado na base de mil nascidos

vivos (KEEN *et al.*, 1987; LEAL E SZWARCOWALD, 1997; SILVA *et al.*, 2011). Apesar de desejável, nem sempre é possível obter a taxa como, por exemplo, na ausência da informação sobre o número de nascidos vivos, nos recortes.

Tendo em vista o que foi abordado anteriormente, o objetivo deste artigo é comparar os mapas das estimativas para o número de casos de mortalidade infantil, na região urbana de Alfenas, MG, gerados pelos métodos da geoestatística clássica e geoestatística baseada em modelos e, assim, identificar o padrão espacial subjacente aos dados na região.

2. MATERIAL E MÉTODOS

2.1 Dados

O estudo foi conduzido no município de Alfenas, MG, situado, geograficamente, a 21°25'45" de latitude Sul e 45°56'50" de longitude Oeste. Ocupa uma área de 849,2 km² de extensão, com altitude máxima de 888 metros acima do nível do mar. Os dados, provenientes do Sistema de Informação sobre Mortalidade (SIM), são as contagens de mortes infantis, ocorridas em 68 setores censitários da região urbana desse município, entre os anos de 2000 e 2004. O cartograma dos 68 setores censitários e os limites da área considerada no estudo, com centróides representando as coordenadas espaciais da informação constante em cada setor censitário, encontram-se ilustrados na Fig. 1.

Além das contagens de mortes infantis ocorridas, foram ainda pesquisadas seis variáveis, utilizando-se o banco de dados do IBGE do ano 2000, para os 68 setores censitários. São elas: "número de mulheres

em idade fértil”, “número de mulheres em idade de risco gestacional”, “número de mulheres em idade fértil e analfabetas”, “número de residências com mais de seis moradores”, “soma da renda mensal das mulheres”, “soma da renda mensal dos homens” e “densidade demográfica do setor censitário”. Todos os dados foram coletados por um dos autores do trabalho (D.A.N).

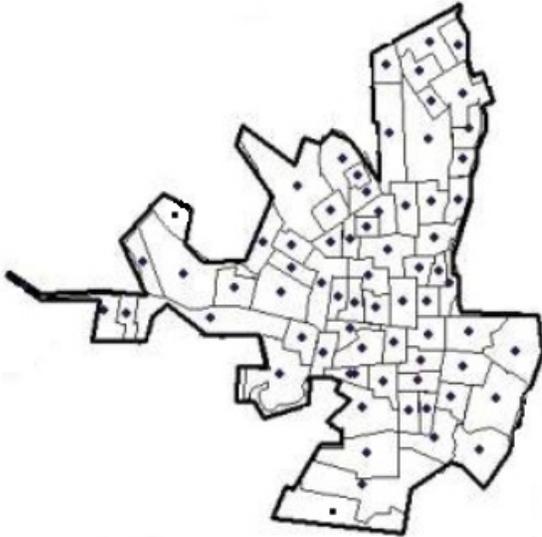


Fig. 1 - Cartograma do perímetro urbano da cidade de Alfenas, MG, com os 68 setores censitários considerados no estudo, com os limites de área e respectivos centroides.

2.2 Análise estatística

Inicialmente, foi conduzida uma análise exploratória dos dados sem considerar a posição geográfica. Assim, foram aplicadas estatísticas descritivas para verificar características da distribuição dos casos de mortalidade infantil e o teste de aderência de qui-quadrado, para verificar a aderência dos dados à distribuição de Poisson (SIEGEL, 1975). Em seguida, foi utilizado o teste de correlação de Spearman para verificar a existência de correlação entre as seis variáveis e a mortalidade infantil (SIEGEL, 1975).

Para conduzir a análise geoestatística clássica, consideraram-se as respostas Y_i , $i=1, \dots, n$, medidas nas localizações x_i , $i=1, \dots, n$ como sendo as contagens das mortes infantis, na região estudada $A \subset \mathbb{R}^2$, sendo que nesse caso, A corresponde à região urbana do município de Alfenas – MG. O conjunto de respostas $Y(\cdot) = \{Y(x): x \in A \subset \mathbb{R}^2\}$, conhecido por processo de medição (DIGGLE E RIBEIRO JR., 2007), contém o número de mortes infantis em cada localização amostrada.

Admitindo a hipótese de estacionariedade espacial, a dependência espacial dos dados pode ser caracterizada por meio de um semivariograma empírico (ou experimental) cujo estimador é dado por

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [Y(x) - Y(x+\mathbf{h})]^2$$

em que $N(\mathbf{h})$ é o número de pares de valores medidos e $Y(x)$ e $Y(x+\mathbf{h})$ são os valores medidos na posição x , separados por um vetor distância \mathbf{h} (CRESSIE, 1993).

A partir do variograma experimental obtido, um modelo (variograma teórico) pode ser ajustado aos pontos. Existem vários modelos teóricos que podem ser ajustados, tais como o linear, esférico, exponencial, gaussiano e Matérn (CRESSIE, 1993; ISAACS; SRIVASTAVA, 1989). Esses modelos apresentam, como parâmetros: o patamar, o alcance e o efeito pepita. Cada um desses parâmetros representa características próprias. Assim, o parâmetro patamar representa o valor máximo no qual $\gamma(\mathbf{h})$ se estabiliza, tornando-se, a partir daí, independente à variação da distância \mathbf{h} . O parâmetro alcance é o valor de \mathbf{h} , para o qual essa estabilização ocorre e, finalmente, o parâmetro efeito pepita é um valor positivo, do qual $\gamma(\mathbf{h})$ se aproxima quando \mathbf{h}

tende para zero (CRESSIE, 1993). Esses parâmetros, para os diversos modelos, foram estimados usando o método de mínimos quadrados ordinários (DIGGLE; RIBEIRO JR., 2007).

A escolha do melhor modelo teórico ajustado foi feita utilizando o Critério de Informação de Akaike (AIC). Dentre os vários modelos comparados, é considerado o melhor aquele que tiver o menor valor de AIC (CRESSIE, 1997).

O modelo de semivariograma escolhido permite interpretar a estrutura de correlação espacial a ser utilizada no procedimento de estimativas dos valores da variável estudada em locais não amostrados, conhecido como krigagem. Neste trabalho, foi utilizado o método de interpolação conhecido como krigagem ordinária, que considera que a média oscila localmente, assim, estacionaridade é limitada a áreas locais (ISAAKS; SRIVASTAVA, 1989).

Para conduzir a análise da geoestatística baseada em modelos, assume-se que os elementos do processo de medição, condicionais ao processo espacial $S(\cdot)$, são realizações de variáveis aleatórias Poisson independentes e identicamente distribuídas (DIGGLE; RIBEIRO JR., 2007). Considera-se, também, que o processo gaussiano $S(\cdot)$, conhecido como processo sinal, apresenta $E[S(x)] = 0$, $\text{Var}[S(x)] = \sigma^2$ e função de correlação dada por $\rho(h) = C\{S(x), S(x')\}$, em que $h = \|x - x'\|$ é a distância euclidiana entre x e x' . Assim, tem-se que: $Y_i|S(x_i) \sim \text{Poisson}(\lambda_i)$, $i = 1, \dots, n$, em que a média de mortes infantis na localização x_i amostrada, condicional a $S(\cdot)$ é $\lambda_i = E[Y_i|S(x_i)]$. O modelo de preditores é, então, dado por $h(\lambda_i) = \alpha + S(x_i)$, sendo $h(\cdot)$ a

função de ligação logarítmica dada por $\log \lambda_i = \alpha + S(x_i)$. Assim, sendo o número de mortes infantis $Y(x_i)|S(x_i) = s$, com distribuição de Poisson com média $\exp[\alpha + S(x_i)]$, $i = 1, \dots, n$, tem-se que

$$P[Y(x) = y | S(x) = s] = \frac{e^{-e^{\alpha+S(x_i)}} \left(e^{\alpha+S(x_i)} \right)^{y(x_i)}}{y(x_i)!}.$$

A função de verossimilhança, para esse modelo, não tem solução explícita e, assim, utilizaram-se métodos de Monte Carlo via Cadeias de Markov (MCMC), para obter uma solução aproximada. Por intermédio desse procedimento, obteve-se uma cadeia com 50000 iterações, tendo sido descartadas as 1000 primeiras iterações (*burn-in*) e sendo armazenada uma amostra a cada 20 amostras geradas, logo, obteve-se um total efetivo de 2500 amostras (DIGGLE; RIBEIRO JR., 2007). As cadeias obtidas, para cada parâmetro foram analisadas para verificar a convergência do algoritmo de MCMC. Para a simulação condicional pelo algoritmo MCMC, os valores iniciais, para comporem o modelo, foram obtidos a partir da análise geoestatística clássica (DIGGLE; RIBEIRO JR., 2007).

O modelo ajustado foi utilizado para obter as estimativas dos casos de mortalidade infantil em locais não amostrados, utilizando krigagem clássica (DIGGLE; RIBEIRO JR., 2007). Para comparar os valores preditos da variável mortalidade infantil obtidos pela geoestatística clássica e pela geoestatística baseada em modelos, utilizou-se o AIC.

Todas as análises do estudo foram realizadas utilizando-se o *software* R (R DEVELOPMENT CORE TEAM, 2012) e as bibliotecas *geoR* (DIGGLE; RIBEIRO JR., 2001) e *geoRglm* (CHRISTENSEN *et al.*,

2002).

3. RESULTADOS E DISCUSSÃO

O Brasil está incluído no grupo dos países que têm reduzido, consideravelmente, as taxas de mortalidade infantil, mas ainda há muito que fazer, visto que a divergência entre regiões brasileiras é muito grande (IBGE, 2008). Na região Nordeste, por exemplo, a taxa de mortalidade infantil (por mil nascidos vivos) é de 52,5 enquanto que na região sul, não passa de 17,1 (DUARTE *et al.*, 2002). Existem, também, indícios de que as taxas de mortalidade infantil podem ser diferentes mesmo em escalas menores como estados e cidades e até mesmo em bairros (KEEN *et al.*, 1987; LEAL; SZWARCOWALD, 1997). Assim, é importante que os dados georeferenciados de mortalidade infantil de uma determinada região sejam analisados por meio de técnicas de estatística espacial que permitam uma boa resolução espacial sobre o problema. Nesse sentido, este trabalho procurou analisar contagens do número de casos de mortalidade infantil, por setor censitário, para obter mapas das estimativas desses casos em locais não

amostrados obtidos pelos métodos da geoestatística clássica e geoestatística baseada em modelos.

O estudo foi realizado na região urbana da cidade de Alfenas, MG, localizada na região sul do estado de Minas Gerais e que tem, aproximadamente, 100 mil habitantes. A cidade é um importante pólo agro-industrial e universitário do estado e que observou, no período de abrangência deste estudo (2000-2004), uma queda acentuada nos índices de mortalidade infantil, caindo de 31,4 para 7,9 (PREFEITURA MUNICIPAL DE ALFENAS, 2011). O objetivo básico foi utilizar os métodos estatísticos descritos anteriormente para identificar o padrão espacial subjacente aos dados de mortalidade infantil no perímetro urbano da cidade.

Durante a primeira fase do estudo, foi conduzida uma análise exploratória das variáveis envolvidas no estudo, sem considerar a informação espacial dos setores censitários. Assim, utilizaram-se apenas estatísticas descritivas clássicas. A Tabela 1 e a Fig. 1 apresentam uma ideia geral da distribuição dos casos de mortalidade infantil na cidade de Alfenas, MG.

Tabela 1 - Análise descritiva para os dados de mortalidade infantil de 68 setores censitários da cidade de Alfenas –MG, ocorridos entre os anos de 2000 e 2004

MEDIDAS DESCRITIVAS	VALORES
Casos de mortes	134
Valor mínimo	0
Valor máximo	7
Média	1,97
Mediana	2
Desvio Padrão	1,74
Coeficiente de variação	0,88
Coeficiente de assimetria	0,82
1º Quartil	1
3º Quartil	3

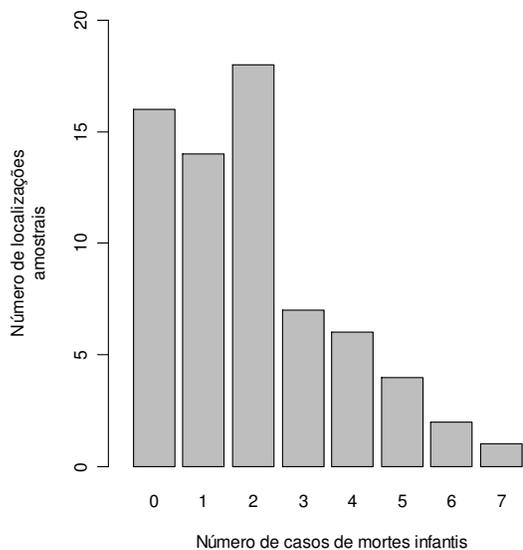


Fig. 2- Distribuição dos casos de óbitos infantis ocorridos em 68 setores censitários da cidade de Alfenas – MG, entre 2000 e 2004.

Pode-se observar, nas Tabela 1 e Fig. 1, que nos 68 setores censitários ocorreram 134 casos de mortalidade infantil. Observa-se, também, que existem muitos setores censitários nos quais não ocorreram óbitos, enquanto em um setor ocorreram sete óbitos. O terceiro quartil mostra que 75%

dos setores censitários tiveram três ou menos casos de óbitos. Pode-se considerar que os setores censitários da região urbana de Alfenas apresentam uma alta variabilidade no número de óbitos, em relação à média, o que pode ser confirmado pelo coeficiente de variação de 88,1%. Esses resultados corroboram estudos conduzidos em outras cidades que mostraram existir diferenças significativas da mortalidade infantil, em pequenas escalas, dentro de um mesmo perímetro urbano (KEEN *et al.*, 1987; LEAL E SZWARCOWALD, 1997). O coeficiente de assimetria de 0,82 sugere uma não simetria dos dados, o que poderia ser característica de uma distribuição de Poisson. Uma inspeção visual da Fig. 2 também indica que o número de casos de mortalidade infantil parece seguir uma distribuição de Poisson, o que foi confirmado pelo teste de aderência de qui-quadrado ($\chi^2 = 11,62$, $P = 0,23$). A Tabela 2 apresenta algumas estatísticas descritivas para as outras variáveis envolvidas no estudo.

Tabela 2 - Estatísticas descritivas para as variáveis: número de mulheres em idade fértil (A); número de mulheres em idade de risco gestacional (B); número de mulheres em idade fértil, analfabetas (C); número de residências com mais de seis moradores (D); soma da renda mensal das mulheres (E); soma da renda mensal dos homens (F) e densidade demográfica (G) nos 68 setores censitários da região urbana, da cidade de Alfenas, MG, entre 2000 e 2004

	Média	Mediana	Mínimo	Máximo	Desvio	Coeficiente
A	259,0	272,5	4,0	370,0	71,51	27,6
B	142,6	146,0	1,0	243,0	40,34	28,3
C	2,8	2,0	0,0	10,0	2,59	92,5
D	27,82	25,0	0,0	63,0	13,66	49,1
E	32750	23110	0,0	122000	26278,9	80,2
F	161700	128600	971	537400	105344,8	65,1
G	902,9	940	17,0	1396,0	258,64	28,6

Observa-se, pelos coeficientes de variação (CV) apresentados na Tabela 2, que as variáveis “Número de mulheres em idade fértil, analfabetas”, “Número de

residências com mais de seis moradores”, “Soma da renda mensal da mulher” e “Soma da renda mensal do homem”, que apresentam uma variabilidade bem alta, enquanto as variáveis “Número de mulheres em idade fértil”, “Número de mulheres em idade de risco gestacional” e “Densidade demográfica do setor censitário” apresentam uma variabilidade menor. Estes resultados indicam que, mesmo em uma cidade de médio

porte, existem diferenças socioeconômicas marcantes entre os diversos setores censitários, o que pode explicar a alta variabilidade dos casos de mortalidade infantil, observada no perímetro urbano da cidade.

A Tabela 3 apresenta os valores do coeficiente de correlação de Spearman para a verificação da correlação existente entre a variável mortalidade infantil e as demais covariáveis.

Tabela 3 - Valores do coeficiente de correlação de Spearman (R_s) para a relação entre cada covariável e a variável mortalidade infantil, nos 68 setores censitários da cidade de Alfenas, MG, no período de 2000 a 2004

COVARIÁVEIS	R_s
Número de mulheres em idade fértil	0,55
Número de mulheres em idade de risco gestacional	0,21
Soma da renda mensal do homem	-0,12
Soma da renda mensal da mulher	-0,09
Densidade demográfica do setor censitário	-0,08
Número de residências com mais de seis moradores	-0,07
Número de mulheres em idade fértil analfabetas	0,03

Observando-se os dados da Tabela 3, pode-se verificar que o “número de mulheres em idade fértil” é a variável mais correlacionada, positivamente, com o número de casos de mortalidade infantil ($R_s = 0,55$, $P < 0,001$), indicando que quanto maior o número de mulheres em idade fértil, maior os casos de mortalidade infantil. O aumento no número de mulheres em idade de risco gestacional também parece aumentar o número de óbitos infantis ($R_s = 0,21$, $P < 0,001$), enquanto uma diminuição da soma da renda mensal do homem ($R_s = -0,12$, $P < 0,001$) e da mulher ($R_s = -0,09$, $P < 0,001$), apesar da baixa correlação, embora significativa, também pode indicar aumento no número de casos de mortalidade infantil. Esses resultados parecem indicar que as

condições socioeconômicas da população são aspectos que influenciam os casos de mortalidade infantil em uma região, o que corrobora diversas pesquisas sobre o assunto (IBGE, 2008; DUARTE *et al.*, 2002; FARAH, 2004; SILVA *et al.*, 2011; KEEN *et al.*, 1987).

Apesar das análises apresentadas anteriormente representarem uma fotografia da mortalidade infantil, na região urbana da cidade de Alfenas, no período de 2000 a 2004, acredita-se que a descrição da distribuição espacial desse fenômeno possa enriquecer a análise em vários aspectos. A incorporação da dimensão espacial permite não apenas delimitar geograficamente o problema, mas também fazer uma análise

integrada que inclui o estudo da dependência no espaço da mortalidade infantil. Além disso, ao se detectar a existência de um componente espacial por meio da estatística espacial, a análise passa a verificar os fatores ou processos que geram a não dependência da distribuição espacial da mortalidade infantil e a identificação de áreas de maior interesse para posteriores investigações (LEAL; SZWARCOWALD, 1997). Assim, na segunda etapa do estudo, foi conduzida a análise da dependência espacial usando geoestatística e, portanto, passou-se a considerar, explicitamente, na análise, a informação espacial dos setores censitários associados aos casos de mortalidade infantil.

Deve-se observar que, para utilizar a geoestatística, os limites de cada setor censitário foram ignorados, visto que essa técnica pressupõe continuidade espacial (CRESSIE, 1993; ISAACS; SRIVASTAVA, 1989). Assim, considerou-se que as mortes infantis poderiam ter ocorrido em qualquer região da área em estudo.

Para analisar a dependência espacial dos casos de mortalidade infantil, foi obtido o semivariograma empírico, ao qual foram ajustados os modelos teóricos Matérn com kappa igual a 0,5 (conhecido como modelo exponencial), Matérn com kappa igual a 1 e o modelo esférico. O modelo mais bem ajustado foi o esférico (AIC = 272,4), enquanto que o modelo exponencial (AIC = 272,9) e o Matérn com kappa igual a 1 (AIC = 272,8) também produziram bons ajustes, mas com qualidade inferior ao primeiro.

A Fig. 3 apresenta o modelo esférico ajustado ao semivariograma empírico dos dados de mortalidade infantil, cujos parâmetros estimados foram $\sigma^2 = 2,598$; $\phi =$

400,5 e $\tau^2 = 0,38$. Assim, tem-se um patamar de 2,598, a partir do qual $\gamma(h)$ é constante e um alcance da dependência espacial dos dados de 400,5 metros, ou seja, pressupõe-se que as mortes ocorridas, até essa distância, são espacialmente dependentes.

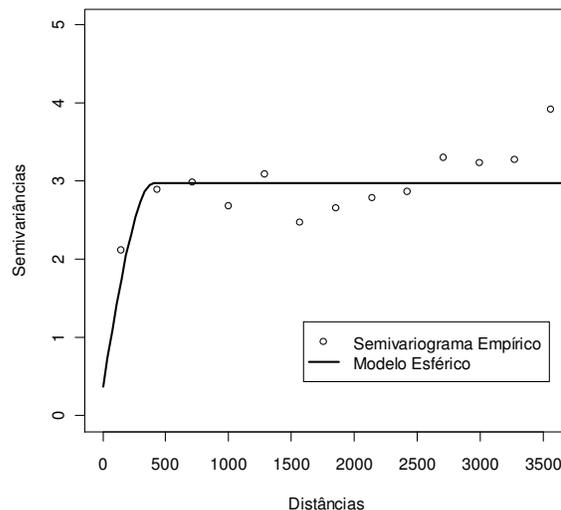


Fig. 3 - Semivariograma empírico (círculos), ajustado pelo modelo esférico (linha contínua).

Tendo em vista a obtenção de um método capaz de simular valores em locais não amostrados, considerando que os dados originais não apresentam uma distribuição Normal, a geoestatística baseada em modelos utiliza-se do método MCMC que, por meio dos valores iniciais $\sigma^2 = 2,598$; $\phi = 400,5$; $\tau^2 = 0,38$ e $\beta = 1,9721$, provenientes do modelo esférico ajustado por meio da geoestatística clássica, procurou encontrar um modelo que se ajustasse aos dados, considerando que estes provêm de uma distribuição de Poisson. Na tentativa do ajuste mais adequado, foram testados os modelos esférico e Matérn, com valores de kappa iguais a 0,5; 1,0; 1,5 e 2,0 e selecionou-se, com o auxílio do critério de Akaike (AIC = -88,84), o modelo Matérn com kappa igual a

0,5 (ou exponencial). A partir do ajuste do modelo selecionado, foram obtidos os parâmetros $\sigma^2 = 0,6595$; $\phi = 109,6954$; $\tau^2 = 0$ e $\beta = 0,4690$, os quais foram utilizados na obtenção de um mapa de predição.

O principal objetivo de um estudo geoestatístico é obter estimativas de valores (neste caso, do número de mortes infantis) em locais não amostrados, considerando o padrão de distribuição espacial e integrar informações a partir de pontos observados. Para tal, a geoestatística oferece uma grande variedade de métodos que podem fornecer estimativas dos casos de

mortalidade infantil em locais não amostrados (fora do centróide do setor censitário) e, assim, possibilitar a construção de uma superfície suavizada da mortalidade infantil na região urbana da cidade. Esses métodos de interpolação são conhecidos como krigagem e dependem do modelo escolhido para a tendência da função aleatória (CRESSIE, 1993; ISAAKS; SZWARDCWALD, 1989; DIGGLE; RIBEIRO JR., 2007). A Fig. 4 mostra os mapas de predição gerados pela geoestatística clássica e pela geoestatística baseada em modelos.

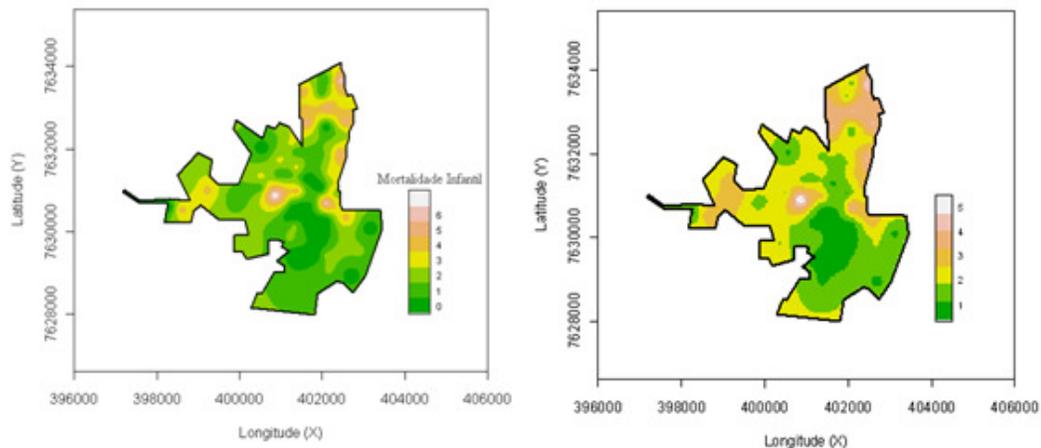


Fig. 4 - Mapas de predição, gerados por meio da geoestatística clássica (à esquerda) e geoestatística baseada em modelos (à direita), para a cidade de Alfenas, MG, entre 2000 e 2004.

Uma inspeção visual da Fig. 4 mostra que os dois métodos apontam a mesma região (em branco) com o maior número de óbitos no perímetro urbano da cidade de Alfenas. Observa-se, também, que o mapa de predição proveniente da geoestatística baseada em modelos parece ser mais discriminativo das regiões com altas e baixas ocorrências de mortalidade infantil.

Deve-se observar que a geoestatística baseada em modelos considera,

explicitamente, na composição do modelo, a distribuição de probabilidade dos dados (neste artigo, a distribuição de Poisson), simula valores a partir deles e, assim, encontra estimativas em locais não amostrados. Isso contribui para a obtenção de estimativas mais precisas, o que não acontece na geoestatística clássica, que usa os dados originais para encontrar valores em locais não amostrados, sem considerar a distribuição probabilística dos dados (DIGGLE; RIBEIRO JR., 2007). Além disso,

embora a normalidade não seja um pré-requisito para a krigagem, é uma propriedade desejável. Krigagem somente vai gerar as melhores estimativas absolutas se a função aleatória se ajusta a uma distribuição normal (CRESSIE, 1993).

Na comparação dos dois métodos, obteve-se um $AIC = -88,84$, para a geoestatística baseada em modelos, enquanto a geoestatística clássica apresentou $AIC = 272,4$. Assim, a geoestatística baseada em modelos

apresentou-se como a mais adequada ao estudo espacial da mortalidade infantil, visto ter apresentado o menor valor de AIC.

A Tabela 3 mostrou que a variável “número de mulheres em idade fértil” é a mais correlacionada, positivamente, com os casos de mortalidade infantil. Para tentar verificar se essa correlação é verificada espacialmente, produziram-se mapas de predição, para essas variáveis, utilizando-se a geoestatística baseada em modelos, conforme podem ser visualizados na Fig. 5.

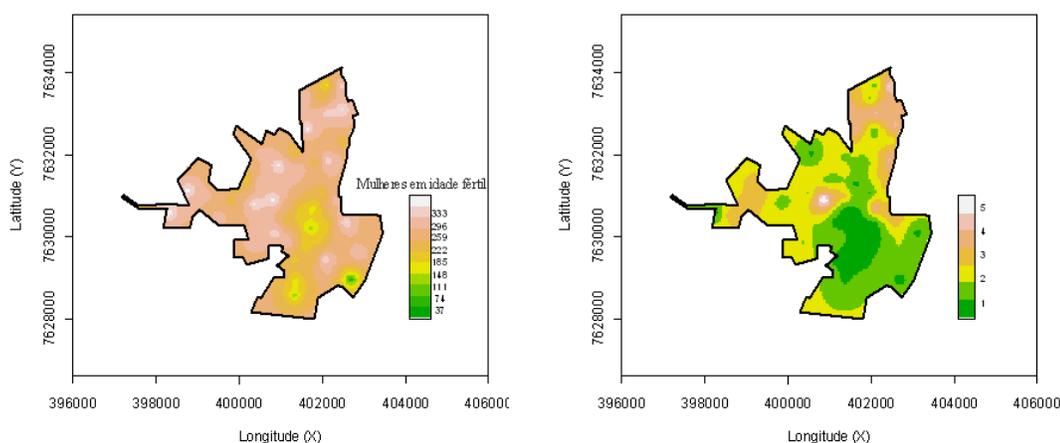


Fig. 5 - Mapas de predição para as variáveis “número de mulheres em idade fértil” (à esquerda) e “mortalidade infantil” (à direita), para a cidade de Alfenas, MG, entre 2000 e 2004.

Percebe-se, visualmente, pela Fig. 5 (ambas as legendas mostram valores que crescem da cor verde para a cor branca), que, na maior parte das regiões onde há uma maior incidência de mulheres em idade fértil, há, também, maior incidência de mortes infantis.

4. CONCLUSÕES

A geoestatística baseada em modelos, utilizando a distribuição de Poisson, mostrou-se mais adequada que a

geoestatística clássica para estudar a distribuição espacial dos casos de mortalidade infantil. O estudo permitiu identificar áreas do perímetro urbano da cidade de Alfenas, MG, em que há maiores incidências de mortes infantis, correlacionadas ao número de mulheres em idade fértil e ao número de mulheres em idade de risco gestacional. Espera-se que esses resultados possam servir para gerir formas de intervir nessas regiões, de maneira a minimizar as ocorrências de mortalidade infantil.

AGRADECIMENTOS

Aos dois avaliadores pelas sugestões e comentários que melhoraram a qualidade do trabalho.

REFERÊNCIAS

CHRISTENSEN, O. F.; RIBEIRO JR., P. J. GeoRglm: a package for generalized linear spatial models. *R-News*, v. 2, pp. 26-28, 2002.

COSTA, G. N. *Mortalidade perinatal, determinantes biológicos, de atenção à saúde materno infantil e socioeconômicos: uma análise das desigualdades entre bairros do Recife*. 2008. 156f. Tese (Doutorado em Saúde Pública) - Centro de Pesquisas Aggeu Magalhães / Fundação Oswaldo Cruz, Recife, 2008.

CRESSIE, N. *Statistics for spatial data*. New York: J. Wiley, 1993.

DIGGLE, P. J.; RIBEIRO JR., P. J. GeoR: A Package for geostatistical analysis. *R-News*, v. 1, pp. 15-18, 2001.

DIGGLE, P. J.; RIBEIRO JR., P. J. *Model-based geostatistics*. New York: Springer, 2007.

DUARTE, E. C. *et al.* Expectativa de vida ao nascer e mortalidade no Brasil em 1999: análise exploratória dos diferenciais regionais. *Revista Panamericana de Salud Pública*, v. 12, p. 436-444, 2002.

FARAH, M. F. S. Gênero e políticas públicas. *Revista de Estudos Feministas*, v. 12, p. 47-71, 2004.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). *Projeção da população do Brasil por sexo e idade para 1980-2050*. Rio de Janeiro, 2008. 93p. (Séries Estudos e Pesquisas, v. 24).

ISAAKS, E. H.; SRIVASTAVA, R. *An introduction to applied geostatistics*. New

York: Oxford University, 1989.

KEEN, C. S.; RIP, M. R.; WOODS, D. L. Intraurban variations of neonatal and post-neonatal mortality in a developing city. *Social Science and Medicine*, v. 25, p. 889-894, 1987.

LEAL, M. C.; SZWARCOWALD, C. L. Características da mortalidade neonatal no Estado do Rio de Janeiro na década de 80: uma visão espaço-temporal. *Revista de Saúde Pública*, v. 31, p. 457-465, 1997.

PREFEITURA MUNICIPAL DE ALFENAS. *Matriz de mortalidade para Alfenas*. Disponível em: http://epidemiologia.alfenas.mg.gov.br/news/arquivos/Matriz_mortalidade_alfenas.pdf. Acesso em: 20 dez. 2011.

R DEVELOPMENT CORE TEAM. *R: a language and environment for statistical computing*. Viena: R Foundation for Statistical Computing, 2012. Disponível em: <http://www.R-project.org>.

SIEGEL, S. *Estatística não paramétrica para ciências do comportamento*. São Paulo: McGraw-Hill, 1975.

SHIMAKURA, S. E. *et al.* Distribuição espacial do risco: modelagem da mortalidade infantil em Porto Alegre, Rio Grande do Sul, Brasil. *Cadernos de Saúde Pública*, v. 77, p. 1251-1261, 2001.

SILVA, S. L. C. *et al.* Visualização dos padrões da variação da taxa de mortalidade infantil no Rio Grande do Sul: comparação entre as abordagens bayesiana empírica e totalmente bayesiana. *Cadernos de Saúde Pública*, v. 27, p. 1423-1432, 2011.

SIMÕES, C. C. S. *Perfis de saúde e de mortalidade no Brasil: uma análise de seus condicionantes em grupos populacionais específicos*. Brasília: Organização Pan-Americana da Saúde, 2002.