

A IMPORTÂNCIA DA CORREÇÃO DA ANISOTROPIA EM ANÁLISES GEOESTATÍSTICAS

Diogo Francisco Rossoni^{1,2}, Renato Ribeiro de Lima^{1,2}

Resumo: *A anisotropia é uma característica na qual dados espacialmente dependentes apresentam comportamento diferenciados em determinadas direções. Quando o comportamento da variável aleatória é similar em todas as direções, dizemos que o fenômeno é isotrópico. Muitos autores têm identificado (análises exploratórias e visuais) e corrigido a anisotropia nos mais variados experimentos: poluição atmosférica, variáveis químicas e físicas do solo, dispersão de populações nativas de planta, variáveis geológicas, etc. Todavia, em muitos trabalhos, constata-se que os autores pressupõe que a variável é isotrópica e procedem com a modelagem do fenômeno - e posterior predição - sem efetuarem a devida correção da anisotropia. O intuito principal desse trabalho é evidenciar a melhora significativa da predição de um determinado fenômeno para diversas configurações de dependência espacial, bem como para distintos graus de anisotropia. Para tal, sessenta populações foram simuladas. De cada população, mil amostras aleatórias de tamanho mil foram selecionadas. A análise abordou a correção e não correção da anisotropia. A qualidade da predição foi medida através do erro quadrático médio de predição, e o conjunto de respostas foi comparado através do intervalo de confiança Monte Carlo. Para todas as populações, os intervalos do EQMP para os dados que tiveram a anisotropia corrigida foram menores - e estatisticamente diferentes - do que para os dados que não tiveram a correção da anisotropia.*

Palavras-chave: *Anisotropia, Geoestatística, Erro quadrático médio de predição, Intervalo de confiança Monte Carlo.*

1 Introdução

Uma variável regionalizada é isotrópica quando o padrão de continuidade espacial é idêntico em todas as direções. A isotropia é de suma importância para a estimação em locais não amostrados. É importante na construção de mapas de variabilidade, pois, por meio da identificação de que a variável analisada é isotrópica, o estudo da dependência espacial pode ser feito por meio de um único semivariograma, chamado de omnidirecional.

Entretanto, quando o fenômeno em estudo revela diferentes padrões de dependência espacial, ou seja, apresenta uma variabilidade que não é a mesma em todas as direções, o fenômeno em estudo é chamado de anisotrópico.

Os principais tipos de anisotropia estudados na geoestatística são: geométrica, zonal e combinada. A anisotropia geométrica é aquela em que existe uma direção com maior continuidade espacial, isto é, maior valor de alcance no semivariograma empírico; a anisotropia zonal ocorre quando existe uma direção com maior valor de patamar nos semivariogramas empíricos em

¹DEX-Universidade Federal de Lavras, diogo.rossoni@gmail.com

²Agradecimento à FAPEMIG e ao CNPq pelo apoio financeiro.

relação às demais direções; e a anisotropia combinada, quando houver determinadas direções com diferentes valores de alcance e patamar nos semivariograma empíricos.

Muitos autores têm identificado, via análises exploratórias e visuais, e corrigido a anisotropia nos mais variados experimentos: poluição atmosférica, variáveis químicas e físicas do solo, dispersão de populações nativas de planta, variáveis geológicas, etc. Contudo, em muitos trabalhos, constata-se que os autores pressupõe que a variável é isotrópica e procedem com a modelagem do fenômeno - e posterior predição - sem efetuarem a devida correção da anisotropia.

2 Objetivos

O intuito principal deste trabalho é evidenciar a melhora significativa da predição de um determinado fenômeno para diversas configurações de dependência espacial (com distintos graus de anisotropia), quando procede-se com a correção da anisotropia, ante a predição sem correção anisotrópica; através da estimação do erro quadrático médio de predição (EQMP).

3 Metodologia

O procedimento de simulação do campo aleatório estacionário Gaussiano seguiu o procedimento proposto por Wood e Chan (1997), descrito como um algoritmo de simulação rápido e eficiente para malhas finas (gride com grande densidade de pontos referenciados), dada uma função de covariância prescrita. É baseado na idéia da “incorporação circulante”, proposto implicitamente por Davies e Harte (1987) (uma dimensão) e estendido para n -dimensões, independentemente, por Dietrich e Newsam (1993) e Wood e Chan (1994).

Para simular o vetor aleatório Gaussiano com matriz de covariância C , inicialmente simula-se um vetor Y de variáveis Gaussianas independentes padrão, e então calcula-se $Z = C^{1/2}Y$ em que a matriz de incorporação circulante C , simétrica e não-negativa, é dada por $C = \{C_{ij} = c(i-j); i, j \in Y^2\}$, no qual $c(i-j)$ é a função de covariância, sendo que neste trabalho, a mesma é definida como a função de covariância esférica, definida por Isaaks e Sirivastava (1989) como

$$c(h) = \begin{cases} 1 - 1,5\frac{h}{\phi} + 0,5\frac{h}{\phi} & \text{se } h < \phi \\ 0 & \text{outros casos} \end{cases} \quad (1)$$

em que h é o vetor distância e ϕ é o alcance da dependência espacial.

Wood e Chan (1994) apresentam maiores detalhes sobre a decomposição de $C^{1/2}$.

Bonat (2011) indica mais três parâmetros relevantes para a simulação do processo espacial Z : a variância do processo espacial ($\sigma^2 > 0$), a variância do ruído ($\tau^2 > 0$) e a média geral do processo (β). Logo, com a combinação dos parâmetros da Tabela 1, foi obtido um total de 60 populações.

Tabela 1: parâmetros das simulações

β	τ^2	ϕ	σ^2	α	F_α
100	20	250	150	0°	3,5
				45°	3
	75			90°	2,5
	120			135°	2
					1,5

Para a simulação de um processo anisotrópico geométrico é necessário uma série de transformações das coordenadas. Seja $X_{n \times 2}$ a matriz de coordenadas, R a matriz de rotação e T a matriz de dilatação.

$$R = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \quad (2)$$

$$T = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{F_\alpha} \end{pmatrix} \quad (3)$$

O algoritmo de simulação seleciona um grid regular isotrópico e o transforma em um grid anisotrópico, através de $X^* = X(RT)^{-1}$. Após a geração das variáveis aleatórias, faz-se a transformação inversa $X = X^*RT$.

Os parâmetros fator de anisotropia F_α e o ângulo da direção de maior continuidade espacial (α) foram calculados considerando $\alpha = 0^\circ, 45^\circ, 90^\circ, 135^\circ$. Para $F_\alpha = \frac{\phi_1}{\phi_2}$, ϕ_2 e ϕ_1 são, respectivamente, os alcances nas direções de maior e menor continuidade espacial.

Após gerar 60 populações de tamanho 10.000 em um grid de tamanho $1000m \times 1000m$, 1000 amostras aleatórias de tamanho 1000 foram selecionadas. Para cada amostra os seguintes passos foram efetuados:

- i) As semivariâncias omnidirecionais sem correção anisotrópica foram estimadas através do método clássico (ISAAKS; SRIVASTAVA, 1989). O estimador das semivariâncias omnidirecionais com correção anisotrópica geométrica é o mesmo, salvo que o vetor h é substituído por $h' = x_i \times R \times T$;
- ii) Modelos teóricos de semivariogramas foram ajustados às semivariâncias empíricas (sem correção anisotrópica e com correção anisotrópica) através do método de mínimos quadrados ordinários.
- iii) A população de tamanho 10.000 foi estimada através do preditor krigagem ordinária, com base nos modelos ajustados no passo anterior (ii);
- iv) O erro quadrático médio de predição (EQMP) foi calculado para a população estimada com correção de anisotropia e sem correção de anisotropia.

Após efetuar os quatro passos para todas as amostras as distribuições dos EQMP foram comparadas através do intervalo de confiança Monte Carlo (BUCKLAND, 1984) a 5% de significância.

Todas as análises foram feitas no software R (R DEVELOPMENT CORE TEAM, 2012).

4 Resultados

Para facilitar a interpretação dos resultados define-se o Erro1, como sendo o intervalo de confiança Monte Carlo do EQMP com correção da anisotropia e o Erro2, como sendo o intervalo de confiança Monte Carlo do EQMP sem correção de anisotropia. Os modelos de simulação são representados por *modelo de covariância* (τ^2, σ^2, ϕ).

Na Figura 1 percebe-se que em todos os casos simulados o Erro 1 foi menor do que o Erro 2, e estatisticamente distintos, através do intervalo de confiança Monte Carlo a 5% de significância.

Percebe-se que o EQMP é diretamente proporcional ao valor do F_α . À medida que o fator de anisotropia aumenta - para todas as configurações e direções - o EQMP tende a ser maior, tanto para o Erro1, quanto para o Erro2.

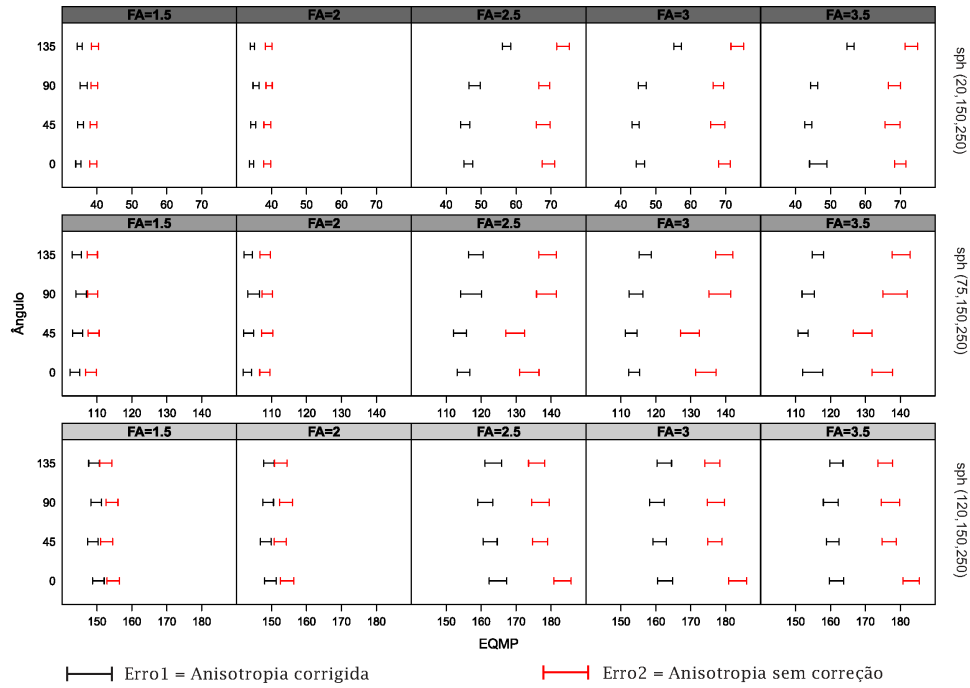


Figura 1: Intervalos de confiança Monte Carlo dos EQMP.

Esse erro também aumenta à medida que a dependência espacial decai, isto é, $sph(20,150,250)$ que apresenta forte dependência espacial, possui EQMP menores do que as configurações de moderada dependência espacial ($sph(75,150,250)$) e fraca dependência espacial ($sph(120,150,250)$), em comparações diretas entre mesmos ângulos e fatores de anisotropia.

À medida que o valor de F_α aumenta, constata-se que os intervalos do Erro 1 e Erro 2 ficam cada vez mais distantes, ou seja, quanto maior o fator de anisotropia, maior o EQMP cometido pelo modelo que não considera a anisotropia.

5 Conclusões

Em todas as configurações de população simuladas, os modelos que utilizaram a correção da anisotropia obtiveram EQMP menores do que os modelos que não corrigiram a anisotropia. À medida que o fator de anisotropia aumenta, a melhora na predição com correção torna-se mais evidente.

Referências

- [1] BONAT, W. H., RIBEIRO JR., P. J., ZEVIANI, W. M., Comparando predições por modelos geoestatístico e aditivo generalizado para reconstituição de superfícies contínuas gaussianas, *Revista Engenharia na Agricultura*, **26(2)**, 119-128, 2011.
- [2] BUCKLAND, S. T., Monte Carlo confidence intervals, *Biometrics*, **40**, 811-817, 1984.
- [3] DAVIES, R. B., HARTE, D. S., Tests for Hurst effect, *Biometrika*, **74**, 95-101, 1987.

- [4] DIETRICH, C. R., NEWSAM, G. N., A fast and exact method for multidimensional Gaussian stochastic simulations, *Water Resources Research*, **29(829)**, 2861-2869, 1993.
- [5] ISAAKS, E. H., SRIVASTAVA, R. M., An introduction to applied geostatistics, New York: Oxford University, 1989.
- [6] R DEVELOPMENT CORE TEAM. R: A language and environment for statistical computing. Versão 2.15.1. Vienna: R Foundation for Statistical Computing, 2012. Disponível em: <http://www.r-project.org>. Acesso em: 22 jun. 2012.
- [7] WOOD, A. T. A., CHAN, G., Simulation of stationary gaussian processes, *Journal of Computational and Graphical Statistics*, **3**, 409-432, 1994.
- [8] WOOD, A. T. A., CHAN, G., An Algorithm for Simulating Stationary Gaussian Random Fields, *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **46(1)**, 171-181, 1997.