

Pós-processamento das Estimativas de Krigagem com Deriva Externa
Post-processing Kriging with External Drift Estimates

Jorge Kazuo Yamamoto (jkyamamo@usp.br)

¹Departamento de Geologia Sedimentar e Ambiental - Instituto de Geociências - USP
R. do Lago 562, CEP 05508-080, São Paulo, SP, BR

Recebido em 28 de abril de 2009; aceito em 08 de setembro de 2009

RESUMO

A krigagem com deriva externa tem sido usada para a estimativa de uma variável pobremente amostrada, baseada em uma variável secundária intensamente amostrada. Mesmo considerando a informação da variável secundária, as estimativas apresentam menor variância que a variância amostral, devido ao efeito de suavização. O algoritmo para pós-processamento das estimativas da krigagem ordinária é adaptado para a krigagem com deriva externa. Conforme os resultados mostrados neste artigo, o algoritmo proposto não somente reproduz o histograma amostral, mas também mantém a relação linear observada entre as variáveis primária e secundária. Para amostras muito pequenas, em termos de número de pontos de dados, o pós-processamento não mantém a precisão local da krigagem suavizada, pois a correção é baseada nos erros derivados da validação cruzada. Contudo, quando o tamanho da amostra é de 3% dos dados completos, o pós-processamento dá resultados melhores que a krigagem suavizada.

Palavras-chave: Efeito de suavização; Incerteza; Variância de interpolação; Reprodução de padrão; Krigagem com deriva externa.

ABSTRACT

Kriging with an external drift has been used for estimating a poorly sampled variable based on an exhaustively sampled secondary variable. Even considering the information coming from secondary data, the estimates present smaller variance than the sample variance due to the smoothing effect. A post-processing algorithm for correcting the smoothing effect of ordinary kriging estimates is adapted for kriging with an external drift. According to results shown in this paper, the proposed algorithm not only reproduces the sample histogram but also keeps the linear relationship observed between primary and secondary variables. For very small samples, in terms of number of data points, the post-processing did not preserve the local precision as smoothed kriging does, because the correction is based on errors derived from the cross-validation. However, when available sample comprises 3% of the exhaustive data set, the post-processing algorithm gives better results than smoothed kriging.

Keywords: Smoothing effect; Uncertainty; Interpolation variance; Pattern reproduction; Kriging with an external drift.

INTRODUÇÃO

Diferentes variáveis espaciais amostradas em um mesmo domínio podem apresentar correlações em graus variáveis. Dentre essas variáveis, algumas podem ser amostradas a um baixo custo, enquanto outras apresentam um elevado custo de aquisição. Este é o caso da indústria de petróleo em que levantamentos sísmicos são realizados a um custo relativamente baixo e proporcionando uma grande quantidade de informações. Por outro lado, a perfuração de poços de petróleo é uma operação extremamente cara e, por isso, poucos furos são feitos na fase exploratória e, conseqüentemente, pequena quantidade de dados.

Em termos de precisão, as informações sísmicas apresentam baixa precisão, enquanto os dados de poços boa precisão. Quanto à correlação, a amplitude sísmica apresenta correlação com a porosidade de formações sedimentares. A ideia é então estimar a porosidade, pobremente amostrada, com base na informação sísmica, fortemente amostrada.

A geoestatística proporciona diferentes técnicas para a solução desse problema. Por exemplo, a cokrigagem é uma técnica tradicionalmente utilizada na estimativa de uma variável primária com base na informação de uma variável secundária. Embora tenha uma formulação teórica perfeita; na prática, a cokrigagem nem sempre proporciona resultados satisfatórios, devido, principalmente à instabilidade do sistema de equações de cokrigagem. Além disso, a solução desse sistema requer o cálculo e a modelagem de variogramas diretos e cruzados entre todas as variáveis primária e secundária(s). A cokrigagem pode ser considerada uma técnica de aplicação geral que não requer condições especiais na amostragem. Por exemplo, que se conheça em todos os pontos a informação primária e secundária, bem como a informação secundária ser conhecida em todos os pontos da malha regular que se quer estimar. Caso a informação disponível esteja neste formato, há outras técnicas que fazem uso disso, com vantagens. Nesse sentido, pode-se citar a krigagem com deriva externa e a cokrigagem colocalizada.

A krigagem com deriva externa é uma técnica que permite fazer a estimativa de uma variável aleatória em um ponto não amostrado com base no padrão espacial descrito por uma variável secundária. Embora a krigagem com deriva externa proporcione uma razoável reprodução do padrão espacial, as estimativas resultantes apresentam variância reduzida devido ao efeito de suavização. O histograma amostral não é reproduzido e, portanto, as estatísticas da estimativas serão muito diferentes das características reais. Evidentemente, não se sabe nada a respeito das características da população, mas as informações fornecidas pela amostra disponível devem ser utilizadas para inferir os parâmetros populacionais. Então, se o histograma amostral for

reproduzido no processo de estimativa, as estatísticas amostrais permitirão inferir com maior segurança os parâmetros populacionais. A correção do efeito de suavização das estimativas por krigagem garante a reprodução do histograma amostral como demonstrado por Yamamoto (2005, 2007, 2008). As estimativas por krigagem com deriva externa podem ser pós-processadas para correção do efeito de suavização, como se pretende demonstrar neste trabalho.

KRIGAGEM COM DERIVA EXTERNA

Seja $Z(x)$ a variável primária e $Y(x)$ a variável secundária, onde $Z(x)$ é pobremente amostrada e a secundária fartamente amostrada dentro do domínio. $Z(x)$ e $Y(x)$ apresentam uma relação linear (Wackernagel, 1995, p. 190):

$$E[Z(x)] = a_o + b_1 Y(x) \quad (1)$$

O estimador da krigagem com deriva externa, de acordo com Wackernagel (1995, p. 191):

$$Z_{KED}^*(x_o) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (2)$$

Para garantir o não enviesamento das estimativas, a esperança do erro deve ser igual a zero:

$$E[Z(x_o) - Z_{KED}^*(x_o)] = 0,$$

que resulta na condição de não viés:

$$\sum_{i=1}^n \lambda_i = 1 \quad (3)$$

Aplicando-se o operador da esperança matemática ao estimador da krigagem com deriva externa (2) e substituindo-se a média pelo lado direito da equação (1), segundo Wackernagel (1995, p. 191), tem-se:

$$\begin{aligned} E[Z_{KED}^*(x_o)] &= \sum_{i=1}^n \lambda_i E[Z(x_i)] \\ &= a_o + b_1 \sum_{i=1}^n \lambda_i y(x_i) \\ &= a_o + b_1 y(x_o) \end{aligned}$$

Deste último desenvolvimento, tem-se a segunda condição de não viés:

$$y(x_o) = \sum_{i=1}^n \lambda_i y(x_i) \quad (4)$$

Esta condição de restrição faz com que os pesos da krigagem com deriva externa descrevam a forma média dada pela variável secundária.

Os pesos que minimizam a variância do erro, sujeita às duas condições de não viés (equações 3 e 4), resultam da solução do seguinte sistema de equações lineares (Wackernagel, 1995, p. 191):

$$\begin{cases} \sum_{j=1}^n \lambda_j C_r(x_i - x_j) - \mu_1 - \mu_2 y(x_i) = C_r(x_i - x_o) \\ \sum_{j=1}^n \lambda_j = 1 \\ \sum_{j=1}^n \lambda_j y(x_j) = y(x_o) \end{cases} \quad \text{para } i = 1, n \quad (5)$$

onde μ_1 e μ_2 são os multiplicadores de Lagrange e $C_r(x_i - x_j)$ é a covariância dos resíduos entre os pontos x_i e x_j .

CORREÇÃO DO EFEITO DE SUAVIZAÇÃO DAS ESTIMATIVAS POR KRIGAGEM COM DERIVA EXTERNA

A solução para correção do efeito de suavização das estimativas de krigagem ordinária, proposta originalmente por Yamamoto (2005), foi baseada na transferência de erros verificados na validação cruzada. A validação cruzada é uma poderosa ferramenta em que é possível derivar o erro verdadeiro a partir da estimativa de um ponto de dado usando a informação dos pontos vizinhos:

$$\text{ErroVerdadeiro} = Z_{KED}^*(x_o) - Z(x_o)$$

O erro verdadeiro traz uma informação referente à suavização, ou seja, quando ele é maior que zero significa que houve uma superestimativa dos valores baixos e quando menor que zero uma subestimativa dos valores altos. O erro verdadeiro poderia ser transferido diretamente para a etapa da krigagem ordinária para correção do efeito de suavização, mas esta variável pode apresentar extremos, que poderia prejudicar a correção. Assim, ao invés de transferir o erro, a solução foi padronizar o erro verdadeiro dividindo-se pelo desvio padrão de interpolação (Yamamoto, 2000):

$$NS_o = \frac{-\text{ErroVerdadeiro}}{S_o}$$

É importante ressaltar que o desvio padrão de interpolação é sempre diferente de zero na validação cruzada, pois o próprio valor do ponto validado é retirado do conjunto de pontos de dados.

Essa nova variável foi denominada número de desvios padrão de interpolação. Como se pode verificar, esta variável combina erro com incerteza. O sinal negativo, efetivamente, atua na correção da suavização, pois quando há subestimativa dos valores altos (o erro é negativo) significa que uma quantidade de correção deve ser adicionada. Por outro lado, quando ocorre a superestimativa (erro positivo) significa que uma quantidade de correção deve ser subtraída.

A validação cruzada de um conjunto de pontos para krigagem com deriva externa é feita retirando-se a informação da variável $Z(x_o)$, mas mantendo-se o valor da variável secundária $y(x_o)$, pois esta é necessária para garantir a segunda condição de não viés (equação 4), necessária para a solução do sistema de equações (5) da krigagem com deriva externa.

Assim, após o processo de validação cruzada, tem-se para cada ponto uma variável adicional NS_o . Essa variável é interpolada na etapa da krigagem com deriva externa, propriamente dita, juntamente com a variável primária $Z_{KED}^*(x_o)$ e a incerteza associada S_o . Multiplicando-se a variável interpolada NS_o pelo desvio padrão de interpolação S_o obtém-se uma quantidade de correção, que adicionada à estimativa por krigagem com deriva externa dá origem à estimativa corrigida:

$$Z_{KED}^{**}(x_o) = Z_{KED}^*(x_o) + NS_o * S_o \quad (6)$$

onde $Z_{KED}^{**}(x_o)$ é a estimativa corrigida e $NS_o * S_o$ é a quantidade de correção.

A quantidade de correção pode ser excessiva fazendo com que a estimativa corrigida caia fora do intervalo de valores dos n pontos de dados vizinhos [$zmin$, $zmax$], ou seja, quando $Z_{KED}^{**}(x_o) < zmin$ ou $Z_{KED}^{**}(x_o) > zmax$, então a quantidade de correção é substituída por $delta = Z_{KED}^*(x_o) - zmin$ se o sinal de NS_o é negativo e $delta = zmax - Z_{KED}^*(x_o)$ se o sinal de NS_o é positivo:

$$Z_{KED}^{**}(x_o) = Z_{KED}^*(x_o) + delta \quad (7)$$

A quantidade de correção nas equações (6) e (7) podem ser substituídas por uma nova variável $Z_{NS_o}^*(x_o)$ e, portanto, a estimativa corrigida torna-se:

$$Z_{KED}^{**}(x_o) = Z_{KED}^*(x_o) + Z_{NS_o}^*(x_o) \quad (8)$$

Como o objetivo é a reprodução do histograma amostral, então é necessário que a variância da estimativa cor-

rigida $Var[Z_{KED}^{**}(x_o)]$ seja igual à variância amostral $Var[Z(x)]$. A simples multiplicação da estimativa corrigida por uma constante K poderia corrigir para a variância amostral, mas com consequências em relação à média da estimativa corrigida, que também seria transladada K vezes de sua posição original e, portanto, enviesada:

$$Var[Z(x)] = Var[K.Z_{KED}^{**}(x_o)] = K^2 Var[Z_{KED}^{**}(x_o)]$$

$$E[K.Z_{KED}^{**}(x_o)] = K.E[Z_{KED}^{**}(x_o)]$$

Então, a solução é multiplicar apenas a quantidade de correção por um fator constante que garanta $Var[Z_{KED}^{**}(x_o)] = Var[Z(x)]$:

$$Z_{KED}^{**}(x_o) = Z_{KED}^*(x_o) + Z_{NS_o}^*(x_o) * fator \quad (9)$$

Aplicando-se o operador da variância à expressão (9) e igualando-se à variância amostral chega-se a uma equação de 2º grau (Yamamoto, 2007):

$$Var[Z_{NS_o}^*(x_o)]fator^2 + 2Cov(Z_{KED}^*(x_o), Z_{NS_o}^*(x_o))fator + (Var[Z_{KED}^*(x_o)] - Var[Z(x)]) = 0 \quad (10)$$

A solução da equação quadrática (10) resulta em duas raízes, uma negativa e outra positiva, que é o fator ótimo procurado:

$$fator = \frac{-2Cov(Z_{OK}^*(x_o), Z_{NS_o}^*(x_o)) + \sqrt{\Delta}}{2Var[Z_{NS_o}^*(x_o)]}$$

onde

$$\Delta = 4Cov(Z_{KED}^*(x_o), Z_{NS_o}^*(x_o))^2 - 4Var[Z_{NS_o}^*(x_o)](Var[Z_{KED}^*(x_o)] - Var[Z(x)])$$

é o discriminante da equação de 2º grau.

MATERIAIS E MÉTODOS

Dois conjuntos de dados sintéticos para as variáveis primária e secundária foram gerados em computador. Esses conjuntos, compostos por 3404 pontos de dados, distribuídos em uma malha regular de 74 x 46 nós, constituirão os dados populacionais, doravante denominados conjuntos completos. Os parâmetros populacionais das variáveis

primária e secundária, denominadas VP e VS1, respectivamente, encontram-se na Tabela 1.

As imagens das variáveis primária e secundária, do conjunto completo, encontram-se na Figura 1.

As projeções em perspectiva permitem visualizar melhor a variação espacial das variáveis em estudo (Figura 2) e mostram que a variável secundária apresenta uma variabilidade maior em pequena escala (ruído branco adicionado).

As variáveis primária e secundária, apresentam uma correlação igual a 0,825 (Figura 3).

No diagrama de dispersão (Figura 3) fica clara a correlação existente entre as duas variáveis, mas com uma dispersão em torno da reta de regressão devido ao ruído branco existente na variável VS1, como se pode verificar na Figura 2B.

Do conjunto completo foram extraídas duas amostras com 40 e 104 pontos, por amostragem aleatória estratificada, conforme os mapas de localização apresentados na Figura 4. As estatísticas amostrais são apresentadas na Tabela 2.

As amostras extraídas com 40 e 104 pontos representam 1,18 e 3,06% do conjunto completo, respectivamente. Considerando essas porcentagens, as amostras extraídas representam razoavelmente o conjunto completo, conforme a Figura 5.

Como se pode observar nesta figura, há falta de aderência nas caudas superiores das distribuições de frequência amostrais. Evidentemente, a amostra com 104 pontos é mais representativa do conjunto completo, como mostra o diagrama P-P (Figura 5). As medidas de distância média em relação à bisettriz (diagrama P-P) iguais a 2,35 e 1,53, respectivamente para as amostras com 40 e 104 pontos confirmam a observação anterior.

As correlações entre as variáveis primária e secundária das amostras extraídas do conjunto completo encontram-se na Figura 6.

Semivariogramas omnidirecionais residuais foram computados conforme modelos ilustrados na Figura 7. A solução prática adotada para o cálculo de variogramas re-

Tabela 1. Parâmetros populacionais (variáveis primária e secundária) do conjunto completo.

| | VP | VS1 |
|--------------------------|-------|--------|
| Nº de dados | 3404 | 3404 |
| Média | 3,270 | 7,371 |
| Desvio padrão | 1,224 | 2,321 |
| Coefficiente de variação | 0,374 | 0,315 |
| Máximo | 9,167 | 18,194 |
| Quartil superior | 3,694 | 8,501 |
| Mediana | 2,894 | 6,966 |
| Quartil inferior | 2,457 | 5,816 |
| Mínimo | 1,683 | 2,471 |

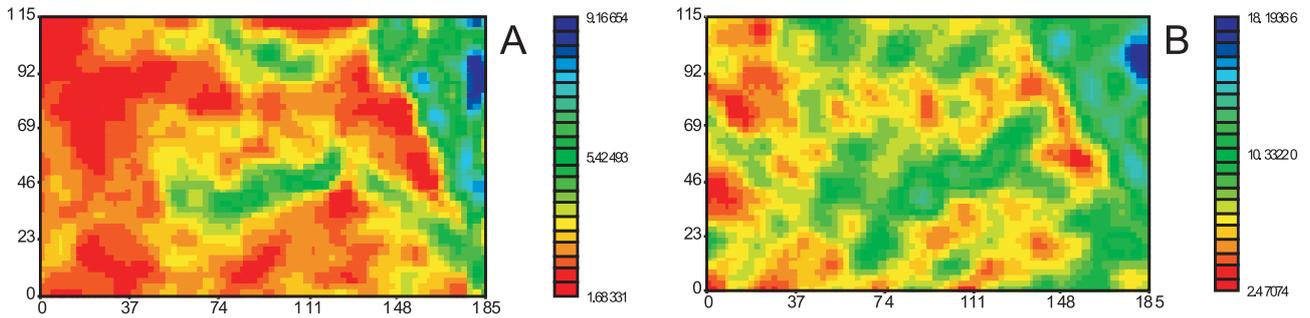


Figura 1. Imagens do conjunto completo: variável primária - VP (A) e variável secundária - VS1 (B). Coordenadas em metros.

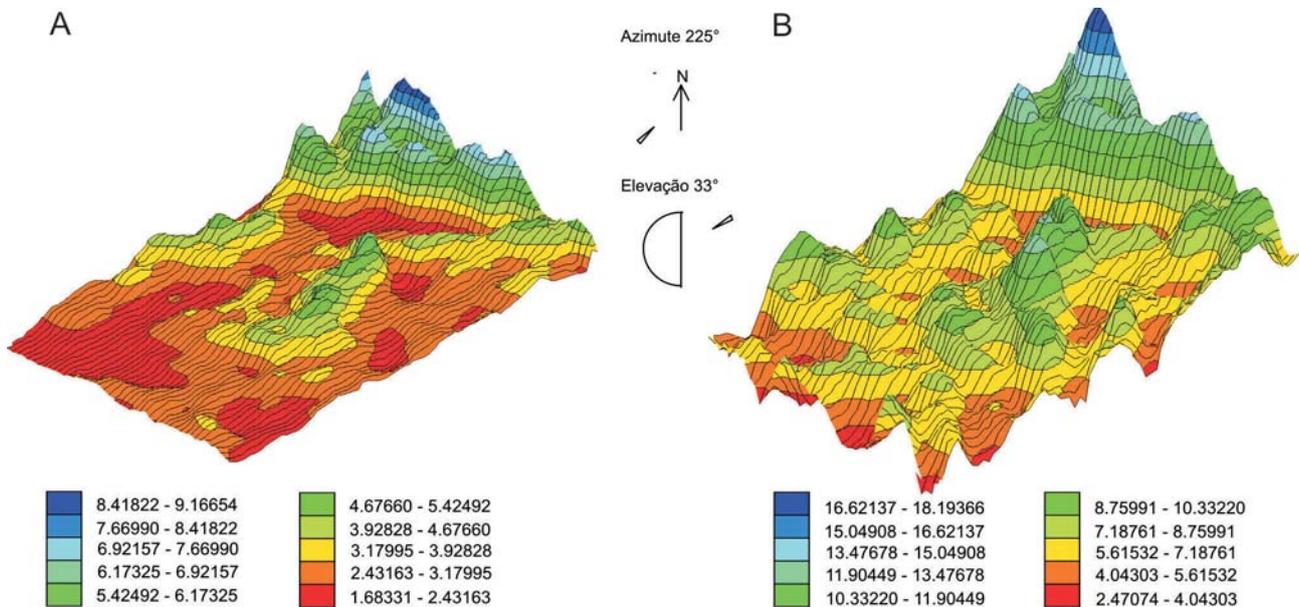


Figura 2. Projeções em perspectiva da variável primária - VP (A) e da variável secundária - VS1 (B) para um observador situado a 225° N e elevação de 33°.

siduais foi baseada nos resíduos da regressão da variável primária em função da secundária.

O conjunto completo e as amostras dele extraídas constituem os materiais desse estudo. Estimativas por krigagem com deriva externa serão realizadas conforme a equação (2), as quais serão pós-processadas para correção do efeito de suavização (equação 9). O objetivo desse estudo é mostrar que as estimativas da krigagem com deriva ex-

terna pós-processadas para correção do efeito de suavização reproduzem o histograma amostral e mantêm a relação linear inicial entre as variáveis primária e secundária.

Os processamentos foram realizados em um programa denominado CROSSORDKDE2 (Yamamoto, 2001) que permite obter em uma única execução, as estimativas de krigagem com deriva externa pós-processadas para eliminação do efeito de suavização.

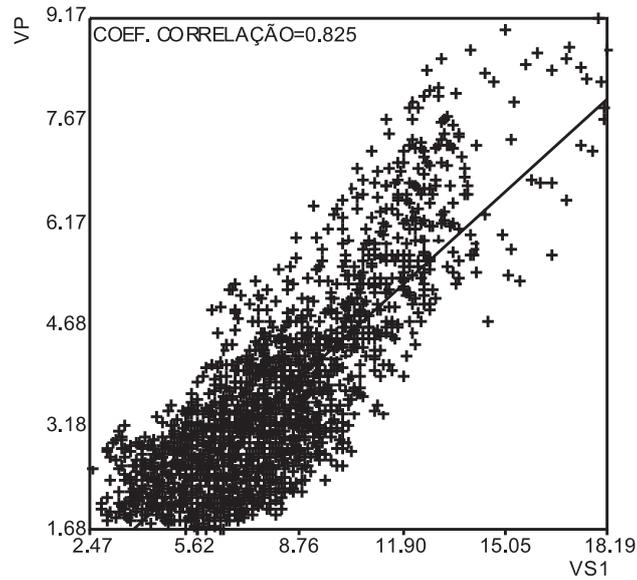


Figura 3. Correlação entre as variáveis VP e VS1, do conjunto completo.

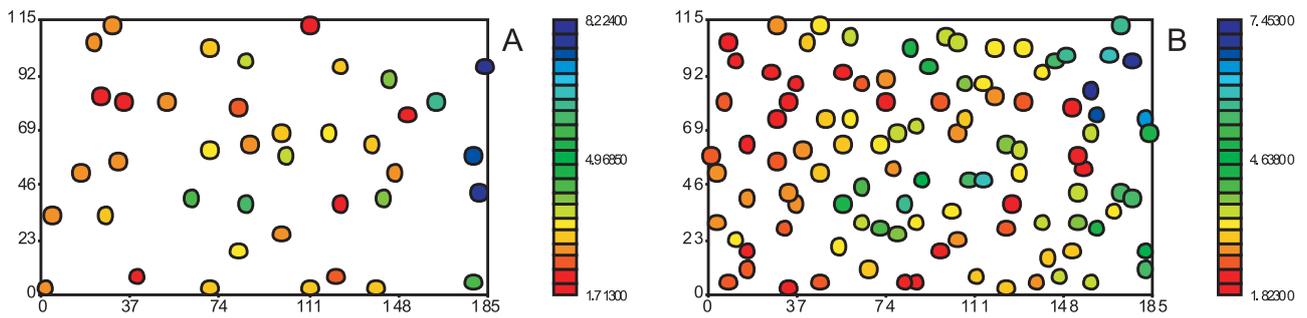


Figura 4. Mapa de localização de pontos das amostras com isotopia: 40 pontos (A) e 104 pontos (B).

Tabela 2. Estatísticas descritivas das variáveis primária e secundária das amostras em estudo.

| Estatísticas | Amostra40 | | Amostra104 | |
|------------------|-----------|--------|------------|--------|
| | VP | VS1 | VP | VS1 |
| N | 40 | 40 | 104 | 104 |
| Média | 3,368 | 7,207 | 3,311 | 7,308 |
| Desvio padrão | 1,540 | 2,902 | 1,212 | 2,264 |
| Coef. variação | 0,457 | 0,403 | 0,366 | 0,310 |
| Máximo | 8,224 | 17,941 | 7,453 | 15,841 |
| Quartil superior | 3,554 | 8,310 | 3,594 | 8,620 |
| Mediana | 2,845 | 6,371 | 3,017 | 6,851 |
| Quartil inferior | 2,543 | 5,421 | 2,457 | 5,585 |
| Mínimo | 1,713 | 2,498 | 1,823 | 3,503 |

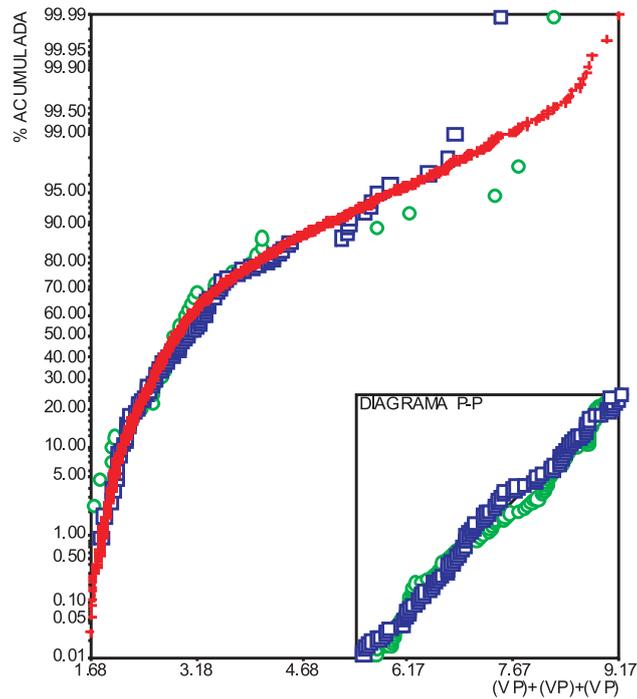


Figura 5. Distribuições de frequência acumulada do conjunto completo (cruz vermelha) das amostras com 40 pontos (círculo verde) e com 104 pontos (quadrado azul). No diagrama P-P; no eixo das abscissas está representada a probabilidade estimada e no eixo das ordenadas a probabilidade de referência.

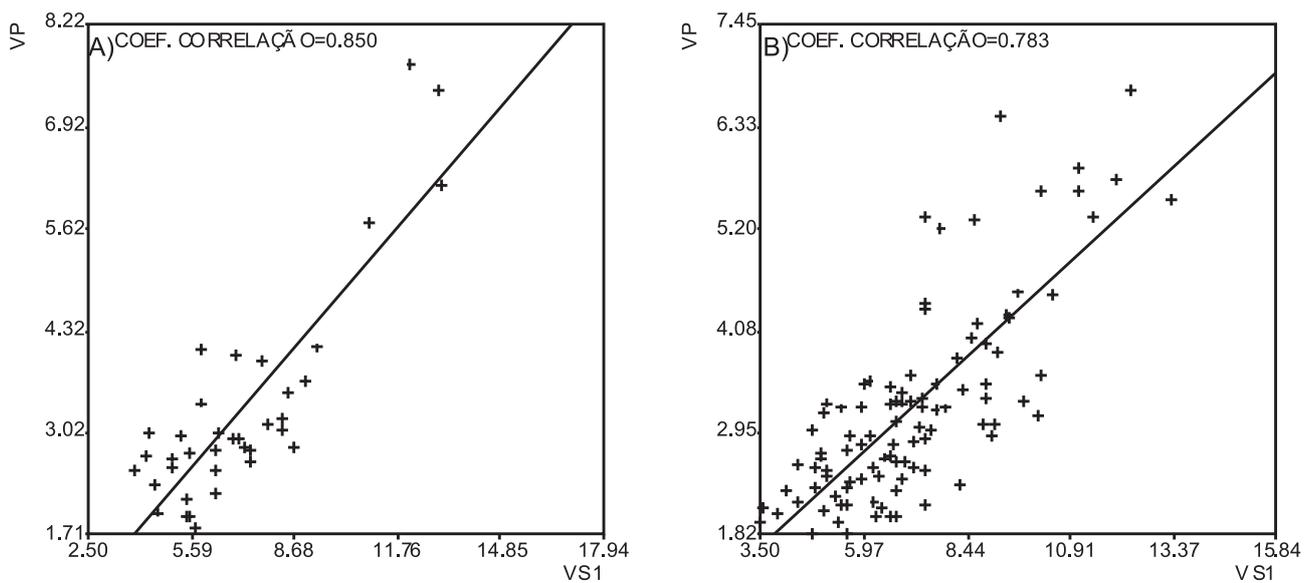


Figura 6. Correlação entre as variáveis primária e secundária para as amostras: 40 pontos (A) e 104 pontos (B).

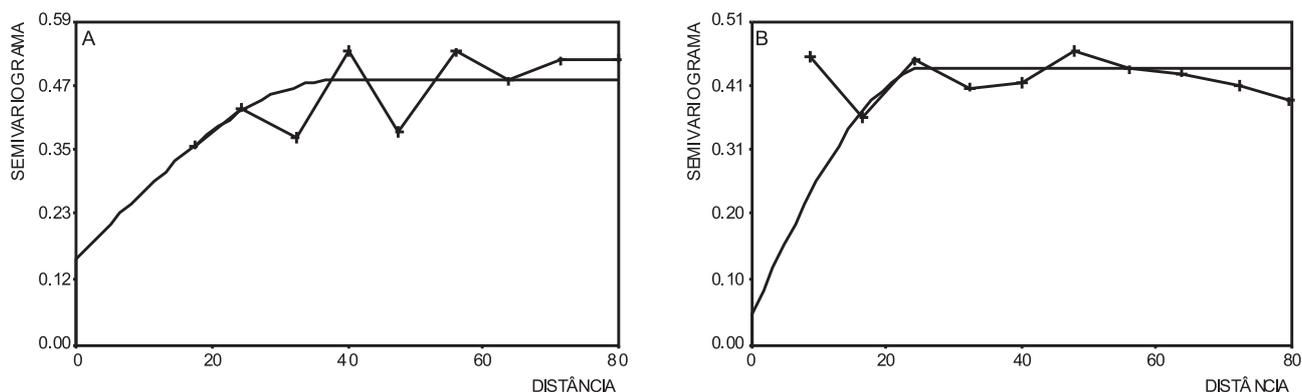


Figura 7. Modelos de semivariogramas residuais para as amostras em estudo: 40 pontos (A) e 104 pontos (B).

RESULTADOS E DISCUSSÃO

As imagens, da variável primária, estimadas para as amostras com 40 e 104 pontos, encontram-se nas Figuras 8 e 9, respectivamente. As estatísticas descritivas das estimativas encontram-se na Tabela 3.

As imagens estimadas com a amostra de 104 pontos (Figuras 9A e 9B) apresentam maior correlação com a imagem real (Figura 1A), nas quais o padrão espacial é reproduzido. Na imagem corrigida, para a amostra com 40 pontos (Figura 8B), é possível verificar que algumas feições inexistentes são criadas pela união entre pontos de locais com baixa densidade de amostragem.

As imagens estimadas podem ser comparadas quantitativamente com a imagem real (Figura 1A) através do cálculo do coeficiente de correlação de Pearson. Os diagramas de dispersão e os valores do coeficiente de Pearson encontram-se nas Figuras 10 e 11, respectivamente para as amostras com 40 e 104 pontos.

Nos diagramas de dispersão, pode-se verificar que há uma perda de precisão local na imagem corrigida para a amostra com 40 pontos (Figura 10B). Nesse caso, a krigagem com deriva externa mesmo suavizada apresenta correlação superior e reta de regressão muito próxima à bissetriz do diagrama de dispersão. Isso mostra que para amostras muito pequenas (apenas 1,18%), estimativas suavizadas ainda produzem resultados melhores que as estimativas pós-processadas. Na realidade, deve-se levar em conta que o processo de correção da suavização implica em transferir erros da validação cruzada para a etapa de estimativa por krigagem com deriva externa. Portanto, se a amostra não for representativa do conjunto completo, há o risco de transferir erros indesejáveis e assim diminuir a precisão local.

Os diagramas de dispersão da Figura 11 mostram que o pós-processamento não só manteve a precisão local, como também aumentou de 0,879 para 0,895. Além disso, os diagramas mostram os ângulos da reta de regressão maior que 1 para as estimativas suavizadas e menores que 1 para as estimativas corrigidas, como apresentado por Yamamoto (2005).

As distribuições de frequência acumuladas para as amostras e estimativas, bem como os diagramas P-P, encontram-se na Figura 12. As estatísticas por krigagem com deriva externa e estimativas pós-processadas encontram-se na Tabela 3.

Observando a Figura 12, verifica-se que as distribuições de frequência das estimativas suavizadas ficam distantes das respectivas distribuições amostrais, ao contrário das estimativas corrigidas que se aproximam bem das distribuições de frequência amostrais. Como seria de se esperar, o melhor ajuste é verificado para a amostra com 104 pontos (Figura 12B). As distâncias médias medidas nos diagramas P-P confirmam as observações anteriores (Tabela 4).

Como na krigagem com deriva externa a variável secundária é conhecida em todos os nós da malha regular interpolada, torna-se possível verificar se a relação linear entre a variável estimada e secundária foi mantida, tal como observada nas amostras (Figuras 5A e 5B). As Figuras 13 e 14 apresentam as relações entre as variáveis estimadas e a variável secundária, respectivamente para as amostras de 40 e 104 pontos.

Nessas figuras, verifica-se que somente as estimativas corrigidas da suavização mantêm as regressões lineares observadas nas amostras, enquanto as estimativas originais estão fortemente enviesadas em relação às regressões amostrais.

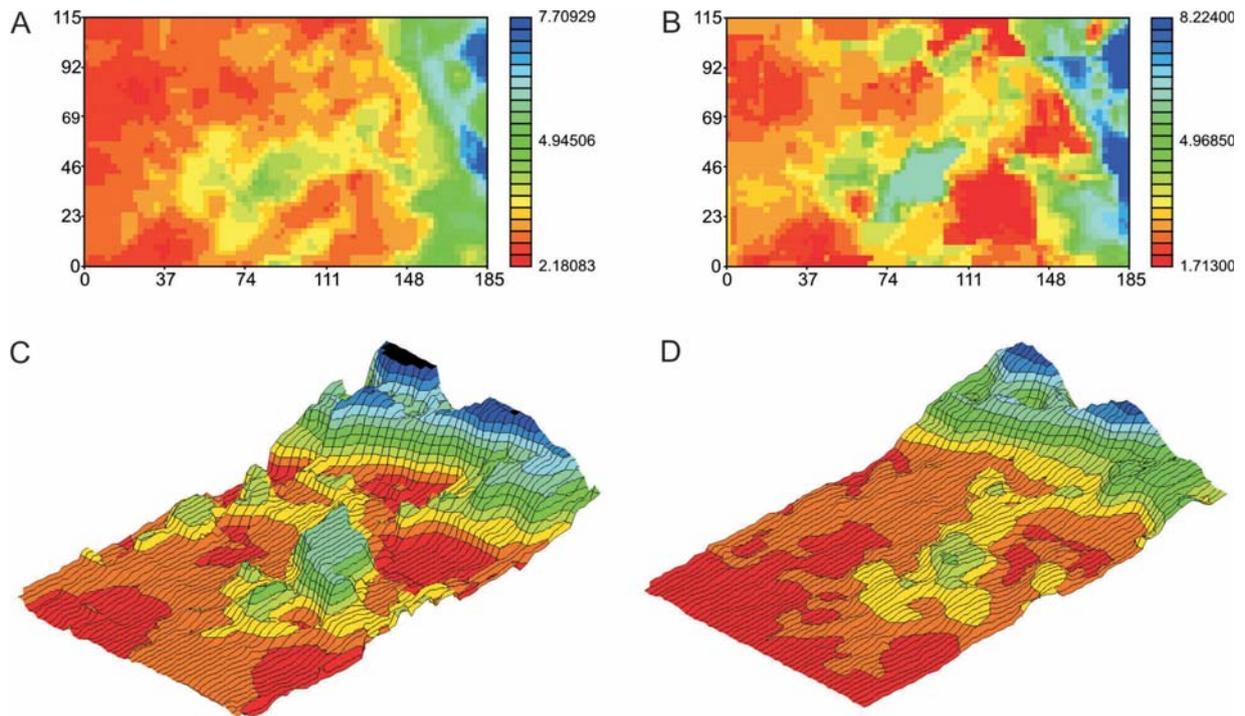


Figura 8. Imagens estimadas para a amostra com 40 pontos: krigagem com deriva externa (A e C) e após correção da suavização (B e D).

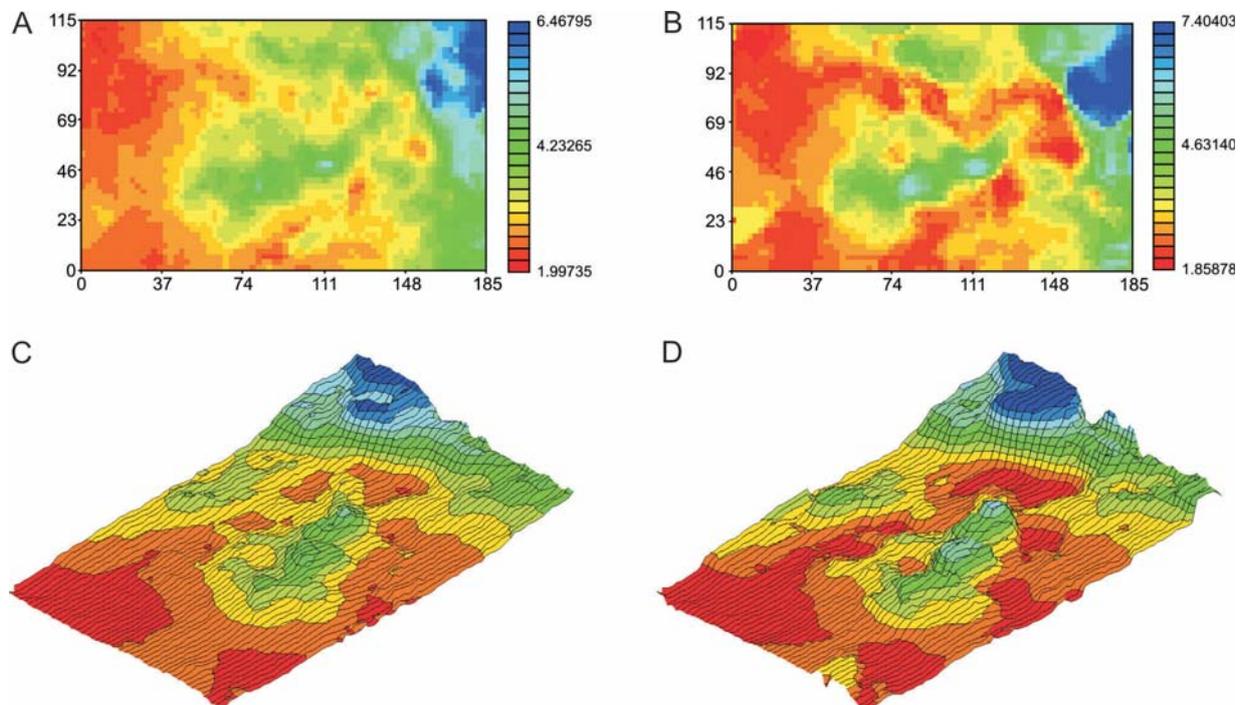


Figura 9. Imagens estimadas para a amostra com 104 pontos: krigagem com deriva externa (A e C) e após correção da suavização (B e D).

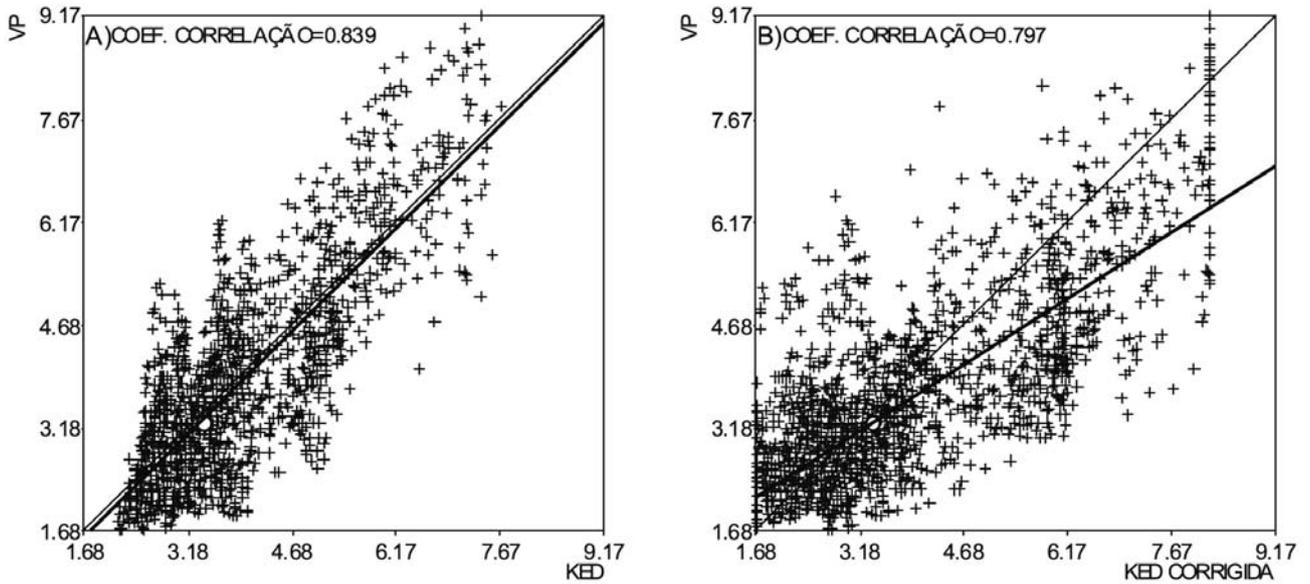


Figura 10. Diagramas de dispersão comparando valores reais e estimados pela krigagem com deriva externa (A) e após correção da suavização (B) para a amostra com 40 pontos.

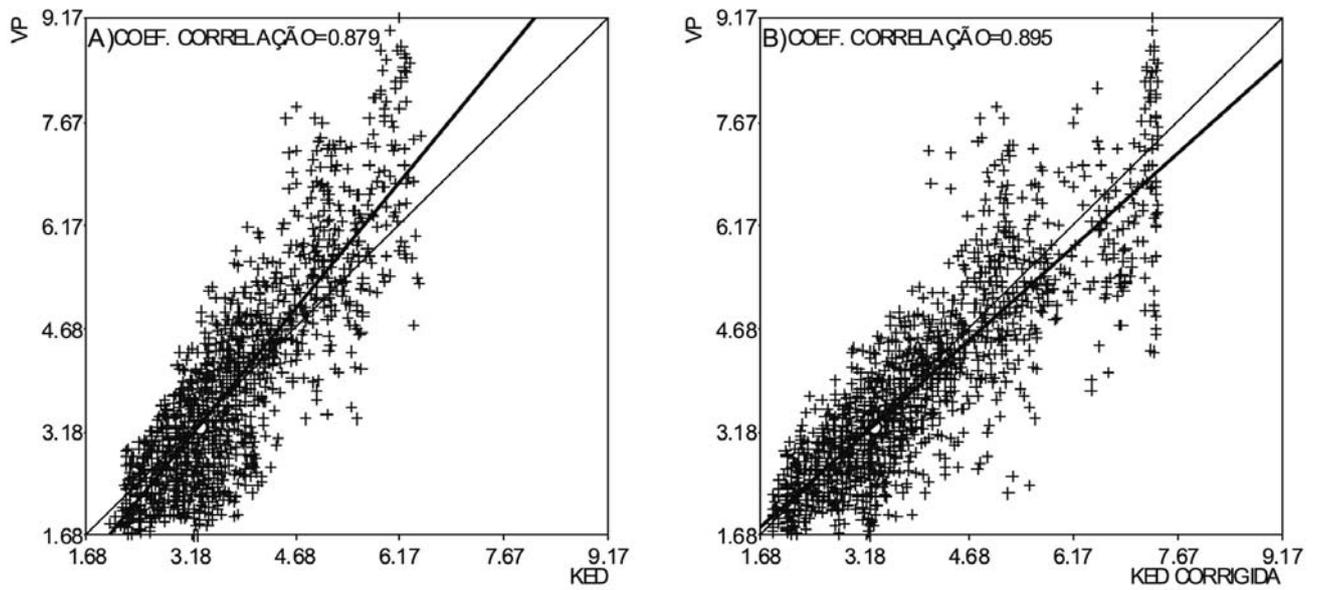


Figura 11. Diagramas de dispersão comparando valores reais e estimados pela krigagem com deriva externa (A) e após correção da suavização (B) para a amostra com 104 pontos.

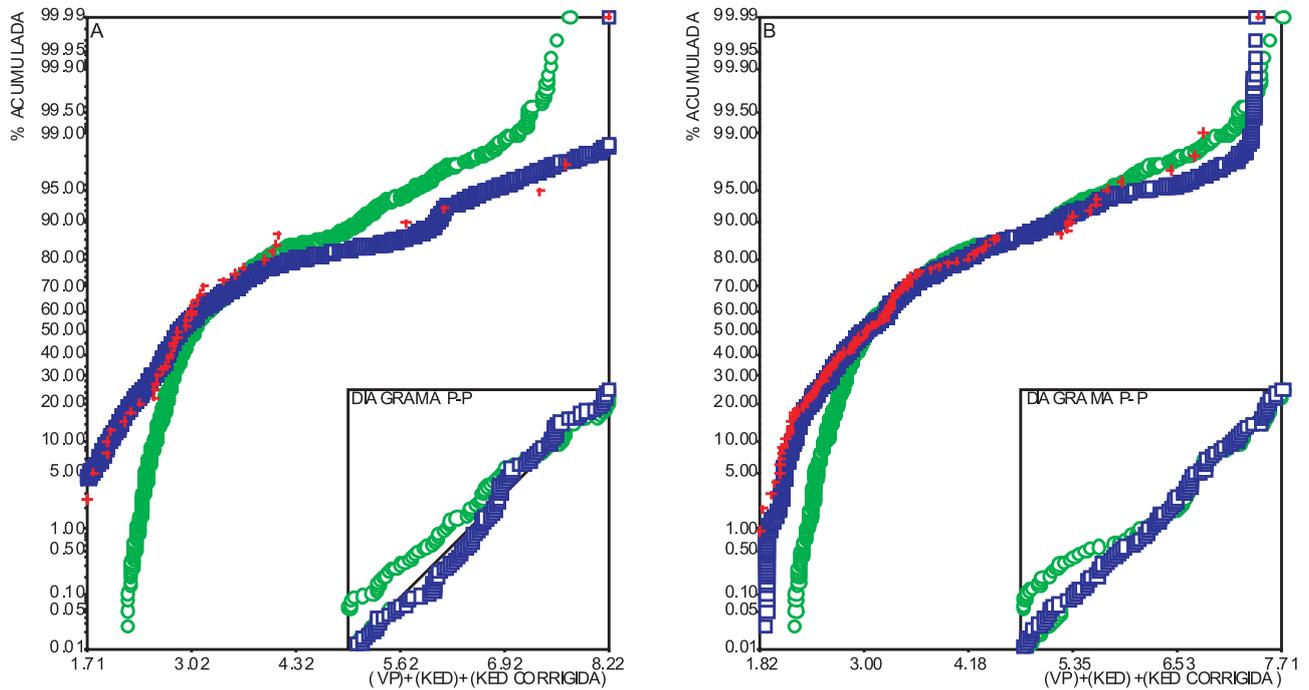


Figura 12. Distribuições de frequência e diagramas P-P para amostras com 40 pontos (A) e 104 pontos (B). *Legenda:* cruz vermelha = amostra; círculo verde = krigagem com deriva externa e quadrado azul = krigagem com deriva externa corrigida. No diagrama P-P; no eixo das abscissas está representada a probabilidade estimada e no eixo das ordenadas a probabilidade de referência.

Tabela 3. Estatísticas descritivas das estimativas para a variável primária para as amostras em estudo.

| Estatísticas | Amostra40 | | Amostra104 | |
|------------------|-----------|-----------|------------|-----------|
| | KED | KED corr. | KED | KED corr. |
| N | 3404 | 3404 | 3404 | 3404 |
| Média | 3,379 | 3,372 | 3,297 | 3,311 |
| Desvio padrão | 1,026 | 1,522 | 0,878 | 1,212 |
| Coef. variação | 0,304 | 0,451 | 0,266 | 0,366 |
| Máximo | 7,709 | 8,224 | 6,468 | 7,404 |
| Quartil superior | 3,647 | 3,775 | 3,620 | 3,701 |
| Mediana | 2,990 | 2,822 | 3,103 | 2,955 |
| Quartil inferior | 2,709 | 2,384 | 2,668 | 2,462 |
| Mínimo | 2,181 | 1,713 | 1,997 | 1,859 |

Tabela 4. Distâncias médias nos diagramas P-P para estimativas KED e KED corrigidas.

| Amostra | KED | KED corr. |
|---------|------|-----------|
| 40 | 5,03 | 2,52 |
| 104 | 4,54 | 1,18 |

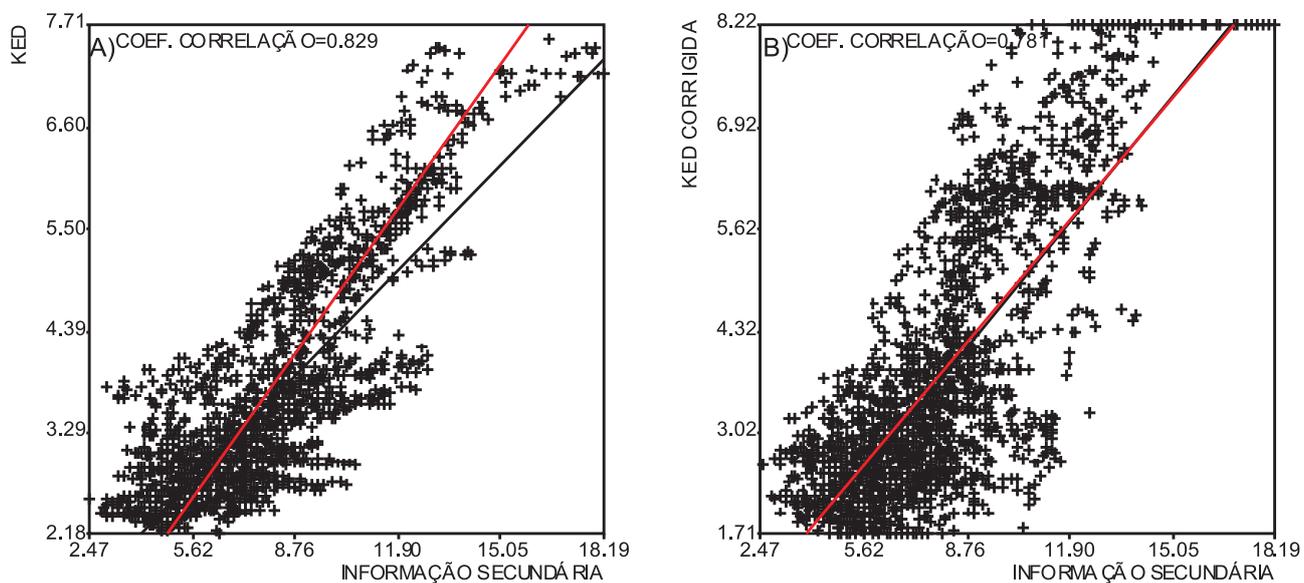


Figura 13. Correlação entre estimativas por krigagem com deriva externa e variável secundária (A) e entre estimativas corrigidas e variável secundária (B) para a amostra com 40 pontos. *Legenda:* linha vermelha = regressão amostral; linha preta = regressão de estimativas vs. variável secundária.

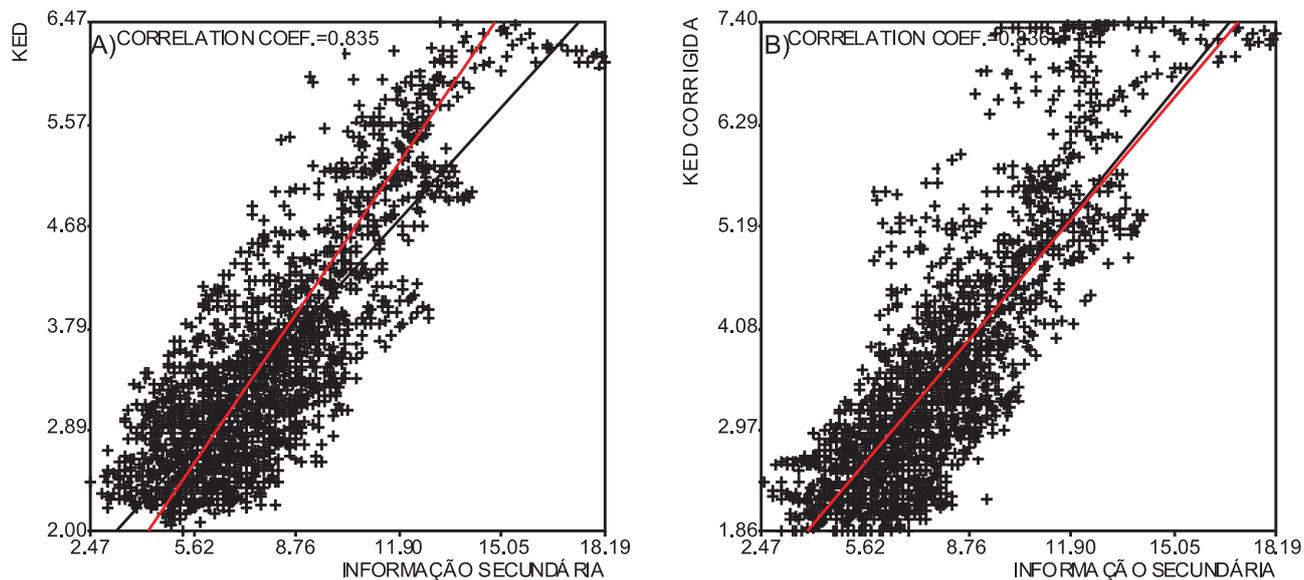


Figura 14. Correlação entre estimativas por krigagem com deriva externa e variável secundária (A) e entre estimativas corrigidas e variável secundária (B) para a amostra com 104 pontos. *Legenda:* linha vermelha = regressão amostral; linha preta = regressão de estimativas vs. variável secundária.

CONCLUSÕES

A krigagem com deriva externa é uma técnica que permite estimar uma variável primária com suporte de uma variável secundária. Como todos os demais métodos da krigagem, a krigagem com deriva externa está sujeita ao efeito de suavização. A correção do efeito de suavização, baseada na transferência do erro verdadeiro verificado na validação cruzada, torna-se efetivo quando a amostra for suficientemente grande, de tal modo a atenuar a propagação de erros. No caso em estudo, a amostra com 3% já foi suficiente para garantir a efetividade do método. Para amostras muito pequenas, a correção do efeito de suavização das estimativas por krigagem com deriva externa não é recomendável, pois o erro local calculado, com base em pequeno número de pontos de dados, é transferido para uma grande vizinhança, diminuindo a precisão local das estimativas corrigidas.

Assim, quando a amostra for representativa do fenômeno espacial em estudo, a correção das estimativas por krigagem com deriva externa é efetiva e os resultados apresentam distribuição de frequência muito próxima da distribuição amostral. Além disso, as estimativas corrigidas e a variável secundária mantêm a relação linear amostral.

AGRADECIMENTOS

O autor agradece ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq pelo apoio à presente pesquisa, através da concessão de bolsa de produtividade em pesquisa (Processo CNPq 303505/2007-9). Aos revisores Prof. Dr. Paulo Milton Barbosa Landim e Prof. Dr. João Felipe C. L. Costa pelas sugestões que melhoraram o manuscrito original.

REFERÊNCIAS

YAMAMOTO, J. K. An alternative measure of the reliability of ordinary kriging estimates. *Mathematical Geology*, v. 32, p. 489-509, 2000.

YAMAMOTO, J. K. *Sistema Geovisual 2.2: guia do usuário*. São Paulo: Instituto de Geociências - USP, 2001. 82 p.

YAMAMOTO, J. K. Correcting the smoothing effect of ordinary kriging estimates. *Mathematical Geology*, v. 37, p. 69-94, 2005.

YAMAMOTO, J. K. On unbiased backtransform of lognormal kriging estimates. *Computational Geosciences*, v. 11, p. 219-234, 2007.

YAMAMOTO, J. K. Estimation or simulation? That is the question. *Computational Geosciences*, v. 12, p. 573-591. 2008.

WACKERNAGEL, H. *Multivariate Geostatistics: An Introduction with applications*. Berlin: Springer-Verlag, 1995. 256 p.