

ESTUDO DA INFLUÊNCIA DA CORRELAÇÃO INICIAL ENTRE VARIÁVEIS NOS RESULTADOS DE CO-ESTIMATIVAS

Jorge WATANABE ¹, Jorge Kazuo YAMAMOTO ²,
Marcelo Monteiro da ROCHA ², Priscilla Pinto da FONSECA ²

(1) Companhia Vale do Rio Doce – Avenida de Ligação, 3580. CEP 34000-000. Nova Lima, MG.

Endereço eletrônico: jorge.watanabe@vale.com

(2) Instituto de Geociências, Universidade de São Paulo, Rua do Lago, 562. CEP 05508-080. São Paulo, SP.

Endereços eletrônicos: jkyamamo@usp.br; mmrocha@usp.br; priscilla.fonseca@usp.br

Introdução
Cokrigagem Ordinária
Cokrigagem Ordinária Colocalizada
Krigagem com Deriva Externa
Correção do Efeito de Suavização
Materiais e Métodos
Apresentação e Discussão dos Resultados
Cokrigagem Ordinária
Cokrigagem Ordinária Colocalizada
Krigagem com Deriva Externa
Conclusões
Agradecimentos
Referências Bibliográficas

RESUMO – Ainda não se sabe se a correlação amostral constitui fator restritivo à utilização de algum método de co-estimativa. Por isso, este artigo visa analisar as variações que ocorrem nos resultados de diferentes métodos de co-estimativa à medida que o coeficiente de correlação entre as variáveis primária e secundária diminui. Para isso, cokrigagem ordinária, cokrigagem ordinária colocalizada e krigagem com deriva externa com e sem correção do efeito de suavização foram aplicadas a cinco conjuntos de dados apresentando diferentes correlações entre a variável primária e a secundária. Todos os métodos tornaram-se menos eficientes à medida que o coeficiente de correlação entre as variáveis diminuiu. Em todos os casos, a cokrigagem ordinária colocalizada apresentou resultados com melhor precisão local e preservação da correlação inicial entre as variáveis. Porém, se o objetivo do estudo também for reproduzir o histograma amostral, a krigagem com deriva externa com correção do efeito de suavização constitui melhor alternativa para conjuntos de dados com alta correlação. Por fim, a utilização da cokrigagem ordinária deve se restringir aos casos em que as variáveis são altamente correlacionadas e escassamente amostradas.

Palavras-chave: co-estimativas; coeficiente de correlação; geoestatística multivariada; krigagem com deriva externa; correção do efeito de suavização.

ABSTRACT – *J. Watanabe, J.K. Yamamoto, M.M. da Rocha, P.P. da Fonseca - Study of influence of initial correlation between variables in the co-estimates results.* It is not known yet if sample correlation constitutes a constrained factor to the application of co-estimation techniques. So, this paper aims to analyze the variations that occur in results of different co-estimation methods as the correlation coefficient between primary and secondary variables diminishes. For this, ordinary cokriging, ordinary collocated cokriging and external drift kriging with or without smoothing effect correction were applied to five data sets presenting different correlation coefficients between primary and secondary variables. All techniques became less efficient as the correlation coefficient between variables diminished. In all cases, ordinary collocated cokriging showed the best local precision and maintenance of initial correlation coefficient between variables. However, if the aim of the study is also to reproduce sample histogram, external drift kriging with smoothing effect correction is the best alternative for highly correlated variables. Finally, the usage of ordinary cokriging must be constrained to situations in which variables are poorly sampled and highly correlated.

Keywords: co-estimation; coefficient of correlation; multivariate geostatistics; external drift kriging; smoothing effect correction.

INTRODUÇÃO

Estudos a respeito do melhor método de co-estimativa a ser empregado devem considerar diversos aspectos relacionados à reprodução do histograma amostral, à preservação da correlação inicial entre as variáveis primária, ou seja aquela de interesse, e secundária, isto é, aquela que vai auxiliar no enten-

dimento do comportamento espacial da variável primária, e à obtenção de precisão local.

Esses aspectos, por sua vez, dependem da correlação entre as variáveis primária e secundária e, por isso, torna-se necessário averiguar como os mesmos variam com a diminuição dessa correlação. Embora

todos os autores que se dedicam ao assunto são unânimes em afirmar a necessidade de correlação entre variável primária e secundária, nunca fica claro qual o valor exato desse coeficiente.

Dessa maneira, as técnicas da cokrigagem ordinária, cokrigagem colocalizada e krigagem com deriva externa com e sem correção do efeito de

suavização foram aplicadas a cinco conjuntos de dados, cada qual com um determinado coeficiente de correlação entre a variável primária e a secundária.

O objetivo desse artigo é, portanto, analisar a influência da correlação inicial entre a variável primária e a secundária nos resultados obtidos para assim determinar o melhor método a ser aplicado em cada caso.

COKRIGAGEM ORDINÁRIA

A estimativa do valor Z da variável p_0 no ponto x_0 mediante cokrigagem ordinária resulta da combinação linear de dados $Z_p(x_i)$ de N variáveis, cada qual multiplicado por um peso λ_i^p . Assim, o estimador da cokrigagem ordinária é (Wackernagel, 1998):

$$Z_{p_0}(x_0) = \sum_{p=1}^N \sum_{i=1}^{n_p} \lambda_i^p Z_p(x_i)$$

em que n_p é o número de amostras e depende do índice p das variáveis. Ainda de acordo com esse autor, os ponderadores são obtidos por meio da resolução do sistema de cokrigagem ordinária definido como:

$$\begin{cases} \sum_{p=1}^N \sum_{\beta=1}^{n_p} \lambda_{\beta}^p \gamma_{ij} (x_{\alpha} - x_{\beta}) + \mu_i = \gamma_{i0} (x_{\alpha} - x_0) & \text{para } i=1, \dots, N; \quad \alpha=1, \dots, n_i \\ \sum_{\beta=1}^{n_p} \lambda_{\beta}^p = \delta_{pp_0} & \text{para } i=1, \dots, N \end{cases}$$

$$\text{com } \delta_{pp_0} = \begin{cases} 1 & \text{se } p = p_0 \\ 0 & \text{se } p \neq p_0 \end{cases}$$

A função semivariograma cruzado é o modelo de correlação espacial considerado pela cokrigagem ordinária, podendo ser modelada a partir do variograma cruzado experimental que é uma generalização natural do variograma (Matheron, 1965) e pode ser expresso como:

$$\gamma_{ij}(h) = \frac{1}{2} E \{ [Z_i(x+h) - Z_i(x)] * [Z_j(x+h) - Z_j(x)] \}$$

COKRIGAGEM ORDINÁRIA COLOCALIZADA

A cokrigagem ordinária colocalizada pode ser utilizada nos casos em que a vizinhança é multi-colocalizada (Figura 1) a qual, segundo Chilès & Delfiner (1999), caracteriza-se pela existência de dados primários e secundários(s) em todos os pontos amostrais

e de dados secundários nos locais onde a variável primária será estimada.

De acordo com Goovaerts (1997), o estimador da cokrigagem ordinária colocalizada $Z_{CCK}^*(x_0)$ é:

$$Z_{CCK}^*(x_0) = \sum_{i=1}^n \lambda_i^Z Z(x_i) + \lambda_2^S [S(x_0) - m_1 + m_2]$$

Neste estimador, o ponderador e a média local da variável primária $Z(x_i)$ são λ_i^Z e m_1 , enquanto os da variável secundária $S(x_i)$ são λ_2^S e m_2 , respectivamente. A resolução do sistema de equações lineares a seguir resulta nos ponderadores já referidos:

$$\begin{cases} \sum_{i=1}^n \lambda_i^P C_{11}(x-x+h) + \lambda_j^S C_{12}(x_0-x+h) + \mu = C_{11}(x-x_0) \\ \sum_{i=1}^n \lambda_i^P C_{21}(x-x+h) + \lambda_j^S C_{22}(x_0-x+h) + \mu = C_{21}(x-x_0) \quad j=1, \dots, m \\ \sum_{i=1}^n \lambda_i^P(x_0) + \sum_{j=1}^m \lambda_j^S(x_0) = 1 \end{cases}$$

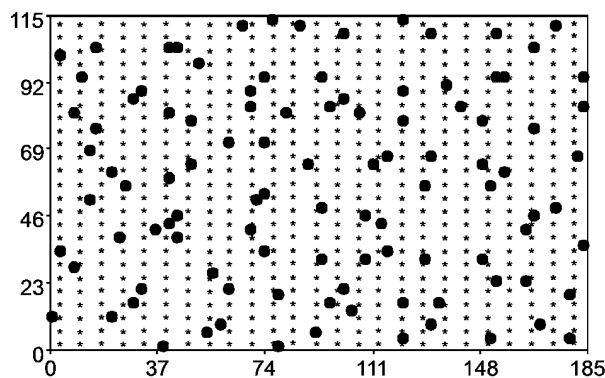


FIGURA 1. Mapa ilustrativo da multi-colocalização da amostragem. Os pontos cheios são locais onde os valores das variáveis primária e secundária foram obtidos. Os asteriscos são os locais onde a variável primária será estimada e há apenas informação da variável secundária.

A utilização deste método reduz a vizinhança considerada na co-estimativa, mas ainda requer o cálculo dos variogramas cruzados. Assim, assume-se a hipótese de efeito tela do tipo Markoviano (Xu et al., 1992):

$$E\{Z_2(x)|z_1(x), z_1(x+h)\} = E\{Z_2(x)|z_1(x)\} \quad \forall z_1(x+h)$$

De acordo com esta hipótese, o dado primário $z_1(x)$

oculta a influência de qualquer informação da variável primária $z_1(x+h)$ localizada na vizinhança da variável secundária colocalizada $Z_2(x)$. Com isso, as covariâncias cruzadas $C_{12}(h) = C_{21}(h)$ podem ser calculadas a partir da covariância da variável primária $C_1(h)$ multiplicada pelo coeficiente de correlação entre as variáveis primária e secundária, resultando na simplificação do algoritmo de cokrigagem.

KRIGAGEM COM DERIVA EXTERNA

A utilização da krigagem com deriva externa também requer que os dados sejam multi-colocalizados. De acordo com Wackernagel (2003), o estimador da krigagem com deriva externa resulta da somatória dos valores da variável regionalizada Z nos pontos x_i multiplicados por um peso λ_i :

$$Z^{KDE}(x_0) = \sum_{i=1}^n \lambda_i Z(x_i)$$

Os pesos λ_i provêm da resolução do sistema de equações (Wackernagel, 2003):

$$\begin{cases} \sum_{j=1}^n \lambda_j C_r(x_i - x_j) - \mu_1 - \mu_2 s(x_i) = C_r(x_i - x_0) & \text{para } i = 1, n \\ \sum_{j=1}^n \lambda_j = 1 \\ \sum_{j=1}^n \lambda_j s(x_j) = s(x_0) \end{cases}$$

A aplicação deste método considera a função variograma residual como modelo de correlação espacial. O cálculo do variograma residual exige que as variáveis primária e secundária sejam conhecidas em todos os pontos amostrais e correlacionadas entre si para que se possa obter primeiramente a reta de regressão $V_p^* = a + bV_s$ e, posteriormente, os resíduos $V_p^* - V_p$.

CORREÇÃO DO EFEITO DE SUAVIZAÇÃO

O efeito de suavização das estimativas é inerente a todos estimadores baseados na média ponderada e se caracteriza pela subestimativa de valores mais altos e superestimativa de valores mais baixos, resultando na perda de precisão global, ou seja, na não reprodução do histograma e variograma amostrais. Yamamoto (2005) propôs um algoritmo de pós-processamento baseado no número de desvios padrão de interpolação $Ns_0(x_0)$ e no desvio padrão de interpolação S_0 para cada ponto, conforme:

$$Z_{KO}^{**}(x_0) = Z_{KO}^*(x_0) + Ns_0(x_0) * S_0$$

no qual $Z_{KO}^{**}(x_0)$ é o valor corrigido e $Z_{KO}^*(x_0)$ é o valor estimado.

Informações detalhadas sobre esse algoritmo podem ser obtidas em Yamamoto (2005, 2007) e detalhes sobre o cálculo e características do desvio padrão de interpolação são apresentados em Yamamoto (2000). Neste artigo apenas as co-estimativas feitas por krigagem com deriva externa foram apresentadas com e sem correção do efeito de suavização, já que o algoritmo referido requer que os pesos atribuídos às amostras sejam positivos, o que não ocorre nos casos de cokrigagem ordinária e colocalizada.

MATERIAIS E MÉTODOS

Os dados completos das variáveis primária e secundária originais (Figura 2) apresentam coeficiente de correlação igual a 0,993 (Figura 3). A obtenção dos conjuntos de dados com menores coeficientes de correlação entre variáveis envolveu duas etapas. A primeira consistiu em introduzir ruído branco em níveis crescentes aos dados da variável secundária original obtendo-se

quatro variáveis secundárias com correlação decrescente com a variável primária original (Figura 4).

Na segunda etapa, foi realizada amostragem aleatória estratificada dos dados da variável primária e das cinco variáveis secundárias de modo a simular uma situação real de um empreendimento em geociências em que os dados iniciais são geralmente redu-

zidos. Assim sendo, foram amostrados 104 pontos de dados (Figura 5) para cada variável.

Uma vez obtidos os dados, foram calculadas as estatísticas descritivas dos mesmos. Seguiu-se, então, a análise geoestatística e a validação cruzada dos dados. Finalmente, foram realizadas as co-estimativas via cokrigagem ordinária, cokrigagem colocalizada e krigagem com deriva externa com e sem correção do efeito de suavização. Todos os processamentos foram efetuados através dos programas do Sistema GeoVisual (Yamamoto, 2006).

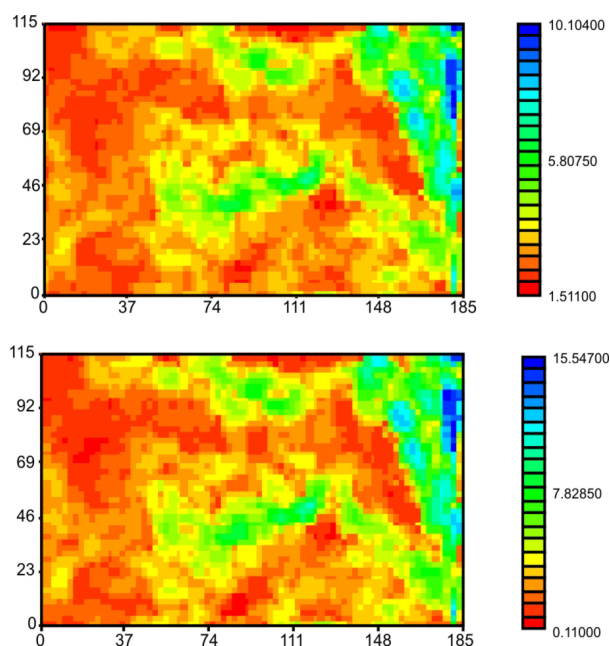


FIGURA 2. Mapa imagem da variável primária (acima) e da secundária (abaixo).

Os resultados são mapas imagem e estatísticas descritivas dos valores co-estimados, além de diagramas probabilidade – probabilidade, utilizados, respectivamente, para o estudo da reprodução do padrão original e da precisão global dos resultados. Também foram obtidos diagramas de dispersão entre a variável primária real e a estimada e entre as variáveis primária estimada e secundária a partir do quais foram analisadas a precisão local das co-estimativas e a preservação da correlação amostral, respectivamente.

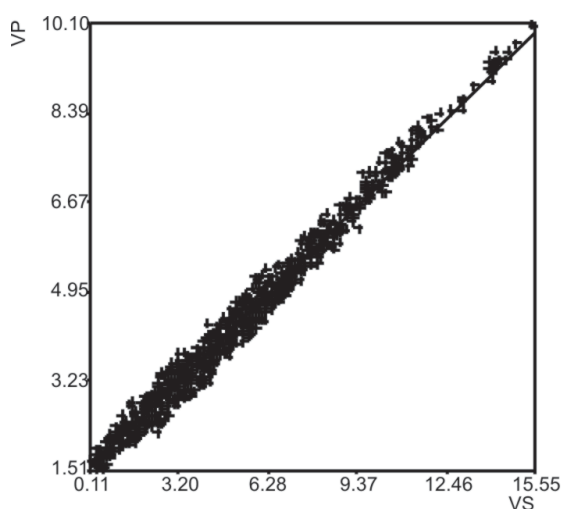


FIGURA 3. Diagrama de dispersão mostrando a alta correlação entre a variável primária e a variável secundária originais.

APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

COKRIGAGEM ORDINÁRIA

Na Tabela 1, o valor médio das co-estimativas não reproduz a média amostral e, à medida que o coeficiente de correlação inicial entre as variáveis primária e secundária diminui, as médias aumentam, muito provavelmente por causa da influência crescente da variável secundária nas estimativas. Ainda nesta tabela, verifica-se que o desvio padrão apresenta comportamento semelhante ao da média, sugerindo, neste caso, que as co-estimativas não se apresentam suavizadas, ao contrário do que se esperaria de um estimador baseado em média ponderada.

Os valores de máximo e mínimo das co-estimativas estão bastante distantes dos amostrais, havendo inclusive co-estimativas negativas conforme indica o valor mínimo para coeficiente de correlação igual a

0,993 (Tabela 1). A ampliação do intervalo de valores das co-estimativas reflete o maior intervalo de valores dos dados secundários em relação aos primários, influenciando diretamente o estimador da cokrigagem ordinária, o qual considera variáveis com dimensões distintas, sem as diferenciar. Já a presença de valores co-estimados negativos decorre da condição de não viés da cokrigagem ordinária segundo a qual a somatória dos pesos da variável secundária deve ser igual a zero, o que resulta obrigatoriamente em pesos negativos e consequentemente em co-estimativas negativas. Entretanto, deve-se salientar que nem sempre co-estimativas negativas são passíveis de ocorrer na realidade.

Os mapas imagem dos valores co-estimados a partir dos dados amostrais com maiores coeficientes

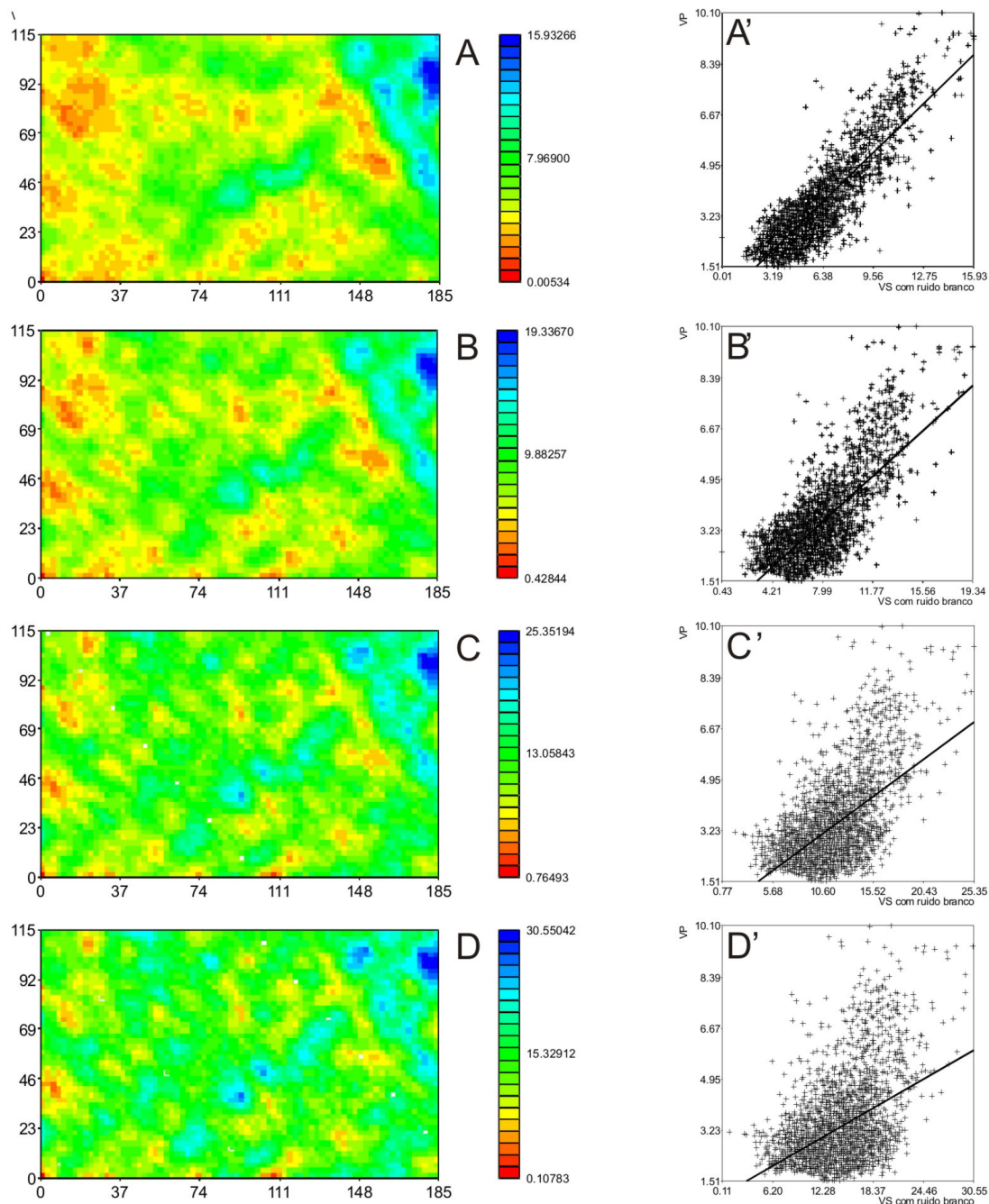


FIGURA 4. Mapa imagem da variável secundária com ruído branco (esquerda) e diagrama de dispersão entre a variável primária e a variável secundária deteriorada (direita). Coeficiente de correlação igual a 0,870 (A e A'), 0,752 (B e B'), 0,588 (C e C') e 0,461 (D e D').

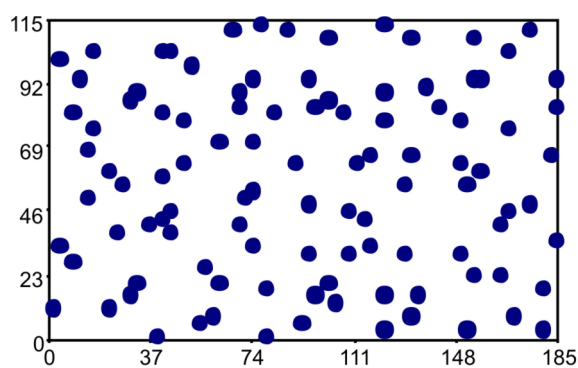


FIGURA 5. Mapa de localização dos pontos de dados com informação primária e secundária.

TABELA 1. Estatísticas amostrais e das co-estimativas feitas por cokrigagem ordinária.

Estatística	Am. 104	$r = 0,993$	$r = 0,870$	$r = 0,752$	$r = 0,588$	$r = 0,461$
No. Dados	104	3116	3116	3116	3116	3116
Média	3,146	3,376	5,400	7,073	10,498	13,018
Desvio	1,370	2,514	2,449	2,889	3,956	4,984
CV	0,435	0,745	0,454	0,408	0,377	0,383
Máx.	7,541	16,546	15,903	19,243	25,606	30,880
QS	3,272	4,161	6,589	8,683	12,685	16,118
Med.	2,646	2,750	4,921	6,905	10,730	13,485
QI	2,287	1,780	3,675	5,167	8,467	10,531
Mín.	1,787	-2,577	0,530	1,061	1,330	1,533

de correlação (Figura 6 A, B, C) são mais heterogêneos e refletem, embora muito parcialmente, tanto a distribuição de valores da variável primária (Figura 2 esquerda) quanto da secundária (Figura 2 direita e Figura 4 A, B). Os outros mapas (Figura 6 D, E) reproduzem apenas a distribuição de valores dos dados secundários (Figura 4 C, D), estando ausentes as características do mapa da variável primária. Esse resultado indica uma maior influência da variável secundária nas co-estimativas quando o coeficiente de correlação amostral é mais baixo.

Observa-se ainda que qualquer que seja o coeficiente considerado, há faixas de valores constantes em todos os mapas, sendo que à medida que o coeficiente diminui, a cor das mesmas tende àquela representativa dos valores mínimos das co-estimativas. A ocorrência dessas faixas está associada a co-estimativas iguais para diferentes valores reais da variável primária conforme pode ser observado no canto inferior esquerdo dos diagramas da Figura 6.

Com relação à precisão local das co-estimativas, verifica-se nos diagramas da Figura 6 que a regressão estimada não coincide com a reta bissetriz estando abaixo da mesma, exceto a porção inferior da reta de regressão da Figura 6 A. Assim, na maioria dos casos, os resultados foram superestimados por causa da influência da variável secundária cujos valores são mais elevados. Ainda nesta figura, observa-se que o coeficiente de correlação entre os dados primários reais e estimados diminui com a redução da correlação amostral, indicando que a influência da variável secundária nas co-estimativas é tanto maior quanto menor for a correlação amostral. Essa influência também pode ser inferida a partir dos valores máximos estimados, os quais aumentam com a diminuição da correlação inicial entre as variáveis e estão muito próximos dos valores máximos da variável secundária (Figura 2, abaixo e Figura 4).

Na Figura 7, observa-se que em todos os casos estudados as retas de regressão da estimativa e da amostra não coincidem, sendo que esta se encontra sempre abaixo daquela indicando uma melhor correlação entre as co-estimativas e a variável secundária justamente pela influência desta última na estimativa da primária. Além disso, o ângulo e a distância entre as retas aumentam conforme a diminuição da correlação inicial entre as variáveis primária e secundária. Assim sendo, a correlação amostral se mantém parcialmente apenas quando o coeficiente de correlação inicial entre as variáveis primária e secundária é elevado, como no caso do diagrama 7 A em que o coeficiente amostral é de 0,993.

COKRIGAGEM ORDINÁRIA COLOCALIZADA

Na Tabela 2, para qualquer coeficiente de correlação amostral, as médias das co-estimativas são superiores à média amostral, mas se encontram mais próximas desta quando comparadas com as médias das co-estimativas feitas por cokrigagem ordinária (Tabela 1). Ainda, ao contrário do que ocorreu com os resultados da cokrigagem ordinária, houve uma diminuição da média com a redução da correlação amostral, sugerindo que a cokrigagem colocalizada apresenta um melhor controle sobre a influência da variável secundária nas estimativas.

O desvio padrão das co-estimativas obtidas é inferior ao desvio padrão amostral em todos os casos, diminuindo com o decréscimo da correlação inicial entre os dados (Tabela 2). Dessa maneira, as estimativas obtidas mediante cokrigagem colocalizada resultam suavizadas e esta suavização é função inversa do coeficiente de correlação amostral.

Observando os valores de máximo e mínimo da Tabela 2, verifica-se ainda que a ampliação do intervalo de valores, por influência da variável secundária, também ocorre para qualquer correlação amostral. No

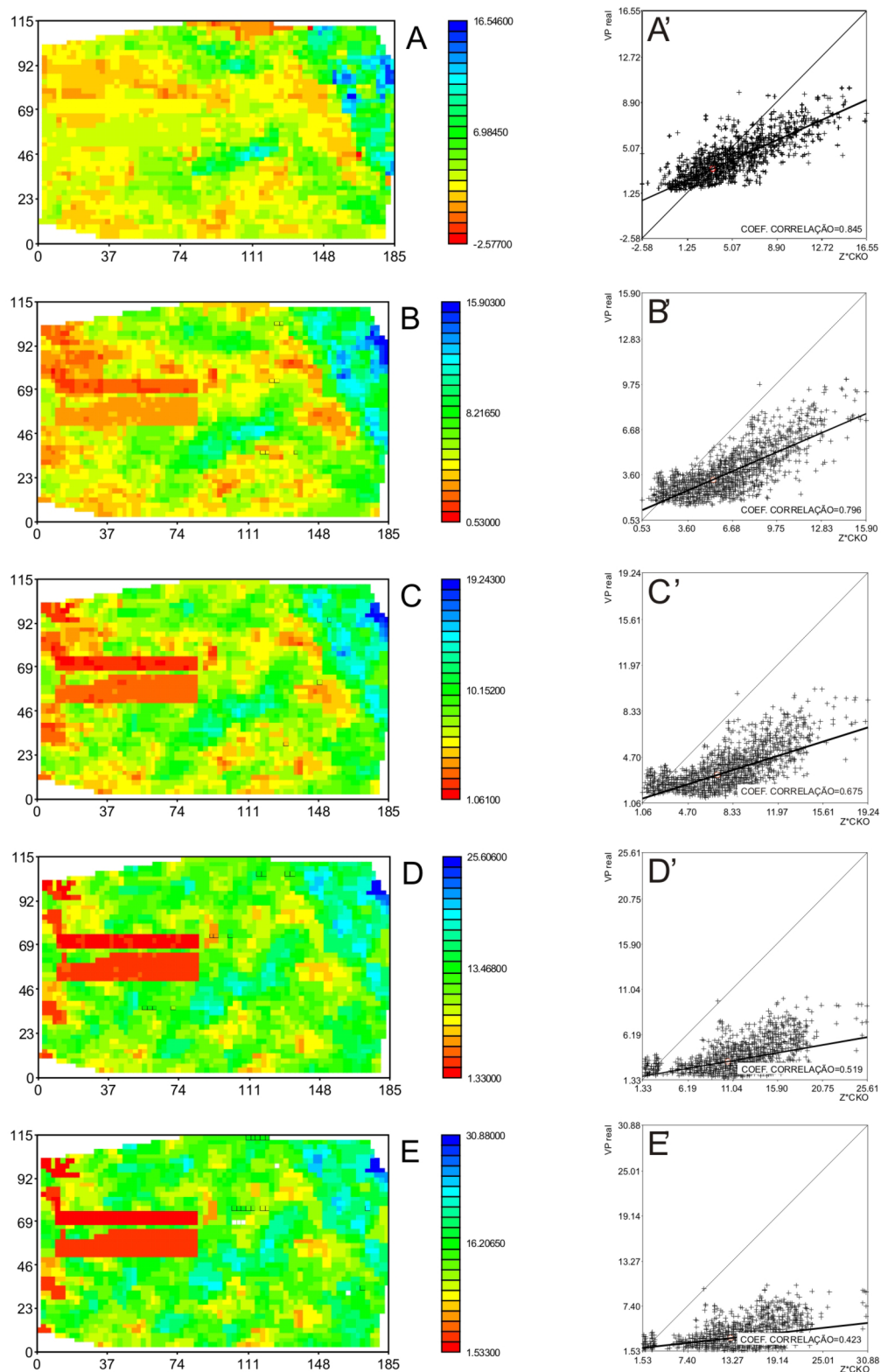


FIGURA 6. Mapas imagens (esquerda com índices de A até E) e diagramas de dispersão (direita índices de A' até E') das co-estimativas obtidas por cokrigagem ordinária. Os coeficientes de correlação inicial entre as variáveis primária e secundárias são: A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

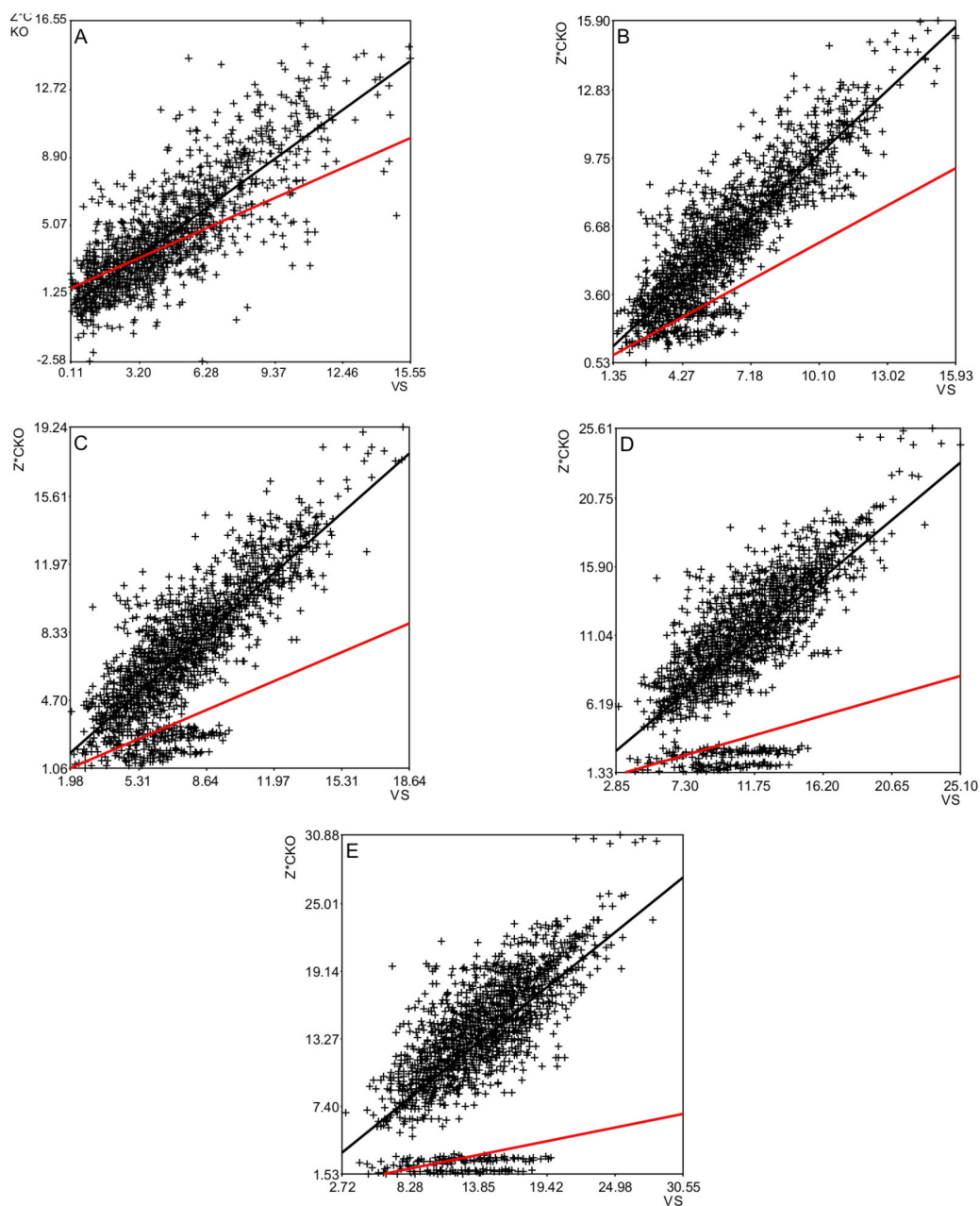


FIGURA 7. Diagramas de dispersão entre os valores estimados por cokrigagem ordinária e a variável secundária. Linha cheia preta = regressão da estimativa e linha vermelha cheia = regressão amostral. Os coeficientes de correlação inicial entre as variáveis primária e secundárias são: A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

TABELA 2. Estatísticas amostrais e das co-estimativas obtidas por cokrigagem ordinária colocolizada.

Estatística	Am. 104	$r = 0,993$	$r = 0,870$	$r = 0,752$	$r = 0,588$	$r = 0,461$
No. Dados	104	3404	3404	3404	3404	3404
Média	3,146	3,281	3,264	3,250	3,230	3,215
Desvio	1,370	1,342	1,311	1,313	1,276	1,241
CV	0,435	0,409	0,402	0,404	0,395	0,386
Máx.	7,541	9,931	8,849	8,796	8,263	7,829
QS	3,272	3,742	3,708	3,662	3,573	3,499
Med.	2,646	2,915	2,880	2,895	2,868	2,840
QI	2,287	2,375	2,408	2,399	2,424	2,445
Mín.	1,787	1,541	0,488	0,511	0,643	0,946

entanto, essa ampliação é menor relativamente àquela que ocorre com os resultados da cokrigagem ordinária (Tabela 1) e diminui conforme decresce a correlação inicial entre as variáveis primária e secundária. A menor influência da variável secundária pode estar relacionada à condição que restringe a soma dos pesos das variáveis primária e secundária a um, impedindo que o dado secundário colocalizado tenha um peso muito grande na estimativa.

Os diagramas da Figura 8 mostram como a distri-

buição estimada se ajusta à distribuição amostral. Observa-se que, para qualquer coeficiente de correlação amostral, o ajuste ocorre apenas para as caudas inferior e superior, sendo que à medida que a correlação diminui, a aderência diminui para a cauda inferior e aumenta para a superior. Esse resultado está relacionado ao fato da assimetria positiva das distribuições das co-estimativas ser muito mais acentuada relativamente à amostral, conforme indicam as diferenças entre os valores médio e mediano de cada distribuição.

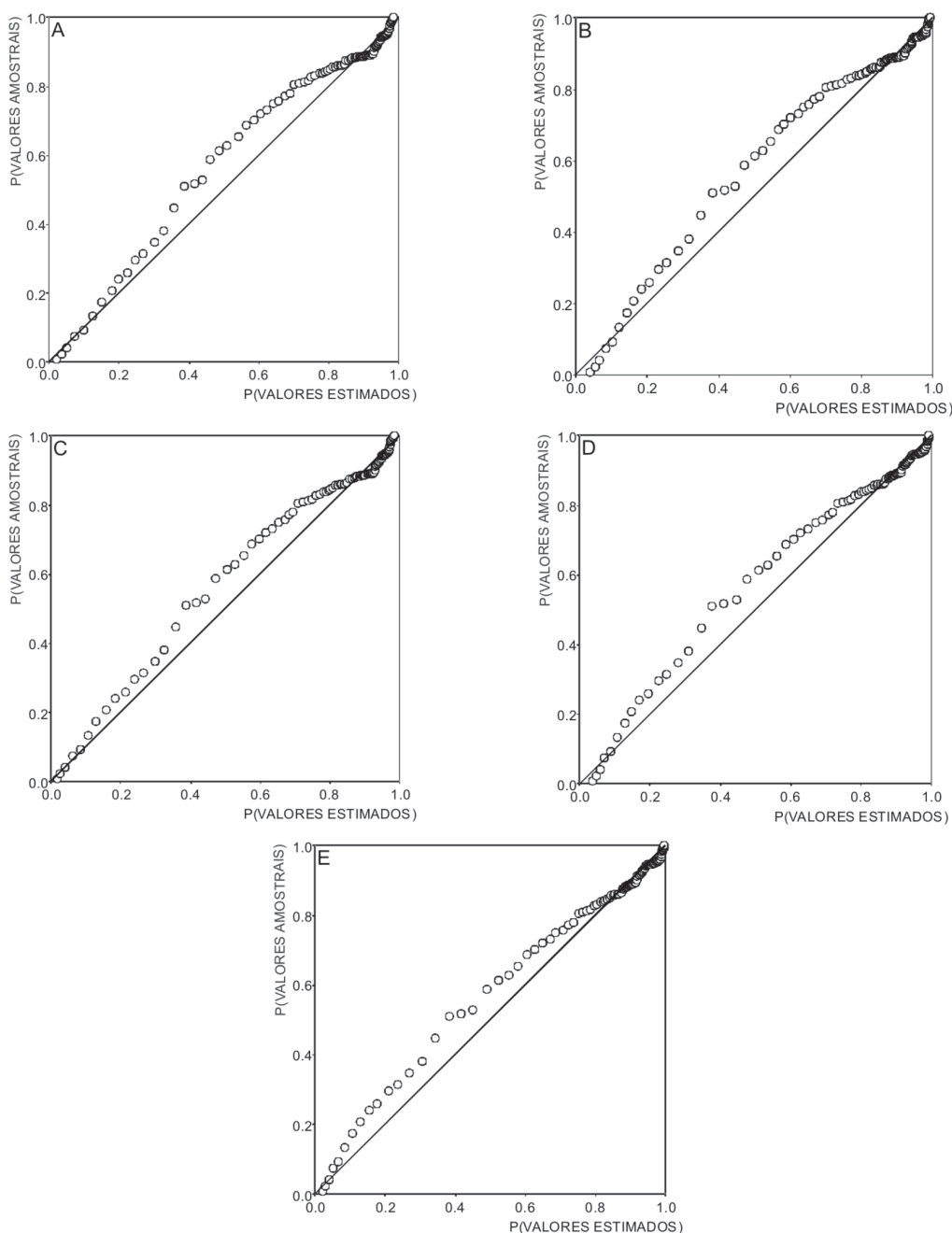


FIGURA 8. Diagramas P-P comparando a distribuição amostral com a distribuição da variável primária estimada pela cokrigagem ordinária colocalizada. Os coeficientes de correlação inicial entre as variáveis primária e secundárias são: A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

Na Figura 9 A, o mapa imagem obtido para coeficiente de correlação igual a 0,993 exibe um padrão de distribuição de valores praticamente igual ao da variável primária e da secundária (Figura 2 esquerda e direita, respectivamente). À medida que diminui a correlação inicial entre os dados, os demais mapas (Figura 9 B, C, D, E) preservam muito pouco o padrão do mapa da variável secundária (Figura 4 A, B, C, D) assemelhando-se mais àquele da variável primária não obstante haja alguma suavização. Além disso, as feições do mapa da variável primária estão mais bem preservadas nos mapas da Figura 9 do que naqueles obtidos por cokrigagem ordinária (Figura 6), sobretudo para menores coeficientes de correlação amostral.

A correlação entre a variável primária real e as co-estimativas da mesma (Figura 9) é alta para qualquer caso, sendo sempre maior em relação àquelas obtidas pela cokrigagem ordinária (Figura 6) e pouco diminuindo com o decréscimo da correlação inicial entre os dados primário e secundário. Assim, a influência da variável secundária nas co-estimativas feitas por cokrigagem colocalizada é menor do que na cokrigagem ordinária. Nesses mesmos diagramas, verifica-se que a porção superior da reta de regressão das co-estimativas encontra-se um pouco abaixo da reta bisetriz para correlações amostrais iguais e menores que 0,870 (Figura 9 B, C, D, E). Entretanto, não se pode afirmar que houve superestimativa haja vista o deslocamento entre as retas ser mínimo.

A regressão da estimativa coincide completamente com a regressão amostral quando os coeficientes de correlação amostral são mais altos (Figura 10 A, B) e parcialmente para coeficientes mais baixos (Figura 10 C, D, E). A dispersão da nuvem de pontos aumenta conforme diminui a correlação inicial entre os dados, indicando que apesar do bom ajuste entre as retas, a correlação entre as co-estimativas e os valores da variável secundária diminui.

KRIGAGEM COM DERIVA EXTERNA

As médias dos valores co-estimados não corrigidos (Tabela 3) são sempre maiores que a média amostral, o que pode ser decorrente da influência, mesmo que indireta, da variável secundária nas co-estimativas. Essa influência, entretanto, deve ser menor que aquela observada nos resultados dos métodos anteriores, haja visto os valores médios obtidos por krigagem com deriva externa estarem mais próximos do amostral. Há uma tendência de queda dos valores médios conforme diminui o coeficiente de correlação amostral sinalizando para uma menor influência da variável secundária.

Os desvios padrão obtidos pela krigagem com deriva externa são sempre menores que o desvio padrão amostral, o que mostra que os resultados encontram-

se suavizados. Essa suavização foi menor e maior do que aquelas obtidas, respectivamente, pela cokrigagem colocalizada e pela cokrigagem ordinária. Novamente, observa-se que o desvio padrão tende a diminuir com a redução da correlação amostral, sinalizando para o aumento da suavização dos resultados.

Em relação às co-estimativas não corrigidas (Tabela 3), nota-se que a média e o desvio padrão das co-estimativas corrigidas (Tabela 4) variam menos com a diminuição da correlação amostral e estão mais próximas dos equivalentes amostrais, sendo inclusive iguais aos mesmos para os coeficientes mais baixos. Dessa maneira, o algoritmo de correção do efeito de suavização proposto por Yamamoto (2005, 2007) reproduz de forma mais exata as estatísticas amostrais.

Os valores co-estimados máximo e mínimo (Tabelas 3 e 4) são respectivamente menores e maiores que os amostrais, indicando redução do intervalo de valores a qual está relacionada a dois fatores. Primeiro, com a suavização dos resultados e, segundo, com o fato deste método considerar apenas os dados da variável primária diretamente no estimador, diferentemente dos outros métodos. Os diagramas da Figura 11 mostram que, assim como as distribuições obtidas pela cokrigagem colocalizada, as distribuições das co-estimativas não corrigidas apresentam-se ajustadas à distribuição amostral apenas na cauda superior. Por outro lado, a distribuição das estimativas corrigidas do efeito de suavização da krigagem com deriva externa se aproxima bastante de toda a distribuição amostral, independentemente do coeficiente de correlação inicial entre as variáveis primária e secundária. Comparando-se as diferenças entre média e mediana em cada caso, confirma-se esse resultado já que a assimetria positiva das distribuições das co-estimativas corrigidas é mais próxima à amostral do que aquela das não corrigidas.

Os mapas imagem corrigidos e não corrigidos da Figura 12 são muito semelhantes entre si quando se considera um mesmo coeficiente de correlação amostral. As diferenças ocorrem nos locais onde são registrados valores máximos e mínimos os quais são, respectivamente, superiores ou inferiores e definem manchas maiores nos mapas corrigidos, o que está de acordo com o esperado já que a correção do efeito de suavização resulta no aumento do desvio padrão dos resultados. Todos os mapas desta figura exibem o padrão geral de distribuição de valores do mapa da variável primária inicial (Figura 2), havendo poucas diferenças entre eles.

Nota-se que há pouca modificação no padrão dos mapas da Figura 12 à medida que o coeficiente de correlação amostral diminui, ao contrário do que ocorre com os mapas obtidos a partir da cokrigagem ordinária

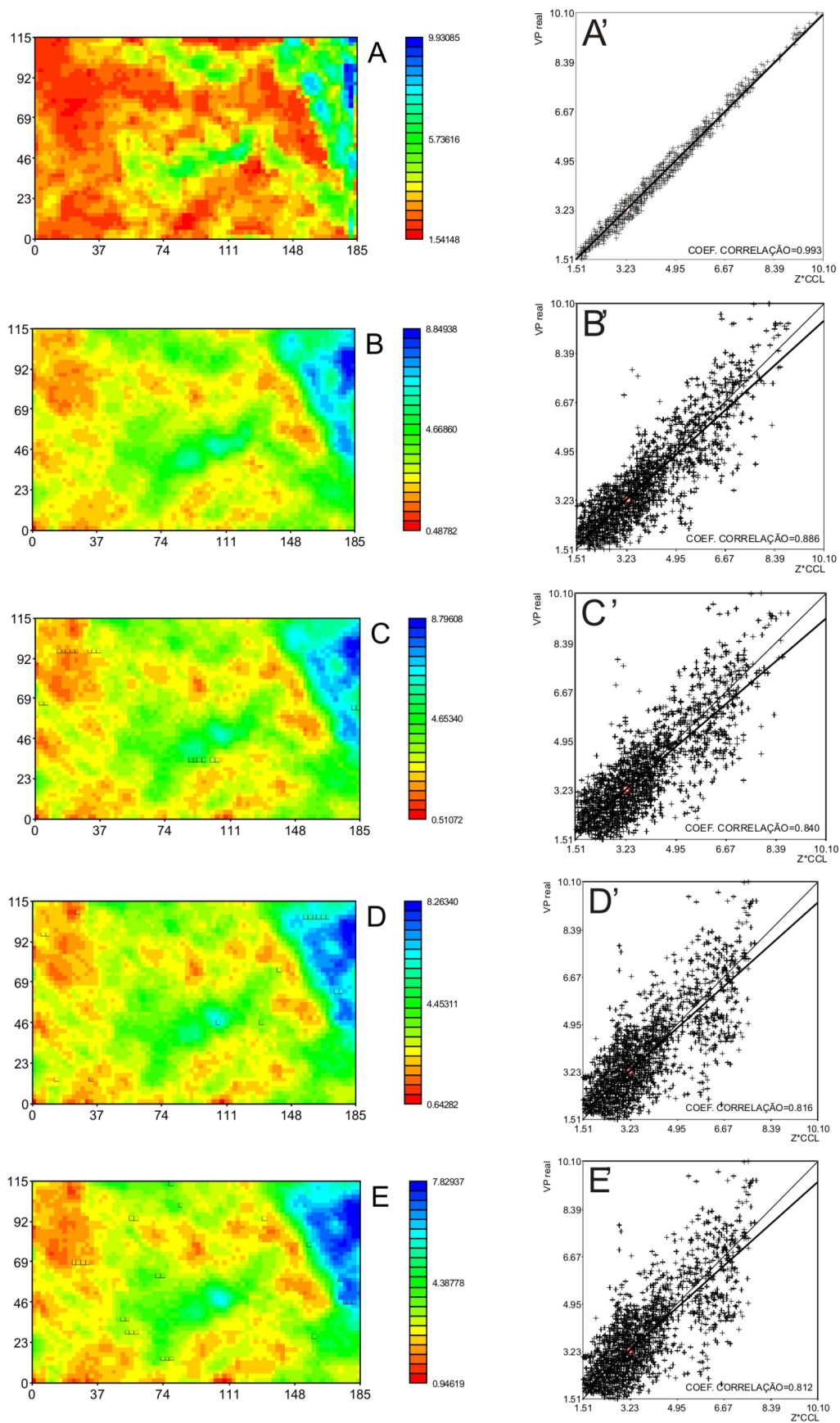


FIGURA 9. Mapas imagens (esquerda) e diagramas de dispersão (direita) das co-estimativas obtidas por cokrigagem ordinária colocalizada. Os coeficientes de correlação inicial entre as variáveis primária e secundárias são: A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

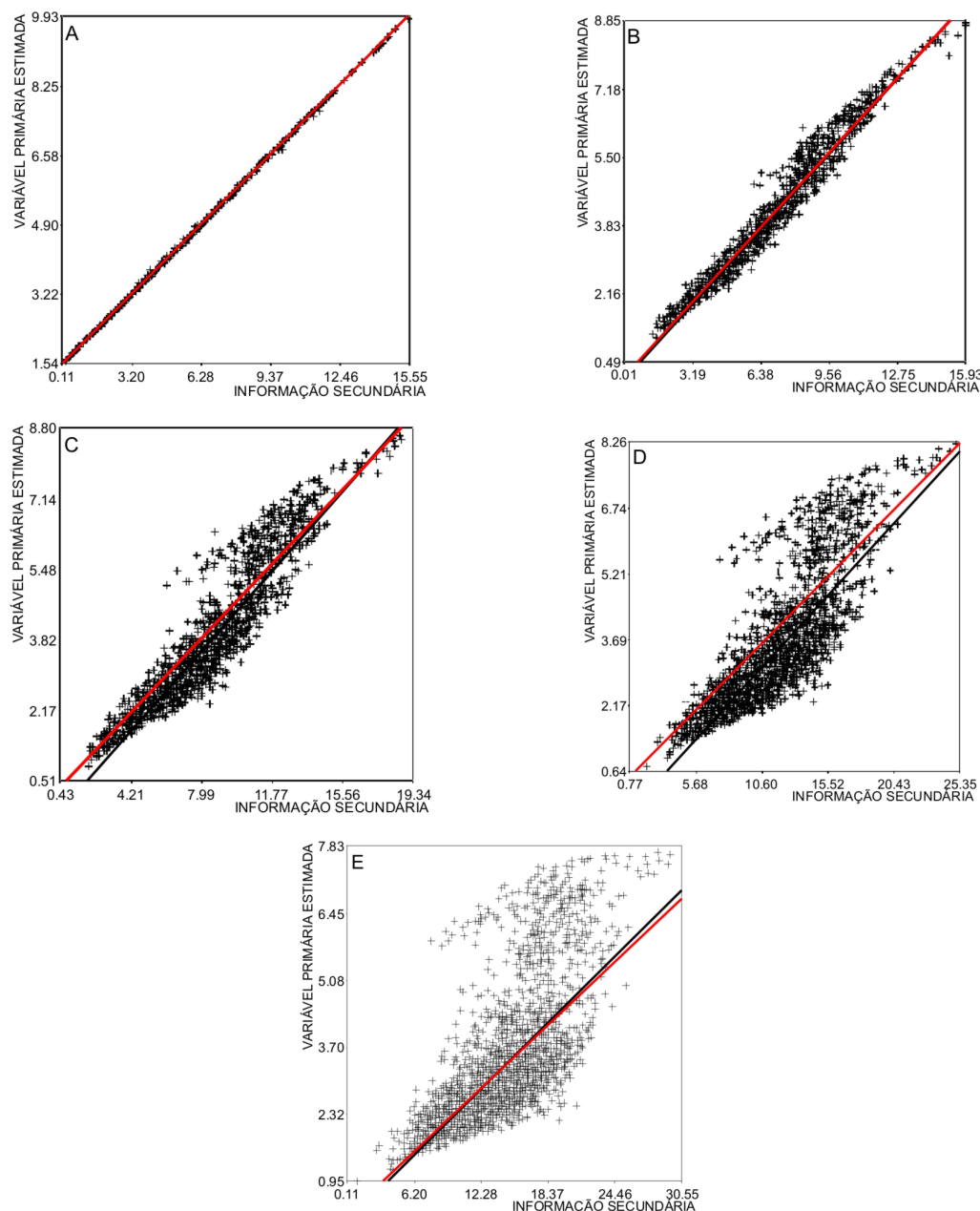


FIGURA 10. Diagramas de dispersão entre a variável primária estimada pela cokrigagem ordinária colocalizada e a variável secundária. Linha cheia preta = regressão da estimativa e linha vermelha cheia = regressão amostral. Os coeficientes de correlação inicial entre as variáveis primária e secundárias são: A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

TABELA 3. Estatísticas amostrais e dos valores estimados pela krigagem com deriva externa.

Estatística	Am. 104	$r = 0,993$	$r = 0,870$	$r = 0,752$	$r = 0,588$	$r = 0,461$
No. Dados	104	3404	3404	3404	3404	3404
Média	3,146	3,181	3,200	3,190	3,182	3,174
Desvio	1,370	1,049	1,107	1,099	1,081	1,069
CV	0,435	0,330	0,346	0,345	0,340	0,337
Máx.	7,541	7,477	7,308	7,262	7,227	7,166
QS	3,272	3,367	3,371	3,339	3,315	3,288
Med.	2,646	2,846	2,812	2,795	2,798	2,806
QI	2,287	2,524	2,539	2,541	2,555	2,558
Mín.	1,787	1,826	1,893	1,935	1,936	1,938

TABELA 4. Estatísticas amostrais e dos valores corrigidos do efeito de suavização da krigagem com deriva externa.

Estatística	Am. 104	$r = 0,993$	$r = 0,870$	$r = 0,752$	$r = 0,588$	$r = 0,461$
No. Dados	104	3404	3404	3404	3404	3404
Média	3,146	3,149	3,147	3,146	3,146	3,146
Desvio	1,370	1,366	1,369	1,370	1,370	1,370
CV	0,435	0,434	0,435	0,435	0,435	0,435
Máx.	7,541	7,541	7,398	7,398	7,412	7,412
QS	3,272	3,452	3,372	3,302	3,270	3,278
Med.	2,646	2,695	2,663	2,670	2,659	2,655
QI	2,287	2,232	2,281	2,302	2,317	2,319
Mín.	1,787	1,787	1,787	1,787	1,787	1,787

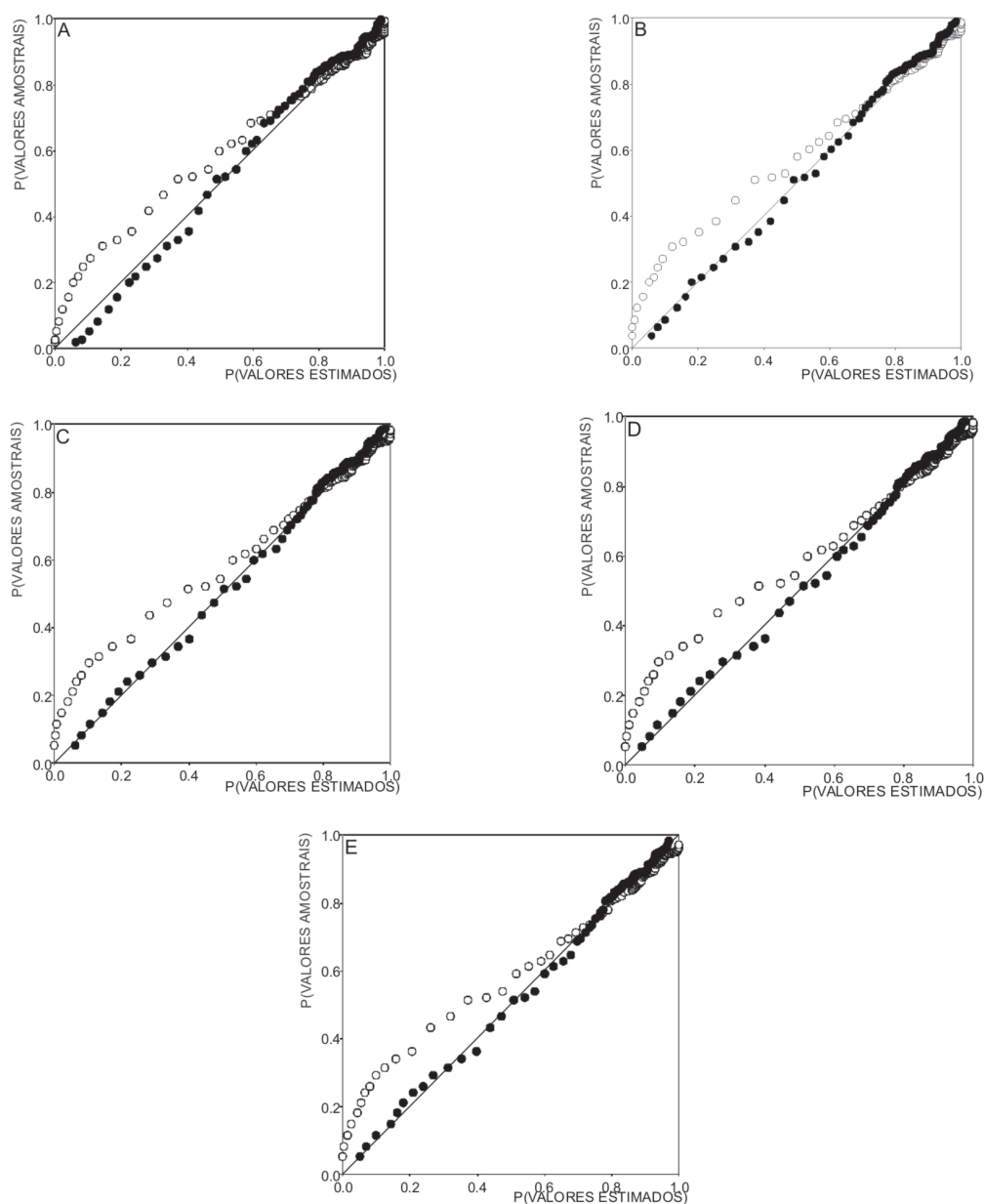


FIGURA 11. Diagramas P-P comparando a distribuição amostral com a distribuição da variável primária estimada pela krigagem com deriva externa. Os coeficientes de correlação inicial entre as variáveis primária e secundárias são: A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. Círculos vazios=krigagem com deriva externa; círculos cheios=estimativas corrigidas do efeito de suavização.

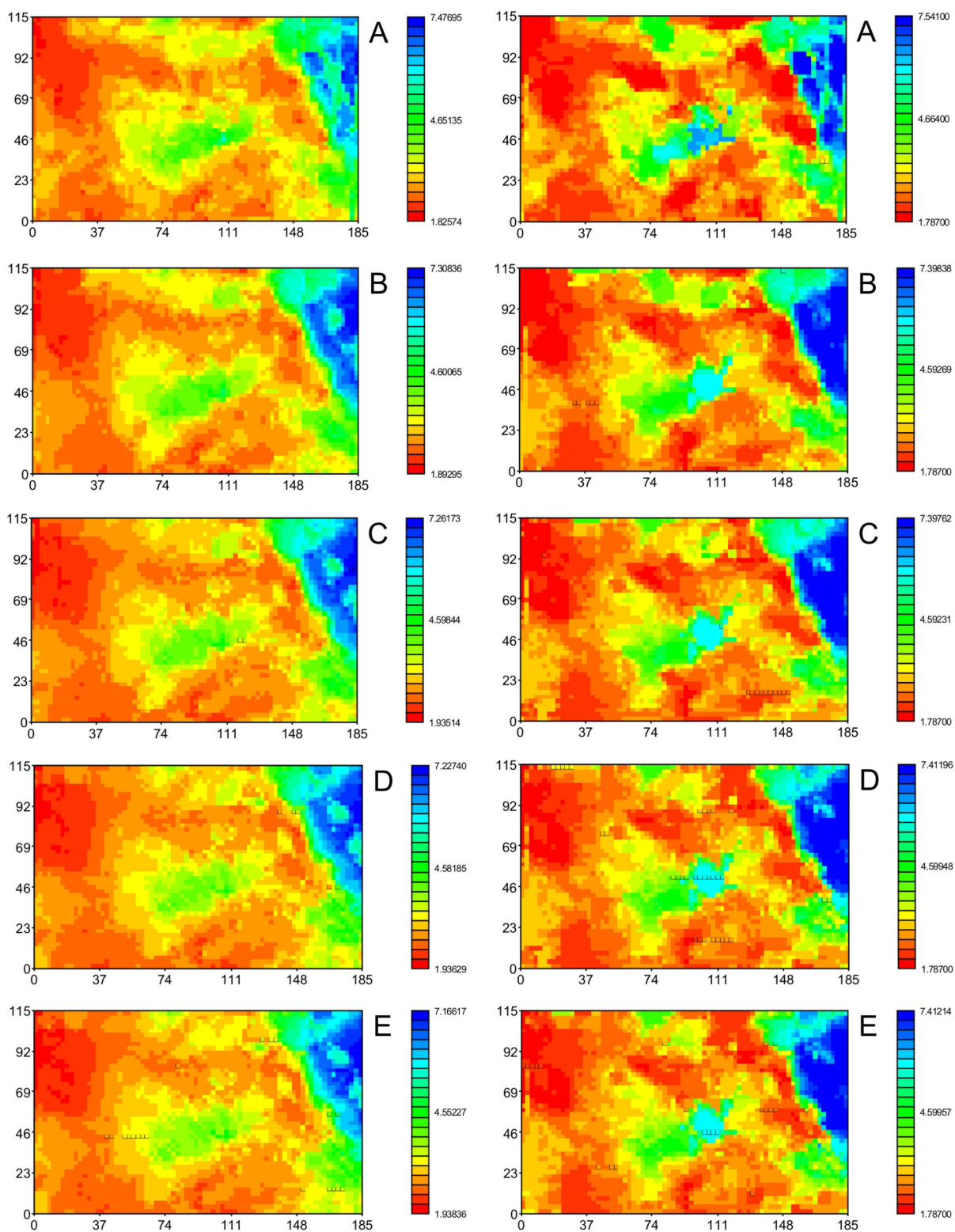


FIGURA 12. Mapas imagens das co-estimativas obtidas por krigagem com deriva externa (esquerda) e das co-estimativas corrigidas do efeito de suavização (direita). Os coeficientes de correlação inicial entre as variáveis primária e secundárias são: A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

(Figura 6) e da cokrigagem colocalizada (Figura 8) os quais consideram os valores da variável secundária (Figuras 3 e 4) diretamente no estimador. A krigagem com deriva externa, por sua vez, apenas imprime a geometria dessa variável em seus resultados.

Os coeficientes de correlação entre as co-estimativas e os dados primários reais (Figura 13) são maiores e menores que os obtidos, respectivamente, pela cokrigagem ordinária (Figura 6) e pela cokrigagem colocalizada (Figura 8).

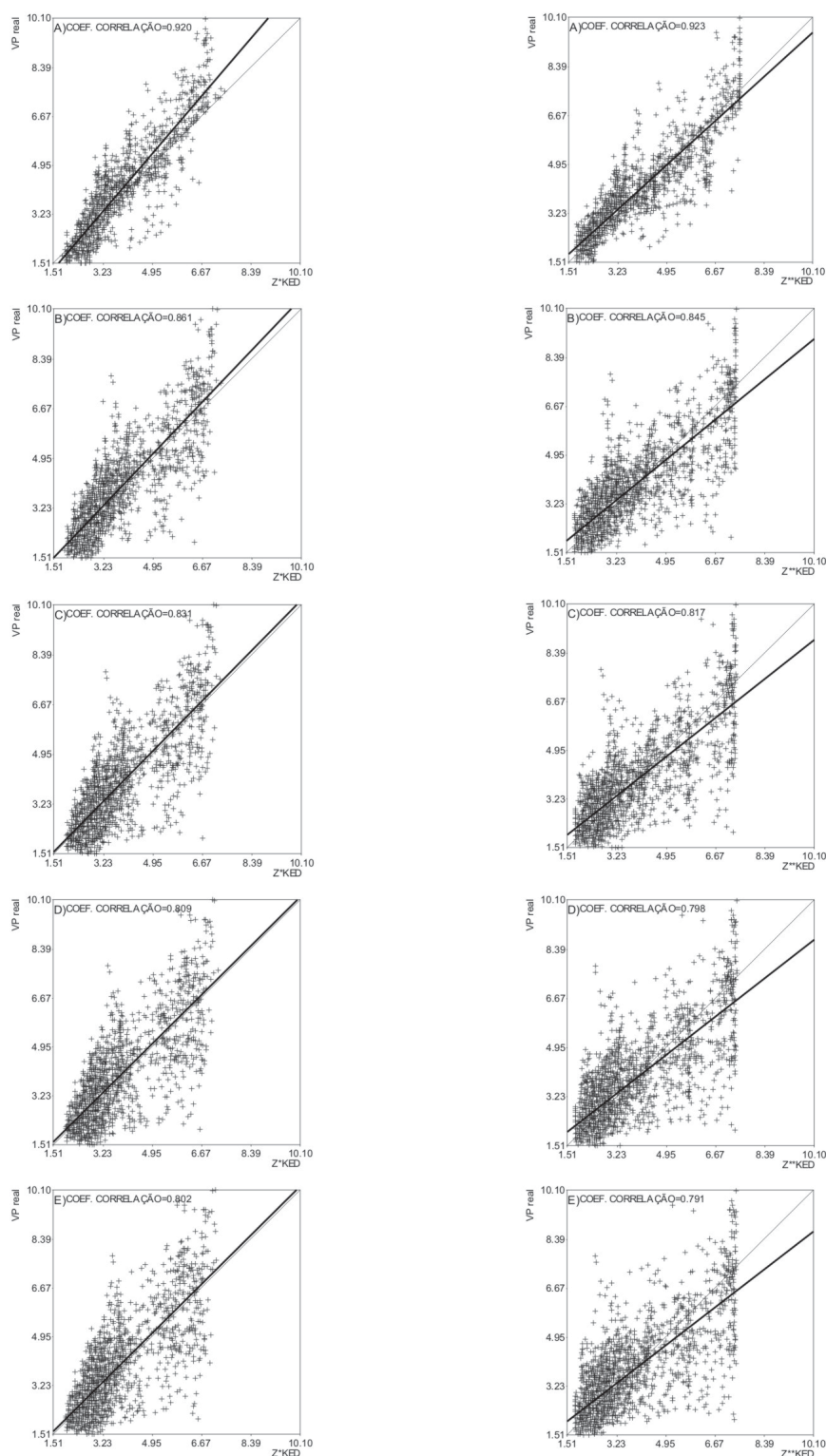


FIGURA 13. Diagramas de dispersão entre valores reais e valores estimados pela krigagem com deriva externa (esquerda) e dos valores corrigidos do efeito de suavização (direita). Os coeficientes de correlação inicial entre as variáveis primária e secundárias são: A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

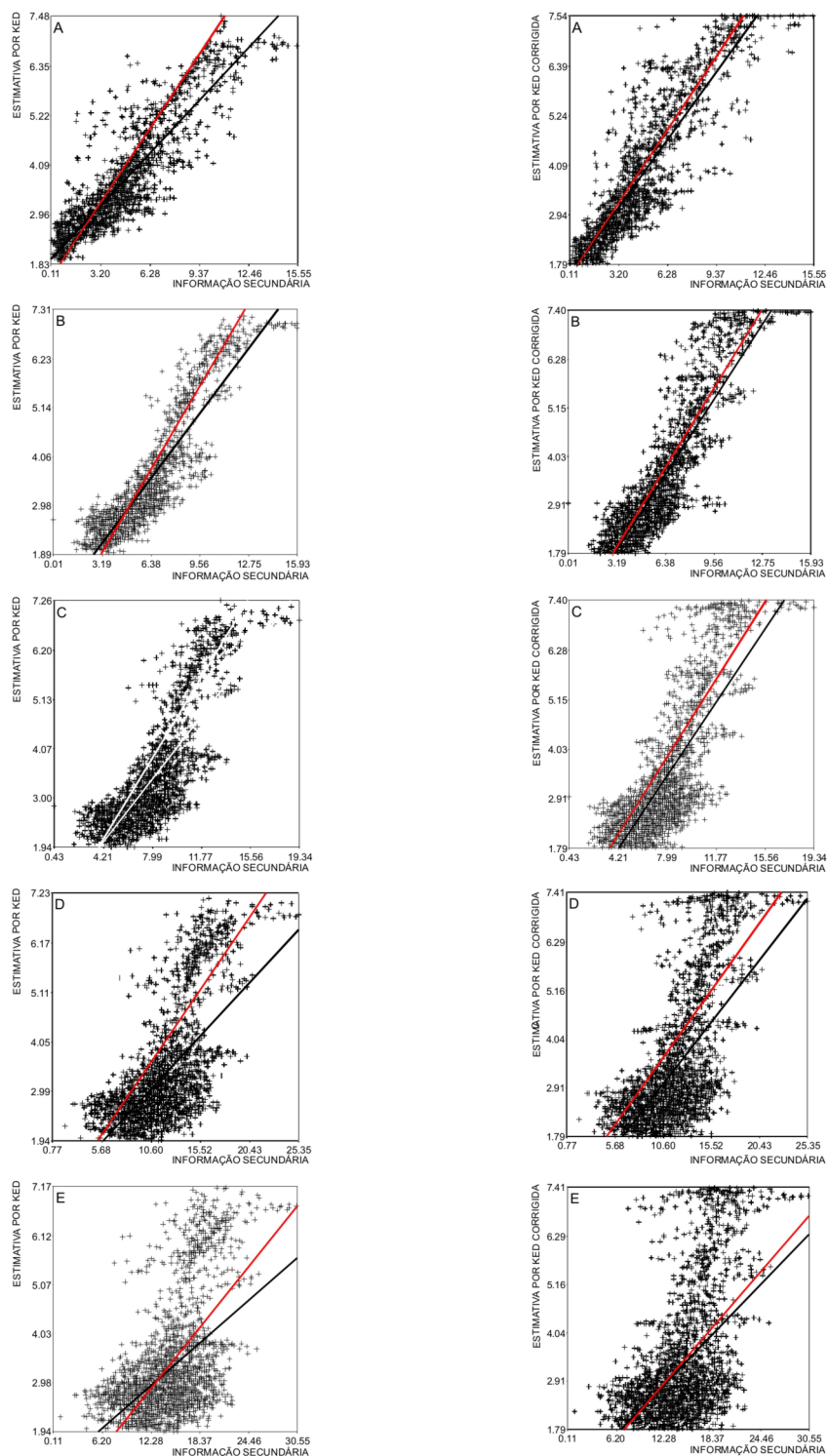


FIGURA 14. Diagramas de dispersão entre a informação secundária e a variável primária estimada pela krigagem com deriva externa (esquerda) e as estimativas corrigidas (direita). Linha cheia preta = regressão da estimativa e linha vermelha cheia = regressão amostral. Os coeficientes de correlação inicial entre as variáveis primária e secundárias são: A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

Dessa maneira, em termos de precisão local, este método apresenta melhores resultados apenas em relação à cokrigagem ordinária. A melhor precisão local obtida pela cokrigagem colocalizada deve-se provavelmente a utilização mais criteriosa da informação secundária, sobretudo quando a correlação entre as variáveis é baixa. No caso da krigagem com deriva externa, a precisão local das co-estimativas fica comprometida, porque os pesos utilizados em seu sistema refletem a correlação inicial entre as variáveis por estarem restritos à seguinte condição:

$$\sum_{j=1}^n \lambda_j s(x_j) = s(x_0)$$

Conforme indica a diminuição do coeficiente de correlação entre as co-estimativas e os valores primários reais (Figura 13), o comprometimento dos resultados é tanto maior quanto menor for a correlação entre as variáveis primária e secundária.

As co-estimativas corrigidas do efeito de suavização apresentam-se menos correlacionadas com os valores reais se comparadas com as co-estimativas não corrigidas para um mesmo coeficiente de correlação amostral, exceto quando esta última é igual a 0,993

(Figura 13). Esses resultados mostram que o algoritmo de correção do efeito de suavização melhora a precisão local apenas quando a correlação inicial entre as variáveis é alta, amplificando o problema decorrente da condição de restrição imposta aos pesos da krigagem com deriva externa nos casos de média e baixa correlação amostral.

Comparando-se os diagramas da Figura 14 para um mesmo coeficiente de correlação amostral, verifica-se que a regressão das co-estimativas não corrigidas encontra-se mais distante da regressão amostral relativamente àquela das co-estimativas corrigidas. A dispersão da nuvem de pontos, no entanto, pouco se altera de um diagrama para outro. Dessa maneira, a correção do efeito de suavização resultou em melhor preservação da correlação inicial entre as variáveis primária e secundária embora tenha havido perda de precisão local.

Ainda para uma mesma correlação amostral, as regressões das co-estimativas corrigidas e das não corrigidas (Figura 14) estão mais próximas da regressão amostral do que aquelas obtidas por cokrigagem ordinária (Figura 7) e mais distantes da regressão amostral se comparadas com as obtidas mediante cokrigagem colocalizada (Figura 9).

CONCLUSÕES

A escolha do método de co-estimativa mais adequado deve ser feita considerando o tipo de amostragem das variáveis e o objetivo do estudo, por exemplo, reprodução do histograma amostral ou obtenção de melhor precisão local. A correlação inicial entre as variáveis primária e secundária, por sua vez, influencia apenas na eficácia dos métodos aplicados.

A cokrigagem ordinária apresentou os piores resultados relativamente àqueles dos demais métodos. Por isso, aconselha-se que sua aplicação seja limitada apenas a situações em que as variáveis são muito bem correlacionadas (correlação igual ou maior que 0,993) e cuja amostragem é restrita. Já em situações de baixa correlação entre variáveis com amostragens restritas, é recomendável, se possível, melhorar a amostragem para aplicar um dos outros métodos.

Nos casos em que os dados são multicolocalizados, independentemente da correlação entre

as variáveis, a cokrigagem ordinária colocalizada constitui a melhor alternativa, pois preserva a correlação amostral e ainda apresenta ganho quanto à precisão local. Entretanto, este método não reproduz totalmente o histograma amostral.

A utilização da krigagem com deriva externa sem correção do efeito de suavização deve-se limitar aos casos de variáveis com alta correlação, pois, conforme mencionado, os pesos utilizados em seu sistema de equações dependem dessa correlação.

Finalmente, a opção pela krigagem com deriva externa corrigida do efeito de suavização deve ser feita quando o interesse for não apenas obter boa precisão local, mas também garantir a reprodução do histograma amostral. Deve-se restringir aos casos em que as variáveis apresentam alta correlação, já que à medida que o coeficiente de correlação entre as variáveis diminui, a correção da suavização apenas agrava os resultados no que tange à precisão local das estimativas.

AGRADECIMENTOS

Os autores agradecem à Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP (processo nº 03/10367-7) pelo auxílio financeiro concedido para o desenvolvimento da pesquisa de mestrado que originou este artigo.

REFERÊNCIAS BIBLIOGRÁFICAS

1. CHILÈS, J-P. & DELFINER, P. **Geostatistics: Modeling Spatial Uncertainty**. New York: John Wiley & Sons, Inc., 695 p., 1999.
2. GOOVAERTS, P. **Geostatistics for natural resources evaluation**. New York: Oxford University Press, 483 p., 1997.
3. MATHERON, G. Les Principes de la Geostatistique. **École Nationale Supérieure des Mines de Paris**. Fontainebleau, França. Rapport 65, 26 p., 1965.
4. WACKERNAGEL, H. **Multivariate Geostatistics**. Berlin: Springer-Verlag, 291 p., 1998.
5. WACKERNAGEL, H. **Multivariate Geostatistics**. Berlin: Springer-Verlag, 387 p., 2003.
6. YAMAMOTO, J.K. An Alternative Measure of the Reliability of Ordinary Kriging Estimates. **Mathematical Geology**, v. 32, p. 489-509, 2000.
7. YAMAMOTO, J.K. Correcting the Smoothing Effect of Ordinary Kriging Estimates. **Mathematical Geology**, v. 32, p. 69-94, 2005.
8. YAMAMOTO, J. K. **Sistema GeoVisual Versão 3**, São Paulo, 2006.
9. YAMAMOTO, J.K. On unbiased backtransform of lognormal kriging estimates. **Computational Geosciences**, v. 11, p. 219-234, 2007.
10. XU, W.; TRAN, T.T.; SRIVASTAVA, R.M.; JOURNEL, A.G. Integrating Seismic Data in Reservoir Modelling: The Collocated Cokriging alternative. In: ANNUAL TECHNICAL CONFERENCE OF THE SOCIETY OF PETROLEUM ENGINEERS, 1992. **Proceedings...** v. 67, p. 833-842.

*Manuscrito Recebido em: 30 de setembro de 2009
Revisado e Aceito em: 20 de janeiro de 2010*