UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE GEOCIÊNCIAS

MÉTODOS GEOESTATÍSTICOS DE CO-ESTIMATIVAS: ESTUDO DO EFEITO DA CORRELAÇÃO ENTRE VARIÁVEIS NA PRECISÃO DOS RESULTADOS

Jorge Watanabe

Orientador: Prof. Dr. Jorge Kazuo Yamamoto

DISSERTAÇÃO DE MESTRADO

Programa de Pós-Graduação em Recursos Minerais e Hidrogeologia

São Paulo

2008

Agradecimentos

Agradeço ao Prof. Dr. Jorge Kazuo Yamamoto pela inestimável orientação, pelo auxílio nos momentos necessários e firmeza dos momentos devidos.

Agradeço à CAPES e à FAPESP, processo 03/10367-7, pelo apoio financeiro concedido para elaboração desta dissertação de mestrado.

Ao Departamento de Geologia Sedimentar e Ambiental (GSA) agradeço o apoio e acolhida, bem como a possibilidade e infra-estrutura necessárias para a conclusão desta dissertação.

Agradeço ao Prof. Dr. Marcelo Monteiro da Rocha pelo auxílio nos momentos difíceis, pelos momentos de descontração e pelas conversas e discussões que tornaram o percurso desta etapa mais produtivo e agradável.

Ao Prof. Dr. Teodoro Isnard Ribeiro de Almeida agradeço pela orientação, no início da pós-graduação e pelo suporte fornecido em diversos momentos.

À Ana Paula Cabanal e Magali Poli Fernandes Rizzo pelo auxílio, pela atenção e dedicação oferecidas durante todo o curso.

Aos amigos Antonio Tadashi Kikuda, Frabrício Baú Dalmas e Sidney Schaberle Gouveia agradeço pelos bons momentos que tivemos no laboratório e pelo apoio oferecido sempre que necessário.

Agradeço também ao pessoal da Seção de Gráfica e Encadernação, Henrique, Claudionor, Edmir e José.

Gostaria finalmente de agradecer à Andreia C. Teodoro pelo apoio incondicional durante todo o período desta pesquisa, e aos meus pais, Shigueru e Aparecida, que indiscutivelmente possibilitaram tudo.

ABSTRACT

This master dissertation presents the results of a survey into co-estimation methods commonly used in geostatistics. These methods are ordinary cokriging, collocated cokriging and kriging with an external drift. Besides that ordinary kriging was considered just to illustrate how it does work when the primary variable is poorly sampled. As we know co-estimation methods depend on a secondary variable sampled over the estimation domain. Moreover, this secondary variable should present linear correlation with the main variable or primary variable. Usually the primary variable is poorly sampled whereas the secondary variable is known over the estimation domain. For instance in oil exploration the primary variable is porosity as measured on rock samples gathered from drill holes and the secondary variable is seismic amplitude derived from processing seismic reflection data. It is important to mention that primary and secondary variables must present some degree of correlation. However, we do not know how they work depending on the correlation coefficient. That is the question. Thus, we have tested co-estimation methods for several data sets presenting different degrees of correlation. Actually, these data sets were generated in computer based on some data transform algorithms. Five correlation values have been considered in this study: 0.993; 0.870; 0.752; 0.588 and 0.461. Collocated simple cokriging was the best method among all tested. This method has an internal filter applied to compute the weight for the secondary variable, which in its turn depends on the correlation coefficient. In fact, the greater the correlation coefficient the greater the weight of secondary variable is. Then it means this method works even when the correlation coefficient between primary and secondary variables is low. This is the most impressive result that came out from this research.

RESUMO

Esta dissertação de mestrado apresenta os resultados de uma investigação sobre os métodos de co-estimativa comumente utilizados em geoestatística. Estes métodos são: cokrigagem ordinária; cokrigagem colocalizada e krigagem com deriva externa. Além disso, a krigagem ordinária foi considerada apenas a título de ilustração como esse método trabalha quando a variável primária estiver pobremente amostrada. Como sabemos, os métodos de co-estimativa dependem de uma variável secundária amostrada sobre o domínio a ser estimado. Adicionalmente, esta variável deveria apresentar correlação linear com a variável principal ou variável primária. Geralmente, a variável primária é pobremente amostrada enquanto a variável secundária é conhecida sobre todo o domínio a ser estimado. Por exemplo, em exploração petrolífera, a variável primária é a porosidade medida em amostras de rocha retiradas de testemunhos e a variável secundária é a amplitude sísmica derivada de processamento de dados de reflexão sísmica. É importante mencionar que a variável primária e a variável secundária devem apresentar algum grau de correlação. Contudo, nós não sabemos como eles funcionam dependendo do grau de correlação. Esta é a questão. Assim, testamos os métodos de co-estimativa para vários conjuntos de dados apresentando diferentes graus de correlação. Na verdade, esses conjuntos de dados foram gerados em computador baseado em algoritmos de transformação de dados. Cinco valores de correlação foram considerados neste estudo: 0,993, 0,870, 0,752, 0,588 e 0,461. A cokrigagem colocalizada foi o melhor método entre todos testados. Este método tem um filtro interno que é aplicado no cálculo do peso da variável secundária, que por sua vez depende do coeficiente de correlação. De fato, quanto maior o coeficiente de correlação, maior é o peso da variável secundária. Então isso significa que este método funciona mesmo quando o coeficiente de correlação entre a variável primária e a variável secundária é baixo. Este é o resultado mais impressionante desta pesquisa.

Índice

Índice ____ *iii*

Índice de Figuras ____ *iv*

Índice de Tabelas ____ *viii*

CAPÍTULO 1 ____ *1*

INTRODUÇÃO ____ *1*

1.1 Objetivos ____ **2**

1.2 Justificativas ____ **3**

CAPÍTULO 2 ____ *4*

REVISÃO DA METODOLOGIA CORRENTE ____ *4*

2.1 Modelos de correlação espacial ____ **4**

2.1.1 A função semivariograma ____ 4

2.1.2 A função semivariograma cruzado ____ 5

2.1.2.1 Modelo linear de corregionalização ____ 6

2.1.3 A função variograma residual ____ 7

2.2 Amostragem de dados multivariados ____ **8**

2.3 Krigagem Ordinária ____ **10**

2.3.1 Variância de Krigagem ____ 11

2.3.2 Variância de Interpolação ____ 12

2.3.3 Correção do Efeito de Suavização ____ 13

2.4 Cokrigagem ____ **17**

2.4.1 Cokrigagem Simples ____ 20

2.4.2 Cokrigagem Ordinária ____ 23

2.4.3 Cokrigagem Colocalizada ____ 26

2.4.3.1 Cokrigagem colocalizada simples ____ 27

2.4.3.2 Cokrigagem colocalizada ordinária ____ 29

2.4.3.2.1 Utilização do Modelo Markoviano ____ 31

2.5 Krigagem com deriva externa ____ **32**

CAPÍTULO3 ____ *35*

MATERIAIS E MÉTODOS ____ *35*

CAPÍTULO 4 ____ *44*

APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS ____ *44*

4.1 Krigagem Ordinária ____ **44**

4.2 Cokrigagem Ordinária ____ **46**

4.3 Cokrigagem Colocalizada Simples ____ **51**

4.4 Krigagem com deriva externa ____ **64**

REFERÊNCIAS BIBLIOGRÁFICAS ____ *78*

Índice de Figuras

Figura 1: Configurações possíveis da distribuição espacial das variáveis amostradas: isotopia (esquerda), heterotopia parcial (centro) e heterotopia total (direita). Os círculos azuis representam a pontos de amostragem da variável primária, os vermelhos da variável secundária e os asteriscos representam os pontos em que as duas variáveis são conhecidas. 8

Figura 2: Comparação do desempenho da krigagem ordinária (B), regressão linear (C) e cokrigagem (D), segundo Doyen (1988) 19

Figura 3: Histograma amostral (esquerda) e dos resultados da cokrigagem ordinária (direita). 26

Figura 4: Mapas da diferença entre a topografia densamente amostrada e o nível de água estimado, utilizando-se krigagem ordinária (A) e cokrigagem colocalizada ordinária (B), conforme Boezio *et al.* (2006)..... 29

Figura 5: Mapa imagem da variável primária. 35

Figura 6: Mapa imagem da variável secundária (esquerda) e diagrama de dispersão mostrando a alta correlação entre a variável primária e a variável secundária (direita). 36

Figura 7: Mapa imagem da variável secundária com ruído branco (esquerda) e diagrama de dispersão entre a variável primária e a variável secundária com coeficiente de correlação 0,870 (direita)..... 37

Figura 8: Mapa imagem da variável secundária com ruído branco (esquerda) e diagrama de dispersão entre a variável primária e a variável secundária com coeficiente de correlação 0,752 (direita)..... 37

Figura 9: Mapa imagem da variável secundária com ruído branco (esquerda) e diagrama de dispersão entre a variável primária e a variável secundária com coeficiente de correlação 0,588 (direita)..... 37

Figura 10: Mapa imagem da variável secundária com ruído branco (esquerda) e diagrama de dispersão entre a variável primária e a variável secundária com coeficiente de correlação 0,461 (direita)..... 38

Figura 11: Diagramas de dispersão para as amostras retiradas do conjunto completo Composto20.txt com 60 unidades (A) e com 104 unidades (B). 40

Figura 12: Diagramas de dispersão para as amostras retiradas do conjunto completo Composto40.txt com 60 unidades (A) e com 104 unidades (B). 40

Figura 13: Diagramas de dispersão para as amostras retiradas do conjunto completo Composto100.txt com 60 unidades (A) e com 104 unidades (B). 40

Figura 14: Diagramas de dispersão para as amostras retiradas do conjunto completo Composto120.txt com 60 unidades (A) e com 104 unidades (B). 41

Figura 15: Diagramas de dispersão para as amostras retiradas do conjunto completo Composto121.txt com 60 unidades (A) e com 104 unidades (B). 41
Figura 16: Mapa de localização de pontos para arquivos isotópicos com 60 pontos. 43
Figura 17: Mapa de localização de pontos para arquivos isotópicos com 104 pontos. 43
Figura 18: Mapas imagens da krigagem ordinária (esquerda) e da krigagem ordinária corrigida (direita) para amostras com 60 unidades. 44
Figura 19: Mapas imagens da krigagem ordinária (esquerda) e da krigagem ordinária corrigida (direita) para amostras com 104 unidades. 45
Figura 20: Diagramas de dispersão entre a variável primária real e a estimativa pela krigagem ordinária (A) e a estimativa corrigida do efeito de suavização (B) para a amostra com 60 unidades. 46
Figura 21: Diagramas de dispersão entre a variável primária real e a estimativa pela krigagem ordinária (A) e a estimativa corrigida do efeito de suavização (B) para a amostra com 60 unidades. 46
Figura 22: Resultados da cokrigagem ordinária: mapas imagens à esquerda e diagramas de dispersão à direita para amostras com 60 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. 49
Figura 23: Resultados da cokrigagem ordinária: mapas imagens à esquerda e diagramas de dispersão à direita para amostras com 104 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. 50
Figura 24: Diagramas de dispersão entre a variável primária estimada e a variável secundária para amostras com 60 unidades à esquerda e com 104 unidades à direita. Linha cheia preta = regressão da estimativa e linha vermelha cheia = regressão amostral. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. 52
Figura 25: Resultados da cokrigagem colocalizada simples: mapas imagens à esquerda e diagramas de dispersão à direita para amostras com 60 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. 56
Figura 26: Resultados da cokrigagem colocalizada simples: mapas imagens à esquerda e diagramas de dispersão à direita para amostras com 104 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. 57
Figura 27: Diagramas de dispersão entre a variável primária estimada pela cokrigagem colocalizada simples e a variável secundária para amostras com 60 unidades à esquerda e com 104 unidades à direita. Linha cheia preta = regressão da estimativa e linha vermelha cheia = regressão amostral. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. 58
Figura 28: Diagramas P-P comparando a distribuição amostral com a distribuição da variável primária estimada pela cokrigagem colocalizada simples para amostras com 60

unidades à esquerda e com 104 unidades à direita. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. 59
Figura 29: Mapas imagens da soma dos pesos da variável primária (esquerda) e do peso da variável secundária (direita) para amostras com 60 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. 60
Figura 30: Mapas imagens da soma dos pesos da variável primária (esquerda) e do peso da variável secundária (direita) para amostras com 104 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. 61
Figura 31: Histogramas da soma dos pesos da variável primária à esquerda e do peso da variável secundária à direita para amostras com 60 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. 62
Figura 32: Histogramas da soma dos pesos da variável primária à esquerda e do peso da variável secundária à direita para amostras com 104 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. 63
Figura 33: Relação entre o coeficiente de correlação e a média da soma dos pesos da variável primária (A) e com a média do peso da informação secundária (B) para a amostra com 60 unidades. 64
Figura 34: Relação entre o coeficiente de correlação e a média da soma dos pesos da variável primária (A) e com a média do peso da informação secundária (B) para a amostra com 104 unidades. 64
Figura 35: Mapas imagens dos valores estimados pela krigagem com deriva externa (esquerda) e dos valores corrigidos do efeito de suavização (direita) para amostras com 60 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. 68
Figura 36: Mapas imagens dos valores estimados pela krigagem com deriva externa (esquerda) e dos valores corrigidos do efeito de suavização (direita) para amostras com 104 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. 69
Figura 37: Diagramas de dispersão entre valores reais e valores estimados pela krigagem com deriva externa (esquerda) e dos valores corrigidos do efeito de suavização (direita) para amostras com 60 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. .. 70
Figura 38: Diagramas de dispersão entre valores reais e valores estimados pela krigagem com deriva externa (esquerda) e dos valores corrigidos do efeito de suavização (direita) para amostras com 104 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. 71
Figura 39: Diagramas de dispersão entre a variável primária estimada pela krigagem com deriva externa (esquerda) e os mesmo para as estimativas corrigidas (direita) para amostras com 60 unidades. Linha cheia preta = regressão da estimativa e linha vermelha cheia = regressão amostral. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. 72

Figura 40: Diagramas de dispersão entre a variável primária estimada pela krigagem com deriva externa (esquerda) e os mesmo para as estimativas corrigidas (direita) para amostras com 104 unidades. Linha cheia preta = regressão da estimativa e linha vermelha cheia = regressão amostral. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. 73

Figura 41: Diagramas P-P comparando a distribuição amostral com a distribuição da variável primária estimada pela krigagem com deriva externa para amostras com 60 unidades à esquerda e com 104 unidades à direita. 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. 74

Índice de Tabelas

Tabela 1: Conjuntos completos e subconjuntos obtidos por amostragem aleatória estratificada. 39
Tabela 2: Estatísticas da variável primária para as amostras selecionadas. 39
Tabela 3: Retas de regressão da variável primária e a variável secundária para os subconjuntos amostrados. 41
Tabela 4: Subconjuntos amostrados adicionados com 33% dos nós da malha regular com valores da informação secundária. 42
Tabela 5: Estatísticas amostrais e dos valores estimados pela krigagem ordinária. 45
Tabela 6: Estatísticas amostrais e dos valores estimados pela cokrigagem ordinária para a amostra com 60 unidades. 47
Tabela 7: Estatísticas amostrais e dos valores estimados pela cokrigagem ordinária para a amostra com 104 unidades. 47
Tabela 8: Estatísticas amostrais e dos valores estimados pela cokrigagem colocalizada simples para a amostra com 60 unidades. 53
Tabela 9: Estatísticas amostrais e dos valores estimados pela cokrigagem colocalizada simples para a amostra com 104 unidades. 53
Tabela 10: Estatísticas amostrais e dos valores estimados pela krigagem com deriva externa para a amostra com 60 unidades. 65
Tabela 11: Estatísticas amostrais e dos valores corrigidos do efeito de suavização da krigagem com deriva externa para a amostra com 60 unidades. 65
Tabela 12: Estatísticas amostrais e dos valores estimados pela krigagem com deriva externa para a amostra com 104 unidades. 66
Tabela 13: Estatísticas amostrais e dos valores corrigidos do efeito de suavização da krigagem com deriva externa para a amostra com 104 unidades. 66

CAPÍTULO 1

INTRODUÇÃO

Na natureza, muitas variáveis podem apresentar correlação, seja por afinidade química, seja por propriedade física. Existem inúmeros exemplos na geoquímica, um deles é a correlação positiva entre sílica e alumina. Na indústria do petróleo, a correlação entre propriedades físicas é explorada em trabalhos de cubagem de reservatórios de hidrocarbonetos. O exemplo mais comum é a utilização da correlação da amplitude sísmica com a porosidade da rocha. Trata-se em correlacionar um dado de sísmica de reflexão, obtido em levantamentos sísmicos, com um dado obtido em laboratório (porosidade). O problema é que poços de petróleo são muito caros e, portanto, escassos, mesmo considerando campos em avançados estágios de desenvolvimento. Por exemplo, o Campo de Namorado (agora disponibilizado pela ANP) tem apenas 60 furos exploratórios e de desenvolvimento. Isso é muito pouco se compararmos com um prospecto em mineração onde há centenas de furos de pesquisa e outros milhares de furos de desenvolvimento (furos de desmonte) que também são utilizados para a reavaliação local de tonelagens e teores. Para fins do presente texto, vamos considerar a informação primária como objeto de investigação e a informação secundária, a informação adicional que pode ser utilizada para auxiliar a estimativa da primária.

Na geoestatística, existem duas técnicas que fazem uso da informação secundária, mais abundante, para prover a informação primária, mais escassa, ou seja, estas técnicas utilizam-se de variável(is) secundária(s) para estimar a variável primária. Elas são: cokrigagem e krigagem com deriva externa. A cokrigagem é uma técnica consagrada na literatura geoestatística e permite fazer a estimativa de uma variável primária com base na informação da variável secundária. Na indústria do petróleo, há um arranjo especial de dados em que a variável secundária é conhecida não só nos pontos de amostragem da variável primária, como também em quaisquer outros pontos. Assim, é possível obter a informação secundária sobre os nós da malha regular que se deseja estimar. Quando esta informação está disponível, pode-se fazer uso da

cokrigagem colocalizada com grande vantagem em relação a cokrigagem ordinária, pois elimina a informação secundária redundante em torno do ponto que se quer estimar. Além da cokrigagem colocalizada, pode-se utilizar também da krigagem com deriva externa, que também necessita da informação secundária sobre os nós da malha regular. Como se sabe, a aplicação das técnicas mencionadas depende do grau de correlação, entre as variáveis primária e secundária, medido em um diagrama de dispersão. Isso significa que a variável primária pode ser obtida através de uma relação funcional com a variável secundária. O grande desafio que se coloca é chegar à mesma relação funcional da variável primária estimada com a variável secundária disponível.

1.1 Objetivos

Este estudo tem por objetivo realizar uma comparação entre cokrigagem ordinária, cokrigagem colocalizada e krigagem com deriva externa em conjuntos de pontos de dados mostrando diversas correlações, bem como fazer uma análise crítica dos métodos e resultados. Além das técnicas mencionadas, deve-se mostrar que a krigagem ordinária não pode ser empregada em situações de escassez de informação sem prejuízo na interpretação dos resultados obtidos. Assim, inicialmente será mostrada a técnica da krigagem ordinária e, em seguida, as técnicas da cokrigagem ordinária, cokrigagem colocalizada e krigagem com deriva externa.

Com relação à técnica da cokrigagem ordinária, pretende-se fazer um estudo utilizando um arranjo de dados com a informação secundária amostrada em uma grande quantidade de nós da malha regular a ser estimada.

Tanto a cokrigagem colocalizada como a krigagem com deriva externa fazem uso de um arranjo especial de pontos de dados, em que a variável secundária é conhecida em todos os pontos de amostragem, bem como sobre os nós da malha regular.

As técnicas de estimativa geoestatísticas apresentam um sério problema que é a variabilidade reduzida das estimativas, devido ao efeito de suavização da krigagem. O problema da suavização não é inerente às técnicas geoestatísticas, mas a qualquer método de estimativa baseado na média ponderada. Todas as técnicas mencionadas

apresentam o problema citado e, portanto, deve-se analisar a conseqüência deste efeito na preservação da correlação inicial entre as variáveis.

1.2 Justificativas

A obtenção da informação geológica implica em custos elevados, pois envolve trabalhos de campo, laboratório e experimentos diversos. Na indústria do petróleo, a informação geológica tem um custo ainda maior, pois envolve levantamentos geofísicos terrestres ou marinhos e perfuração de poços ao custo de dezenas de milhões de dólares. Portanto, a informação geológica ainda que escassa deve ser trabalhada de modo a proporcionar as melhores estimativas em pontos não amostrados. A utilização das diversas técnicas geoestatísticas para diversos graus de correlação, entre as variáveis primária e secundária, irá mostrar como se pode obter os melhores resultados para um mesmo conjunto de dados geológicos.

CAPÍTULO 2

REVISÃO DA METODOLOGIA CORRENTE

Nesta pesquisa, serão consideradas as técnicas geoestatísticas: krigagem ordinária, cokrigagem ordinária, cokrigagem ordinária colocalizada e krigagem com deriva externa, as quais são baseadas em diferentes modelos de correlação espacial, apresentados nos itens seguintes. Após isso, serão tratados os métodos da krigagem ordinária (este mais sucintamente), a cokrigagem ordinária, a cokrigagem ordinária colocalizada e a krigagem com deriva externa.

2.1 Modelos de correlação espacial

Os métodos geoestatísticos de estimativa ou simulação estocástica caracterizam-se pela utilização de um modelo de correlação espacial subjacente, obtido experimentalmente a partir da informação disponível.

2.1.1 A função semivariograma

Uma variável aleatória pode apresentar correlação espacial quando valores vizinhos são próximos e apresentam tendência geral de aumento ou diminuição em uma determinada direção. Em outra situação, a variável aleatória não apresenta nenhuma tendência, ou seja, de um ponto a outro há aumento, mas em seguida diminuição e assim sucessivamente. Portanto, neste caso não há nenhum padrão espacial e, portanto, a variável não apresenta correlação espacial, caracterizando um fenômeno espacial completamente aleatório. Assim, caso haja correlação espacial é possível prosseguir a análise geoestatística.

A função semivariograma pode ser obtida pela seguinte aproximação (e.g. Journel & Huijbregts, 1978):

$$\gamma(h) = \frac{1}{2} E\left[\left(Z(x) - Z(x+h)\right)^2\right]$$

onde $\gamma(h)$ é a função semivariograma, $Z(x)$ é o valor variável no ponto x e $Z(x+h)$ a mesma variável, porém no ponto (x+h).

A função semivariograma mede variâncias para pontos separados por uma distância h. Essas variâncias plotadas em um gráfico de distância originam o semivariograma, ou simplesmente variograma. O variograma é então utilizado para inferir o valor da variância espacial para uma determinada distância h. Para isso é necessário ajustar um modelo de variograma que permita obter para qualquer valor de h, o valor de $\gamma(h)$.

2.1.2 A função semivariograma cruzado

A função semivariograma cruzado pode ser modelada através do variograma cruzado experimental que, segundo Matheron (1965), é uma generalização natural do variograma e pode ser escrito como:

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

Para funções aleatórias intrínsecas multivariadas (de ordem zero) deve satisfazer (Chilès & Delfiner 1999):

$$\begin{cases} E[Z_i(x+h) - Z_i(x)] = 0 & para\ i = 1, \cdots, p \\ Cov[Z_i(x+h) - Z_i(x), Z_j(x+h) - Z_j(x)] = 2\gamma_{ij}(h) & existe\ e\ depende\ apenas\ de\ h \end{cases}$$

Ainda segundo Chilès & Delfiner (1999), o variograma cruzado possui duas vantagens sobre o covariograma cruzado, a saber:

1. não assume variâncias finitas e;
2. a estimativa do variograma cruzado não é contaminada pela estimativa das médias. Quando existe, a relação entre o variograma cruzado e a covariância cruzada é:

$$\gamma_{ij}(h) = C_{ij}(0) - \frac{1}{2}\left[C_{ij}(h) + C_{ij}(-h)\right] + \frac{1}{2}\left[C_{ij}(h) - C_{ij}(-h)\right] \quad (2)$$

Wackernagel (1998) descreve o variograma cruzado como uma função obviamente ímpar que satisfaça a desigualdade:

$$\gamma_{ii}(h)\gamma_{jj}(h) \geq \left|\gamma_{ij}(h)\right|^2$$

pois o quadrado da covariância dos incrementos de duas variáveis é limitado ao produto das variâncias do incremento correspondente. Em outras palavras, a covariância obtida no variograma cruzado não pode ser maior que o produto das variâncias dos variogramas diretos.

Assim como o variograma experimental, o variograma cruzado é uma função direcional e diferencia-se daquele por considerar duas variáveis diferentes no cálculo da variância. Os outros procedimentos seguem exatamente aqueles para o cálculo de variogramas experimentais.

Algumas questões práticas devem ser mencionadas para o variograma cruzado (Wackernagel, 1998):

a. uma dada estrutura (por exemplo: efeito pepita) não pode estar presente no variograma cruzado (γ_{12}) se não estiver presente nos variogramas diretos (γ_1 e γ_2);
b. uma dada estrutura (por exemplo: efeito pepita) pode estar presente nos variogramas diretos (γ_1 e γ_2) e não estar presente no variograma cruzado (γ_{12})

2.1.2.1 Modelo linear de corregionalização

Em geoestatística, quando duas ou mais variáveis regionalizadas são definidas em um domínio aleatório, denomina-se corregionalização (Olea, 1999).

Segundo Wackernagel (1998), a regionalização multivariada de um conjunto de funções aleatórias pode ser representada por um modelo linear espacial multivariado. O modelo linear de corregionalização será a soma de modelos de covariâncias

proporcionais. A matriz de corregionalização B_i de ordem N x N pode ser ajustada em função do modelo de covariância multivariada como:

$$C(h) = \sum_{i=0}^{S} B_i \rho_i(h)$$

em que B_i é a matriz de corregionalização e ρ_i é a função de correlação.

O modelo de variograma associado com o modelo linear de funções aleatórias estacionárias é:

$$\Gamma(h) = \sum_{i=0}^{S} B_i g_i(h)$$

com $g_i(h)$ sendo os variogramas normatizados

Segundo Wackernagel (1998) no caso de duas variáveis, inicialmente ajustam-se os modelos de variograma diretos. A primeira e menor estrutura comum dos dois variogramas é inicialmente ajustada para que o variograma cruzado não seja equivalente ao modelo linear de corregionalização trivial. As estruturas presentes nos dois modelos de variograma podem ser utilizadas no ajuste do modelo linear, desde que respeitem a desigualdade descrita em Wackernagel (1998).

2.1.3 A função variograma residual

A krigagem com deriva externa que utiliza duas variáveis simultaneamente é baseada em um modelo de correlação espacial obtido a partir da função variograma residual. O variograma residual é simplesmente o variograma calculado a partir dos resíduos de uma função linear $V_p^* - V_p$, onde $V_p^* = a + bV_s$. Assim, como se conhece em todos os pontos amostrais tanto a variável primária como a secundária é possível proceder ao ajuste de uma função dos mínimos quadrados obtendo-se uma reta de regressão descrita como $V_p^* = a + bV_s$. Portanto, a condição necessária é que as variáveis estejam correlacionadas entre si, de modo a permitir o cálculo dos resíduos.

2.2 Amostragem de dados multivariados

Antes da definição e discussão dos métodos de geoestatística multivariada é importante definir as possíveis configurações espaciais que uma base de dados pode assumir, no que diz respeito às variáveis analisadas em cada ponto amostral.

Existem cinco configurações espaciais, possíveis, quando observa-se a presença de informações das variáveis primária e secundária: isotopia, heterotopia parcial e heterotopia total (Figura 1), bem como arranjos especiais: co-localização e multi co-localização.

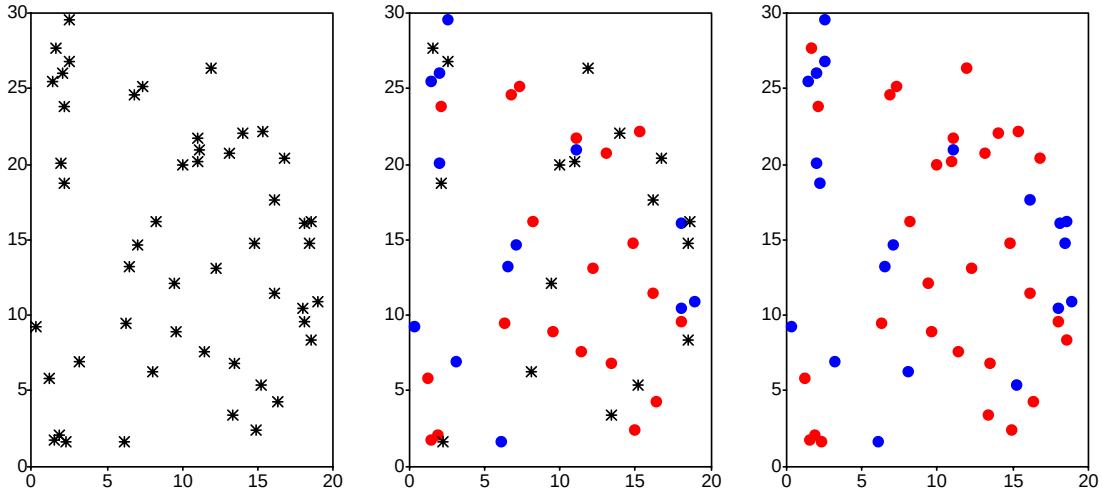

Figura 1: Configurações possíveis da distribuição espacial das variáveis amostradas: isotopia (esquerda), heterotopia parcial (centro) e heterotopia total (direita). Os círculos azuis representam a pontos de amostragem da variável primária, os vermelhos da variável secundária e os asteriscos representam os pontos em que as duas variáveis são conhecidas.

Pode-se sucintamente descrever cada uma destas configurações conforme segue:

Isotopia – nesta situação, tanto a variável primária quanto as auxiliares são conhecidas em todos os pontos amostrados. Esta é a melhor condição para o cálculo do variograma cruzado; entretanto, para conjuntos de variáveis intrinsecamente correlacionadas, o resultado da cokrigagem é equivalente ao da krigagem (Wackernagel, 2003), uma vez que os pesos das variáveis secundárias assumem pequena influência na estimativa;

Heterotopia total – esta configuração espacial de dados implica que o ponto amostral que teve a variável primária medida obrigatoriamente não terá as secundárias e vice-versa. Deste modo, o variograma cruzado não pode ser calculado diretamente o

que inviabiliza a estimativa trivial por cokrigagem. Wackernagel (2003) destaca que embora o variograma cruzado seja impossível de se calcular, as covariâncias cruzadas não a são, porém esta solução continua problemática uma vez que os valores de covariância direta correspondentes remeterão a diferentes conjuntos de pontos. Myers (1991) sugere a utilização de pseudo-variogramas cruzados para poder estimar os dados por cokrigagem, nesta configuração de dados. Pseudo-variogramas cruzados são definidos por Clark *et al*. (1989) como:

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

Maiores detalhes sobre este tipo de variograma podem der obtidos nos artigos supra mencionados.

Heterotopia parcial – neste caso existem pontos amostrais onde tanto a variável primária quanto a secundária foram medidas e, também, existem pontos onde apenas a variável primária está disponível e pontos onde apenas a secundária está presente. Embora para esta heterotopia, possa existir dificuldade para o cálculo do variograma cruzado (a depender da contribuição da variável primária no conjunto de dados), é nesta configuração espacial, segundo Wackernagel (1998) e Olea (1999), que o método de cokrigagem pode apresentar seu melhor resultado, uma vez que a contribuição da variável secundária na estimativa não será sobreposta pela informação primária.

Colocalização – nesta situação, os pontos em que a característica de interesse é conhecida são restritos e estão esparsamente distribuídos pelo domínio. Por outro lado, a variável secundária (quasi-exaustiva) é conhecida em, pelo menos, cada ponto da malha que será estimada. Neste tipo de conjunto de dados a escolha da vizinhança utilizada na estimativa é uma etapa crítica, pois a alta densidade de informação secundária pode resultar na instabilidade no sistema de equações lineares da cokrigagem ordinária (Wackernagel, 2003; Goovaerts, 1997).

Multi-colocalização – esta configuração espacial dos dados é um caso particular de colocalização, em que se conhece o valor da variável secundária na malha a ser estimada e nos pontos em que a variável primária é medida (Olea, 1999).

2.3 Krigagem Ordinária

Os algoritmos de krigagem são um grupo consagrado de métodos de predição. Neste grupo, destaca-se a krigagem ordinária, cuja principal característica é a estimativa do valor de uma variável aleatória regionalizada $Z(x)$, em um ponto não amostrado, com mínima variância do erro. Esta estimativa é suportada pelos pontos vizinhos em que a variável é medida. O estimador da krigagem ordinária:

$$Z^*_{KO}(x_0) = \sum_{i=1}^{n} \lambda_i \, z(x_i),$$

em que $Z^*_{KO}(x_0)$ é o valor da estimativa por krigagem ordinária no ponto x_0, λ_i é o *i-ésimo* ponderador e $z(x_i)$ é o valor da variável regionalizada $Z(x)$ no *i-ésimo* local da vizinhança composta por n amostras.

A condição de não viés da krigagem ordinária é alcançada quando a igualdade a seguir é satisfeita (Journel & Huijbregts, 1978):

$$E\left[Z^*_{KO}(x)\right] - E\left[Z(x)\right] = 0$$

Considerando que:

$$E\left[Z(x)\right] = m,$$

$$E\left[Z^*_{KO}(x)\right] = E\left[\sum_{i=1}^{n} \lambda_i Z(x_i)\right] = \sum_{i=1}^{n} \lambda_i \, E\left[Z(x_i)\right] = \sum_{i=1}^{n} \lambda_i \, m$$

Desta maneira, para que:

$$E\left[Z^*_{KO}(x)\right] = \sum_{i=1}^{n} \lambda_i \, m,$$

Deve-se então assegurar que a soma dos ponderadores λ_i seja igual a um:

$$\sum_{i=1}^{n} \lambda_i = 1.$$

Esta restrição é satisfeita utilizando um lagrangiano no sistema de equações lineares da krigagem ordinária:

$$\begin{cases} \sum_{i=1}^{n} \lambda_i \gamma(x_i - x_j) - \mu = \gamma\left(x_j - x_0\right) \quad j = 1, \ldots, n \\ \\ \sum_{i=1}^{n} \lambda_i = 1 \end{cases}$$

em que os ponderadores são $\{\lambda_i,\ i = 1, n\}$, a variância espacial $\gamma(.)$ é fornecida pelo modelo teórico de variograma, x_i e x_j representam a vizinhança e μ é o multiplicador de lagrange.

Em termos matriciais, o sistema de equações lineares da krigagem ordinária pode ser representado, em função da variância espacial, da seguinte forma:

$$\begin{bmatrix} \gamma(x_1, x_1) & \gamma(x_1, x_2) & \ldots & \gamma(x_1, x_n) & 1 \\ \gamma(x_2, x_1) & \gamma(x_2, x_2) & \ldots & \gamma(x_2, x_n) & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \gamma(x_n, x_1) & \gamma(x_n, x_2) & \ldots & \gamma(x_n, x_n) & 1 \\ 1 & 1 & & 1 & 0 \end{bmatrix} . \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma(x_0 - x_1) \\ \gamma(x_0 - x_2) \\ \vdots \\ \gamma(x_0 - x_n) \\ 1 \end{bmatrix}$$

2.3.1 Variância de Krigagem

Um estimador que minimiza o erro, como a krigagem ordinária, pode fornecer uma medida de variância do erro no local estimado. Neste sentido, a variância de krigagem é uma medida consagrada de incerteza associada à estimativa por krigagem sendo expressa em função do semivariograma como:

$$\sigma_{KO}^2 = \sum_{i}^{n} \lambda_i \gamma\left(x_0 - x_i\right) + \mu .$$

em que σ_{KO}^2 é a variância de krigagem, $\gamma\left(x_0 - x_i\right)$ a variância espacial entre a vizinhança e o local estimado, $\{\lambda_i,\ i = 1, n\}$ e μ são os ponderadores e o Multiplicador de Lagrange, respectivamente.

Entretanto, como se pode observar na expressão da variância de krigagem ela não depende dos valores da variável regionalizada $Z(x)$ e é portanto, homoscedástica. Ou seja, considera apenas a configuração espacial das informações vizinhas com relação ao local estimado, negligenciando os valores da função aleatória tanto na vizinhança $z(x_i)$, quanto no local estimado $Z^*_{KO}(x_0)$.

2.3.2 Variância de Interpolação

Proposta por Yamamoto (2000), a variância de interpolação é uma medida heterocedástica de incerteza associada à estimativa por krigagem ordinária, uma vez que considera tanto a configuração espacial das amostras utilizadas na estimativa de um local x_0, quanto o valor estimado $Z^*_{KO}(x_0)$ e os valores medidos na vizinhança $z(x_i)$, fornecendo assim uma medida de incerteza mais sensível e realista que a variância de krigagem.

A variância de interpolação é apresentada como (Yamamoto, 2000):

$$s_o^2 = \sum_{i=1}^{n} \lambda_i \left[z(x_i) - z^*_{KO}(x_0) \right]^2$$

onde n é o número de amostras utilizadas na estimativa do local x_0, $\{\lambda_i, i = 1, n\}$ são os pesos da krigagem ordinária, $z(x_i)$ são as amostras vizinhas utilizadas na estimativa e $z^*_{KO}(x_0)$ é o resultado da estimativa.

Como pode ser observado na equação acima, a configuração espacial das amostras utilizadas na estimativa é considerada por intermédio dos pesos $\{\lambda_i, i = 1, n\}$, resultantes da resolução do sistema de equações da krigagem. Nota-se também, que para garantir que a variância de inteporlação seja positiva é imperativo que os pesos sejam positivos. Para satisfazer esta condição é necessário que se utilize um algoritmo de correção dos pesos da krigagem(*e.g.* Journel & Rao, 1996).

Segundo Yamamoto (2000), as principais propriedades da variância de interpolação são:

i. corresponde à exatidão da krigagem ordinária, pois se o local a ser estimado $Z^*_{KO}(x_0)$ coincide com o ponto em que se conhece o valor da variável

aleatória $z(x_i)$ este valor será honrado, uma vez que seu ponderador λ_i será um enquanto os outros iguais a zero. Logo, a variância de interpolação será nula ($S_0^2 = 0$);

ii. É proporcional à dispersão dos pontos amostrais;

iii. Utiliza indiretamente a informação estrutural fornecida pelo variograma, através do ponderador λ_i, que atribuirá maior influência a um ponto baseado no peso utilizado na krigagem.

2.3.3 Correção do Efeito de Suavização

O estimador da krigagem ordinária, bem como outros estimadores baseados em ponderação, geralmente apresentam um sério defeito: o efeito de suavização da estimativa. Este problema consiste basicamente da tendência da estimativa ao valor da média, ou seja, a subestimativa de valores mais altos e superestimativa de valores mais baixos. Este comportamento resulta na perda de precisão global da estimativa, refletido pelo fato do histograma e variograma da estimativa não reproduzirem os correspondentes amostrais. Este comportamento do estimador da krigagem ordinária é dramático em situações em que o interesse esteja nos valores extremos da variável aleatória. Por exemplo, na predição da concentração em depósitos de metais raros, são os altos teores que viabilizam o empreendimento. Outro caso em que os valores extremos devem ser considerados é na modelagem de porosidade de um reservatório.

Buscando solucionar esta limitação dos algoritmos de estimativa, diversos métodos de simulação estocástica foram propostos. Entretanto, o produto final desta família de métodos são *n* imagens ou mapas equiprováveis, ou seja, *n* realidades que não permite a escolha da melhor, pois são equiprováveis e a reprodução do modelo de covariância depende daquele conjunto.

Segundo Yamamoto (2005), o déficit de variância resultante do efeito de suavização da krigagem ordinária pode ser expresso em função da variância de interpolação como:

$$Var[Z(x)] - Var[Z^*_{KO}(x)] = E[S_0^2] \geq 0$$

que pode ser interpretada como a variância suavizada do estimador da krigagem $Z^*_{KO}(x_0)$.

Esta expressão faz sentido como medida de variância de suavização, uma vez que o maior valor de S_0^2 é tomado considerando todos os valores possíveis de uma determinada configuração de amostras. Nota-se também que a variância amostral é igual à variância da estimativa por krigagem ordinária somada à variância de suavização. Desta maneira, se todos os locais contendo amostras forem estimados, então a variância da estimativa por krigagem ordinária será igual à variância amostral, uma vez que o estimador da krigagem ordinária honra os dados, conseqüentemente a variância de suavização será nula.

Utilizando a expressão da variância total em função da variância da estimativa acrescida da variância de interpolação, Yamamoto (2005) descreve o algoritmo de correção de efeito de suavização da krigagem ordinária em quatro etapas, a saber:

i. Executar a validação cruzada para calcular o erro verdadeiro e o desvio padrão de interpolação para cada ponto:

$$erro\ verdadeiro = \left[Z^*_{KO}(x_0) - Z(x_0)\right].$$

Estas duas variáveis serão utilizadas para o cálculo de uma nova variável, o número de desvios padrão de interpolação:

$$Ns_0 = \frac{-\,erro\ verdadeiro}{S_0^2} \quad para\ S_0^2 \neq 0$$

Nota-se que Ns_0 pode ser positivo ou negativo. O sinal negativo compensa o efeito de suavização. Se o erro verdadeiro é maior que zero, então o valor estimado naquele ponto é superior ao valor amostral, assim um erro deve ser decrescido do valor estimado. Por outro lado, se o erro verdadeiro for menor que zero, um erro deve ser somado ao valor estimado, uma vez que este é menor que o valor amostral.

O resultado desta etapa será um Ns_0 para cada local.

ii. Executar a krigagem ordinária para obter o número de desvios padrão de interpolação (Ns_0) para todos os nós (x_o) que serão corrigidos. É importante

mencionar que Ns_0 deve ser estimado com o menor número possível de vizinhos, para evitar a suavização.

iii. Executar novamente a krigagem ordinária para estimar $Z^*_{KO}(x_0)$ em todos os nós x_o que serão corrigidos. Os parâmetros de vizinhança devem ser os mesmos utilizados na etapa *i*.

iv. Executar o procedimento de pós-processamento descrito a seguir:

$$Z^{**}_{KO}(x_0) = Z^*_{KO}(x_0) + Ns_o(x_0) * S_0$$

A expressão anterior reduz consideravelmente o efeito de suavização da krigagem ordinária, entretanto, é importante verificar se o valor do ponto estimado está coerente com o intervalo dos dados vizinhos. Para esta verificação, deve-se considerar:

$$\begin{cases} Z_{\min} = \min[Z(x_i), i = 1, \ldots, n] \\ Z_{\max} = \max[Z(x_i), i = 1, \ldots, n] \end{cases}$$

em que Z_{min} e Z_{max} são os valores de máximo e mínimo da vizinhança utilizada na estimativa.

Desta maneira, deve-se realizar a verificação e correção dos valores estimados segundo Yamamoto (2005):

$$\begin{cases} se \quad Z^{**}_{KO}(x_0) < Z_{\min} \qquad Z^{**}_{KO}(x_0) = Z^*_{KO} + Ns_0(x_0).So.factor \\ se \quad Z^{**}_{KO}(x_0) > Z_{\max} \qquad Z^{**}_{KO}(x_0) = Z^*_{KO} + Z^*_{NS}(x_0).factor \end{cases}$$

em que $Z^*_{NS}(x_0) = Ns_0(x_0).So$.

Yamamoto (2007) desenvolve a equação da variância dos dados em função da variância da estimativa e de $Z^*_{NS}(x_0)$ para calcular o *factor* de correção para a estimativa da seguinte forma:

$$Var[Z(x)] = Var[Z^*_{KO}(x_o) + factor\ Z^*_{Ns_0}(x_0)]$$

que após rearranjo será uma função quadrática, onde o fator (*factor*) será a raiz:

$$Var\left[Z^{*}_{Ns_0}(x_0)\right].factor^2 + 2Cov\left(Z^{*}_{KO}(x_0), Z^{*}_{Ns_0}(x_0)\right).factor + Var\left[Z^{*}_{KO}(x_0)\right] - Var\left[Z(x_0)\right] = 0$$

e o discriminante será:

$$\Delta = 4Cov\left(Z^{*}_{KO}(x_0), Z^{*}_{Ns}(x_0)\right)^2 - 4Var\left[Z^{*}_{Ns}(x_0)\right]\left(Var\left[Z^{*}_{KO}(x_0) - Z(x_0)\right]\right)$$

em que *Cov(.)* é a covariância e *Var[.]* é a variância espacial.

Devido à variância dos dados sempre ser superior à variância da estimativa (efeito de suavização) o discriminante será sempre maior que zero e a função quadrática terá duas raízes reais. Sabendo que a raiz quadrada do discriminante sempre será maior que zero, o fator ótimo (*factor*) que melhor corrige a variância da estimativa à variância das amostras será a raiz positiva da função quadrática, que pode ser calculada por (Yamamoto, 2007):

$$factor = \frac{-2Cov\left(Z^{*}_{KO}(x_0), Z^{*}_{Ns}(x_0)\right) + \sqrt{\Delta}}{2Var\left[Z^{*}_{Ns}(x_0)\right]}.$$

Após a compensação no déficit de variância, é necessário que a média da estimativa também seja coerente com aquela fornecida pelos dados. Yamamoto (2005) utiliza uma constante K, que somada à estimativa aplica um último ajuste fino:

$$K = E\left[Z(x)\right] - E\left[Z^{**}_{KO}(x)\right]$$

$$Z^{**}_{KO}(x_0) = Z^{**}_{KO}(x_0) + K$$

que não altera a variância, pois:

$$Var\left[Z(x) + K\right] = Var\left[Z(x)\right]$$

2.4 Cokrigagem

A família de técnicas, genericamente, denominadas cokrigagem apresenta-se como uma possibilidade bastante atrativa para a estimativa de uma variável de interesse (variável primária) tendo como complementação informações fornecidas por uma ou mais variáveis (variáveis secundárias).

A cokrigagem é uma extensão natural da krigagem quando os dados, os modelos de covariância e variograma disponíveis são multivariados (Wackernagel, 1998). Genericamente, nesta família de métodos, a variável primária é estimada em um local específico, ponto ou bloco, utilizando as informações fornecidas pelos dados dela mesma e dados de variáveis auxiliares. Estas informações traduzem-se de dois modos: a quantificação das variáveis (i.e. teor, porosidade, permeabilidade) e o modelo de covariância ou variograma acessados para o fenômeno em estudo. No caso do modelo, este será o variograma/covariograma direto, quando comparam-se os dados da variável com ela mesma, mas em diferentes posições no espaço; e o variograma/covariograma cruzado, quando a comparação feita considera duas variáveis distintas também em diferentes posições no espaço, conforme apresentado no item 2.1.

Segundo Olea (1999), este é um método de estimativa verdadeiramente multivariado, uma vez que o modelo lida com dois ou mais atributos em um mesmo domínio. Na estimativa por cokrigagem, quando um determinado ponto da base de dados não possui informação das variáveis primária e secundária, esta ausência não enviesa ou degrada o resultado da estimativa, ao contrário, esta é a situação em que a cokrigagem apresenta-se mais eficaz.

Journel & Huijbregts (1978), apontam a coestimativa de uma variável subamostrada como principal utilização da cokrigagem em mineração. Entretanto, Olea (1991) e Myers (1982) mencionam a redução da variância do erro de estimativa de um atributo subamostrado e a estimativa conjunta de diversas características amostradas em um mesmo domínio como maiores contribuições da coestimativa por cokrigagem.

Conde (2000) utiliza cokrigagem ordinária para estimar variáveis com diferentes coeficientes de correlação (0.95, 0.76 e 0.51) para os casos de isotopia e heterotopia parcial para a avaliação de uma jazida de asbestos. Com a redução sistemática da

correlação da variável primária no conjunto de dados isotópico a média e as medidas de dispersão (desvio padrão, coeficiente de variação e raiz quadrada da variância de cokrigagem) apresentaram pequena variação. Para o caso de heterotopia parcial, a autora conclui que a média apresenta pequenas variações, entretanto, na medida em que a contribuição da variável primária é reduzida, a coestimativa degrada-se sistematicamente com coeficientes de correlação mais baixos, refletindo-se no aumento de valores das medidas de dispersão. A autora sugere coeficientes de correlação intravariáveis iguais ou superiores a 0,5.

Pardo-Igúzquiza *et al.* (2006) utilizaram cokrigagem ordinária, com sucesso, na fusão de imagens orbitais de alta resolução espacial (Pan-cromática) e imagens de alta resolução espectral (estas com baixa resolução espacial) buscando maior detalhamento espacial (downscaling) das imagens multi-espectrais.

Doyen (1988) compara a cokrigagem, como alternativa na integração informação auxiliar de origem sísmica (velocidade de trânsito), aos métodos consagrados de estimativa na estimativa de porosidade de um reservatório, alcançando resultados superiores da cokrigagem, como pode ser observado na Figura 2.

Conforme observado por Isaaks & Srivastava (1989), Conde (2000) e Wackernagel (1995, 2003), entre outros autores, quando não existe correlação espacial entre a variável primária e as variáveis auxiliares o resultado da cokrigagem será idêntico ao da krigagem (variáveis autokrigáveis). Isaaks & Srivastava (1989) mencionam ainda, que a cokrigagem não apresentará ganho com relação à estimativa por krigagem no caso em que as informações das variáveis primária e secundária estão presentes em todos os locais amostrados ou que os modelos de covariância direta e cruzada são similares. Neste sentido, a cokrigagem torna-se menos atrativa, uma vez que o processo de modelagem geoestatística se torna mais complexo, já que $n + \left[n.\left(n-1\right)/2 \right]$ modelos de variogramas diretos e cruzados deverão ser calculados e ajustados segundo um modelo linear de corregionalização válido.

Considerando a situação em que os dados primários $\left\{ z_1\left(x_{\alpha_1}\right), \alpha_1,...,n_1 \right\}$ são complementados pela informação secundária relacionada a $\left(N_v - 1\right)$ atributos contínuos z_i, $\left\{ z_i\left(x_{\alpha_i}\right), x_{\alpha_i} 1,\ldots,n_i, i = 2,\ldots,N_v \right\}$, em qualquer, possivelmente diferente, local. O estimador linear estendido para incorporar aquelas informações adicionais pode ser expresso como (Goovaerts, 1997):

$$Z_1^*(u) - m_1 = \sum_{\alpha_1=1}^{n_i(u)} \lambda_{\alpha_1}(u)\left[Z_1\left(u_{\alpha_1}\right) - m_1\left(u_{\alpha_1}\right)\right] \quad + \quad \sum_{i=2}^{N_v} \sum_{\alpha_1=1}^{n_i(u)} \lambda_{\alpha_i}(u)\left[Z_i\left(u_{\alpha_i}\right) - m_i\left(\left(u_{\alpha_i}\right)\right)\right]$$

em que $\lambda_{\alpha_i}(u)$ são os ponderadores associados aos dados da variável primária $z_1\left(u_{\alpha_1}\right)$ e $\lambda_{\alpha_i}(u)$, $i > 1$, são os ponderadores associados às variáveis secundárias $z_i\left(u_{\alpha_i}\right)$. Os valores da esperança matemática de $Z_1(u)$ e $Z_i\left(u_{\alpha_i}\right)$ são $m_1(u)$ e $m_i\left(u_{\alpha_i}\right)$, respectivamente.

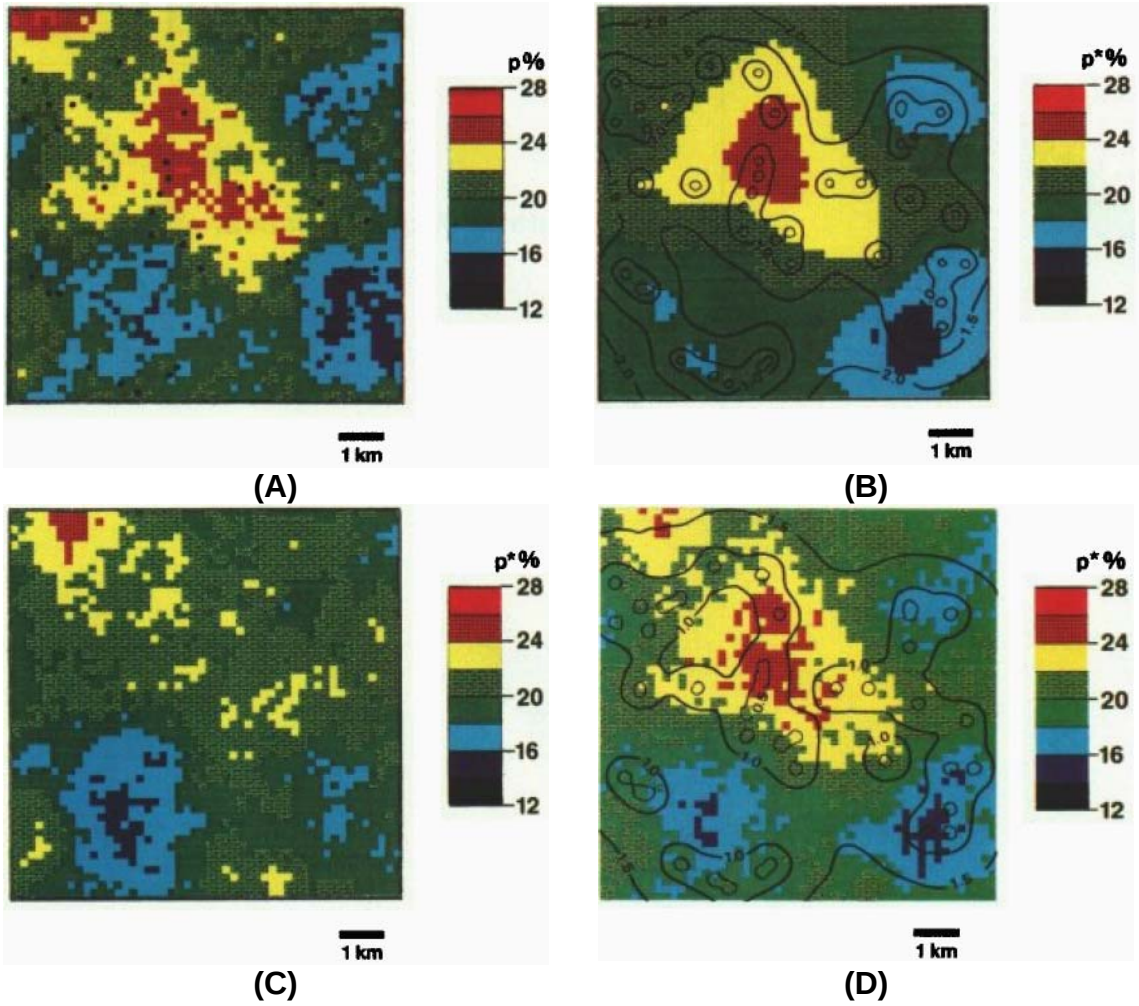

Figura 2: Comparação do desempenho da krigagem ordinária (B), regressão linear (C) e cokrigagem (D), segundo Doyen (1988)

Todos os estimadores de cokrigagem são variações da expressão anterior e devem considerar as seguintes condições de não enviesamento e minimização da variância do erro ($\sigma_E^2(u)$):

$$\sigma_E^2 = Var\{Z_1^*(u) - Z_1(u)\},$$

que é minimizado sob a restrição do erro esperado ser zero:

$$E\{Z_1^*(u) - Z_z(u)\} \quad = \quad 0$$

2.4.1 Cokrigagem Simples

A cokrigagem ordinária não faz sentido quando não há informações da variável de interesse na vizinhança (Wackernagel 2003). Por outro lado, a cokrigagem simples (SCK) é amparada pelo conhecimento *à priori* das médias globais das variáveis, de modo que a estimativa de uma variável pode ser calibrada sem que existam valores de dados desta variável na vizinhança da cokrigagem.

O estimador da cokrigagem simples é composto pela média da variável de interesse somada à uma combinação linear dos pesos λ_α^i com o resíduo das variáveis, que para o caso de duas variáveis será:

$$Z_{SCK}^*(u) \; - m_1 = \sum_{\alpha_1=1}^{n_i} \lambda_{\alpha_1}^{SCK}(u)\left[Z_1(u_{\alpha_1}) - m_1\right] + \sum_{\alpha_2=1}^{n_2} \lambda_{\alpha_2}^{SCK}(u)\left[Z_2(u_{\alpha_2}) - m_2\right],$$

em que $Z_{SCK}^*(u)$ é o resultado da cokrigagem simples no local u, $\lambda_{\alpha_1}^{SCK}$ são os ponderadores da SCK para a variável primária, $\lambda_{\alpha_2}^{SCK}$ são os ponderadores para a variável secundária, m_1 e m_2 são as médias populacionais das características primária e secundária, respectivamente. Os ponderadores ($\lambda's$) da SCK são determinados para assegurar o não enviesamento e mínima variância do erro da coestimativa. O não viés é garantido pela expressão (Wackernagel, 2003):

$$E\{Z_{SCK}^*(u) - Z(u)\} = \sum_{\alpha_1=1}^{n_i(u)} \lambda_{\alpha_1}^{SCK}(u)\left[E\{Z_1(u_{\alpha_1})\} - m_1\right] + \sum_{\alpha_2=1}^{n_2(u)} \lambda_{\alpha_2}^{SCK}(u)\left[E\{Z_2(u_{\alpha_2})\} - m_2\right]$$

$$+ \left[m_1 - E\{Z_1(u)\}\right] \quad = \quad 0.$$

Ainda segundo Wackernagel (2003) considerando $R_i(u_{\alpha_i}) \quad = \quad Z_i(u_{\alpha_i}) - m_1$ como variável primária ($i = 1$) ou secundária ($i = 2$) de uma variável regionalizada em u_{α_i}, o

erro da coestimativa $Z^*_{SCK}(u) - Z(u)$ no local u pode ser expresso como uma combinação linear de variáveis residuais:

$$Z^*_{SCK}(u) - Z_1(u) = \sum_{\alpha_1}^{n_i(u)} \lambda_{\alpha_1}^{SCK}(u) R_1(u_{\alpha_1}) + \sum_{\alpha_2}^{n_2(u)} \lambda_{\alpha_2}^{SCK}(u) R_2(u_{\alpha_2}) - [Z_1(u) - m_1]$$

$$= \sum_{i=1}^{2} \sum_{\alpha_{i=1}}^{n_i(u)} \lambda_{\alpha_1}^{SCK}(u)\, R_i(u_{\alpha_i}) - R_1(u)$$

$$= R^*_{SCK}(u) - R_1(u).$$

O erro da variância σ_E^2 pode então ser expresso como uma combinação linear dos valores de covariância direta e cruzada dos resíduos (Wackernagel, 2003):

$$\sigma_E^2(u) = Var\{R^*_{SCK}(u)\} + Var\{R_1(u)\} - 2\,Cov\{R^*_{SCK}(u), R_1(u)\}$$

$$= \sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{\alpha_{i=1}}^{n_i(u)} \sum_{\alpha_j}^{n_j(u)} \lambda_{\alpha_i}^{SCK}(u) \lambda_{\alpha_j}^{SCK}(u)\, C_{ij}^R(u_{\alpha_i}, u_{\alpha j}) + C_{11}^R(0) - 2 \sum_{i=1}^{2} \sum_{\alpha_i=1}^{n_i(u)} \lambda_{\alpha_i}^{SCK}(u) C_{i1}^R(u_{\alpha_i}, u)$$

$$= Q\left(\lambda_{\alpha_i}^{SCK}(u),\ \alpha_i = 1, \ldots, n_i(u);\ i = 1, 2\right)$$

Os ponderadores que minimizam a variância do erro são obtidos pelo conjunto de zeros nas primeiras derivadas parciais da forma quadrática $(n_1(u) + n_2(u))$ (Wackernagel, 2003):

$$\frac{1}{2} \frac{\partial Q(u)}{\partial \lambda_{\alpha_i}^{SCK}(u)} = \sum_{\beta_1=1}^{n_1(u)} \lambda_{\beta_i}^{SCK}(u) C_{11}^R(u_{\alpha_i}, u_{\beta_i}) + \sum_{\beta_2=1}^{n_{2(u)}} \lambda_{\beta_2}^{SCK}(u) C_{12}^R(u_{\alpha_i}, u_{\beta_2}) - C_{11}^R(u_{\alpha_i}, u)$$

$$\alpha_1 = 1, \ldots, n_1(u)$$

$$\frac{1}{2} \frac{\partial Q(u)}{\partial \lambda_{\alpha_2}^{SCK}(u)} = \sum_{\beta_1=1}^{n_1(u)} \lambda_{\beta_i}^{SCK}(u) C_{21}^R(u_{\alpha_2}, u_{\beta_i}) + \sum_{\beta_2=1}^{n_{2(u)}} \lambda_{\beta_2}^{SCK}(u) C_{22}^R(u_{\alpha_2}, u_{\beta_2}) - C_{21}^R(u_{\alpha_2}, u)$$

$$\alpha_2 = 1, \ldots, n_2(u)$$

Para Wackernagel (2003), assumindo que tanto as funções covariância $C_{11}^R(h)$ e $C_{22}^R(h)$ das características primária e secundária, quanto a função covariância cruzada $C_{12}^R(h)$

são estacionárias e iguais às funções autocovariância e covariância cruzada das variáveis aleatórias $Z_1(u)$ e $Z_2(u)$:

$$C_{ij}^{R}(h) = E\{R_i(u).\, R_j(u+h)\}$$

$$= E\{[Z_i(u) - m_i].\,[Z_j(u+h) - m_j]\} \quad = \quad C_{ij}(h) \qquad i, j = 1, 2$$

Ainda segundo o mesmo autor, pode-se expressar o sistema de equações da cokrigagem simples como:

$$\begin{cases} \sum_{\beta_1=1}^{n_1(u)} \lambda_{\beta_1}^{SCK}(u) C_{11}(u_{\alpha_1} - u_{\beta_1}) + \sum_{\beta_2=1}^{n_2(u)} \lambda_{\beta_2}^{SCK}(u) C_{12}(u_{\alpha_1} - u_{\beta 2}) = C_{11}(u_{\alpha_1} - u) & \alpha_1 = 1, \ldots, n_1(u) \\ \sum_{\beta_1=1}^{n_1(u)} \lambda_{\beta_1}^{SCK}(u) C_{21}(u_{\alpha_2} - u_{\beta_1}) + \sum_{\beta_2=1}^{n_2(u)} \lambda_{\beta_2}^{SCK}(u) C_{22}(u_{\alpha_2} - u_{\beta_2}) = C_{21}(u_{\alpha_2} - u) & \alpha_2 = 1, \ldots, n_2(u) \end{cases}$$

A variância de cokrigagem simples, ou a mínima variância do erro pode ser calculada pela substituição do sistema de equações de cokrigagem simples nas expressões de variância do erro (Wackernagel, 2003):

$$\sigma_{sck}^2 = C_{11}(0) - \sum_{\alpha_1=1}^{n_1(u)} \lambda_{\alpha_1}^{SCK}(u) C_{11}(u_{\alpha_1} - u) - \sum_{\alpha_2=1}^{n_2(u)} \lambda_{\alpha_2}^{SCK}(u) C_{21}(u_{\alpha_2} - u)$$

A variância do erro geralmente decresce quando o local u estimado coincide com a localização dos dados (maiores ponderadores da cokrigagem) ou quando a covariância da variável primária $C_{11}(h)$ ou a covariância cruzada $C_{12}(h)$ aumenta.

O sistema de equações de cokrigagem simples pode ser expresso em notação matricial, conforme (Wackernagel, 2003):

$$\begin{bmatrix} C_{11} & \cdots & C_{1j} & \cdots & C_{1n} \\ \vdots & \ddots & & & \vdots \\ C_{21} & & C_{2j} & & C_{2n} \\ \vdots & & & \ddots & \vdots \\ C_{31} & \cdots & C_{3j} & \cdots & C_{3n} \end{bmatrix} . \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_2 \\ \vdots \\ \lambda_3 \end{bmatrix} = \begin{bmatrix} c_{1i_0} \\ \vdots \\ c_{2i_0} \\ \vdots \\ c_{3i_0} \end{bmatrix},$$

em que a matriz do lado esquerdo é composta por blocos simétricos quadrados $n_i \ x \ n_i$ de C_{ii} na diagonal e com blocos retangulares $n_i \ x \ n_j$ de C_{ij}, em que

$$C_{ij} \ = \ C_{ij}^{\mathrm{T}} .$$

Os blocos C_{ij} contém as covariâncias entre as amostras de um par fixo de variáveis. Os vetores C_{ii_0} listam as covariâncias com a variável de interesse, para uma variável específica do conjunto, entre amostras e locais da estimativa. Os vetores λ_i representam os ponderadores associados às amostras de uma variável.

Wackernagel (2003), observa que na cokrigagem simples:

i. O sistema de cokrigagem simples possui uma única solução com variância de cokrigagem positiva se e somente se a matriz de covariância é definida positiva. Esta condição é satisfeita com um modelo linear de corregionalização autorizado;

ii. se as variáveis primária e secundária não são correlacionadas, o estimador da cokrigagem se reverterá no estimador da krigagem simples;

iii. o estimador da cokrigagem simples é um interpolador exato e honra os dados primários em seus locais de amostragem (é um estimador condicional *a priori*).

2.4.2 Cokrigagem Ordinária

A cokrigagem ordinária é um método de estimativa baseado na combinação linear de ponderadores λ_i^p com dados de diferentes variáveis, possivelmente com amostras de tamanhos diferentes localizadas na vizinhança do ponto x_0 que será estimado. De maneira similar a krigagem ordinária, não é necessário conhecer a média populacional da variável aleatória na estimativa por cokrigagem ordinária, uma vez que a média é local e calculada na vizinhança do ponto que será coestimado. O estimador da cokrigagem ordinária é definido como (Wackernagel, 1998):

$$Z_{po}(x_0) = \sum_{p=1}^{N} \sum_{i=1}^{n_p} \lambda_i^p Z_p(x_1)$$

em que po refere-se a uma variável específica de um conjunto de N variáveis. O índice n_p é o número de amostras e dependerá do índice p das variáveis.

Segundo Wackernagel (1998), no arcabouço da hipótese intrínseca conjunta o objetivo da cokrigagem ordinária é estimar uma variável particular de um conjunto de *N* variáveis e com base em um erro de estimativa que seria nulo em média. Ainda segundo este autor, esta condição é satisfeita pela seleção de pesos cuja soma seja um para a variável de interesse e que tenha soma zero para as variáveis auxiliares, conforme:

$$\sum_{i=1}^{n_p} \lambda_i^p = \delta_{pp_0} = \begin{cases} 1 \; se \; p = p_0 \\ 0 \; se \; p \neq p_0 \end{cases}.$$

Seguindo a expansão da expressão do erro médio proposta por Wackernagel (1998), tem-se:

$$E\left[Z_{p_0}^*(x_0) - Z_{p_0}(x_0)\right] = E\left[\sum_{p=1}^{N}\sum_{i=1}^{n_p} \lambda_i^p Z_p(x_i) - \underbrace{\sum_{i=1}^{n_p} \lambda_i^{p_0}}_{1} Z_{p_0}(x_0) - \sum_{\substack{p=0 \\ p\neq p_0}}^{N} \underbrace{\sum_{i=1}^{n_p} \lambda_i^p}_{0} Z_p(x_0)\right] =$$

$$\sum_{p=1}^{N}\sum_{i=1}^{n_p} \lambda_i^p \underbrace{E\left[Z_p(x_i) - Z_p(x_0)\right]}_{0} = 0$$

A variância do erro médio de estimativa será:

$$\sigma_E^2 = E\left[\left(\sum_{p=1}^{N}\sum_{i=0}^{n_p} \lambda_{i\alpha}^p Z_p(x_\alpha) - Z_p(x_0)\right)^2\right]$$

Introduzindo os pesos λ_0^p definidos como:

$$\lambda_0^p = -\delta_{ii_0} = \begin{cases} -1 & se \; i = i_{0,} \\ 0 & se \; i \neq i_{0,} \end{cases}$$

que são incluídos nas somas e permitem a redução da expressão da variância da estimativa:

$$\sigma_E^2 = E\left[\left(\sum_{p=1}^{N}\sum_{i=0}^{n_p} \lambda_i^p Z_p(x_i)\right)^2\right]$$

Então, inserindo-se variáveis aleatórias fictícias $Z_i(0)$ posicionadas arbitrariamente na origem, pode-se formar incrementos:

$$\sigma_E^2 = E\left[\left(\sum_{i=1}^{N}\left(\sum_{\alpha=0}^{n_i}\lambda_\alpha^i Z_i(x_\alpha) - \underbrace{Z_i(0)\sum_{\alpha=0}^{n_i}\omega_\alpha^i}_{0}\right)\right)^2\right]$$

$$= E\left[\left(\sum_{i=1}^{N}\sum_{\alpha=0}^{n_i}\lambda_\alpha^i \underbrace{\left(Z_i(x_\alpha) - Z_i(0)\right)}_{incremento}\right)^2\right].$$

Definindo a covariância cruzada de incrementos $C_{ij}^I(x_\alpha, x_\beta)$ (que não são uma translação invariante) obtém-se:

$$\sigma_E^2 = \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{\alpha=0}^{n_i}\sum_{\beta=0}^{n_j}\lambda_\alpha^i\,\lambda_\beta^j\,C_{ij}^I(x_\alpha, x_\beta)$$

A fim de converter incrementos de covariância em variogramas, deve-se supor que os incrementos de covariância são simétricos. Com esta hipótese, pode-se obter a translação de quantidade invariante (Wackernagel, 1998):

$$\sigma_E^2 = 2\sum_{i=1}^{N}\sum_{\alpha}^{n_i}\lambda_\alpha^i\,\gamma_{ii_0}(x_\alpha - x_0) - \gamma_{i_0 i_0}(x_0 - x_0) - \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{\alpha=0}^{n_i}\sum_{\beta=0}^{n_j}\lambda_\alpha^i\,\lambda_\beta^j\gamma_{ij}(x_\alpha - x_\beta)$$

Após a minimização, as restrições nos ponderadores geram *N* multiplicadores de Lagrange μ_i, obtendo-se o sistema de cokrigagem ordinária, que segundo (Wackernagel, 1998) é:

$$\begin{cases}\displaystyle\sum_{j=1}^{N}\sum_{\beta=1}^{n_j}\lambda_\beta^j\gamma_{ij}(x_\alpha - x_\beta) + \mu_i = \gamma_{ii_0}(x_\alpha - x_0) & para \quad i = 1,...N; \quad \alpha = 1,...n_i \\ \displaystyle\sum_{\beta=1}^{n_i}\lambda_\beta^i = \delta_{ii_0} & para \quad i = 1,...N,\end{cases}$$

e a variância de cokrigagem será:

$$\sigma_{CK}^2 = \sum_{i=1}^{N}\sum_{\alpha=1}^{n_i}\lambda_\alpha^i\gamma_{ii_0}(x_\alpha - x_0) + \mu_{i_0} - \gamma_{i_0 i_0}(x_0 - x_0)$$

2.4.3 Cokrigagem Colocalizada

Uma situação particular de heterotopia é encontrada na colocalização dos dados, pois, como dito, a variável de interesse é conhecida em poucos pontos do domínio e a variável auxiliar é conhecida em todo o domínio, ou pelo menos, em todos os nós de uma malha de estimativa e nos locais de amostragem da variável de interesse. Com abundância de dados da informação auxiliar, a escolha de uma vizinhança criteriosa é um ponto crucial na estimativa (Wackernagel, 2003).

Segundo Xu *et al*. (1992), Goovaerts (1997) e Wackernagel (2003), se as variáveis secundárias são exaustivamente amostradas, quando comparadas à variável primária, o sistema de cokrigagem pode se tornar instável, devido à correlação dos dados secundários (localmente próximos) ser muito maior que a correlação entre os distantes dados primários. Esta feição gera nos resultados valores negativos, pois parte dos pesos da variável secundária serão negativos para garantir a segunda condição de restrição da cokrigagem ordinária, onde a soma dos pesos deve ser zero. A Figura 3 ilustra esta situação, em que se mostra o histograma do resultado de uma cokrigagem ordinária com a presença de valores menores do que zero, quando o valor mínimo amostral era 1,69. Além disso, dados que estão próximos, ou até mesmo colocalizados com valores primários não conhecidos tendem a esconder a influência de dados secundários que se encontram mais afastados (efeito tela).

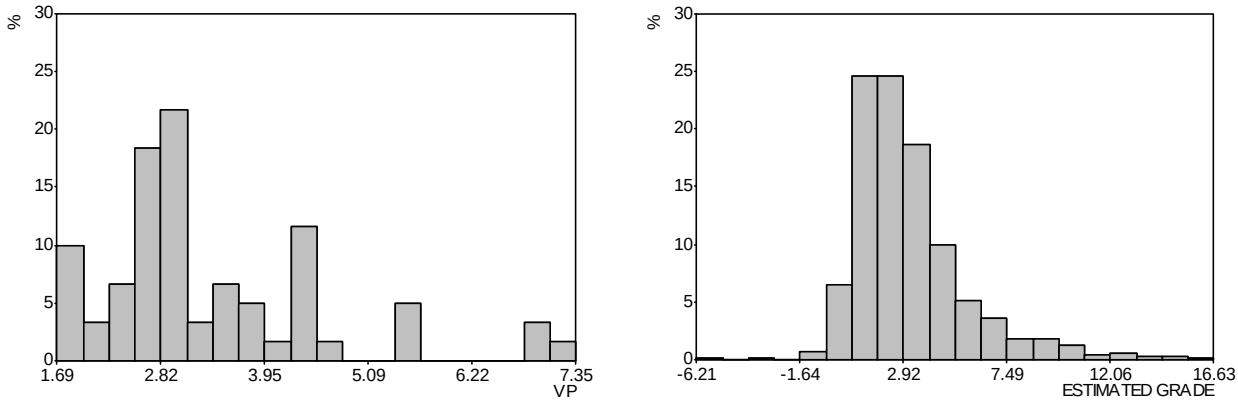

Figura 3: Histograma amostral (esquerda) e dos resultados da cokrigagem ordinária (direita).

Na pesquisa de petróleo, a situação descrita anteriormente é bastante comum. Como exemplo pode-se citar a caracterização do volume poroso de um reservatório utilizando informações de porosidade, medidas esparsamente nos poços, e os dados

de velocidade de trânsito ou impedância acústica obtidos em levantamentos sísmicos tridimensionais. Estas informações são correlacionadas naturalmente com a porosidade e podem ser consideradas quasi-exaustivas.

Outra utilização em que a cokrigagem colocalizada alcança bom desempenho é na hidrogeologia, por exemplo, na predição do nível piezométrico de uma região. Neste contexto, a variável de interesse, nível piezométrico é subamostrada em poucos poços de monitoramento, enquanto a topografia, fortemente correlacionada com piezometria pode ser densamente amostrada em levantamentos ou cartas topográficas.

Boezio *et al.* (2006), aplicaram esta técnica na coestimativa do nível piezométrico utilizando a topografia como informação secundária, alcançando resultados satisfatórios com esta técnica (Figura 4).

Neste contexto, a cokrigagem colocalizada (CCK) (Xu *et al.*, 1992 p. 835) é uma alternativa para integrar informações secundárias na coestimativa, pois as informações de variáveis secundárias densamente amostradas auxiliam na redução da variância do erro sem comprometer a estabilidade do sistema de equações lineares da cokrigagem. Além desta característica pode-se utilizar a cokrigagem colocalizada sem a exigência de colocalização dos dados secundários com as informações da variável primária (como acontece na krigagem com deriva externa).

A situação em que os dados secundários são colocalizados com a variável primária e com a malha que se deseja estimar caracteriza a vizinhança multi-colocalizada (Chilès & Delfiner, 1999) e, neste caso, pode-se utilizar a cokrigagem ordinária colocalizada.

2.4.3.1 Cokrigagem colocalizada simples

A cokrigagem colocalizada consiste basicamente na simplificação da vizinhança que será considerada na estimativa. Quando há colocalização da informação secundária apenas com o ponto x_0 que será estimado, pode-se utilizar cokrigagem simples colocalizada. O principal reflexo desta simplificação é a redução do número de ponderadores e dados da variável secundária a um, conforme expresso a seguir (Xu *et al.*, 1993 p. 835):

$$Z_1^{CCL}(x_o) - m_1 \quad = \quad \sum_{i=1}^{N} \lambda_i^1 \left[z_1(x_i) - m_1 \right] + \lambda^2 \left[z_2(x_0) - m_2 \right],$$

em que $Z_1^{CCL}(x_o)$ é o valor da variável primária calculado por cokrigagem colocalizada simples no ponto x_0, λ_i^1 e $z_1(x_i)$ são os ponderadores e os valores amostrados da variável primária, respectivamente, m_1 e m_2 são as médias globais das variáveis primária e secundária, λ^2 e $z_2(x_0)$ são o ponderador e o valor da variável secundária no ponto x_0 sendo estimado.

Ainda segundo este autor, o sistema de cokrigagem será definido como:

$$\begin{cases} \sum_{i=1}^{n} \lambda_i^1 C_{11}(x_i - x_j) + \lambda^2 C_{21}(x_0 - x_j) = C_{11}(x_0 - x_j), \quad j = 1, \ldots, n \\ \\ \sum_{i=1}^{n} \lambda_i^1 C_{12}(x_i - x_j) + \lambda^2 C_{22}(0) = C_{12}(0) \end{cases}$$

Em que $C_{11}(h)$, $C_{22}(h)$, $C_{12}(h) = C_{21}(h)$ são as covariâncias cruzadas:

$$C_{11}(h) = Cov\{Z_1(x), Z_1(x+h)\}$$

$$C_{22}(h) = Cov\{Z_2(x), Z_2(x+h)\}$$

$$C_{12}(h) = Cov\{Z_1(x), Z_2(x+h)\}$$

$$C_{21}(h) = Cov\{Z_2(x), Z_1(x+h)\}.$$

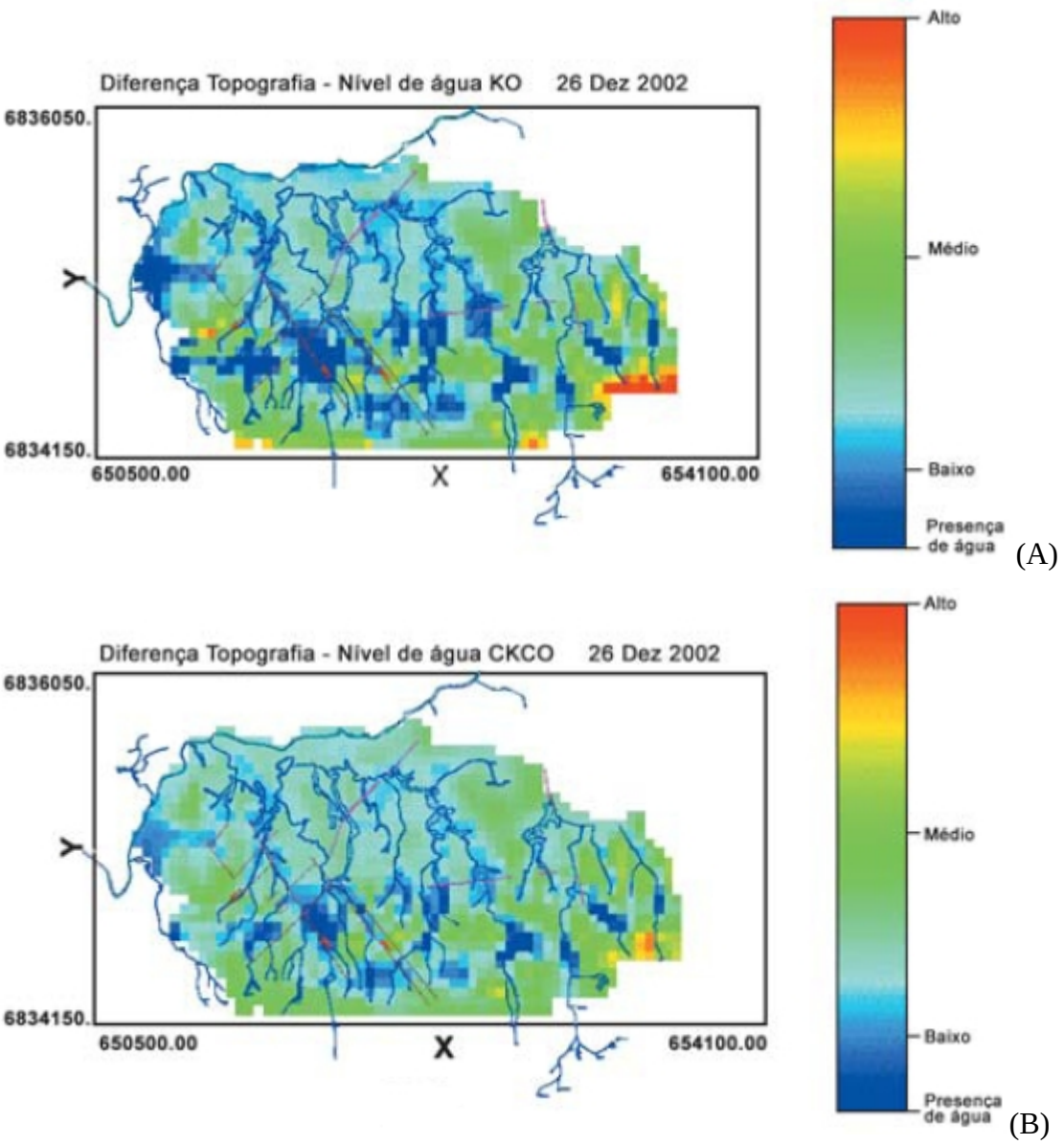

Figura 4: Mapas da diferença entre a topografia densamente amostrada e o nível de água estimado, utilizando-se krigagem ordinária (A) e cokrigagem colocalizada ordinária (B), conforme Boezio *et al.* (2006).

2.4.3.2 Cokrigagem colocalizada ordinária

Assim como a krigagem com deriva externa, a cokrigagem ordinária colocalizada exige que a vizinhança seja multi-colocalizada. Esta premissa é necessária para assegurar que o resultado do sistema de cokrigagem ordinária não seja a solução trivial, pois, para respeitar a condição de não viés o ponderador da variável secundária λ_2 seria zero e sua influência na estimativa não seria considerada (Wackernagel, 2003 p.167).

Segundo Goovaerts (1997), o ponderador da cokrigagem ordinária colocalizada (multi-colocalizada) é expresso como:

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

em que $Z^*_{CCK}(x_0)$ é o estimador da cokrigagem ordinária colocalizada, λ_i^z e λ_i^S são os ponderadores da variável primária e secundária, respectivamente, λ_0 é o ponderador da característica secundária no ponto sendo estimado x_0, $Z(x_i)$ é a variável primária e $S(x_i)$ é a variável secundária e m_1 e m_2 são as médias locais das variáveis primária e secundária.

Como em todos os algoritmos de cokrigagem, na CCK é necessário respeitar as condições de não viés e minimização da variância do erro. Entretanto, não é possível utilizar as restrições embasadas na soma dos ponderadores da variável primária e secundária ser igual a um e zero, respectivamente, uma vez que um peso nulo para a variável secundária na CCK resultaria na estimativa por krigagem. Esta particularidade do coestimador da CCK pode ser contornada pela utilização da restrição simples de não viés, originada do algoritmo de cokrigagem ordinária padronizada (Isaaks & Srivastava, 1989 p. 409 – 416; Deutsch & Journel, 1992 p. 70) que consiste em garantir que a soma de todos os ponderadores seja igual a um, conforme:

$$\sum_{i=1}^{n} \lambda_i^1(x_0) + \sum_{j=1}^{m} \lambda_j^2(x_0) \quad = \quad 1.$$

A utilização de uma única restrição de não viés implica na exigência que a média da variável secundária $Z_2(x)$ seja padronizada ou re-escalonada com a média da variável primária.

Finalmente, os ponderadores da variável primária e da variável secundária na cokrigagem ordinária colocalizada podem ser calculados pela resolução do sistema de equações lineares a seguir:

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

2.4.3.2.1 Utilização do Modelo Markoviano

Considerando a hipótese de efeito tela do tipo Markoviano ter-se-á (Xu *et al.*, 1992):

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

ou seja, o dado primário $z_1(x)$ ocultará qualquer influência de qualquer informação da variável primária $z_1(x+h)$ vizinha a variável secundária colocalizada $Z_2(x)$.

Desta maneira, é possível expressar as covariâncias $C_{12}(h)$ = $C_{21}(h)$ da seguinte forma:

$$C_{12}(h) \quad = \quad \frac{C_{12}(0)}{C_1(0)} C_1(h),$$

que pode ser escrito em função da correlação como:

$$\rho_{12}(h) \quad = \quad \rho_{12}(0)\rho_1(h)$$

em que $\rho_1(h) = \frac{C_1(h)}{C_1(0)}$ é o correlograma de Z_1, $\rho_{12}(h) = \frac{C_{12}(h)}{\sqrt{C_1(0)C_2(0)}}$ é o correlograma cruzado entre Z_1 e Z_2 e $\rho_{12}(0)$ é o coeficiente de correlação entre as variáveis.

A utilização do modelo Markoviano de corregionalização válido é bastante conveniente uma vez que o modelo de covariância $C_{12}(h)$ é derivado e padronizado em

função da covariância da variável primária e indiscutivelmente simplifica o algoritmo de cokrigagem. Outro aspecto positivo na utilização deste modelo é a facilidade para verificar sua validade, uma vez que para tal é necessário apenas que se calcule a covariância cruzada experimental e a compare na forma padronizada (Xu *et al.*, 1992):

$$\frac{C_{12}(h)}{C_{12}(0)} \cong \frac{C_1(h)}{C_1(0)}, \quad \forall h .$$

Estes autores interpretam os métodos de estimativa que utilizam um modelo Markoviano como os consagrados filtros do tipo passa-baixa, em que os resultados serão mapas mais suavizados sem a reprodução do espectro da variância em freqüências mais altas.

2.5 Krigagem com deriva externa

O método de krigagem com deriva externa, bem como krigagem universal ou krigagem regressiva são denominados, por McBratney *et al.* (2000), como métodos híbridos, ou seja, métodos que utilizam funções determinísticas (representando a contribuição não estacionária do fenômeno) e funções probabilísticas (referentes à contribuição estacionária do fenômeno espacial).

Proposta por Wackernagel (1998), a krigagem com deriva externa (*KED*) é um método alternativo, utilizado em situações em que uma característica de interesse é amostrada esparsamente e uma variável auxiliar ou secundária linearmente correlacionada com a variável de interesse é exaustivamente amostrada ou então multi-colocalizada. Este cenário é bastante comum na indústria do petróleo, como mencionado anteriormente.

Bourennane *et al*. (1996, 2000) constataram melhor desempenho da *KED* na estimativa de espessura de um horizonte pedológico tendo como deriva uma mapa de inclinações. Bourennane & Kind (2003) utilizam múltiplas variáveis (topografia, resistividade) como deriva na krigagem universal e krigagem com deriva externa para determinação do topo da camada de um carbonato, alcançando resultados consistentes com a *KED*.

Segundo Wackernagel (2003), as variáveis de interesse/primária e a variável auxiliar/secundária são respectivamente:

- uma medida precisa amostrada no poço ou sondagem. Sua quantidade é reduzida por conta dos altos custos de perfuração. Geralmente é uma função aleatória estacionária de segunda ordem representada como $Z(x)$ e sua função covariância é conhecida;
- uma medida imprecisa e cobrindo o domínio em grande escala. Esta é representada como uma variável regionalizada $s(x)$ e considerada como determinística.

Tanto a variável primária como a secundária são duas formas diferentes de expressar um fenômeno. Desta maneira, espera-se que estas sejam iguais em média. Conforme Wackernagel (2003), assumindo que $Z(x)$ é em média igual a $s(x)$ mais uma constante a_0 e um coeficiente de correlação b_1:

$$E[Z(x)] = a_0 + b_1 s(x)$$

A função determinística $s(x)$ descreve a forma geral e imprecisa da característica de interesse, enquanto os dados da função $Z(x)$ fornecem, em alguns locais, a medida exata da variável. O método mescla as diferentes fontes de informação e utiliza $s(x)$ como uma função determinística de deriva externa para a estimativa de $Z(x)$.

Ainda segundo Wackernagel (2003), o estimador da krigagem com deriva externa é a combinação linear:

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

em que a soma dos pesos λ_i é igual a um, assim:

$$E[Z^{KED}(x_0)] = E[Z(x_0)],$$

que desenvolvido apresenta-se como (Wackernagel, 2003):

$$E\left[Z^{KED}(x_0)\right] \;=\; \sum_{i=1}^{n}\lambda_i\, E\left[Z(x_0)\right] \;=\; a_0 + b_1\sum_{i=1}^{n}\lambda_i s(x_i) \;=\; a_0 + b_1\, s(x_0).$$

Esta última equação implica que os pesos devem ser consistentes com uma interpolação exata de $s(x)$:

$$s(x_0) = \sum_{i=1}^{n}\lambda_i\, s(x_i).$$

Esta última restrição dos pesos $\{\lambda_i, i = 1, n\}$ descreve a forma média fornecida pela variável secundária. Os pesos que minimizam a variância do erro sob as duas restrições resultam da solução do seguinte sistema de equações (Wackernagel, 2003):

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

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

A krigagem com deriva externa, cujo estimador é semelhante ao da krigagem ordinária (não considera a informação secundária diretamente no estimador), permite a correção do efeito de suavização, tal como descrito no item 2.3.3. Esta alternativa será considerada no presente estudo.

CAPÍTULO3

MATERIAIS E MÉTODOS

Dados reais de um campo produtor ou mesmo de um campo já completamente explotado como o Campo de Namorado são muito difíceis de serem disponibilizados para fins de pesquisa acadêmica. No caso desta pesquisa, os dados do Campo Escola Namorado foram disponibilizados pela Agência Nacional do Petróleo, mas após cuidadosa verificação, concluiu-se que os dados não eram confiáveis, pois não apresentavam consistência quanto à localização dos poços, bem como dificuldade em recuperar os dados sísmicos. Estas foram às razões que nos levaram a trabalhar com dados sintéticos. Quando se trabalha com a cokrigagem ou krigagem com deriva externa há necessidade do conjunto apresentar ao menos duas variáveis (no caso da cokrigagem) que estejam correlacionadas entre si. Assim, para a geração de dados, para esta pesquisa, foram considerados os dados completos conforme o mapa imagem da variável primária mostrado na Figura 5 e o mapa imagem da variável secundária e o diagrama de dispersão apresentados na Figura 6. As duas variáveis apresentam alta correlação como mostra a Figura 6 (coeficiente de correlação igual a 0,993).

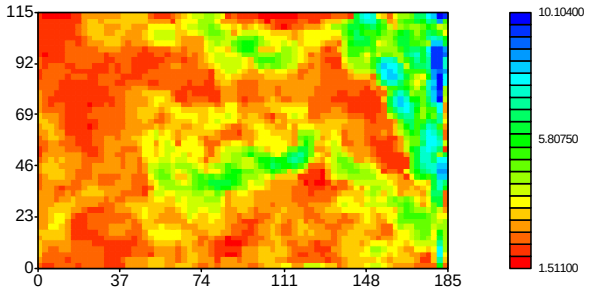

Figura 5: Mapa imagem da variável primária.

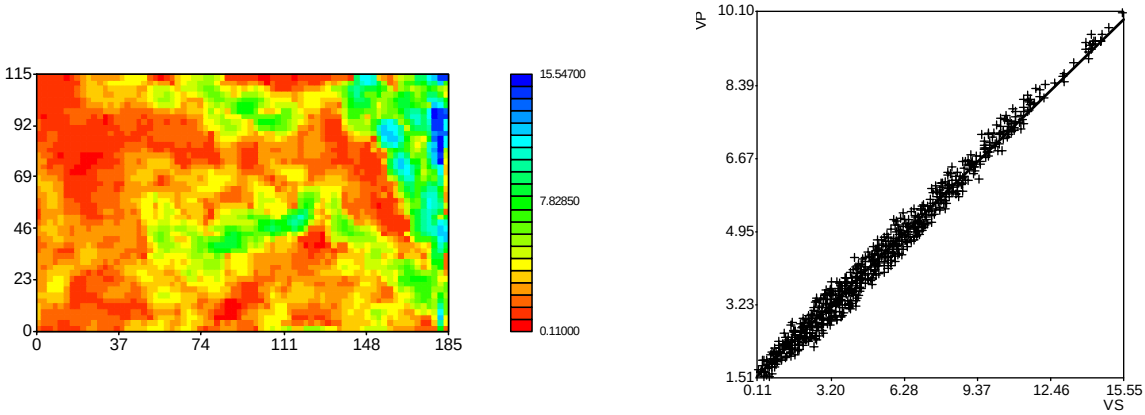

Figura 6: Mapa imagem da variável secundária (esquerda) e diagrama de dispersão mostrando a alta correlação entre a variável primária e a variável secundária (direita).

Quando a correlação é boa, a cokrigagem dá bons resultados. É preciso ressaltar que em qualquer método de estimativa, é necessário que o resultado apresente precisão local. Por precisão local entende-se a correlação das estimativas com os pontos vizinhos utilizados. Contudo, se existir um conjunto completo, pode-se medir a precisão local calculando-se a correlação com os dados conhecidos do conjunto completo. Observe-se que a simulação estocástica não apresenta boa precisão local, pois os valores simulados não estão bem correlacionados com os pontos vizinhos, mesmo que a simulação seja condicional. Evidentemente, nem sempre as variáveis encontram-se bem correlacionadas como na Figura 6. Assim, deve-se testar a eficiência da técnica da cokrigagem com variáveis secundárias que apresentem correlações menores com a mesma variável primária. Da mesma forma como a cokrigagem, a precisão local das estimativas de krigagem com deriva externa também depende da correlação entre as variáveis. Qual a correlação desejada para que as estimativas apresentem precisão local? Para responder a esta pergunta, várias variáveis secundárias foram geradas com correlações gradativamente menores. Estas variáveis secundárias foram geradas com base na variável secundária original (Figura 6), mas deterioradas com a introdução de ruído branco em níveis gradativamente maiores. As imagens dos dados completos de novas variáveis secundárias encontram-se nas Figuras 7 a 10, juntamente com os respectivos diagramas de dispersão.

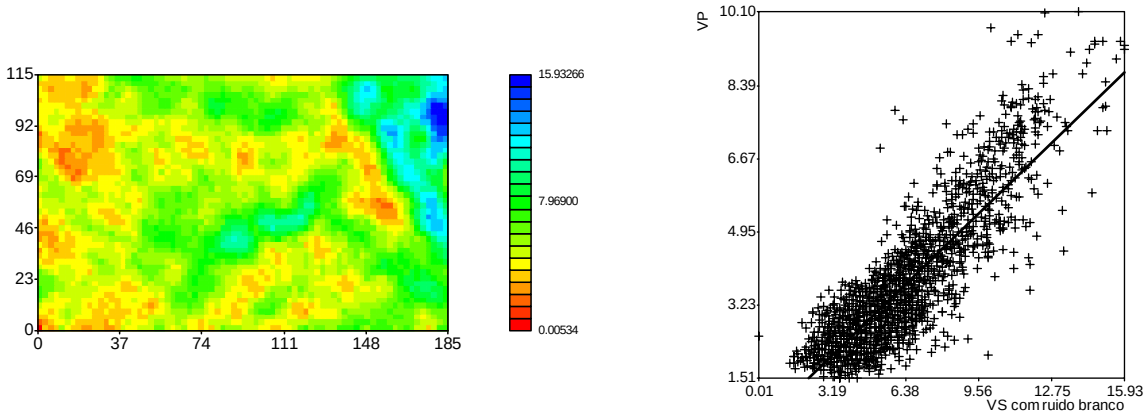

Figura 7: Mapa imagem da variável secundária com ruído branco (esquerda) e diagrama de dispersão entre a variável primária e a variável secundária com coeficiente de correlação 0,870 (direita).

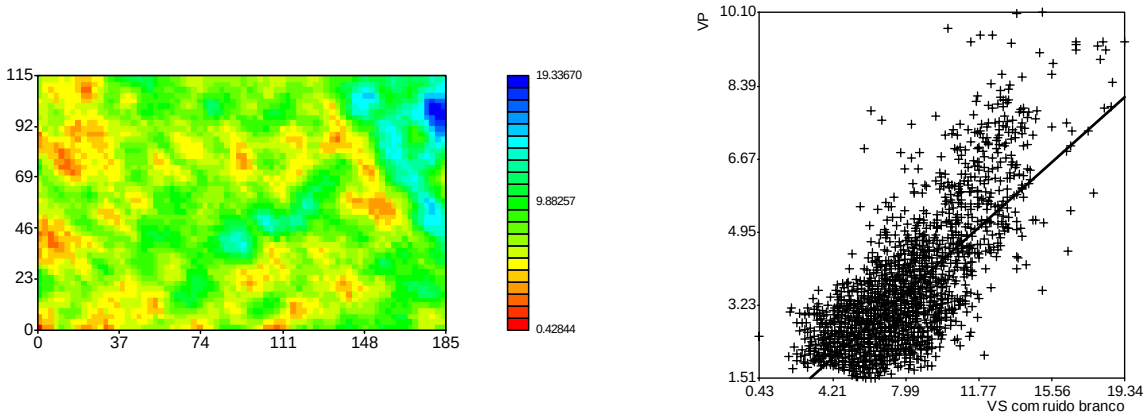

Figura 8: Mapa imagem da variável secundária com ruído branco (esquerda) e diagrama de dispersão entre a variável primária e a variável secundária com coeficiente de correlação 0,752 (direita).

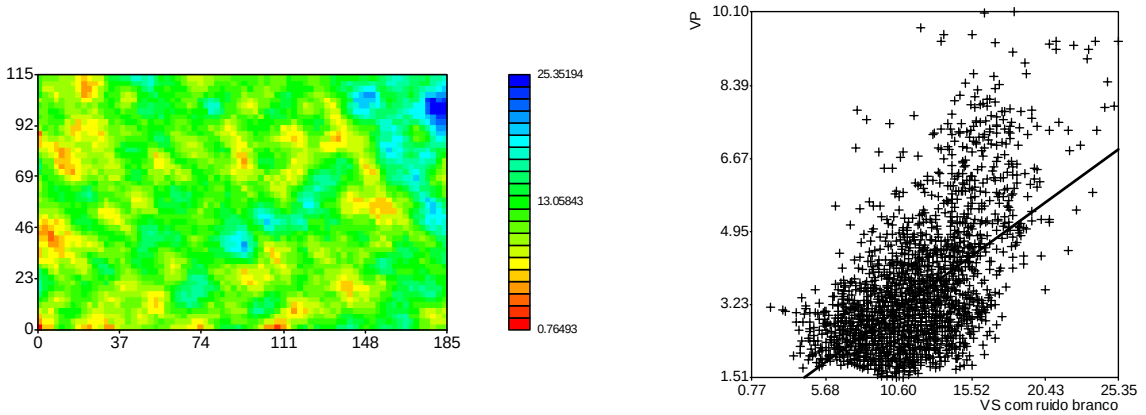

Figura 9: Mapa imagem da variável secundária com ruído branco (esquerda) e diagrama de dispersão entre a variável primária e a variável secundária com coeficiente de correlação 0,588 (direita).

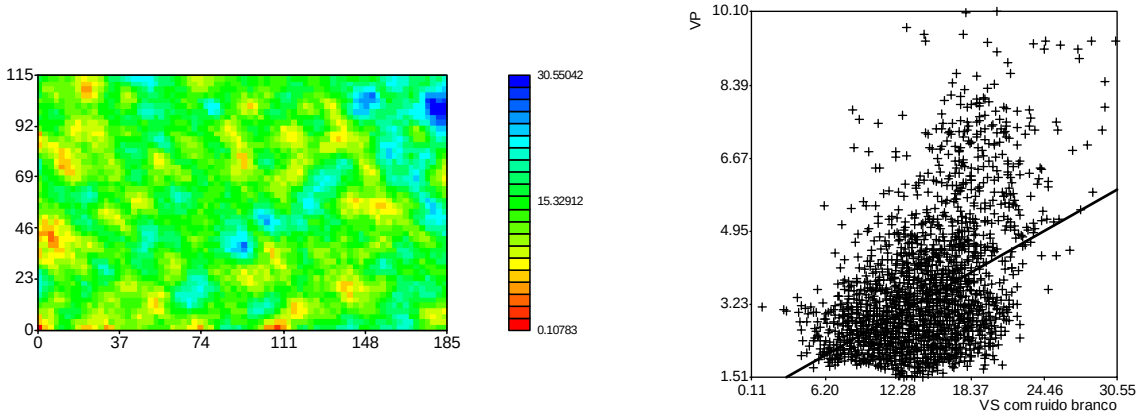

Figura 10: Mapa imagem da variável secundária com ruído branco (esquerda) e diagrama de dispersão entre a variável primária e a variável secundária com coeficiente de correlação 0,461 (direita).

Pode-se observar que à medida que a componente aleatória aumenta, a correlação da variável secundária com a primária diminui. Esses conjuntos de dados completos, nos quais a correlação varia de 0.993 a 0,461, podem ser utilizados para testar a eficiência da técnica da cokrigagem, bem como do método da krigagem com deriva externa.

A partir dos dados completos assim gerados, foram obtidos por amostragem aleatória estratificada os conjuntos de pontos de dados, que constituem os materiais desta pesquisa. Para cada conjunto de dado completo, foram amostrados dois subconjuntos com 60 e 104 pontos, ou respectivamente 1,76 e 3,06% dos dados completos. Geralmente, os dados exploratórios de petróleo são reduzidos frente à grande quantidade de dados sísmicos. Por isso, foram considerados subconjuntos com apenas 1,76% de dados contendo a variável primária. O objetivo de gerar subconjuntos de diferentes tamanhos foi o de testar a precisão local em função do tamanho da amostra.

Essas amostragens foram realizadas tanto para a variável primária como para a variável secundária. Portanto, os subconjuntos obtidos representam dados isotópicos para a cokrigagem. Para organizar melhor os conjuntos de dados, a Tabela 1 apresenta os arquivos de dados que serão utilizados neste trabalho.

Observe-se que para todos esses arquivos, as estatísticas descritivas da variável primária são alteradas somente com o tamanho da amostra, pois a variável secundária em nada influi com relação à distribuição da variável primária, exceto pela correlação. A Tabela 2 apresenta as estatísticas para a variável primária conforme os tamanhos da amostra.

Tabela 1: Conjuntos completos e subconjuntos obtidos por amostragem aleatória estratificada.

Coef. corr.	Completo	Subconjuntos amostrados	
		60	104
0,993	Composto20.txt	Teste2060.txt	Teste20104.txt
0,870	Composto40.txt	Teste4060.txt	Teste40104.txt
0,752	Composto100.txt	Teste10060.txt	Teste100104.txt
0,588	Composto120.txt	Teste12060.txt	Teste120104.txt
0,461	Composto121.txt	Teste12160.txt	Teste121104.txt

Tabela 2: Estatísticas da variável primária para as amostras selecionadas.

Estatísticas	Amostra 60	Amostra 104
No. dados	60	104
Média	3,365	3,146
Desvio padrão	1,262	1,370
Coef. Variação	0,375	0,435
Máximo	7,353	7,541
Quartil superior	3,803	3,272
Mediana	2,949	2,646
Quartil inferior	2,562	2,287
Mínimo	1,685	1,787

Portanto, as estatísticas apresentadas na Tabela 2 constituem uma componente da precisão global, que se pretende reproduzir através dos métodos geoestatísticos.

Para os subconjuntos amostrados, deve-se saber a priori a correlação entre a variável primária e variável secundária, pois esta se altera em função da amostragem. Além da correlação, deve-se saber a equação da reta de regressão. Na verdade, espera-se que a regressão da variável primária estimada com a variável secundária apresente um coeficiente angular próximo da regressão do conjunto amostral. As Figuras 11 a 15 apresentam os diagramas de dispersão para os subconjuntos amostrados. A Tabela 3 lista as retas de regressão entre a variável primária e a variável secundária para os subconjuntos amostrados.

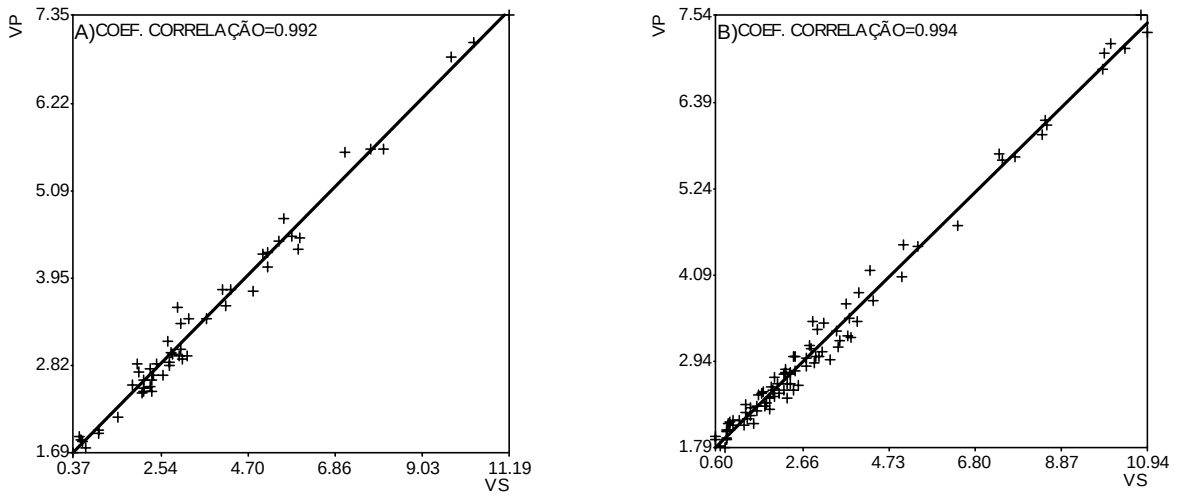

Figura 11: Diagramas de dispersão para as amostras retiradas do conjunto completo Composto20.txt com 60 unidades (A) e com 104 unidades (B).

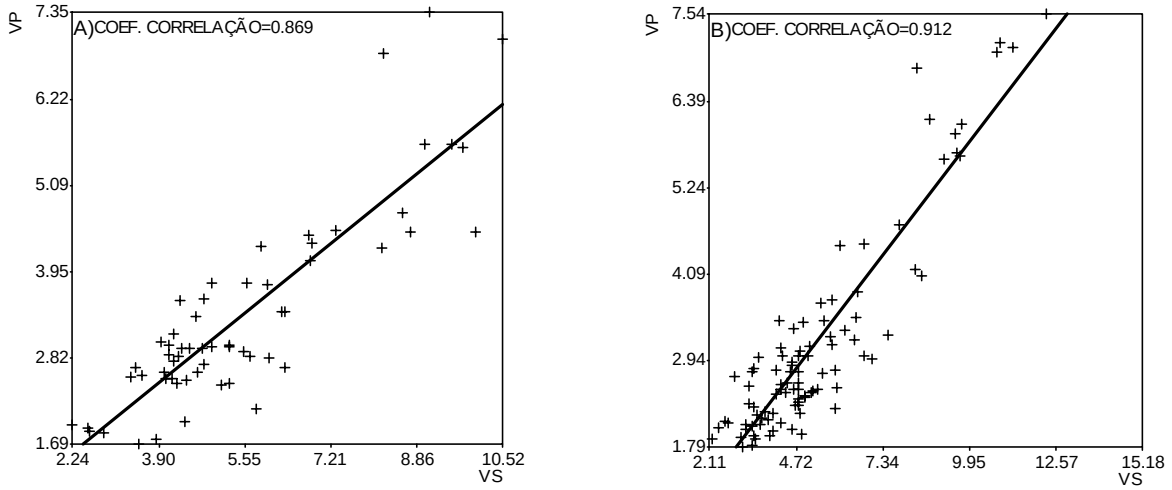

Figura 12: Diagramas de dispersão para as amostras retiradas do conjunto completo Composto40.txt com 60 unidades (A) e com 104 unidades (B).

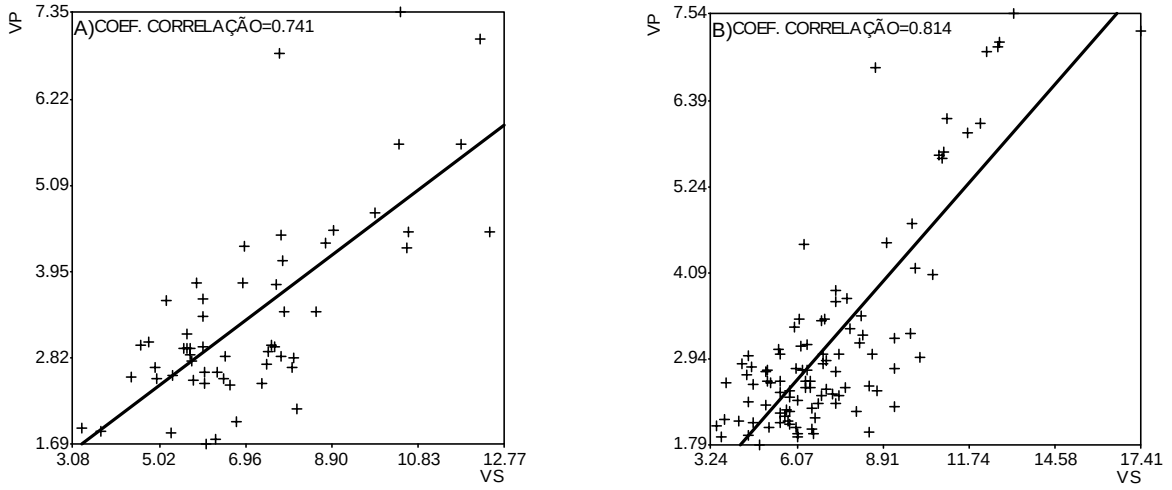

Figura 13: Diagramas de dispersão para as amostras retiradas do conjunto completo Composto100.txt com 60 unidades (A) e com 104 unidades (B).

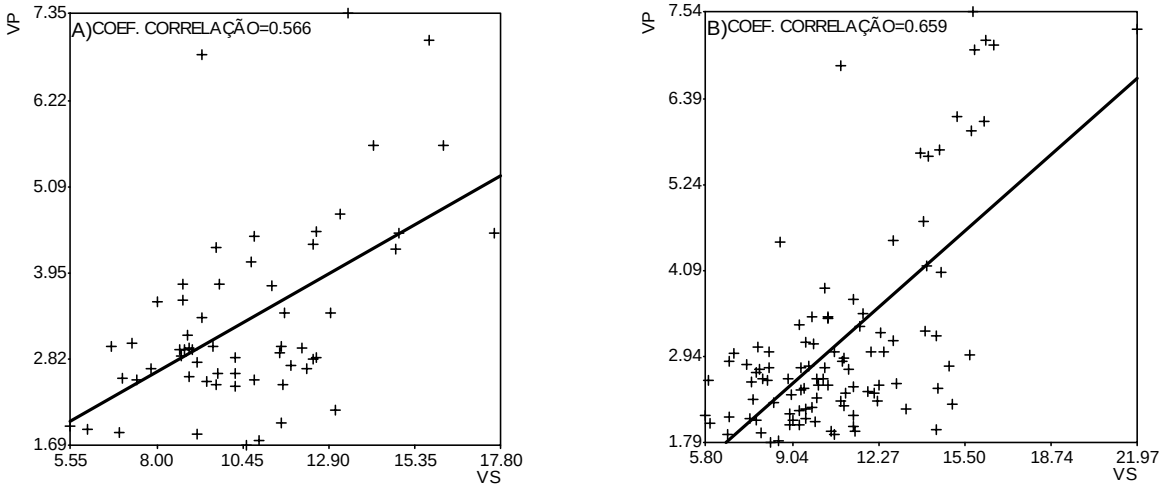

Figura 14: Diagramas de dispersão para as amostras retiradas do conjunto completo Composto120.txt com 60 unidades (A) e com 104 unidades (B).

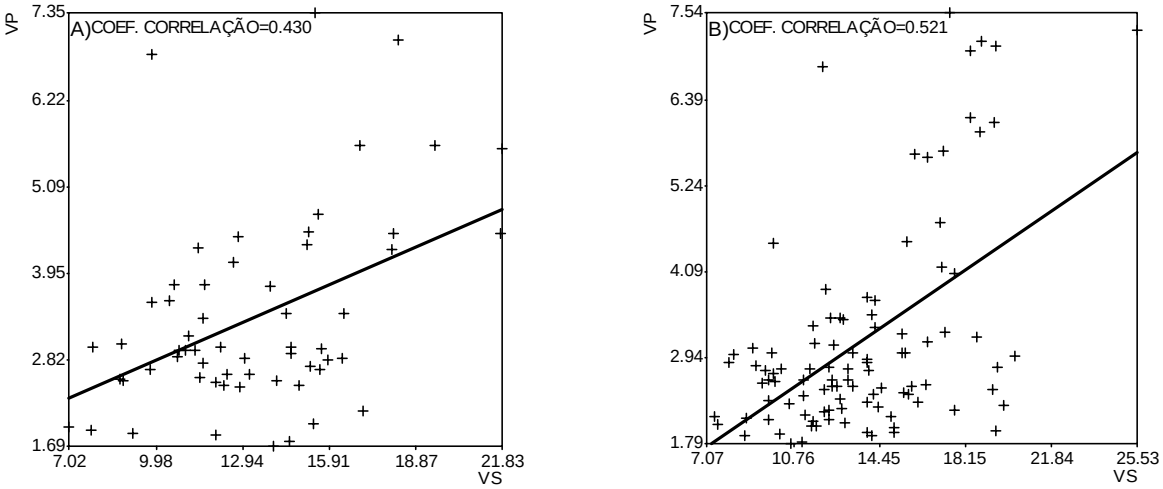

Figura 15: Diagramas de dispersão para as amostras retiradas do conjunto completo Composto121.txt com 60 unidades (A) e com 104 unidades (B).

Tabela 3: Retas de regressão da variável primária e a variável secundária para os subconjuntos amostrados.

Completos	Subconjuntos amostrados	
	60	104
Composto20.txt	VP=1,488+0,530VS	VP=1,473+0,546VS
Composto40.txt	VP=0,326+0,554VS	VP=0,109+0,575VS
Composto100.txt	VP=0,235+0,441VS	VP=-0,153+0,463VS
Composto120.txt	VP=0,552+0,263VS	VP=-0,270+0,315VS
Composto121.txt	VP=1,156+0,166VS	VP=0,248+0,213VS

Os arquivos da Tabela 1 representam dados isotópicos em termos da cokrigagem, pois cada ponto de amostragem tem tanto o valor da variável primária como da secundária. Quando os dados são isotópicos, a cokrigagem não faz uso da

correlação com a variável secundária. Deve-se observar que a cokrigagem parte do princípio que a variável primária é subamostrada e, portanto, precisa de uma informação secundária que esteja mais bem amostrada. Além disso, com esses dados, os resultados da cokrigagem não podem ser comparados com aqueles da krigagem com deriva externa, haja vista esta técnica necessitar da informação secundária em todos os nós da malha regular que vai ser estimada. Assim, há necessidade de introduzir a informação secundária para os dados isotópicos a fim de proporcionar melhores resultados da cokrigagem. Para verificar a influência da informação secundária, foram adicionados valores correspondentes a 33% dos nós da malha regular a ser estimada. Os arquivos assim obtidos, que representam dados com heterotopia parcial, encontram-se na Tabela 4.

Tabela 4: Subconjuntos amostrados adicionados com 33% dos nós da malha regular com valores da informação secundária.

Completo	Subconjuntos amostrados	
	60	104
Composto20.txt	Teste206033.txt	Teste2010433.txt
Composto40.txt	Teste406033.txt	Teste4010433.txt
Composto100.txt	Teste1006033.txt	Teste10010433.txt
Composto120.txt	Teste1206033.txt	Teste12010433.txt
Composto121.txt	Teste1216033.txt	Teste12110433.txt

Os arquivos constantes nas Tabelas 1 e 4 constituem os materiais do projeto de pesquisa da dissertação. Observe-se que são 30 arquivos no total, os quais demandam uma grande quantidade de processamento.

Os métodos a serem empregados referem-se às técnicas consagradas na literatura geoestatística, quais sejam: cokrigagem ordinária, cokrigagem colocalizada e krigagem com deriva externa, conforme apresentados no Capítulo 2 (Revisão da metodologia corrente). Todos os processamentos serão efetuados através dos programas do Sistema GeoVisual (Yamamoto, 2006). Trata-se de um sistema desenvolvido com a finalidade de suportar as atividades práticas em geoestatística aplicada no Instituto de Geociências – USP.

As Figuras 16 e 17 apresentam os mapas de localização de pontos para os arquivos da Tabela 1, respectivamente para 60 e 104.

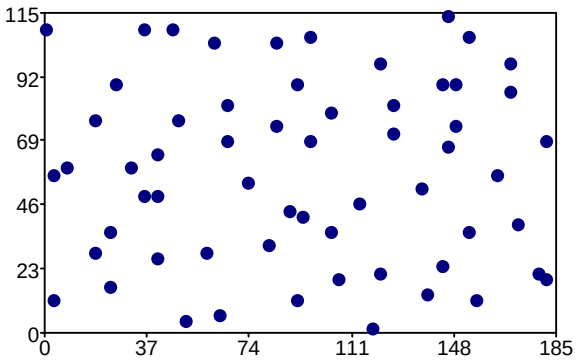

Figura 16: Mapa de localização de pontos para arquivos isotópicos com 60 pontos.

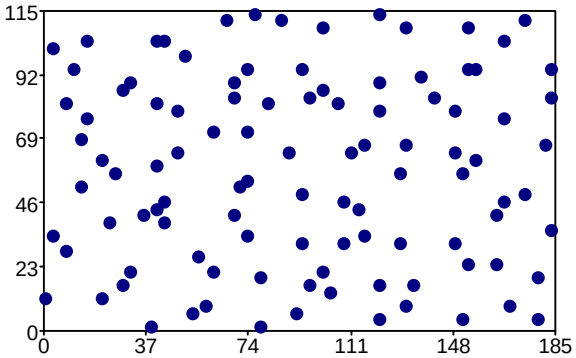

Figura 17: Mapa de localização de pontos para arquivos isotópicos com 104 pontos.

Os arquivos de pontos de dados da Tabela 4 apresentam os mesmos pontos como mostrados nas Figuras 16 e 17, mas somados da variável secundária representando 33% dos nós da malha regular que se pretende estimar.

CAPÍTULO 4

APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

Os resultados obtidos com os métodos de estimativa e coestimativa consagrados, cokrigagem ordinária colocalizada, krigagem com deriva externa e krigagem com deriva externa com correção do efeito de suavização serão apresentados e discutidos nesta ordem:

4.1 Krigagem Ordinária

A krigagem ordinária, por se tratar de um método de estimativa univariado, que considera na estimativa apenas as informações da variável de interesse, apresentou resultados diferentes apenas quando o tamanho da amostra foi alterado (Figuras 18 e 19). As estatísticas encontram-se na Tabela 5. Para a krigagem ordinária de ambas as amostras foram consideradas parâmetros de vizinhança exatamente iguais, quais sejam: 3 pontos por quadrante com um mínimo de três.

Com o acréscimo das amostras, há grande ganho de textura nos resultados, ou seja, melhora na reprodução do padrão, conforme observados nas Figuras 18 (esq.) e 19 (esq.). Contudo, insuficiente quando comparadas com a imagem do dado completo (Figura 5).

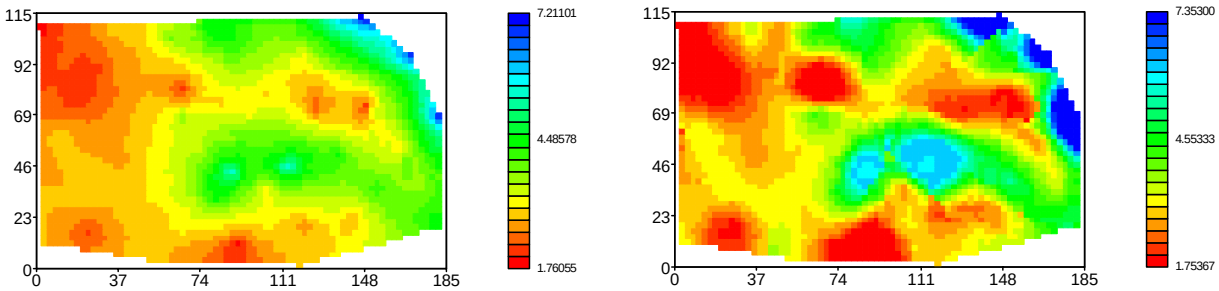

Figura 18: Mapas imagens da krigagem ordinária (esquerda) e da krigagem ordinária corrigida (direita) para amostras com 60 unidades.

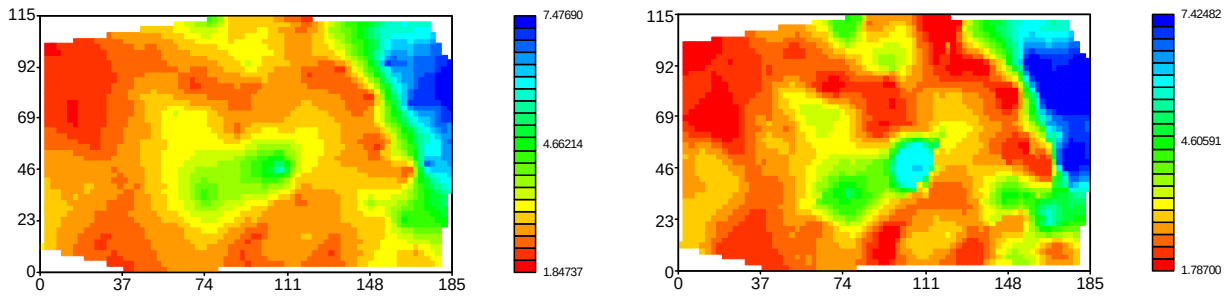

Figura 19: Mapas imagens da krigagem ordinária (esquerda) e da krigagem ordinária corrigida (direita) para amostras com 104 unidades.

Comparando-se as estatísticas da Tabela 5, com relação a krigagem ordinária, verifica-se que a reprodução do histograma é melhor para a amostra com 104 unidades. Isto se deve provavelmente à maior proximidade das amostras em relação ao ponto a ser estimado. No caso da amostra com 60 unidades, a redução da variância é brutal, enquanto para a amostra de 104 unidades a redução é bem menos drástica. Por outro lado, tem-se que a correção do efeito de suavização da krigagem ordinária melhora a reprodução do histograma, conforme se observa na Tabela 5.

Apenas a título de ilustração, vamos calcular as correlações entre a variável estimada e o valor real do conjunto completo (Figuras 20 e 21). Numericamente, as correlações são boas, mas apresentam grandes dispersões em torno da reta de regressão, notadamente após a correção do efeito de suavização (Figuras 20B e 21B). No caso da amostra com 60 unidades, verifica-se que há perda de correlação após a correção do efeito de suavização da krigagem ordinária. Isto se deve à grande distância entre os pontos amostrais, propagando os erros de estimativa, através do número de desvios padrão de interpolação.

Tabela 5: Estatísticas amostrais e dos valores estimados pela krigagem ordinária.

Estatística	Am.60	KO60	KOC60	Am.104	KO104	KOC104
No. dados	60	2934	2934	104	3133	3133
Média	3,365	3,244	3,365	3,146	3,174	2,146
Desvio	1,262	0,761	1,247	1,370	1,108	1,369
CV	0,375	0,235	0,371	0,435	0,349	0,435
Máx.	7,353	7,211	7,353	7,541	7,477	7,425
Q. sup.	3,803	3,728	4,017	3,272	3,273	3,340
Mediana	2,949	3,089	3,023	2,646	2,798	2,687
Q. inf.	2,562	2,699	2,565	2,287	2,541	2,289
Mín.	1,685	1,761	1,754	1,787	1,847	1,787

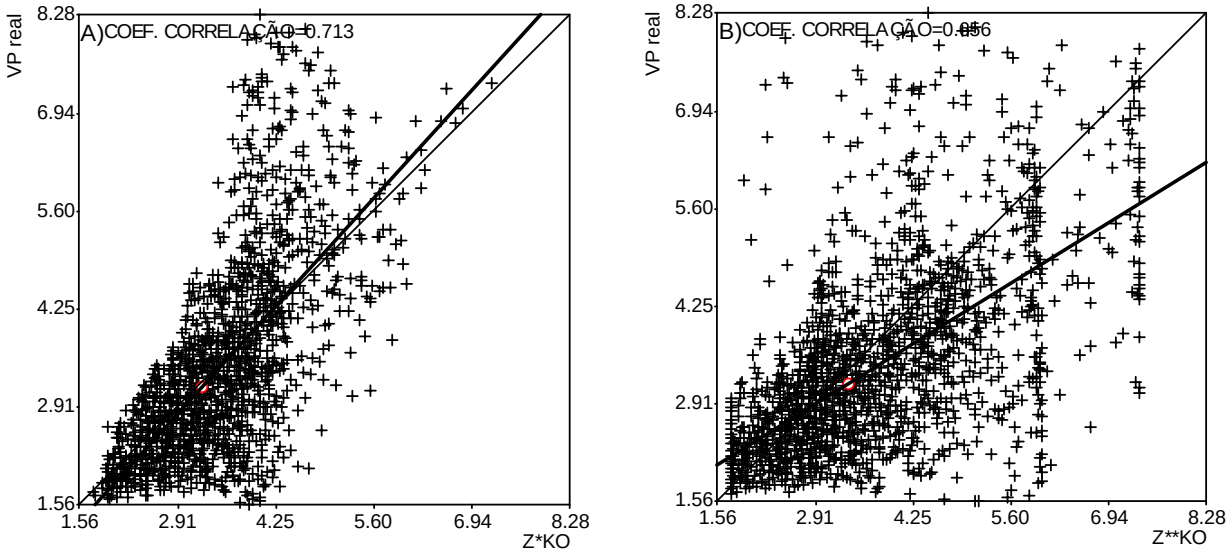

Figura 20: Diagramas de dispersão entre a variável primária real e a estimativa pela krigagem ordinária (A) e a estimativa corrigida do efeito de suavização (B) para a amostra com 60 unidades.

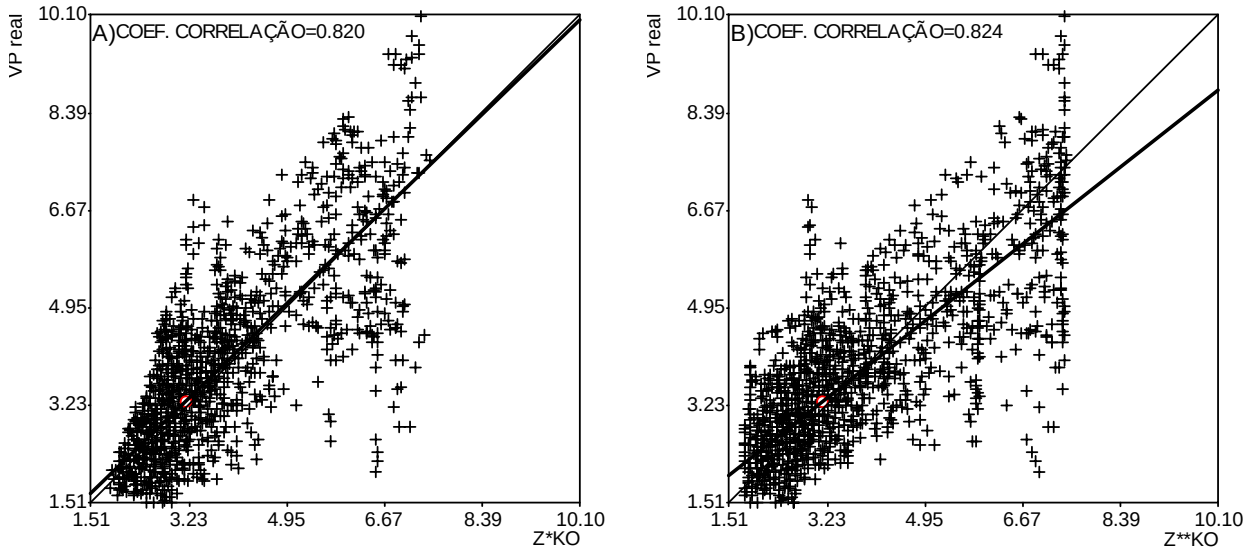

Figura 21: Diagramas de dispersão entre a variável primária real e a estimativa pela krigagem ordinária (A) e a estimativa corrigida do efeito de suavização (B) para a amostra com 60 unidades.

4.2 Cokrigagem Ordinária

A estimativa por cokrigagem ordinária não apresentou bom desempenho, principalmente com o aumento da informação secundária. Mesmo com apenas 33% da malha estimada coberta por informação secundária, a estimativa ficou bastante degradada, apresentando ampliação do intervalo de valores estimados e,

principalmente, com a presença de valores negativos que, fisicamente, é impossível para porosidade. Na realidade, 33% de 3404 nós significam 1123 pontos com informação secundária, o que não é pouco. Então, deve-se analisar primeiro as estatísticas apresentadas nas Tabelas 6 e 7, respectivamente para as amostras com 60 e 104 unidades.

Tabela 6: Estatísticas amostrais e dos valores estimados pela cokrigagem ordinária para a amostra com 60 unidades.

Estatística	Am. 60	CC=0,993	CC=0,870	CC=0,752	CC=0,588	CC=0,461
No. dados	60	2933	2933	2933	2933	2933
Média	3,365	3,165	5,364	7,194	10,952	13,718
Desvio	1,262	2,478	2,202	2,428	3,000	3,717
CV	0,375	0,783	0,411	0,337	0,274	0,271
Máx.	7,353	16,625	13,564	14,697	20,192	24,804
QS	3,803	4,041	6,394	8,581	12,731	16,077
Med.	2,949	2,668	4,975	6,934	10,886	13,735
QI	2,562	1,555	3,865	5,509	8,851	11,149
Mín.	1,685	-6,209	0,292	1,274	1,535	1,657

Tabela 7: Estatísticas amostrais e dos valores estimados pela cokrigagem ordinária para a amostra com 104 unidades.

Estatística	Am. 104	CC=0,993	CC=0,870	CC=0,752	CC=0,588	CC=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

O exame das estatísticas (Tabelas 6 e 7) permite tecer algumas considerações sobre a cokrigagem ordinária. Há uma forte deterioração das estatísticas, qualquer que seja a correlação com a variável secundária, que se reflete no deslocamento da média e aumento dos desvios padrão. Além disso, registra-se a presença de valores negativos. Evidentemente, isso se deve à grande quantidade de informação secundária cobrindo 33% da malha a ser estimada. Portanto, a informação secundária excessiva em nada ajuda a estimativa da variável primária pobremente amostrada. A porcentagem de cobertura com informação secundária igual a 33% foi escolhida com o objetivo de aproximar a comparação com a cokrigagem colocalizada, que requer uma cobertura de 100%.

As imagens estimadas pela cokrigagem ordinária para as amostras com 60 e 104 unidades encontram-se nas Figuras 22 e 23, respectivamente. Analisando os mapas imagens dessa figuras, verifica-se que aqueles gerados a partir de amostras com 104 unidades sempre apresentam artefatos (faixas de valores constantes). Isso se deve, provavelmente, à maior quantidade de pontos com a variável primária que se correlacionam com a variável secundária. Esses artefatos refletem-se também nos respectivos diagramas de dispersão (Figura 23 – canto inferior esquerdo dos diagramas).

Nos diagramas de dispersão das Figuras 22 e 23, verifica-se que existe um forte enviesamento condicional. Os valores estimados através da cokrigagem ordinária estão sempre superestimados em relação aos valores da variável primária real. Isso se deve à influência da variável secundária na cokrigagem ordinária e esta é a principal limitação do método da cokrigagem ordinária. Na verdade, o estimador da cokrigagem ordinária mistura variáveis com dimensões diferentes e isso é muito difícil de controlar para que a estimativa resultante não seja contaminada pela informação secundária.

Como se pode verificar, a cokrigagem ordinária apresenta um sério problema de instabilidade na resolução dos sistemas de equações de cokrigagem ordinária, como está mencionado na literatura corrente (Xu *et al*., 1992; Goovaerts, 1997; Wackernagel, 2003).

Somente em situações muito especiais e com um gerenciamento da vizinhança escolhida tanto para a variável primária como para a secundária e, sobretudo, com uma pequena quantidade de pontos com a informação secundária, se pode obter resultados satisfatórios com a cokrigagem ordinária. Na realidade, o ponto crucial neste método é a quantidade de informação secundária fornecida. Ela não deve ser excessiva, como no caso em estudo, pois deteriora significativamente as estimativas resultantes. Portanto, os dois métodos empregados até o momento não produziram resultados satisfatórios para a pesquisa.

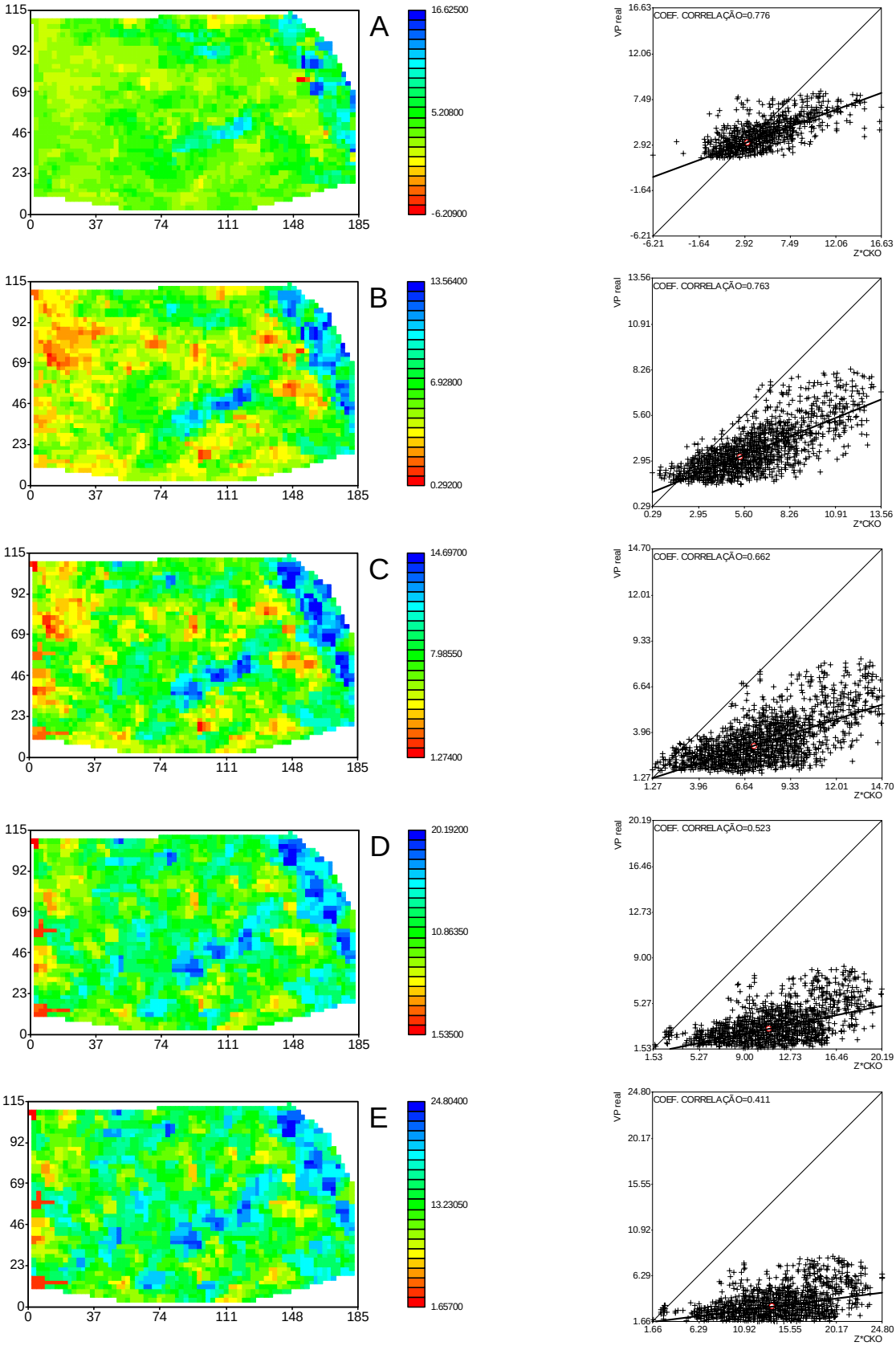

Figura 22: Resultados da cokrigagem ordinária: mapas imagens à esquerda e diagramas de dispersão à direita para amostras com 60 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

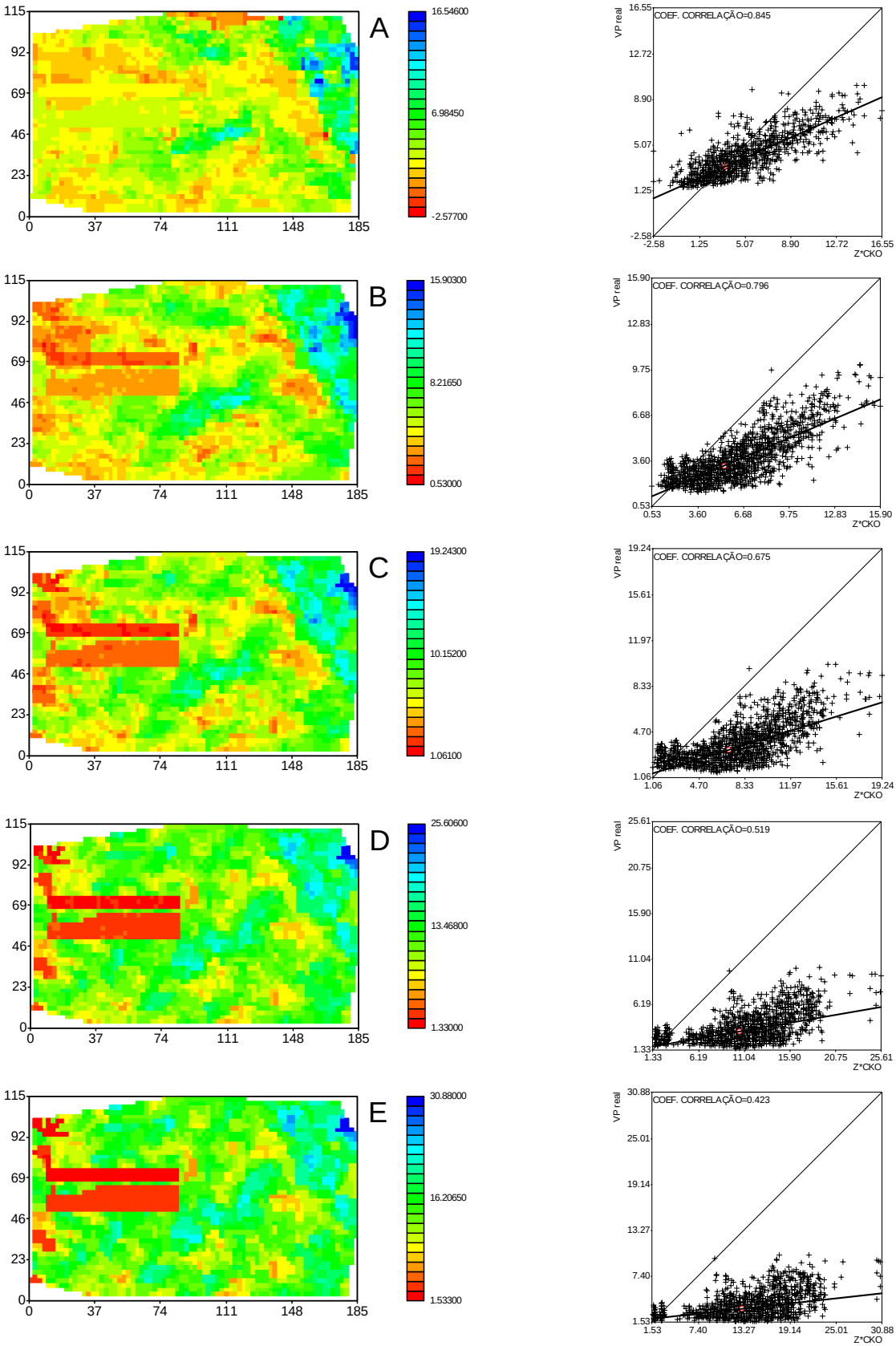

Figura 23: Resultados da cokrigagem ordinária: mapas imagens à esquerda e diagramas de dispersão à direita para amostras com 104 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

As variáveis primária e secundária apresentavam correlações iniciais que deveriam ser respeitadas após o processo de co-estimativa. Assim, com o objetivo de verificar essa correlação, foram construídos diagramas de dispersão entre os valores da variável primária estimada e os valores da variável secundária (Figura 24). Nessa figura, verifica-se que a correlação entre a variável estimada e a variável secundária é muito boa. Contudo, essas correlações não reproduzem as correlações amostrais o que seria importante para a precisão local das estimativas. A boa correlação entre as estimativas e os valores da variável secundária se deve à forte influência dessa variável na estimativa da variável primária, através da cokrigagem ordinária. Observe-se que não há deterioração da correlação com a diminuição do coeficiente de correlação entre as variáveis, como apresentam os conjuntos de dados amostrais. A quantidade de informação secundária considerada foi realmente excessiva neste estudo, mas tinha como objetivo fazer uma comparação com outros métodos de co-estimativas que precisam de 100% da informação secundária na malha a ser estimada.

Por fim, com relação ao efeito do tamanho da amostra na cokrigagem ordinária, observa-se uma pequena melhora na precisão local das estimativas, conforme se verifica nas Figuras 22 (direita) e 23 (direita). Os diagramas de dispersão da Figura 24 também corroboram essa verificação.

4.3 Cokrigagem Colocalizada Simples

Na utilização da técnica de cokrigagem colocalizada simples é imperativo que os dados sejam heterotópicos colocalizados. Para aferir o desempenho deste método, foram utilizadas amostras em que a informação secundária recobre toda a malha estimada e a informação primária contribui com 60 ou 104 pontos amostrais, sendo que os dados primários e secundários apresentam cinco níveis de correlação com a informação secundária, a saber: 0,993, 0,870, 0,752, 0,588 e 0,461, o que resultou em dez amostras distintas. As estatísticas amostrais e dos valores estimados encontram-se nas Tabelas 8 e 9.

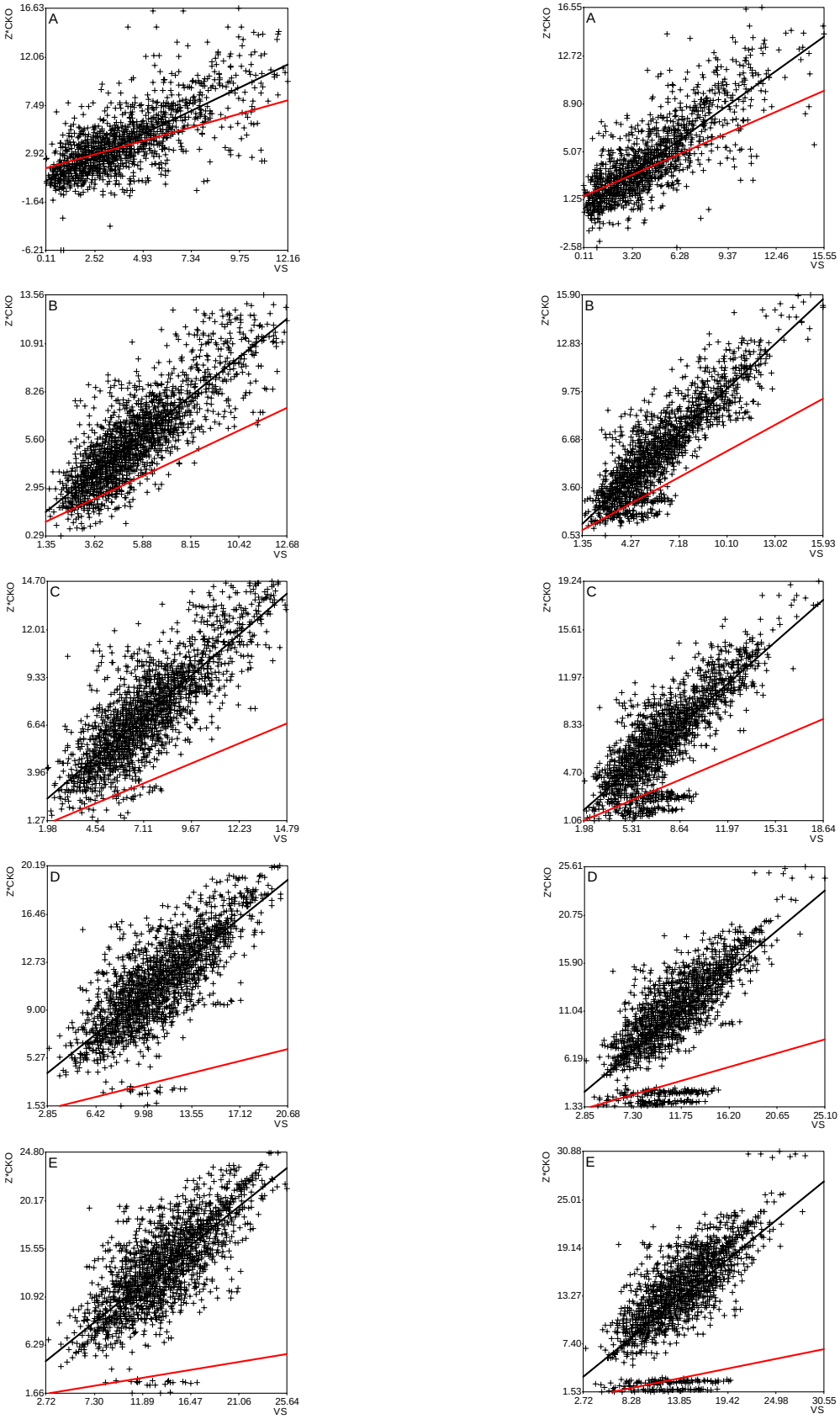

Figura 24: Diagramas de dispersão entre a variável primária estimada e a variável secundária para amostras com 60 unidades à esquerda e com 104 unidades à direita. Linha cheia preta = regressão da estimativa e linha vermelha cheia = regressão amostral. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

Tabela 8: Estatísticas amostrais e dos valores estimados pela cokrigagem colocalizada simples para a amostra com 60 unidades.

Estatística	Am. 60	CC=0,993	CC=0,870	CC=0,752	CC=0,588	CC=0,461
No. dados	60	3404	3404	3404	3404	3404
Média	3,365	3,246	3,381	3,437	3,428	3,411
Desvio	1,262	1,301	1,297	1,206	1,095	1,024
CV	0,401	0,384	0,351	0,319	0,300	0,271
Máx.	7,353	9,722	9,272	8,952	8,025	7,369
QS	3,803	3,701	3,911	3,967	3,938	3,896
Med.	2,949	2,892	3,024	3,170	3,190	3,178
QI	2,562	2,365	2,502	2,610	2,667	2,698
Mín.	1,685	1,546	0,528	0,801	1,020	1,280

Tabela 9: Estatísticas amostrais e dos valores estimados pela cokrigagem colocalizada simples para a amostra com 104 unidades.

Estatística	Am. 104	CC=0,993	CC=0,870	CC=0,752	CC=0,588	CC=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

Com relação à reprodução das estatísticas amostrais, a cokrigagem colocalizada simples é muito superior a cokrigagem ordinária. As médias encontram-se próximas àquelas amostrais. O desvio padrão sofre alteração por conta da expansão do intervalo de valores devido à influência da variável secundária. Até o momento, a cokrigagem colocalizada simples foi o melhor método empregado para os dados em estudo.

Mapas imagens dos resultados da cokrigagem ordinária colocalizada, bem como os respectivos diagramas de dispersão encontram-se nas Figuras 25 e 26. Nessas figuras, é notável a degradação da qualidade à medida que a correlação diminui, mas não tanto quanto a cokrigagem ordinária. Mais importante é que o padrão geral é reproduzido por influência da variável secundária.

Agora, comparando a variável primária estimada pela cokrigagem colocalizada com a informação secundária utilizada pode-se afirmar que, idealmente, a regressão calculada deve estar próxima da regressão amostral inicial. Os resultados desta comparação encontram-se na Figura 27. Os resultados apresentados nesta figura são surpreendentes. As retas de regressão amostral e calculada são muito próximas entre

si. Isto mostra que a técnica preserva a relação inicial entre a variável primária e a variável secundária. Observe-se que estamos trabalhando com amostras compostas por 60 e 104 unidades. Assim, as regressões amostrais refletem as relações mútuas obtidas com 60 e 104 unidades amostrais. Por esse motivo, o resultado obtido é excelente. Esse resultado confirma a conclusão anterior que a cokrigagem colocalizada é o melhor método de coestimativa testado até o momento.

Resta verificar a reprodução do histograma amostral. Para isso, vamos considerar o diagrama P-P (probabilidade – probabilidade) para fazer essa verificação (Figura 28). Ao examinar esse diagrama, verifica-se que as distribuições estimadas não se aproximam em relação às respectivas distribuições amostrais. No caso da amostra com 104 unidades, verifica-se que a variância está sempre acima da variância amostral. Evidentemente, o melhor coestimador deve manter não só a relação inicial entre a variável primária e a variável secundária, mas também procurar reproduzir o histograma. A variância dos valores estimados pode ser maior que a variância amostral devido à influência da variável secundária que tem um peso muito grande na estimativa. Nesse sentido, vale a pena investigar a influência dos pesos aplicados tanto para a variável primária como para a variável secundária na estimativa pela cokrigagem colocalizada simples (Figuras 29 e 30).

Os mapas imagens das Figuras 29 e 30 mostram claramente como ocorre a distribuição da soma dos pesos da variável primária e dos pesos da variável secundária com relação ao coeficiente de correlação. Pode-se observar que em pontos amostrados com a informação primária, ela tem maior relevância, ao passo que a influência da variável secundária é pequena, a qual aumenta para regiões não amostradas. As Figuras 31 e 32 apresentam os histogramas da distribuição da soma dos pesos da variável primária e do peso da variável secundária. Nesses histogramas se observa duas tendências, quais sejam: a soma dos pesos aumenta com a diminuição da correlação; o deslocamento da tendência central do peso da variável secundária tanto quanto diminui a correlação. Portanto, quanto menor a correlação entre a variável primária e a variável secundária, menor a contribuição da variável secundária e maior da variável primária. Isto faz sentido na técnica da cokrigagem colocalizada, pois se a variável secundária é pobremente correlacionada com a variável primária, então a sua contribuição deve ser pequena, para não contaminar a estimativa resultante. Estas relações podem ser vistas nas Figuras 33 e 34.

Como se pode observar nestas figuras, a função que melhor se ajusta é a curva parabólica. O comportamento dos pesos da cokrigagem colocalizada simples justifica a boa precisão local dos resultados obtidos, mesmo em situações de baixa correlação. Evidentemente, há uma deterioração da precisão local com a diminuição da correlação entre as variáveis, mas perfeitamente aceitáveis para dados experimentais. Observe-se inclusive a manutenção da regressão original verificada na amostra.

Com relação ao tamanho da amostra, observa-se apenas uma melhora na reprodução do histograma amostral como mostram dos diagramas P-P da Figura 28.

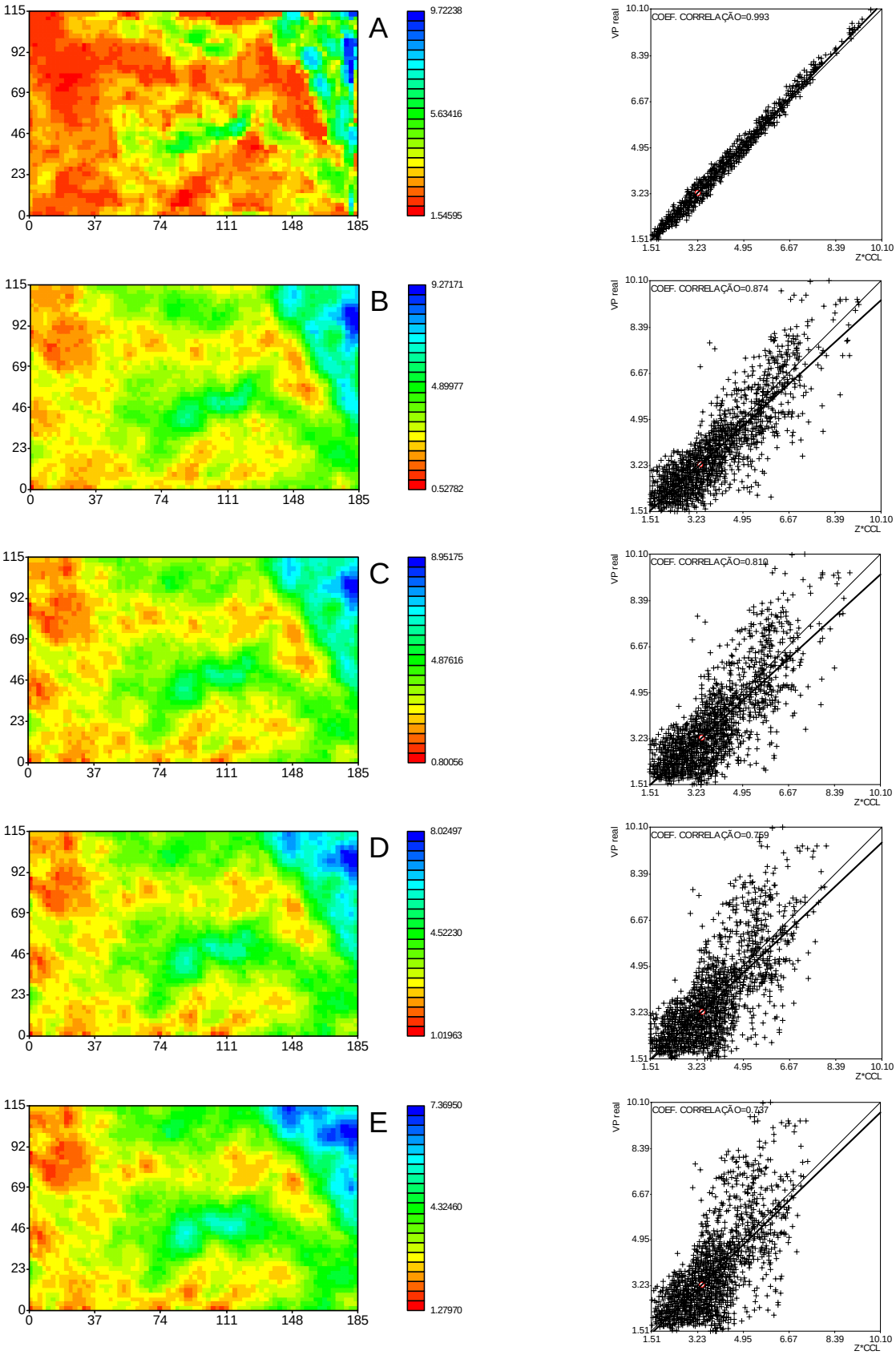

Figura 25: Resultados da cokrigagem colocalizada simples: mapas imagens à esquerda e diagramas de dispersão à direita para amostras com 60 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

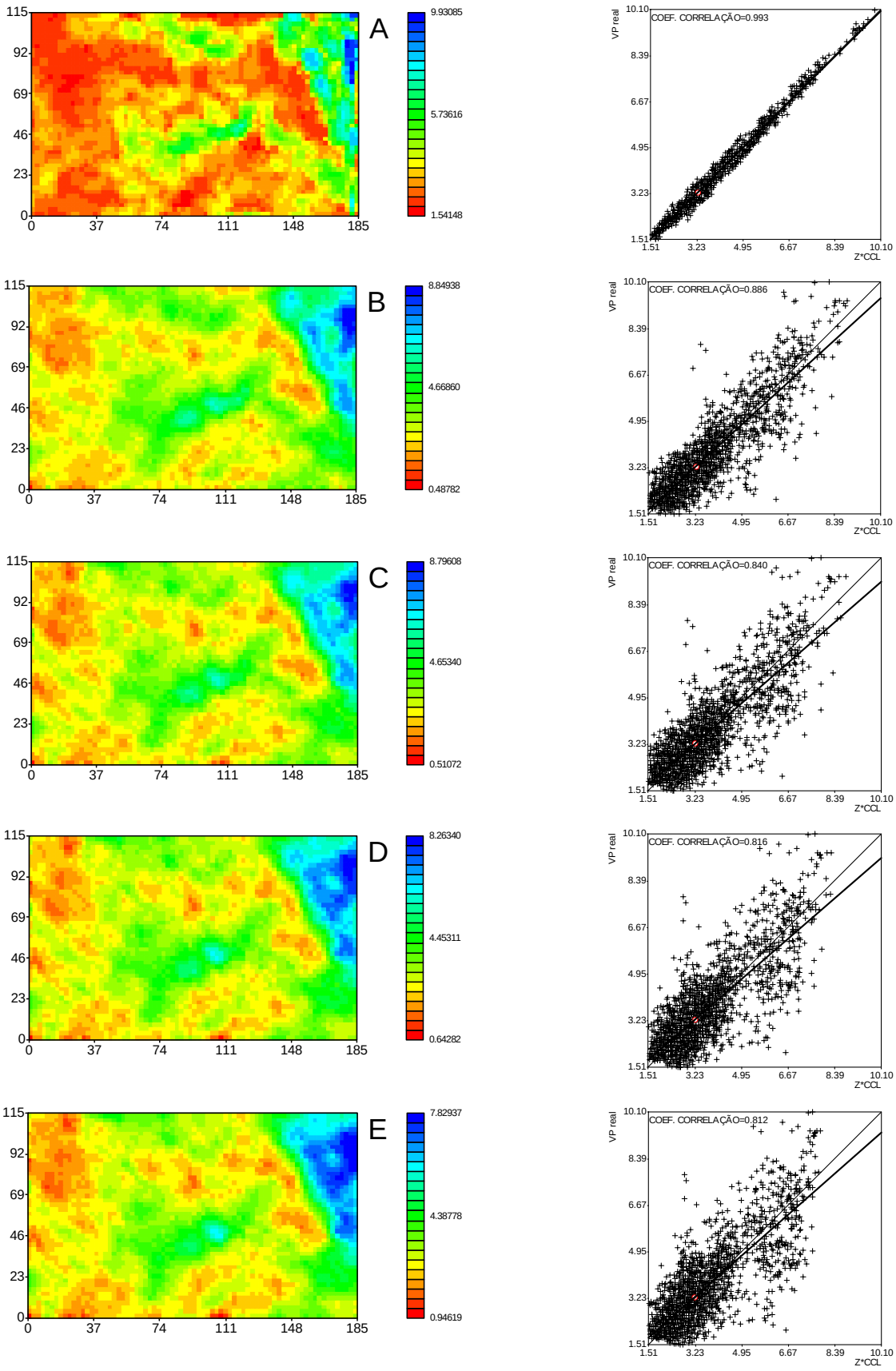

Figura 26: Resultados da cokrigagem colocalizada simples: mapas imagens à esquerda e diagramas de dispersão à direita para amostras com 104 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

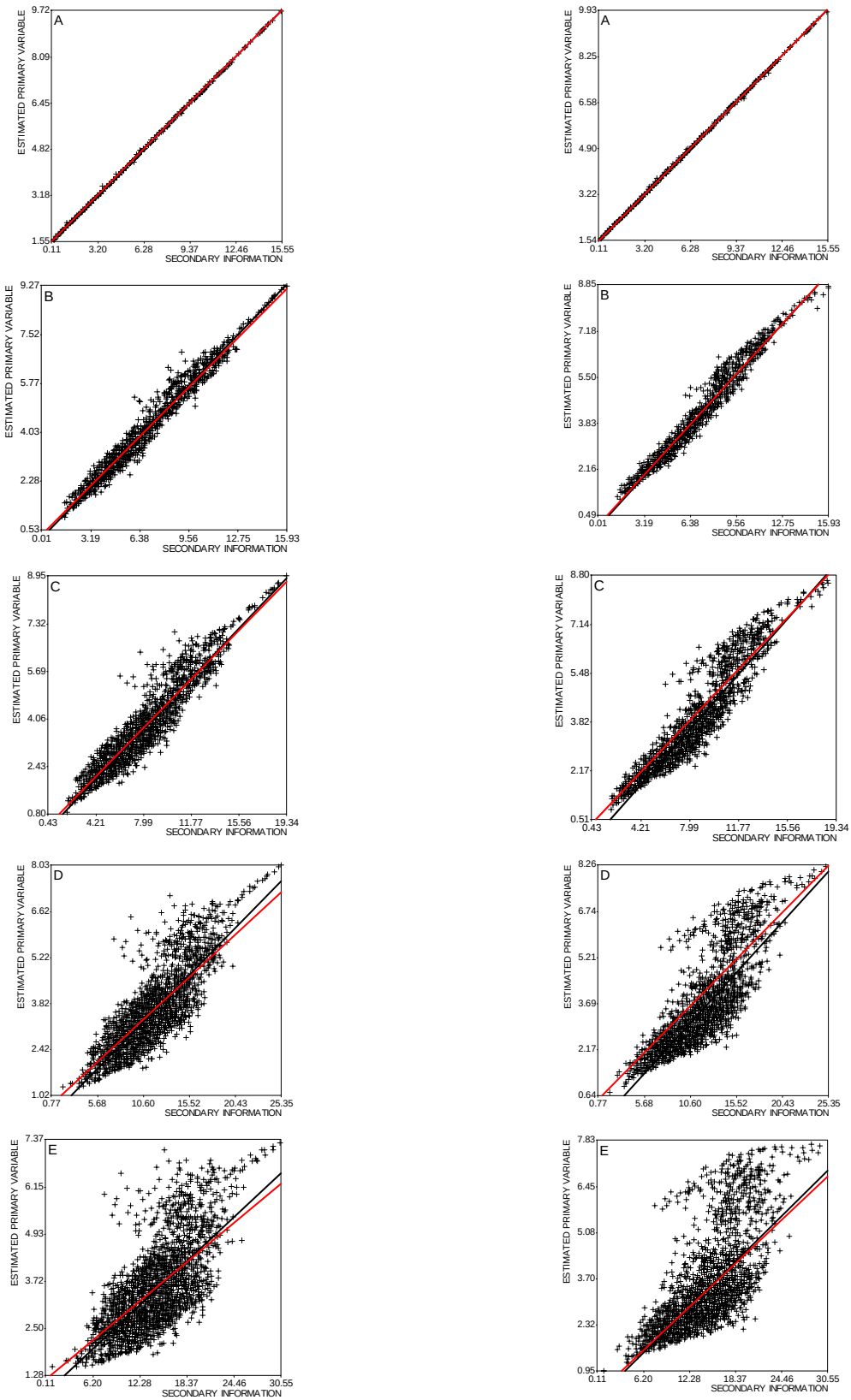

Figura 27: Diagramas de dispersão entre a variável primária estimada pela cokrigagem colocalizada simples e a variável secundária para amostras com 60 unidades à esquerda e com 104 unidades à direita. Linha cheia preta = regressão da estimativa e linha vermelha cheia = regressão amostral. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

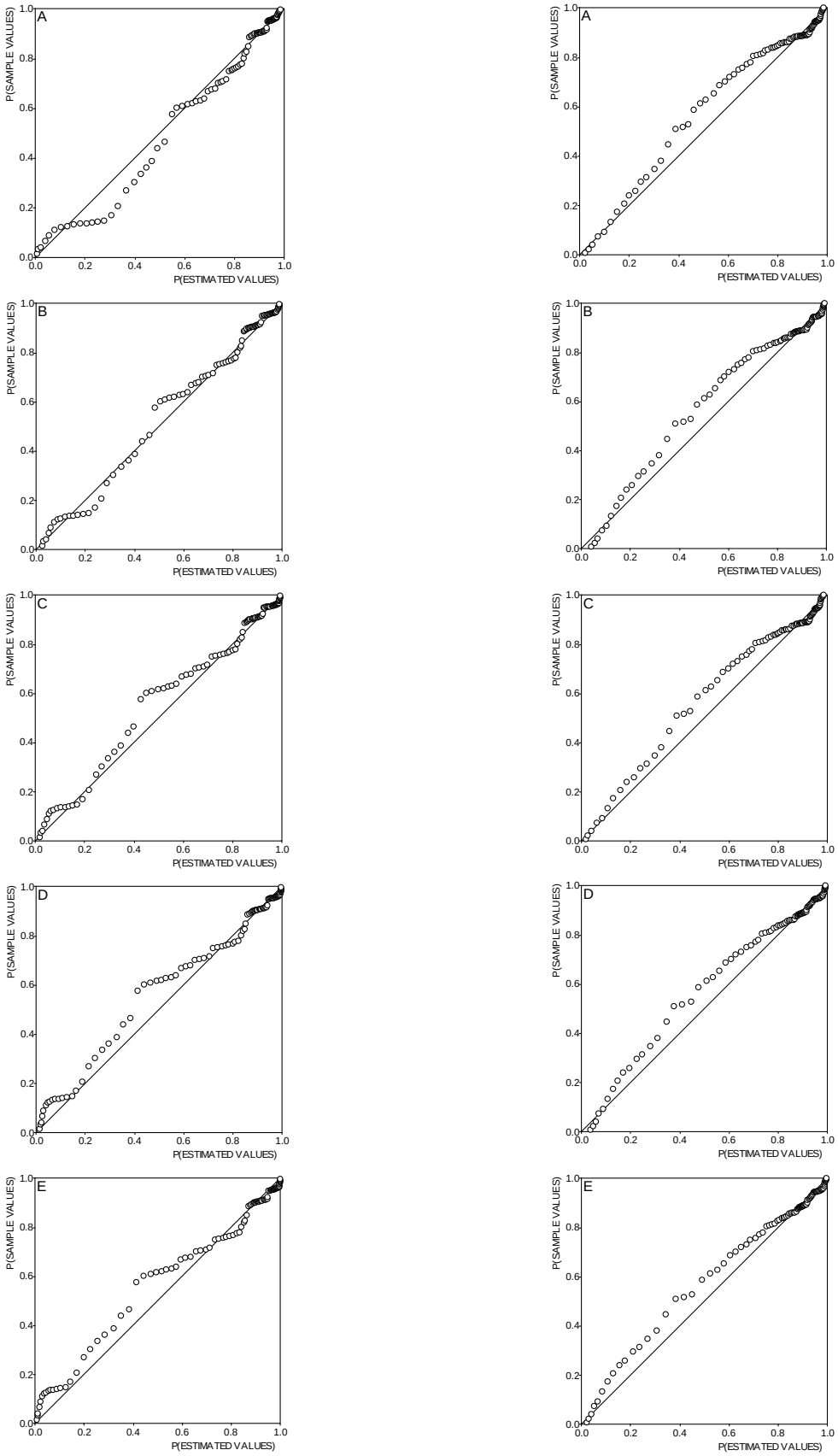

Figura 28: Diagramas P-P comparando a distribuição amostral com a distribuição da variável primária estimada pela cokrigagem colocalizada simples para amostras com 60 unidades à esquerda e com 104 unidades à direita. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

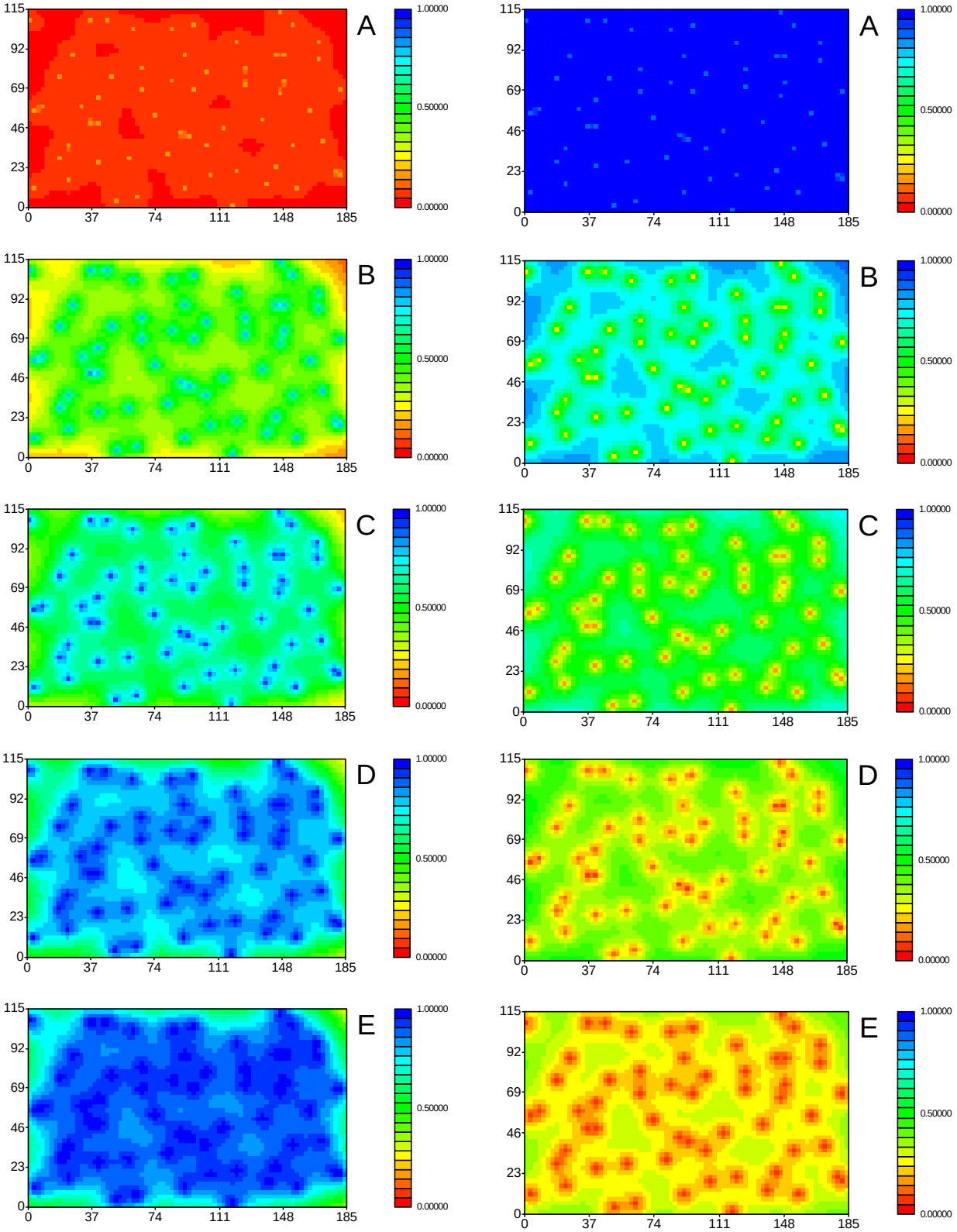

Figura 29: Mapas imagens da soma dos pesos da variável primária (esquerda) e do peso da variável secundária (direita) para amostras com 60 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

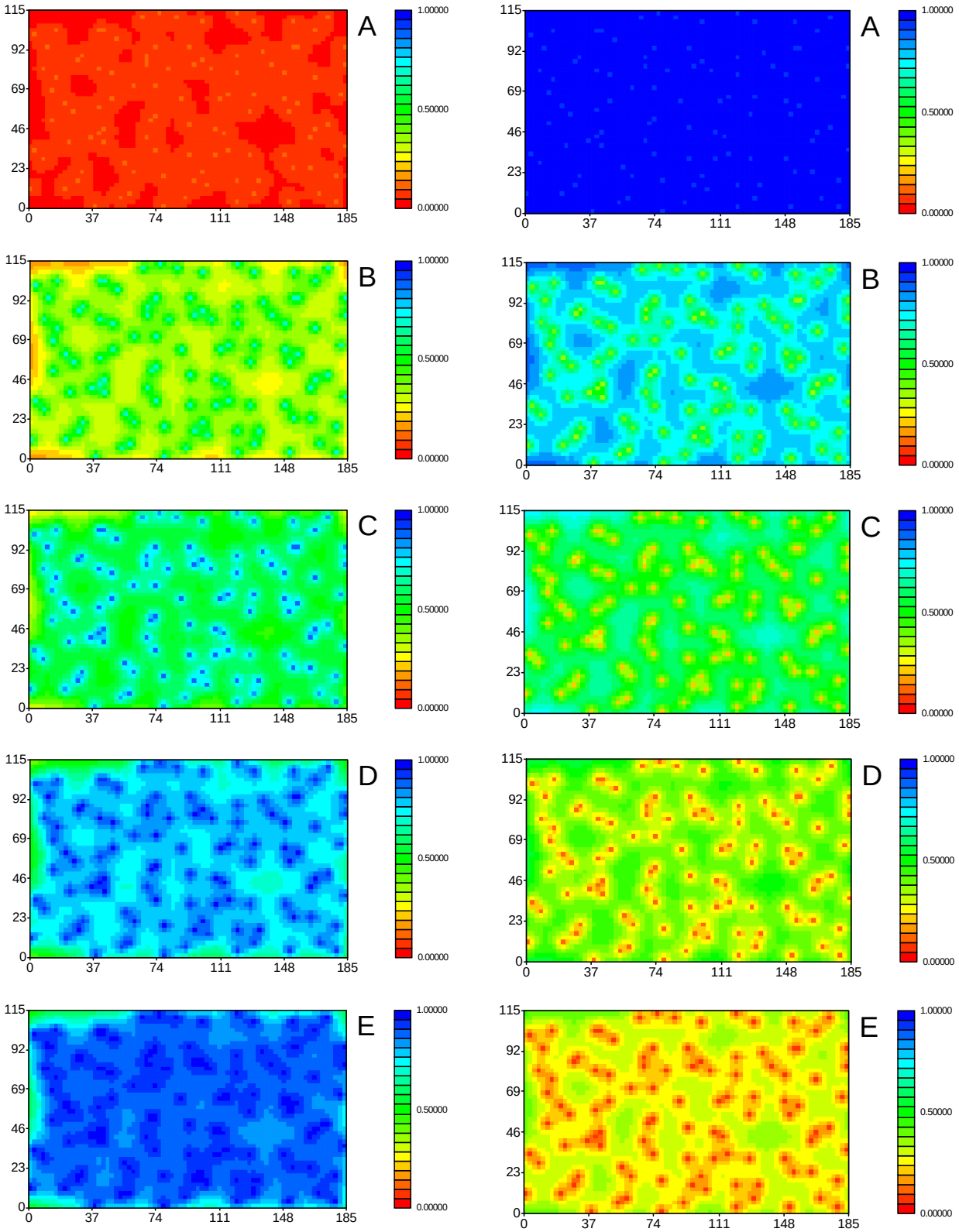

Figura 30: Mapas imagens da soma dos pesos da variável primária (esquerda) e do peso da variável secundária (direita) para amostras com 104 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

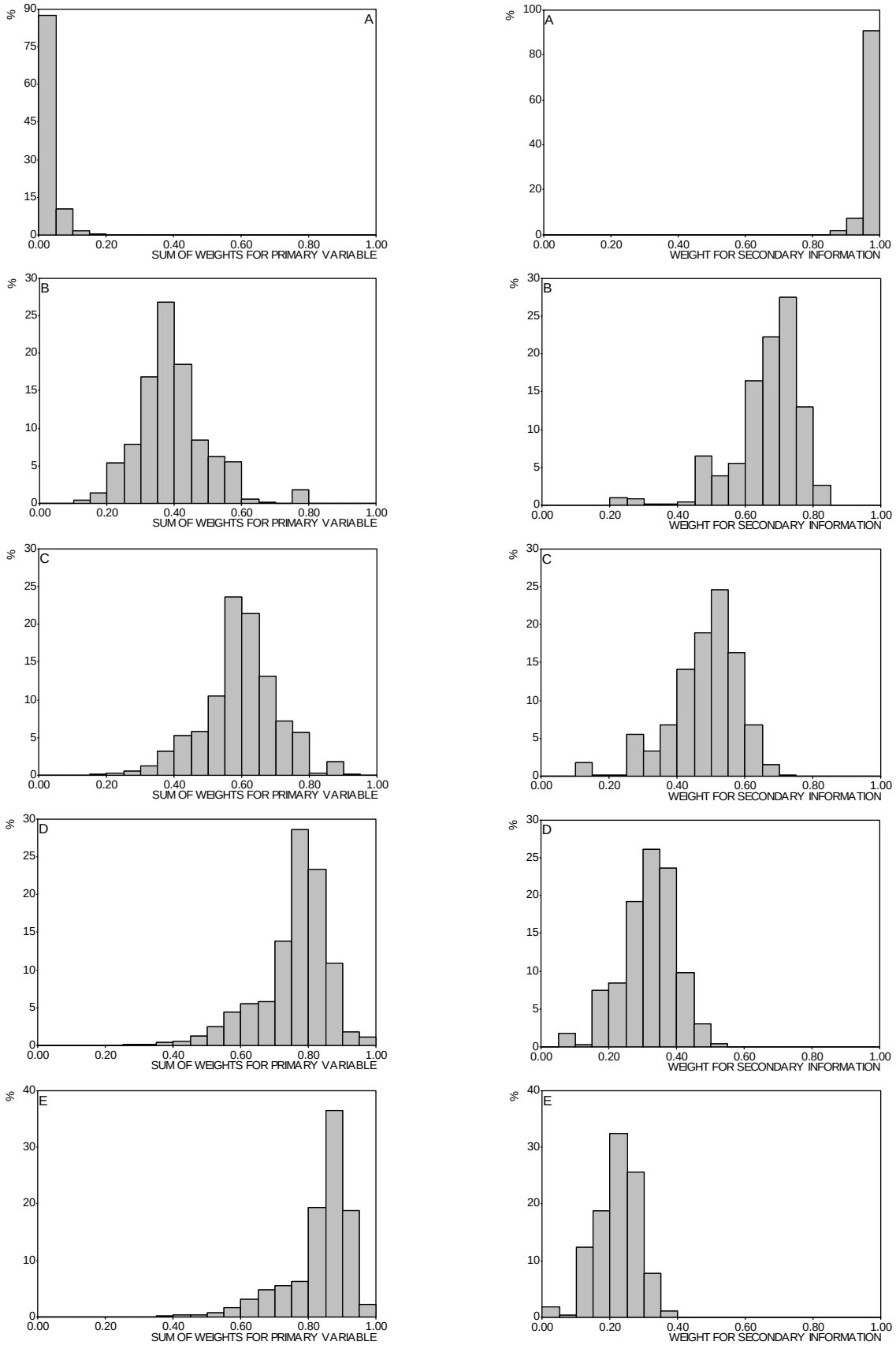

Figura 31: Histogramas da soma dos pesos da variável primária à esquerda e do peso da variável secundária à direita para amostras com 60 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

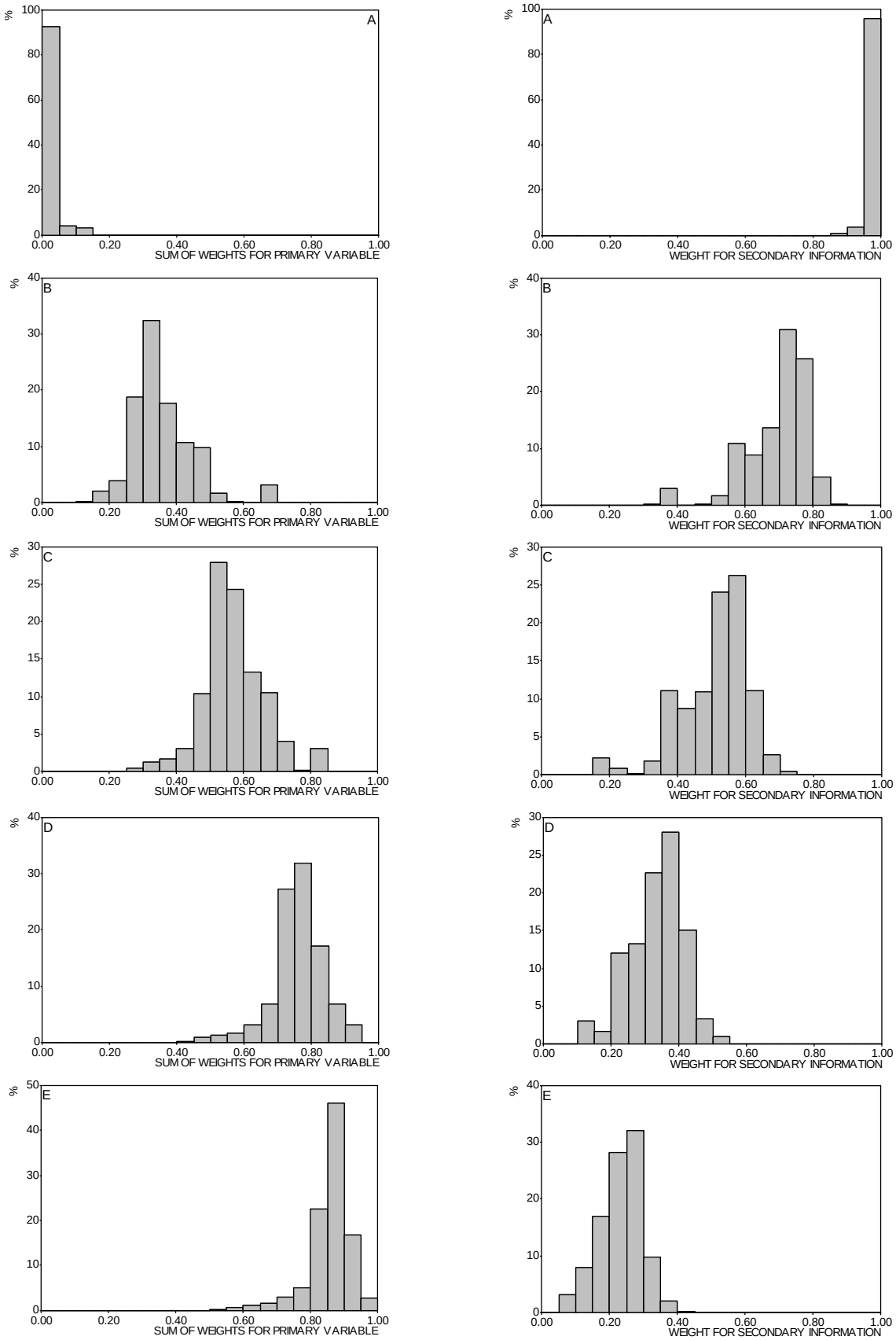

Figura 32: Histogramas da soma dos pesos da variável primária à esquerda e do peso da variável secundária à direita para amostras com 104 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

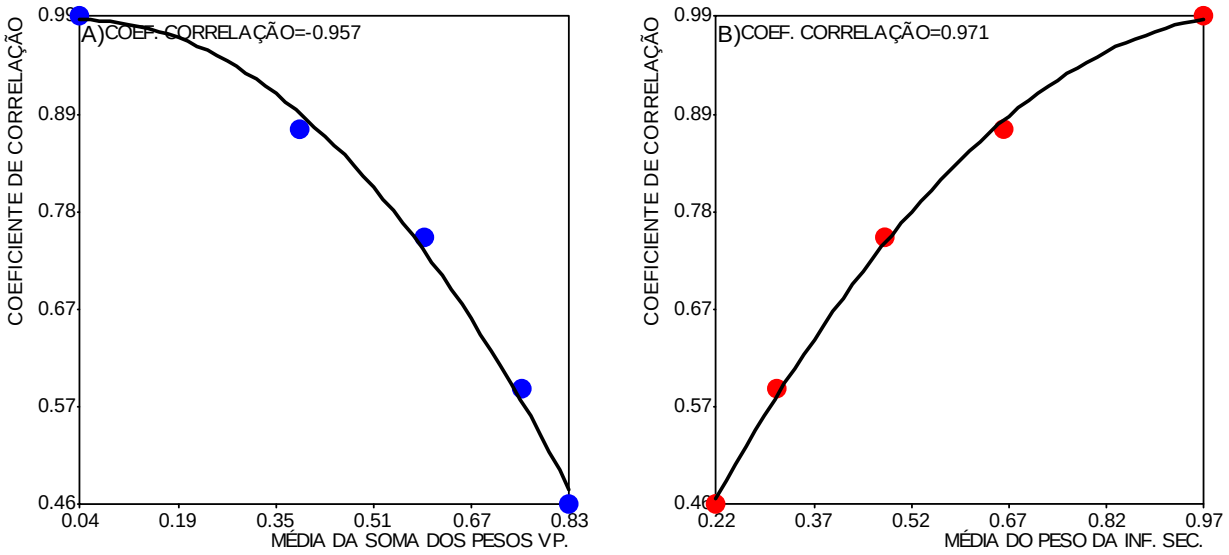

Figura 33: Relação entre o coeficiente de correlação e a média da soma dos pesos da variável primária (A) e com a média do peso da informação secundária (B) para a amostra com 60 unidades.

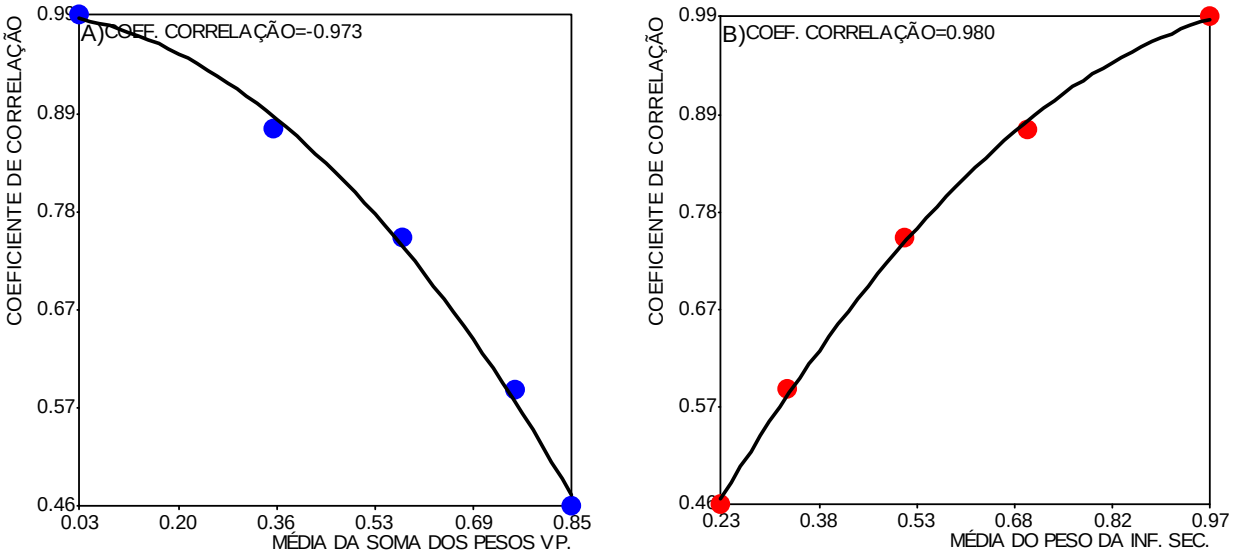

Figura 34: Relação entre o coeficiente de correlação e a média da soma dos pesos da variável primária (A) e com a média do peso da informação secundária (B) para a amostra com 104 unidades.

4.4 Krigagem com deriva externa

A krigagem com deriva externa também faz uso da mesma estrutura de dados que a cokrigagem colocalizada simples. A formulação matemática é diferente da cokrigagem colocalizada, pois os pesos da krigagem com deriva externa são determinados conforme a geometria da função de deriva externa. As vantagens dessa técnica em relação a cokrigagem em geral são inúmeras: a estimativa é baseada tão

somente nos valores da variável primária; possibilidade de cálculo da variância de interpolação e, conseqüentemente, o cálculo da correção do efeito de suavização da krigagem com deriva externa. Inicialmente, são apresentadas as estatísticas para as estimativas com deriva externa, bem como para os valores corrigidos do efeito de suavização da krigagem com deriva externa para a amostra com 60 unidades (Tabelas 10 e 11) e para a amostra com 104 unidades (Tabelas 12 e 13).

Tabela 10: Estatísticas amostrais e dos valores estimados pela krigagem com deriva externa para a amostra com 60 unidades.

Estatística	Am. 60	CC=0,993	CC=0,870	CC=0,752	CC=0,588	CC=0,461
No. dados	60	3404	3404	3404	3404	3404
Média	3,365	3,292	3,358	3,370	3,367	3,368
Desvio	1,262	0,927	0,931	0,893	0,877	0,845
CV	0,401	0,282	0,277	0,265	0,261	0,251
Máx.	7,353	7,204	6,819	7,133	7,165	7,016
QS	3,803	3,615	3,743	3,761	3,792	3,754
Med.	2,949	3,073	3,134	3,182	3,211	3,224
QI	2,562	2,658	2,697	2,729	2,742	2,745
Mín.	1,685	1,753	1,787	1,789	1,782	1,817

Tabela 11: Estatísticas amostrais e dos valores corrigidos do efeito de suavização da krigagem com deriva externa para a amostra com 60 unidades.

Estatística	Am. 60	CC=0,993	CC=0,870	CC=0,752	CC=0,588	CC=0,461
No. dados	60	3404	3404	3404	3404	3404
Média	3,365	3,363	3,365	3,366	3,365	3,366
Desvio	1,262	1,255	1,260	1,261	1,261	1,261
CV	0,401	0,373	0,375	0,375	0,375	0,375
Máx.	7,353	7,353	7,353	7,353	7,305	7,296
QS	3,803	3,802	3,898	3,983	3,964	3,947
Med.	2,949	3,037	3,006	3,001	2,986	3,029
QI	2,562	2,579	2,564	2,547	2,572	2,553
Mín.	1,685	1,781	1,693	1,685	1,685	1,685

Em relação às técnicas anteriores, a krigagem com deriva externa respeita os limites amostrados, pois depende tão somente dos valores da variável primária. A principal diferença entre as estatísticas de uma mesma amostra é a correção do efeito de suavização, onde o desvio padrão das estimativas corrigidas está sempre próximo do desvio padrão amostral.

Observem-se também os mapas imagens obtidos, por krigagem com deriva externa, para as amostras com 60 unidades (Figura 35) e para as amostras com 104 unidades (Figura 36). Invariavelmente, nas imagens corrigidas do efeito de suavização as regiões de valores altos são mais altas e as regiões de valores baixos são mais baixas em relação às imagens da krigagem com deriva externa.

Tabela 12: Estatísticas amostrais e dos valores estimados pela krigagem com deriva externa para a amostra com 104 unidades.

Estatística	Am. 104	CC=0,993	CC=0,870	CC=0,752	CC=0,588	CC=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 13: Estatísticas amostrais e dos valores corrigidos do efeito de suavização da krigagem com deriva externa para a amostra com 104 unidades.

Estatística	Am. 104	CC=0,993	CC=0,870	CC=0,752	CC=0,588	CC=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

Em termos de precisão local, a krigagem com deriva externa só apresenta resultados satisfatórios para casos onde a correlação entre a variável primária e a secundária é alta. Com a diminuição da correlação, há rápida deterioração dos resultados das estimativas via krigagem com deriva externa, como podemos observar nas Figuras 37 e 38, notadamente após a correção do efeito de suavização. Há uma melhora nos níveis de correlação e, portanto, da precisão local com o aumento do tamanho da amostra. Os resultados da cokrigagem colocalizada simples foram muito superiores. Os diagramas de dispersão da Figura 25 são bem melhores que aqueles da Figura 37, acontecendo o mesmo com a Figura 26 quando comparada com a Figura 38. Como mencionado anteriormente, a cokrigagem colocalizada apresenta uma espécie de filtro para dados com baixa correlação e isso melhora significativamente a precisão local das. Por outro lado, a krigagem com deriva externa não tem esse recurso e ao contrário aplica sobre os pesos uma condição que altera os pesos resultantes, conforme:

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

Esta condição de restrição transfere para os pesos da krigagem com deriva externa a pobre correlação existente entre a variável secundária e a variável primária. Assim, a krigagem com deriva externa só pode ser aplicada em condições de alta correlação entre as variáveis.

Com relação à correção do efeito de suavização, os resultados das estimativas corrigidas foram piores para casos com baixa correlação refletindo o problema mencionado anteriormente. Portanto, a correção do efeito de suavização tem efeito positivo somente quando as variáveis encontram-se bem correlacionadas. Observe-se que a perda já ocorre para uma correlação igual a 0,870.

As Figuras 39 e 40 apresentam os diagramas de dispersão entre as variáveis primárias estimadas pela krigagem com deriva externa e a variável secundária. Em geral, a correção do efeito de suavização procura restabelecer a correlação inicial entre a variável primária e a variável secundária observada na amostra, porém isso ocorre com perda de precisão local, o que é inaceitável.

Diagramas P-P da Figura 41 mostram que as estimativas corrigidas do efeito de suavização da krigagem com deriva externa se aproximam bastante da distribuição amostral. Contudo, isso é insuficiente, pois há perda de precisão local na origem das estimativas pela krigagem com deriva externa. Esse resultado apenas confirma que o algoritmo de pós-processamento proposto por Yamamoto (2005) funciona, na reprodução do histograma amostral. Evidentemente, esse algoritmo não pode ser aplicado para correção do efeito de suavização da krigagem com deriva externa. Pelos resultados obtidos até o momento, o melhor método de co-estimativa geoestatística para arranjos especiais de dados (nós da malha regular com a informação secundária) é o da cokrigagem colocalizada.

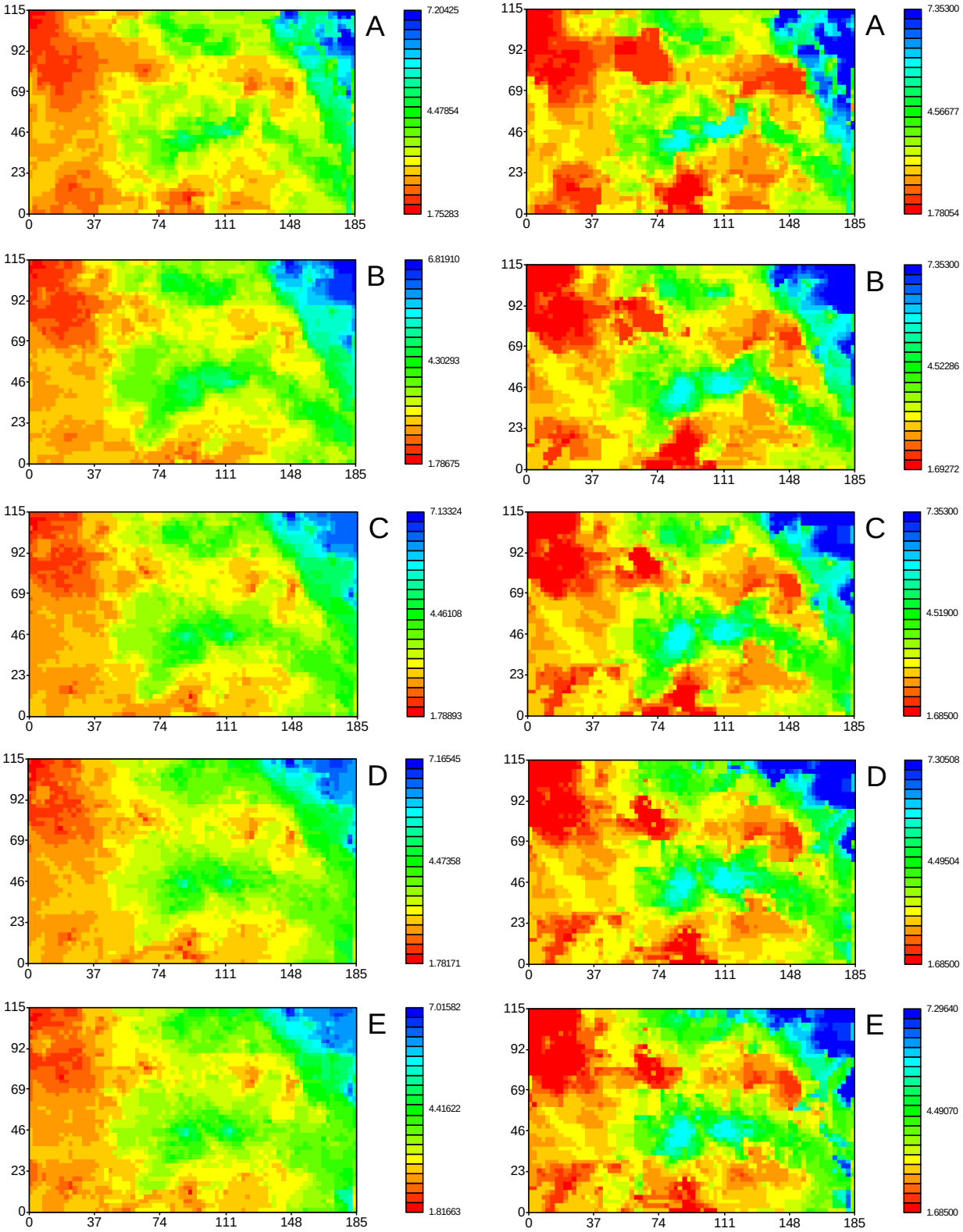

Figura 35: Mapas imagens dos valores estimados pela krigagem com deriva externa (esquerda) e dos valores corrigidos do efeito de suavização (direita) para amostras com 60 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

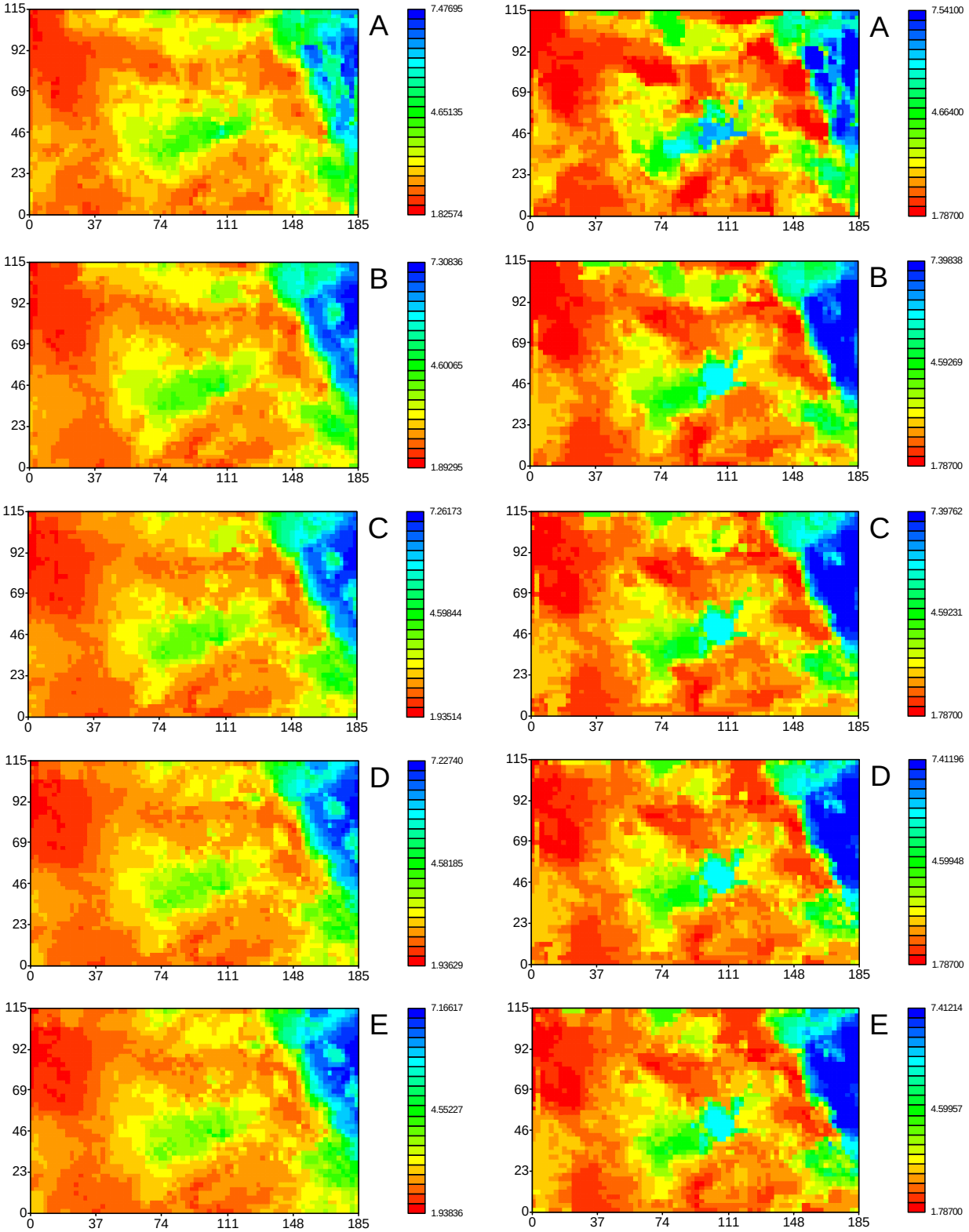

Figura 36: Mapas imagens dos valores estimados pela krigagem com deriva externa (esquerda) e dos valores corrigidos do efeito de suavização (direita) para amostras com 104 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

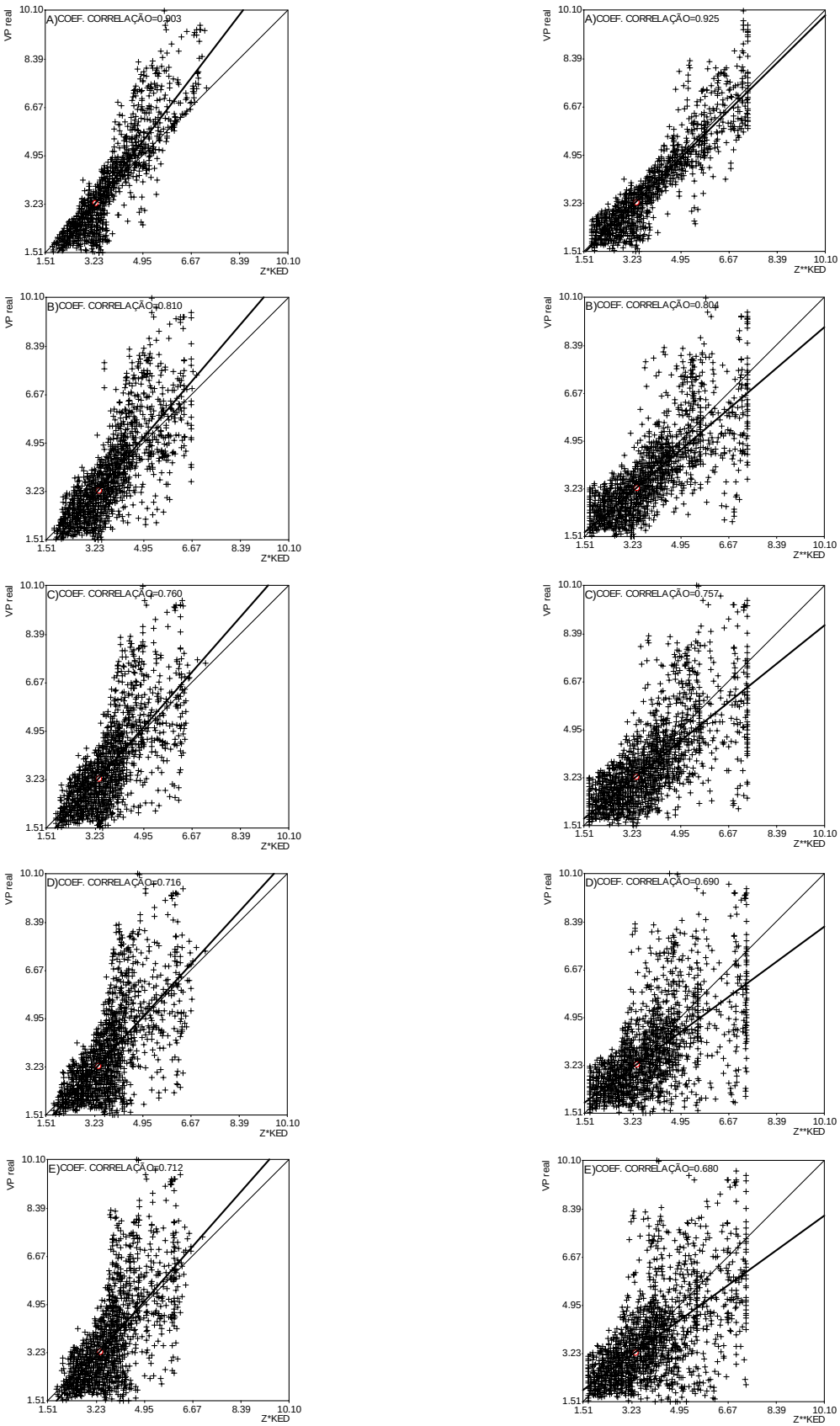

Figura 37: Diagramas de dispersão entre valores reais e valores estimados pela krigagem com deriva externa (esquerda) e dos valores corrigidos do efeito de suavização (direita) para amostras com 60 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

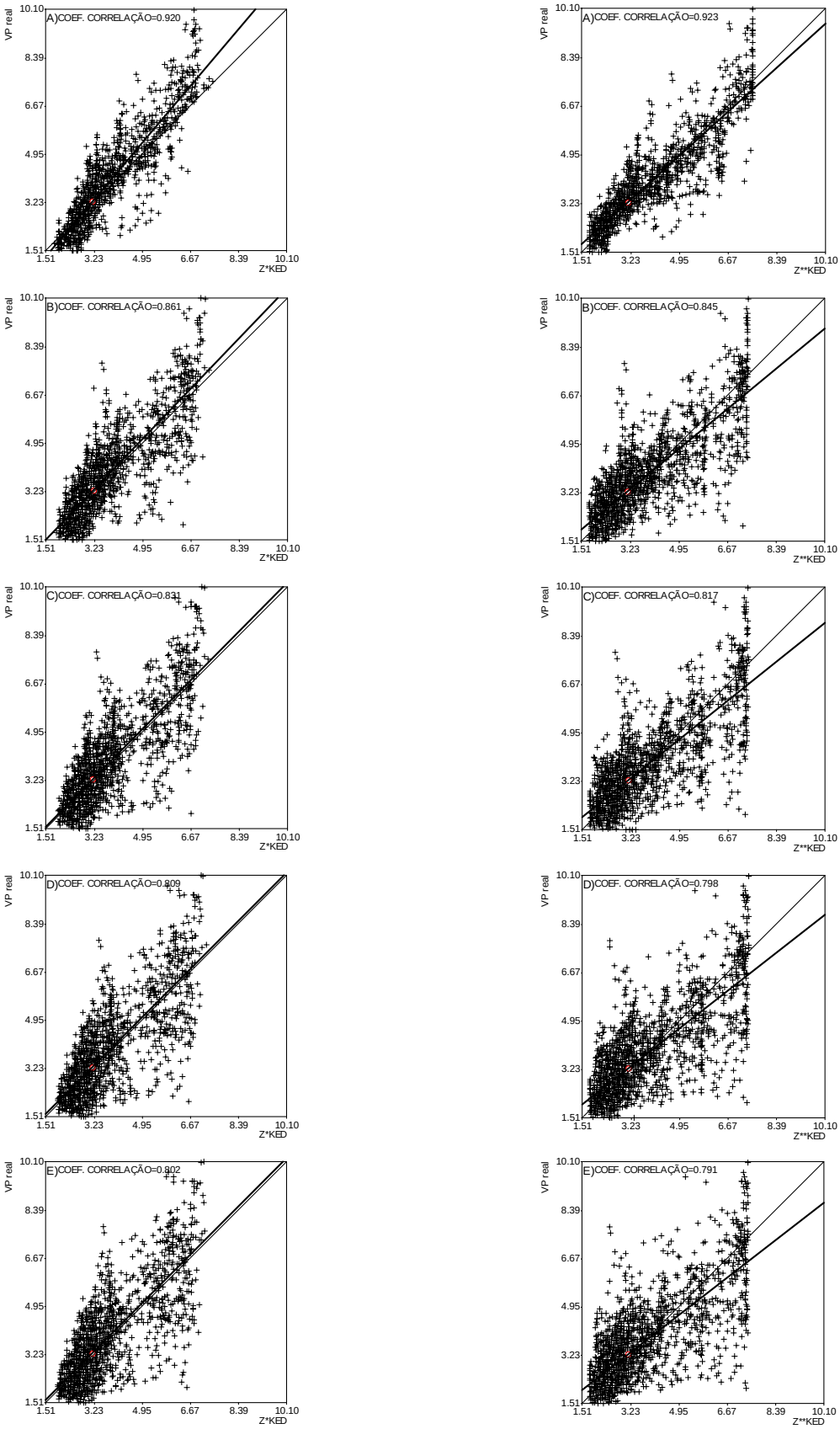

Figura 38: Diagramas de dispersão entre valores reais e valores estimados pela krigagem com deriva externa (esquerda) e dos valores corrigidos do efeito de suavização (direita) para amostras com 104 unidades. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

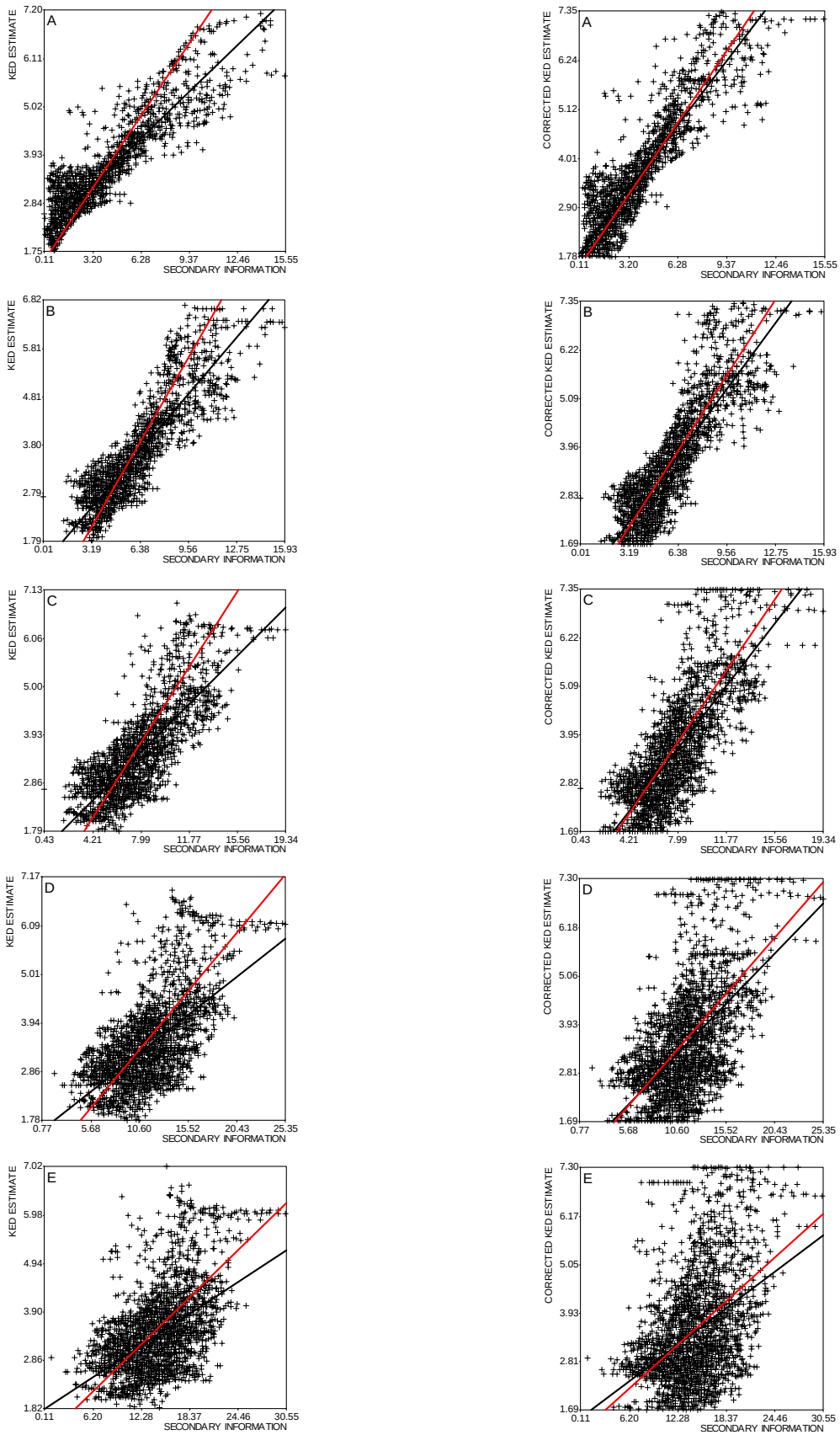

Figura 39: Diagramas de dispersão entre a variável primária estimada pela krigagem com deriva externa (esquerda) e os mesmo para as estimativas corrigidas (direita) para amostras com 60 unidades. Linha cheia preta = regressão da estimativa e linha vermelha cheia = regressão amostral. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

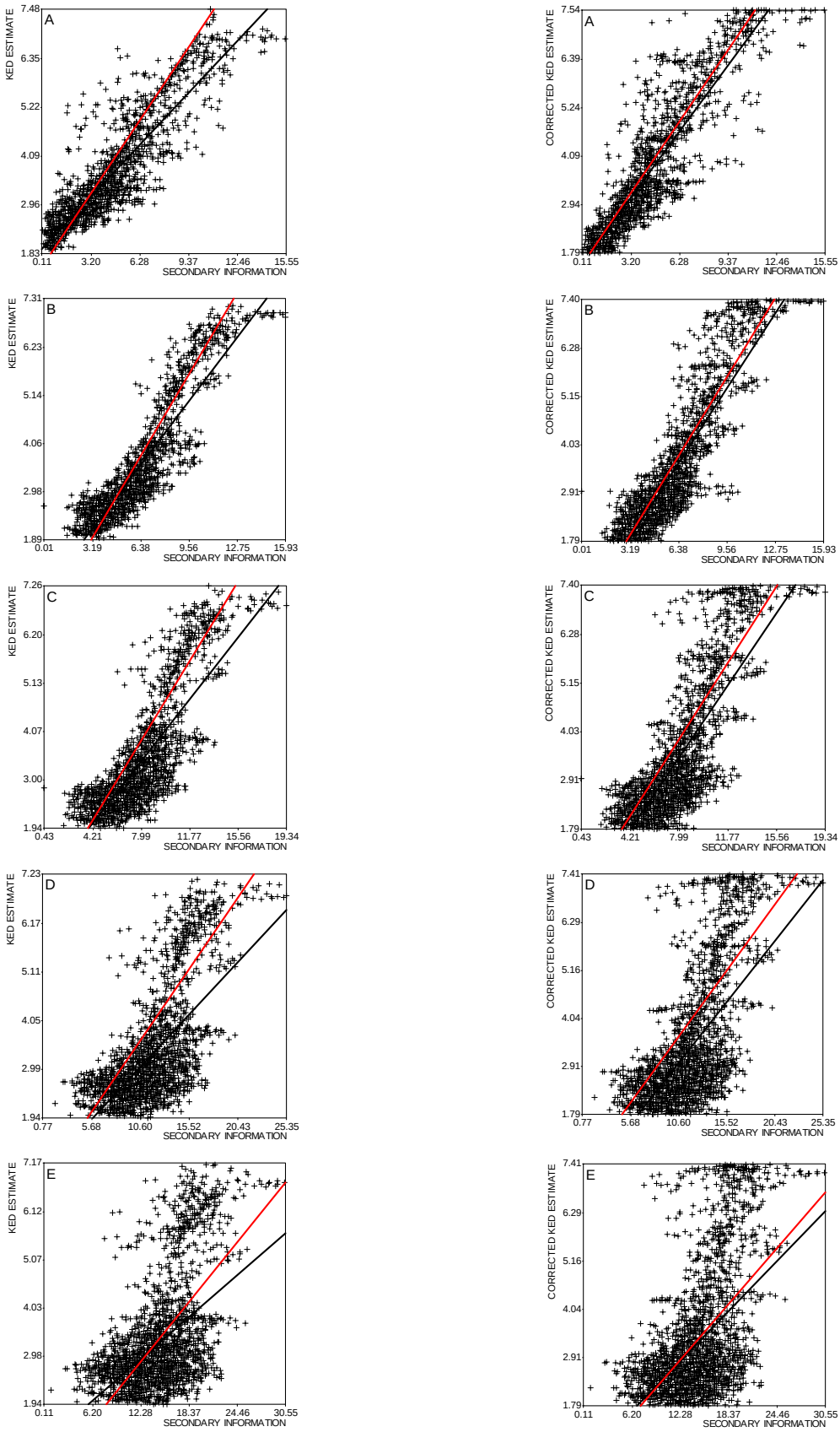

Figura 40: Diagramas de dispersão entre a variável primária estimada pela krigagem com deriva externa (esquerda) e os mesmo para as estimativas corrigidas (direita) para amostras com 104 unidades. Linha cheia preta = regressão da estimativa e linha vermelha cheia = regressão amostral. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461.

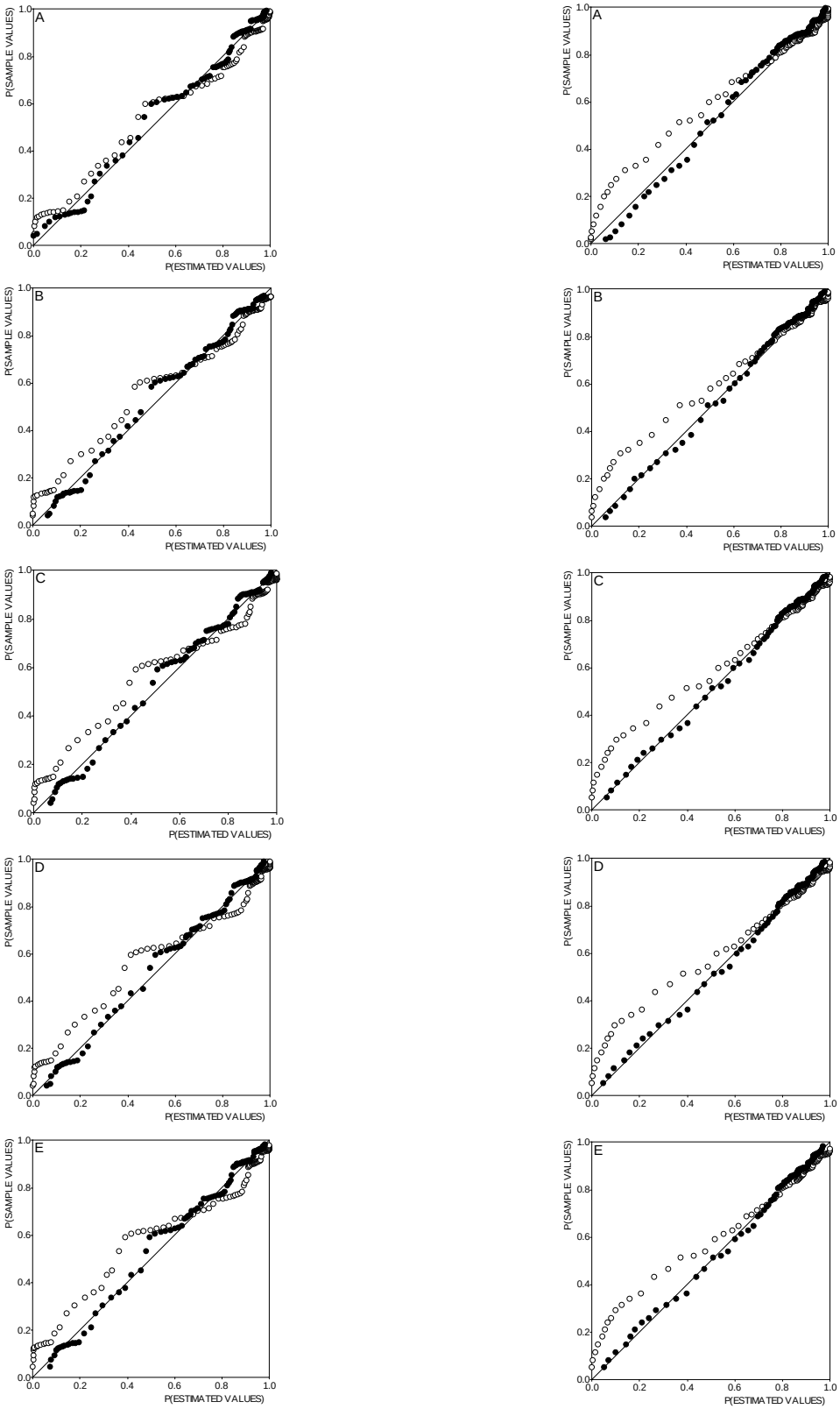

Figura 41: Diagramas P-P comparando a distribuição amostral com a distribuição da variável primária estimada pela krigagem com deriva externa para amostras com 60 unidades à esquerda e com 104 unidades à direita. A=0,993; B=0,870; C=0,752; D=0,588; E=0,461. Círculos vazios=krigagem com deriva externa; círculos cheios=estimativas corrigidas do efeito de suavização.

CAPÍTULO 5

CONSIDERAÇÕES FINAIS

Neste trabalho, foram testados diversos métodos de co-estimativas: cokrigagem ordinária, cokrigagem colocalizada simples e krigagem com deriva externa. Além desses métodos de co-estimativas, foi considerado também o método da krigagem ordinária, apenas a título de comparação com os demais métodos. O objetivo principal desta pesquisa era justamente o de verificar a influência da correlação nos resultados das co-estimativas geoestatísticas, bem como a precisão local em função do tamanho da amostra. Assim, o ponto de partida para o trabalho foi à geração de dados sintéticos tanto para a variável primária como para a variável secundária. Além disso, a variável secundária foi gerada contemplando cinco níveis de correlação com a variável primária, quais sejam: 0,993; 0,870; 0,752; 0,588 e 0,461. Embora correlações mais baixas tivessem sido geradas, elas foram abandonadas devido à dificuldade de calcular e modelar variogramas cruzados para fins de aplicação da cokrigagem ordinária. Assim, foram geradas duas amostras contendo 60 e 104 pontos amostrais, as quais correspondem a 1,76 e 3,06% dos conjuntos completos.

A krigagem ordinária não apresentou resultados excepcionais em termos de reprodução da textura geológica, haja vista não ter utilizado da informação secundária disponível.

No caso da cokrigagem ordinária, tentou-se reproduzir uma situação próxima para comparação com outros métodos de co-estimativa. Assim, foram geradas amostras contendo 33% da informação secundária, ou seja, com 1123 pontos contendo valores da variável primária. Essa quantidade foi excessiva trazendo conseqüências na resolução dos sistemas de equações cokrigagem ordinária, vários dos quais apresentaram problemas de instabilidade. Os diagramas de dispersão mostraram que os resultados apresentam problemas de precisão local, pois as retas de regressão estão distantes das retas bissetrizes. O principal problema verificado neste estudo com respeito a esta técnica se refere aos conjuntos de dados utilizados que apresentaram excessiva quantidade de informação secundária. Conseqüentemente, os valores estimados tiveram grande influência da informação secundária fazendo com que

apresentassem boa correlação, ao contrário do que seria de esperar, visto que a variável secundária foi contaminada progressivamente com ruído branco.

A cokrigagem colocalizada simples que utiliza no estimador a informação de um único valor da variável secundária, qual seja aquela existente sobre o ponto que se está estimando, foi à técnica que apresentou o melhor resultado. É preciso lembrar que ainda existe o problema da influência da variável secundária na estimativa resultante, mas em muito menor escala que a técnica da cokrigagem ordinária clássica. O que mais impressionou nesta técnica foi à filtragem da informação secundária através do peso que é diretamente proporcional à correlação entre a variável primária e a variável secundária, sendo que a relação funcional é descrita por uma equação de segundo grau. A descoberta dessa relação entre a média do peso da variável secundária e o grau de correlação entre a variável primária e a variável secundária é sem dúvida o resultado de maior relevância desta pesquisa. Tendo em vista que a informação secundária utilizada é proporcional ao valor do coeficiente de correlação, as estimativas através da cokrigagem colocalizada simples apresentam boa precisão local. Além disso, elas reproduzem bastante bem as correlações iniciais. O único problema verificado até o momento é a falta de aderência do histograma resultante com o histograma amostral. Este é um assunto para uma pesquisa futura envolvendo o pós-processamento das estimativas da cokrigagem colocalizada visando à reprodução do histograma amostral.

A krigagem com deriva externa que tem uma formulação muito interessante, pois não leva em consideração o valor da variável secundária, mas apenas a conformação geométrica descrita pela variável secundária, na vizinhança do ponto estimado, foi a que apresentou um desempenho baixo. Os resultados foram bons apenas para a situação de alto grau de correlação. Os pesos da krigagem com deriva externa são calculados com base em uma segunda condição de restrição, que leva em consideração a geometria descrita pela variável secundaria. Assim, se esta variável está contaminada com ruído branco, fazendo com que a correlação com a variável primária se deteriore, influenciará diretamente no resultado da krigagem com deriva externa. Além disso, como mencionado no texto, a krigagem com deriva externa apresenta a possibilidade de cálculo da variância de interpolação e, portanto, é passível de correção do efeito de suavização inerente ao método. A correção do efeito de suavização foi efetiva em todos os conjuntos de dados amostrais, mostrando que a técnica é efetiva para reprodução do histograma amostral. Mas, como a estimativa já

vinha contaminada pela baixa correlação apresentada pela informação secundária, o pós-processamento resultou em valores ainda mais enviesados em relação aos valores reais do conjunto completo. Em outras palavras, houve perda de precisão local na estimativa inicial e esta perda foi amplificada no pós-processamento. Portanto, a técnica da krigagem com deriva externa, embora apresente formulação elegante do ponto de vista geoestatístico, tem aplicação restrita.

Com relação ao tamanho da amostra, os resultados, obtidos em todos os métodos aplicados, mostraram uma pequena melhora na precisão local. Assim, conclui-se que é possível utilizar uma variável primária pobremente amostrada e fazer uso da correlação com outra variável auxiliar que apresente alguma correlação.

Finalmente, pelos resultados apresentados nesta pesquisa, a cokrigagem colocalizada simples é a técnica geoestatística de co-estimativa robusta e que pode ser aplicada em qualquer situação real, com alta ou baixa correlação. Recomenda-se investir no aperfeiçoamento desta técnica, principalmente em relação à reprodução do histograma amostral.

REFERÊNCIAS BIBLIOGRÁFICAS

Boezio, M.N.M.; Costa, J.F.C.L.; Koppe, J.C. 2006. Cokrigagem colocada aplicada ao mapeamento do nível de água subterrânea. Revista da Escola de Minas 59(2):159 – 164.

Bourennane, H. & King, D. 2003. Using multiple externals drifts to estimate a soil variable. Geoderma 114, 1– 18.

Bourennane, H.; King, D.; Couturier, A. 2000. Comparison of kriging with external drift and simple linear regression for predicting soil horizon thickness with different sample densities. Geoderma 97 (3– 4), 255– 271.

Bourennane, H.; King, D.; Chery, P.; Bruand, A.. 1996. Improving the kriging of a soil variable using slope gradient as external drift. European Journal of Soil Science 47 (4), 473–483.

Chilès, J-P. & Delfiner, P. 1999. Geostatistics: Modeling Spatial Uncertainty: John Wiley & Sons, Inc. New York, 1st. Ed., 695p.

Clark I., Basinger K. L., Harper, W. V. 1987. MUCK-A Novel Approach to Co-Kriging. Proceedings of the Conference on Geostatistical, Sensitivity, and Uncertainty Methods for Ground-Water Flow and Radionuclide Transport Modeling.. São Francisco, California 21 p

Conde, R.P. 2000. Geoestatística Aplicada a Avaliação de Reservas e Controle de Lavra na Mina de Cana Brava (GO). Tese de Doutoramento apresentada no Instituto de Geociências – USP. 162p.

Deutsch C.V. & Journel, A.G., 1992. GSLIB: Geostatistical Software Library and User's Guide. Oxford University Press. 1st. Ed. 340p.

Doyen, P.M. 1988. Porosity from seismic data: A geostatistical approach. Geophysics 53: 1263 – 1275.

Goovaerts, P. 1997. Geostatistics for natural resources evaluation: Oxford University Press, New York, 483p.

Isaaks, E.H. & Srivastava, R.M. 1989. An introduction to applied geostatistics. New York, Oxford University Press. 561p.

Journel, A.G. & Huijbregts, C.J, 1978. Mining geostatistics: Academic Press. 600p.

Matheron, G. 1965 – Les Principes de la Geostatistique, École Nationale Supérieure des Mines de Paris. Fontainebleau, França. Rapport 65. 26p

McBratney, A.B.; Odeh, I.O.A.; Bishop, T.F.A.; Dunbar, M.S.; Shatar, T.M. 2000. An overview of pedometric techniques for use in soil survey. Geoderma 97:293 – 327.

Myers, D.E., 1991. Pseudo Cross-Variograms, Positive Definiteness and Cokriging, Mathemathical Geology, 23:805-816.

Myers, D.E. 1982. Matrix formulation of Co-Kriging. Mathemathical Geology, 14:249 – 257.

Journel, A.G. & Rao, S.E. 1996. Deriving conditional distribuitions from ordinary kriging: Stanford Certer for Reservoir Forecasting, Stanford, 25p. (Report #9).

Olea R. 1991. *Geostatistical Glossary and Multilingual Dictionary*. 1ª ed. Oxford, Oxford University Press, 175p.

Olea R. 1999. *Geostatistics for Engineers and Earth Scientists*. 1ª ed. Massachusetts, Kluwer Academic Publishers, 303 p.

Pardo-Igúzquiza, E., Chica-Olmo, M. Atkinson, P.M. 2006. Downscaling cokriging for image sharpening. Remote Sensing of Enviroment 102:86 – 98.

Wackernagel, H. 1998. Multivariate Geostatistics: Springer-Verlag, Berlin, 2nd. Ed. 291p.

Wackernagel, H. 2003. Multivariate Geostatistics: Springer-Verlag, Berlin, 3nd. Ed. 387p.

Yamamoto, J.K. 2007. On unbiased backtransform of lognormal kriging estimates. Computational Geosciences,

Yamamoto, J.K. 2005. Correcting the Smoothing Effect of Ordinary Kriging Estimates. Mathemathical Geology, 32:69-94

Yamamoto, J.K. 2000. An Alternative Measure of the Reliability of Ordinary Kriging Estimates, Mathemathical Geology, 32: 489-509.

Xu, W.; Tran, T.T.; Srivastava, R.M.; Journel, A.G. 1992. Integrating Seismic Data in Reservoir Modelling: The Collocated Cokriging alternative in: Annual Technical Conference of the Society of Petroleum Engineers, 67, Proceedings. Washington, paper SPE 24741, 833 – 842.