

A Variação do Suporte Amostral e as Mudanças na Estrutura do Semivariograma

CASSIO F. P. DE ALMEIDA¹, LUIS PAULO VIEIRA BRAGA², CLAUDIO BETTINI³

¹ENCE—Escola Nacional de Ciências Estatísticas - cassiofreitas@ibge.gov.br tamandoa@bol.com.br

²UFRJ - braga@ufrj.br

³Departamento de Geologia - UFRJ - bettini@acd.ufrj.br

Abstract. Na exploração e desenvolvimento de campos de petróleo, um procedimento básico de amostragem, que consiste na perfuração de poços, é muito caro e por isso muito limitado. Assim, o modelo numérico representando as condições geológicas do campo petrolífero terá regiões de grande incerteza devido ao pequeno número de poços. Para atenuar estas incertezas, complementam-se os dados de poços com dados sísmicos, que são menos precisos, porém de ampla distribuição e menos onerosos. Os dados espaciais com que trabalha a Geoestatística nunca têm suas medidas referentes a um ponto, mas sim a um elemento do domínio, denominado *suporte*. Dados de poços e de sísmica, que têm suportes originais distintos, poderão ser combinados através de técnicas geoestatísticas (estimação ou simulação) e, para isso, necessita-se reduzi-los a um mesmo suporte. Esta mudança de suporte, tanto o aumento quanto a redução, para variáveis escalares, como é o caso da porosidade, acarreta uma alteração na estrutura da variabilidade espacial, conseqüentemente nos parâmetros do semivariograma. No procedimento de aumento de escala, é possível calcular a redução da variabilidade (patamar) e estimar a expansão do alcance. Na redução do suporte, a estimação dos parâmetros é ainda um problema aberto. Neste trabalho, avaliaram-se os efeitos da mudança de suporte, restritos a dados de porosidade, admitindo um comportamento espacial isotrópico, ajustável por um modelo variográfico esférico. São propostos dois métodos para predição do modelo de variograma sob a mudança de suporte, nos quais estimam-se todos os parâmetros do modelo de semivariância no procedimento do aumento e redução de escala. Obtiveram-se bons resultados, quando comparados com os métodos mais comuns na literatura, principalmente quando a discrepância entre as escalas aumenta.

1 Introdução

O presente trabalho trata de uma aplicação importante da Geoestatística para o desenvolvimento de campos petrolíferos. A Geoestatística é um conjunto de técnicas estatísticas orientadas para a estimação e simulação de variáveis aleatórias com distribuição espacial, definidas sobre em subconjuntos do \mathfrak{R}^n ; nas situações mais comuns, $n = 1, 2, 3, 4$. Uma das medidas típicas da Geoestatística, talvez a mais importante e difundida, é o semivariograma, informalmente denominado variograma, uma função que mede o grau de dissimilaridade de uma mesma propriedade em dois pontos amostrais, à medida que varia a interdistância e orientação do par de observações.

Os dados espaciais com que trabalha a Geoestatística sempre têm suas medidas referentes a um elemento do domínio. Cada amostra em uma determinada localização tem seu valor medido com referência a um volume, área ou comprimento, que chamaremos *suporte* ou *suporte amostral*. Esta característica não representará qualquer dificuldade, se todas as amostras representarem o mesmo elemento do domínio, ou suporte, e se este for compatível com o objetivo da análise

(CHILÈS [4]). Muitas vezes, o suporte (comprimento, área ou volume) representado pelas amostras é muito pequeno em relação a um outro suporte desejado para se proceder a análise, chamado suporte de referência. Nesse caso, pode-se aproximar o suporte amostral a um ponto, chamando-o de suporte pontual, para efeito de modelagem.

A literatura recente documenta uma prática que se tornou um padrão da indústria do petróleo: logo após a descoberta e a delimitação de um novo campo de petróleo, passa-se por uma etapa de simulação, antes de se perfurar os chamados poços de desenvolvimento. A perfuração de poços e os equipamentos de produção são extremamente onerosos. Assim, a indústria busca meios e métodos para minimizar o número de poços e dispô-los espacialmente de modo que otimizem a produção de petróleo com um custo mínimo.

Antes de pôr em prática um plano de desenvolvimento, que inclui poços produtores de petróleo e injetores (de água, gás, vapor), as equipes responsáveis executam inúmeras simulações do reservatório em duas etapas.

Na primeira etapa, em que se simula o estado da

natureza, necessita-se de muitas informações, tais como: tipo de rocha, porosidade, permeabilidade, saturações de fluidos e outras. Estas informações são as mais precisas, ou seja, provêm de amostras com o menor suporte possível mas estão disponíveis em um número limitado. Consequentemente, nesta primeira etapa de simulação, o modelo numérico representando as condições geológicas do campo petrolífero terá regiões de grande incerteza, devido ao pequeno número de poços. Esta incerteza pode ser atenuada, complementando-se os dados de poços com *dados sísmicos*¹, que são menos precisos, porém de ampla distribuição e menos onerosos.

A segunda etapa, que simula o fluxo, envolve um intenso e complexo processamento iterativo, o que implica uma grande demanda computacional. Isto força a utilização de dados com suporte maior, a fim de reduzir a quantidade de informação a ser processada, viabilizando a simulação. Daí, mais uma vez, a necessidade da mudança de suporte, de que trata o presente trabalho.

Esta mudança de suporte, tanto o aumento quanto a redução, para variáveis escalares, acarreta uma alteração na estrutura da variabilidade espacial, conseqüentemente nos parâmetros do semivariograma.

Os objetivos deste trabalho são: (1) demonstrar as alterações dos parâmetros do variograma em resposta ao aumento do suporte amostral e (2) propor um método alternativo para estimação dos parâmetros do variograma em resposta ao aumento e à diminuição do suporte amostral.

Este estudo estará restrito a dados de porosidade, admitindo um comportamento espacial isotrópico, ajustável por um modelo variográfico esférico². Propriedades tensoriais, como a permeabilidade, não serão tratadas neste trabalho. Estas simplificações tornaram-se necessárias por questões operacionais e não invalidam os aspectos metodológicos aqui desenvolvidos.

Inicia-se este trabalho apresentando as estatísticas e procedimentos envolvidos na mudança de suporte, o método analítico e a correção *afim*. Por fim, desenvolve-se uma aplicação com dados simulados de porosidade, onde são mostradas as alterações do semivariograma na mudança de suporte, bem como algumas aproximações numéricas utilizando rotinas computacionais.

¹Dados provenientes de métodos sísmicos. Com estes dados, é possível obter medidas indiretas das propriedades petrofísicas.

²Sendo o modelo esférico definido como:

$$\gamma(h) = \begin{cases} c_0 + c[1, 5\frac{h}{a} - 0, 5(\frac{h}{a})^3], & \text{se } h < a \\ c_0 + c, & \text{caso } h \geq a \end{cases} \quad (1)$$

onde c_0 = efeito pepita; c = patamar e a = alcance.

2 Mudança de Suporte

2.1 Suporte de Amostra

Nas aplicações em ciências da Terra, em particular na área de Geologia, as unidades amostrais podem ter volumes diferentes. Estas unidades amostrais fisicamente distintas, ou com suportes distintos, geram populações amostrais diferentes, com respeito à variância amostral.

O ideal é empregar, no procedimento de estimação, dados na mesma escala com que foram coletados, porém isto nem sempre é possível. Pode haver a necessidade de se aumentar o suporte dos dados coletados, seja para possibilitar uma simulação, seja para compatibilizar dados de fontes diferentes.

O efeito da mudança de suporte sobre a distribuição de uma variável aleatória, o chamado efeito de suporte (ISAACS [12]), é basicamente a redução da variância quando o suporte aumenta. A relação entre continuidade espacial e efeito de suporte torna o semivariograma uma ferramenta útil para a verificação dos efeitos da mudança de suporte.

2.2 Variância de Estimação

Quando uma medida $z(x_i)$, obtida no centro de um bloco ν , é usada para estimar sua verdadeira média $z_\nu(x_i)$, o erro é: $r(x_i) = z_\nu(x_i) - z(x_i)$. Sendo $z(x_i)$ e $r(x_i)$ realizações da função $Z(x_i)$ e da variável aleatória $R(x_i)$ e sob hipótese de estacionariedade, é sempre possível, apesar de $r(x_i)$ ser desconhecido, calcular a esperança, $m_E = E\{R(x)\}$ e a variância $\sigma_E^2 = Var\{R(x)\}$ da função de distribuição do erro, que proporcionam uma medida de qualidade da estimação.

Dado o enfoque da Geoestatística linear, limitado à classe dos estimadores lineares, sempre será possível calcular a variância do erro a partir da semivariância ou da covariância.

Seja $Z_o(x') = \frac{1}{o} \int_{o(x')} Z(y) dy$ onde o é a área ou volume do suporte amostral.

A variância de estimação é:

$$\begin{aligned} \sigma_E^2 &= E\{[Z_\nu(x) - Z_o(x')]^2\} & (2) \\ &= \frac{1}{\nu^2} \int_{\nu(x)} dy \int_{\nu(x)} C(y - y') dy' \\ &\quad + \frac{1}{o^2} \int_{o(x')} dy \int_{o(x')} C(y - y') dy' \\ &\quad - \frac{2}{\nu o} \int_{\nu(x)} dy \int_{o(x')} C(y - y') dy. \end{aligned}$$

Escrevendo em função da semivariância:

$$\sigma_E^2 = 2\bar{\gamma}(\nu, o) - \bar{\gamma}(o, o) - \bar{\gamma}(\nu, \nu) \quad (3)$$

onde $\bar{\gamma}(\nu, o)$ representa o valor médio da semi-variância, $\gamma(h)$, quando uma extremidade do vetor h está no domínio $\nu(x)$ e a outra no domínio $o(x)$ independentemente (JOURNEL [13]).

3 Variância de Dispersão

Seja um volume V dividido em N partes iguais ν_{x_i} , onde x_i refere-se ao centróide do suporte ν :

$$V = \sum_{i=1}^N \nu_i = N\nu$$

Quando ν é muito pequeno em relação a V , ou seja, $\nu \ll V$, os blocos ν_i são centrados em pontos y_i pertencentes a V . Isto significa que o efeito de borda (unidades ν que ultrapassam os limites de V) podem ser ignorados. Seja $z(y)$ a variável regionalizada no ponto y , onde $y \in V$. A média em cada unidade $\nu(x_i)$, centrada no ponto x_i , é:

$$z_\nu(x_i) = \frac{1}{\nu} \int_{\nu(x_i)} z(y) dy \quad (4)$$

De forma similar, calcula-se a média para o suporte V .

A dispersão dos N valores $z_\nu(x_i)$ em relação à média $z_V(x)$ pode ser caracterizada pelo desvio médio quadrático;

$$s^2(x) = \frac{1}{N} \sum_{i=1}^N [z_\nu(x) - z_\nu(x_i)]^2 \quad (5)$$

Na prática, tanto $z_\nu(x_i)$ quanto $z_V(x)$ não são conhecidos. Assim, o problema é estimar, se não um histograma, ao menos os dois valores que o caracterizam, a média $z_V(x)$ e a variância $s^2(x)$.

Se a variável regionalizada $z(y)$ for interpretada como uma particular realização de uma função aleatória $Z(y)$, o desvio médio quadrático (5) também será uma realização de uma variável aleatória $S^2(x)$ definida no ponto x , centróide do suporte V . Sob a hipótese de estacionariedade (ISAACS [12]) da função aleatória $Z(y)$, a esperança da variável aleatória $S^2(x)$ será, por definição, a variância de dispersão da unidade ν em V .

$$\begin{aligned} D^2(\nu/V) &= E\{S^2(x)\} \\ &= E\left\{\frac{1}{V} \int_{V(x)} [Z_V(x) - Z_\nu(y)]^2 dy\right\} \\ &= \frac{1}{V} \int_{V(x)} \sigma_E^2(V(x), \nu(y)) dy, \quad \nu \ll V \end{aligned} \quad (6)$$

Deve ser observado que a variável aleatória $S^2(x)$ e sua realização $s^2(x)$ dependem da posição x dentro

do suporte V , enquanto, sob a hipótese de estacionariedade, $D^2(\nu/V)$ não mais depende da posição x , mas somente da geometria de ν , V e da covariância $C(h)$.

Utilizando a equação (3), a fórmula (6) pode ser reescrita como:

$$D^2(\nu/V) = \bar{\gamma}(V, V) - \bar{\gamma}(\nu, \nu) \quad (7)$$

Na prática, raramente têm-se dados pontuais, ou seja, valores referentes a um ponto exato. Tem-se normalmente um valor $z_o(x)$ referente a um suporte, ou volume amostral, o , centrado no ponto x . Em alguns casos, onde este volume o é muito menor que o volume ν onde deseja-se inferir, $z_o(x)$ é assumido como um valor pontual, podendo ser chamado quase-pontual. Dado que o ponto regionalizado $z(y)$ é uma particular realização de uma função aleatória estacionária, ao menos de segunda ordem, $Z(y)$, pode-se mostrar que a função aleatória regularizada $Z_\nu(x)$ é estacionária de segunda ordem com média m e variograma $2\gamma_\nu(h)$ (JOURNEL [13]).

O problema está em encontrar o variograma regularizado, $2\gamma_\nu(h)$, a partir do variograma pontual $2\gamma(h)$. Do fato de que o variograma, $2\gamma_\nu(h)$ pode ser interpretado com uma variância de estimação de uma variável $Z_\nu(x)$ por outra $Z_\nu(x+h)$, utilizando a equação da variância estimada, (3), dado que o variograma pontual $\gamma(h)$ é estacionário, então, com a devidas simplificações:

$$\gamma_\nu(h) = \bar{\gamma}(\nu, \nu_h) - \bar{\gamma}(\nu, \nu) \quad (8)$$

onde ν_h denota o suporte ν transladado de ν pelo vetor h , $\bar{\gamma}(\nu, \nu_h)$ representa o valor médio do semivariograma pontual $\gamma(h)$, onde uma extremidade representa o suporte ν e a outra independentemente representa o suporte transladado ν_h . Assim, para distâncias h muito maiores que a dimensão do suporte ν , o valor médio de $\bar{\gamma}(\nu, \nu_h)$ é aproximadamente igual à semi-variância $\gamma(h)$. Então obtém-se uma aproximação da função (8), que pode ser reescrita:

$$\gamma_\nu(h) \simeq \gamma(h) - \bar{\gamma}(\nu, \nu), \quad \text{para } h \gg \nu. \quad (9)$$

4 Correção do efeito de suporte

Há vários procedimentos matemáticos para ajustar uma distribuição estimada, de modo a considerar o efeito de suporte. Todos estes procedimentos têm duas características em comum (ISAACS [12]): não alteram a média da distribuição e ajustam a variância através de um fator denotado f (fator de ajuste de variância)

O método mais comum, e o único abordado aqui, é a Correção Afim.

Esta correção é provavelmente a mais simples entre todos os procedimentos de correção do efeito de suporte e é definida como:

$$q' = \sqrt{f} \cdot (q - m) + m \quad (10)$$

onde f , o fator de ajuste da variância, é definido como:

$$f = 1 - \frac{D^2(\nu/V)}{D^2(\nu/G)} \quad (11)$$

onde, $\nu \subset V \subset G$. Dado que a transformação (10) é feita por uma função linear e mantendo a mesma forma da distribuição original, a simetria não se altera com a mudança de suporte. Um exemplo de cálculo desta transformação pode ser visto em ISAAKS [12].

A principal vantagem deste tipo de correção é a simplicidade. E sua principal desvantagem é que pode gerar um valor mínimo não real.

5 Aplicação

5.1 Os Dados

Dada a dificuldade de se obter dados reais, pois estes normalmente são sigilosos por se referirem a reservas de petróleo, foi decidido que seria utilizado um conjunto simulado segundo um caso real. Assim, tendo como referência o artigo de ALMEIDA [2], foi simulado um conjunto que tivesse características semelhantes à porosidade na direção vertical, lá apresentada.

Para obtenção destes dados de referência, foi utilizado o procedimento de simulação sequencial gaussiana incondicional (DEUTSCH [7]).

O modelo de semivariograma omnidirecional retirado do referido artigo, do gráfico "Core Porosity, normal score, vertical direction". é:

$$\gamma(h) = 0,2 + 0,8 \cdot Sph\left(\frac{h}{18}\right) \quad (12)$$

A malha simulada foi de 300 por 300 pontos distribuídos regularmente de dois em dois metros (em vez de metros, qualquer unidade de distância).

A porosidade é um fenômeno tridimensional. Para simplificação das aplicações, foi utilizado um conjunto de dados representando uma "fatia" do espaço tridimensional, um conjunto bidimensional.

Uma vez criado o conjunto de dados, foram gerados outros conjuntos resultantes do procedimento de aumento de suporte ou *upscaling*. Este procedimento consiste em redefinir o suporte de amostragem ou a

área que cada amostra representa. Desta forma, dentro da malha de 300 por 300 pontos igualmente espaçados de dois em dois metros, para se criar uma nova malha com suporte de 4 por 4 metros, basta agrupar pontos adjacentes contidos na nova área desejada.

Desta forma, vários conjuntos com novos suportes foram criados e analisados quanto às modificações nas suas estruturas espaciais.

5.2 Modelos Ajustados

Para cada um dos conjuntos criados a partir do procedimento de aumento de escala, foi ajustado um variograma através de um procedimento computacional não linear, semelhante o feito por BETTINI [3], utilizando um procedimento de ajuste não linear.

O modelo obtido para o conjunto simulado de porosidade está descrito na equação (13).

$$\gamma(h) = 7,76 + 29,44 \cdot Sph\left(\frac{h}{19,35}\right) \quad (13)$$

Os modelos ajustados para os conjuntos de dados resultantes do procedimento de *upscaling* foram representados também pela função esférica, conforme definida em (1), porém acrescentando os índices que identificam o suporte, ou seja, $\gamma_{i;L}(h)$, onde i e L são os lados do bloco; e os demais parâmetros são definidos de forma correspondente a (1).

Os modelos obtidos foram:

$$\left\{ \begin{array}{l} \gamma_{4;4}(h) = 1,94 + 26,16 \cdot Sph\left(\frac{h}{21,2}\right) \\ \gamma_{6;6}(h) = 0,86 + 23,33 \cdot Sph\left(\frac{h}{23,15}\right) \\ \gamma_{8;8}(h) = 0,48 + 20,68 \cdot Sph\left(\frac{h}{24,29}\right) \\ \gamma_{10;10}(h) = 0,31 + 18,29 \cdot Sph\left(\frac{h}{26,07}\right) \\ \gamma_{12;12}(h) = 0,2 + 16,15 \cdot Sph\left(\frac{h}{27,6}\right) \end{array} \right. \quad (14)$$

Em uma rápida observação dos parâmetros dos modelos em (14), tem-se que c_0 e c (definidos em (1)) decrescem com o aumento do suporte, enquanto o alance, a , aumenta.

5.3 Resolução analítica.

Conforme dito em (8) e (9), a redução na variabilidade é dada pelo valor médio do semivariograma pontual. Uma aproximação simples e discreta desta função analítica é:

$$\bar{\gamma}(\nu, \nu') \simeq \frac{1}{nn'} \sum_{i=1}^n \sum_{j=1}^{n'} \gamma(x_i - x'_j) \quad (15)$$

A função de semivariância utilizada deve ser conhecida ou estimada. Neste caso, como aproximação do modelo pontual, foi utilizada a mesma função obtida com o conjunto de dados simulados, o modelo (13).

Com o cálculo de $\bar{\gamma}(\nu, \nu)$, pode-se obter $\gamma_\nu(h)$; $h > a$ através da função (9).

A fim de tornar mais simples a obtenção de $\bar{\gamma}(\nu, \nu)$, ajustou-se um modelo relacionando a semivariância com os lados que formam a área retangular, pelo método de regressão linear simples com os valores de semivariância transformados. A transformação utilizada foi de BOX & COX (DRAPER [9]).

O modelo obtido na regressão foi:

$$W = 0,094291 - 0,56351 \frac{l}{a} - 0,624713 \frac{L}{a} - 0,044197d \quad (16)$$

onde,

a , L e l são respectivamente o alcance e os lados do suporte retangular.

d é dada por:

$$d = \begin{cases} 1, & \text{se } \left[\frac{l}{a} < 0,2 \text{ ou } \frac{L}{a} < 0,2 \right] \\ 0, & \text{caso contrário.} \end{cases}$$

Outras estatísticas obtidas nesta regressão linear foram: $R^2 = 0,998$ e $F_{0,05} = 8577 \rightarrow p \simeq 0$. E os desvios dos parâmetros, $\beta_0, \beta_1, \beta_2, \beta_3$ foram respectivamente: 0,009958, 0,019038, 0,025112 e 0,012114.

Verificando o ajuste da regressão, na tabela 1 tem-se a comparação de $\bar{\gamma}(\nu, \nu)$ padronizado tabelado, retirado de DAVID [6] ou JOURNAL [13], com os valores obtidos utilizando a função de regressão (16). Observa-se que as diferenças relativas, descritas na quarta coluna da tabela 1, são sempre pequenas.

5.4 Predição dos modelos de Semivariância.

Para a predição do modelo de semivariograma para dados em um outro suporte, é necessário o conhecimento do modelo de semivariância pontual ou quase pontual.

Nesta seção e como na subseção 5.3, o modelo utilizado foi o da criação do conjunto de dados, o modelo (13), considerado quase pontual.

Redução da variabilidade - $\bar{\gamma}(\nu, \nu)$

$\frac{l}{a}$	$\frac{L}{a}$	$\bar{\gamma}_{tab.}(\nu, \nu)$	$\bar{\gamma}_{reg.}(\nu, \nu)$	$\frac{\bar{\gamma}_{tab.}(\nu, \nu)}{\bar{\gamma}_{reg.}(\nu, \nu)}$
0,30	0,10	0,165	0,1655	0,9969
0,60	1,50	0,675	0,6867	0,9830
2,00	1,00	0,800	0,8016	0,9980
0,30	0,15	0,180	0,1910	0,9423
0,20	0,40	0,238	0,2345	1,0235

Tabela 1: $\bar{\gamma}(\nu, \nu)$ padronizado tabelado e estimado pela regressão.

5.5 Método Analítico (Direto)

Conforme mostrado na equação (8), dado o conhecimento do modelo pontual, pode-se calcular a semivariância para quase toda distância h e para um suporte ν .

Neste procedimento de predição do modelo de semivariância para dados sob um determinado suporte, tomaram-se as seguintes regras:

- Modelo : o modelo foi mantido o mesmo, no caso, o modelo esférico. Em DAVID [6] e JOURNAL [13] se subentende que há uma alteração no comportamento da curva próximo de sua origem, quando o suporte aumenta. O semivariograma deixa de ter comportamento linear próximo da origem, sugerindo pelo seu aspecto um modelo de variograma Gaussiano. Nesta aplicação, devido à escala usada, este efeito foi desprezado.
- O patamar *sill* fica reduzido conforme mostrado na relação (9).
- O efeito pepita (*nugget effect*) reflete as microestruturas não captadas pela amostragem, bem como erros na amostragem (CRESSIE [5]). Sabe-se que $\gamma(h) = 0$ quando $h = 0$. Porém, esta função pode ser descontínua na origem, saltando para c_0 , o efeito pepita, que pode ser decomposto em dois termos de forma aditiva, a variação em microescala, c_{MS} , e a variação devida ao erro de medição, c_{EM} . Quando a unidade amostral aumenta, como no procedimento de aumento de escala, c_{MS} tende a desaparecer. O valor atribuído ao centróide do suporte ampliado será a média aritmética dos valores referentes aos suportes menores que o compõem. Considerando que as observações que participam desta estimação têm erro de medição com médias constantes e independentes, a estimativa dada pela média aritmética terá a variabilidade c_{EM}/n , onde n é o número de pontos vizinhos agregados na formação do novo bloco.

Com os resultados obtidos neste algoritmo comparados aos modelos observados, verifica-se que estes têm sempre patamares menores e alcances maiores, conforme pode ser verificado nas figuras 1 e 2 (grafico à esquerda.).

No caso estudado, bidimensional, acreditando no aumento proporcional do alcance em relação ao suporte da amostra, a proposta para estimação deste parâmetro é:

$$\ln \frac{a_2}{a_1} = 0,114591 \ln \frac{A_2}{A_1} \quad (17)$$

$$\ln \frac{a_2}{a_p} = 0,060087 \ln \frac{A_2}{1} \quad (18)$$

Estas funções, (17) e (18), foram obtidas pelo método de regressão linear, a partir dos dados observados, provenientes dos modelos ajustados para os conjuntos de dados simulados nos vários suportes distintos. Os desvios para os parâmetros estimados foram, respectivamente para (17) e (18): 0,00393 e 0,005695. O modelo (17) deve ser aplicado para estimar a razão entre os alcances, para o modelo esférico, quando as áreas envolvidas A_1 e A_2 são maiores que a pontual.

A área pontual foi admitida igual a um, logo o modelo (18) é aplicado para estimar a razão entre o alcance pontual e o da área maior, A_2 , qualquer. Assim o alcance estimado é dado por:

$$a_\nu = \exp\left(\ln \frac{a_2}{a_p}\right) \quad (19)$$

onde ν se refere ao suporte de área A_2 .

A decisão para a utilização de dois modelos foi devido ao fato de o comportamento dos dois conjuntos de dados diferir muito quando a razão é tomada em relação a um, área assumida para o modelo pontual, ou em relação a qualquer outra área maior.

Outro modo de estimação do modelo de semivariância, chamado *Indireto*, agrega a estimação do alcance por (17) e (18), e do patamar e do efeito pepita de modo semelhante ao modo *Direto*.

Os resultados observados, preditos pelos modos direto e indireto (fornecendo a média e um intervalo), estão na tabela 2 e esboçados nas figuras 1 e 2.

Observa-se que, para todos os suportes utilizados, os valores preditos pelo método computacional sempre são menores que os observados, mantendo uma diferença praticamente constante para todas as áreas de bloco.

Esta pequena diferença pode ocorrer pelo fato de se ter utilizado um modelo aproximado ao modelo pontual. Outra causa possível seria a imprecisão devida à

l	L	γ_{Direto}	$\gamma_{Média}$	γ_{Obs}
4	4	26,62	26,66	28,01
6	6	23,23	23,27	24,19
10	10	18,29	18,35	18,60
12	12	16,16	16,23	16,35

Tabela 2: Patamares (*Sill*) para vários suportes

simplificação adotada no efeito pepita. Esta última é mais provável pelo fato de se tomar o efeito pepita somente como erro de medida, desprezando o erro em pequena escala (seção 5.5). Este erro de pequena escala torna-se desprezível quando o suporte é relativamente grande. Observa-se, na figura 1 um comportamento condizente com a idéia de simplificação do efeito pepita. Quando o suporte aumenta, a diferença entre os patamares se reduz. Isto implica que esta simplificação é mais eficiente para estimações em suportes grandes.

Pelo método direto, o resultado obtido é bem próximo ao observado para suportes pequenos, como observado na figura 1. Quando o suporte cresce, aumenta a diferença entre a curva predita pelo modo direto e a observada, como se verifica na figura 2 (figura da esquerda). Isto mostra as restrições deste método para predição de suportes maiores.

Os resultados obtidos na predição pelo método indireto mostram que, para os suportes menores, o modelo observado ficou sempre fora dos limites de confiança construído, um pouco acima do limite superior. Quando o suporte aumenta, os limites de confiança passam a conter a curva observada. Isto se deve, em parte, à diferença entre patamar observado e estimado.

5.6 A Correção Afim

Conforme descrito anteriormente, a correção afim utiliza o fator de ajuste de variância, f , na função (10), de tal modo que a variância seja alterada sem modificação da média.

Para demonstrar um resultado deste procedimento, foi aplicada a correção afim para um suporte de 6 por 6 metros e compararam-se os semivariogramas obtidos deste conjunto transformado com o observado na mudança de suporte (14).

Obteve-se então $f = 0,8073$, que, aplicado a função (10), gerou um novo conjunto de dados corrigido para o suporte desejado (6 x 6 m).

O semivariograma deste conjunto corrigido, juntamente com o semivariograma obtido no procedimento de aumento de suporte (dados simulados) é mostrado na figura 2 (figura da direita).

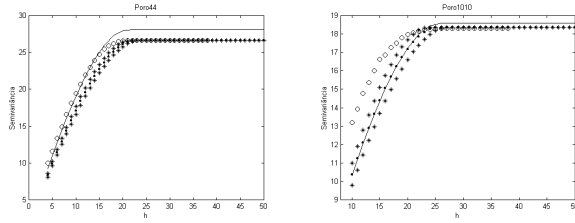


Figura 1: Semivariogramas Preditos (*' - L.I. e L. S.; '.' - média; 'o' - direto) e Observado ("linha") para área 4x4 (à esq.) e área 10x10.

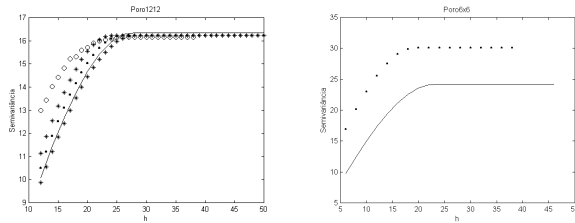


Figura 2: (à esq.) Semivariogramas Preditos (*' - L.I. e L. S.; '.' - média; 'o' - direto) e Observado ("linha") para área 12x12. (à dir.) Semivariogramas obtidos pela correção Afim, '.', e com *upscaling*, "linha".

Observa-se que o patamar do semivariograma para a correção afim é maior do que o obtido no procedimento de *upscaling*, com valores $Sill_{Afim} = 29,78$ e $Sill_{up} = 23,68$ respectivamente. Isto pode ocorrer, pois, no procedimento de aumento de suporte adotado, tem-se uma significativa redução da variabilidade devida à redução do erro de medição, enquanto nesta transformação faz-se simplesmente uma redução na variabilidade como um todo. Verifica-se também que o alcance difere. Isto indica que este procedimento é ineficaz, tendo sua aplicação válida para pequenas mudanças de suporte, quando o fator de ajuste de variância for pequeno e se o valor de interesse ou ponto de corte estiver próximo da média.

6 A Redução do Suporte

Com os resultados obtidos em 5.5, há a possibilidade de fazer as estimações no sentido inverso, ou seja, proceder à redução de escala, o *Downscaling*. Supondo que se conheça a estrutura de variabilidade espacial para um dado suporte, deseja-se estimar os parâmetros para um suporte menor.

Mantendo-se o modelo esférico e alterando somente seus parâmetros, o efeito pepita c_0 mantém-se o mesmo, pois tem-se uma única observação (seção 5.5). E ainda, assumindo que o patamar e o alcance são inde-

pendentes utiliza-se a rotina MATLAB - DOWNSPH, (ALMEIDA [1]) para estimar o semivariograma para um suporte menor, fornecendo os limites de confiança (90%). Dado o conhecimento dos parâmetros do modelo para suporte V e a definição do suporte ν , $\nu < V$ o algoritmo segue os seguintes passos:

1. com a função (18) faz-se a predição dos limites inferior e superior para o alcance pontual;
2. com a função (17) faz-se a predição dos limites inferior e superior para o alcance no suporte desejado; sorteia-se um valor em cada um dos intervalos dos itens anteriores;
3. com a função (16) e o valor sorteado para o alcance pontual, faz-se a predição dos limites inferior e superior para a razão entre a semivariância média para o suporte amostrado e o desejado;
4. sorteia-se um valor do intervalo da razão entre semivariâncias, multiplicando-o pelo patamar observado. O patamar assim estimado, juntamente com o valor sorteado para alcance, possibilita a construção da curva, o variograma.
5. repetem-se os itens anteriores um grande número de vezes e depois, para cada distância, calcula-se a média e os percentis 5 e 95.

Este procedimento foi repetido 2000 vezes e então foram construídos os gráficos do limite inferior e superior (I.C. 90%) e da média, como é mostrado na figura 3.

Observou-se que, em todos os casos, o modelo observado está sempre dentro dos limites calculados, embora, para os suportes menores, as médias calculadas sejam sempre menores, ocorrendo o inverso para suportes grandes. Os intervalos entre os limites são grandes, expressando a incerteza de todos os parâmetros estimados no procedimento de cálculo, o que demonstra a necessidade de melhoria no método, tornando-o mais eficaz, predizendo com menor incerteza.

6.1 Comparação entre modelos selecionados

O procedimento de validação cruzada utilizado foi o da GSLib, a rotina computacional KT3d (DEUTSH [7]), tendo como referência o conjunto de dados com suporte 6x6, conforme descrito na seção 5.1 e utilizando a krigagem ordinária.

(ISAAKS [12]) foi utilizado para comparação entre os modelos: obtido diretamente com os dados observados com suporte 6x6, $\gamma_{obs}(h)$, conforme modelo

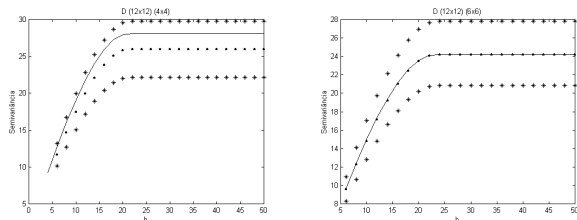


Figura 3: Semivariogramas Predito (L.I. e L.S. '*'; média '. ') e Semivariograma Observado. *Downscaling* $12 \times 12m \rightarrow 4 \times 4m$ (à esq.) e $12 \times 12m \rightarrow 6 \times 6m$.

(14); o predito pelo método analítico, $\gamma_{PredI}(h)$, subseção 5.5; o predito pelo método Direto, $\gamma_{PredD}(h)$ (função (8)) e o modelo obtido da correção Afim, $\gamma_{AFIM}(h)$.

Considerando a raiz média quadrática do erro de predição padronizado,³ verifica-se que o modelo predito pelo modo indireto 'PREDI' apresenta um bom resultado (0,82), mais próximo do observado e da unidade. O modelo observado teve r.m.q.= 0,81. A transformação AFIM, como seria de se esperar, tem o pior resultado(0,61), comparado aos demais.

7 Conclusões

Os resultados deste trabalho referem-se a variáveis escalares, que é o caso da porosidade. As predições para os modelos de semivariância no procedimento de aumento de escala são mais simples, dado o conhecimento do modelo pontual, uma vez que os efeitos são conhecidos. No procedimento de redução de escala esta predição envolve uma incerteza maior, necessitando do conhecimento ou estimação do modelo pontual.

Contudo, como mostrado pela validação cruzada (seção 6.1), os modelos preditos são bons quando comparados com os modelos observados.

Talvez, um modelo que contemple a alteração na forma do semivariograma devido a mudança de suporte, conforme mostrado em JOURNAL [13], proporcione uma redução da incerteza nas predições.

Referências

[1] ALMEIDA, C.F. P.. *A Variação do Suporte Amostral e as Mudanças na Estrutura do Semivariograma*. Rio de Janeiro, 2001, Dissertação de Mestrado, Instituto de Matemática da Universidade Federal do Rio de Janeiro.

- [2] ALMEIDA, A. S.; FRYKMAN, P.. *Geostatistical Modeling of Chalk Reservoir Properties in Dan Field, Danish North Sea*. In Yarus, J.M.; Chambers, R.L., eds., AAPG computer application in geology no. 3, Stochastic modeling and geostatistics: Principles, method and case studies: American Association of Petroleum Geologists, 1994, Tulsa, OK, p. 273-286.
- [3] BETTINI, C.. *Bases Teóricas para desmembrar a proporção de uma categoria litológica em proporções de subcategorias, a partir do semivariograma imbricado de variável indicadora*, Relatório Técnico, n. 94, Laboratório de Estatística, Instituto de Matemática, Universidade Federal do Rio de Janeiro, 1996.
- [4] CHILÈS, J. P.; DELFINER, P.. *Geostatistics: Modeling Spatial Uncertainty*, John Wiley & Sons, Inc.1999.
- [5] CRESSIE, N. A. C. *Statistics for Spatial Data*, John Wiley & Sons, Inc.1993.
- [6] DAVID, M. *Geostatistical Ore Reserve Estimation*, Elsevier Scientific Publishing Company, 1982.
- [7] DEUTSCH, C.V.. JOURNAL, A.G.. *GSLIB: Geostatistics software library and user's guide*. Oxford University Press, 1992.
- [8] DIGGLE, P. J.; RIBEIRO Jr, P. J. *Model Based Geostatistics*. 14 SINAPE, Associação Brasileira de Estatística, 2000.
- [9] DRAPER, N.; SMITH, H.. *Applied Regression Analysis*, second edition, John Wiley & Sons, Inc., New York, 1980.
- [10] GELFAND, A. E.; ZHU, L.; CARLIN, B. P.. *On the Change of Support Problem for Spatial-Temporal Data*, 2000.
- [11] HOHN, M. E.. *Geostatistics and Petroleum Geology*, Computer Methods in Geosciences, Van Nostrand Reinhold, 1988.
- [12] ISAACS, E. H.; SRISVASTAVA, R. M.. *An Introduction to Applied Geostatistics*, Oxford, New York, 1989.
- [13] JOURNAL, A.G; HUIJBREGTS, CH. J.. *Mining Geostatistics*, Sixth Printing, Academic Press, 1993.
- [14] MANN, C. John.. *Uncertainty in Geology*. Computers in Geology, Jonh C. Davis e UTE C. Herzfeld, Oxford University Press, New York, 1993.
- [15] MYERS, J. C.. *Geostatistical Error Management*, Van Nostrand Reinhold, New York, 1996.

$$r.m.q. = \left[\left(\frac{1}{n} \right) \sum_{j=1}^n \left\{ \frac{(Z(s_j) - \bar{Z}(s_j))}{\sigma(s_j)} \right\}^2 \right]^{\frac{1}{2}} \quad (20)$$