Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Estatística

Uma Proposta de Análise Unidimensional para Determinação do Padrão de Transporte de Sedimentos, *Via* o Método Exato Estrutural e o Método de Componentes Principais

P. S. Lucio, E. C. Bodevan, F. O. Garcia, L. A. T. Medrano, H. Dupont e L. V. Ribeiro

Relatório Técnico
RTP-05/2000

Relatório Técnico
Série Pesquisa

UMA PROPOSTA DE ANÁLISE UNIDIMENSIONAL PARA DETERMINAÇÃO DO PADRÃO DE TRANSPORTE DE SEDIMENTOS, *VIA* O MÉTODO EXATO ESTRUTURAL E O MÉTODO DE COMPONENTES PRINCIPAIS:

O CASO DO RIO "WESTERSCHELDE" (HOLANDA/BÉLGICA) .

P. S. LUCIO; E. C. BODEVAN;
F. DE O. GARCIA; L. A. T. MEDRANO
DEPARTAMENTO DE ESTATÍSTICA – ICEX - UFMG
H. DUPONT; L. V. RIBEIRO
DEPARTAMENTO DE GEOLOGIA – IGC - UFMG

RELATÓRIO TÉCNICO DE PESQUISA

DEPARTAMENTO DE ESTATÍSTICA
INSTITUTO DE CIÊNCIAS EXATAS
UNIVERSIDADE FEDERAL DE MINAS GERAIS

JULHO, 2000

RESUMO

Sabe-se que a granulometria de um sedimento é uma característica bastante complexa que requer estudos e definições de parâmetros que vão além de uma simples categorização ou classificação. O estudo granulométrico define a quantidade de sedimento situado em faixas de tamanhos de grão ou a distribuição de freqüências por faixa granulométrica.

Esta distribuição de freqüências do tamanho dos grãos deve ser analisada sob um enfoque estatístico, para que se possa estabelecer os parâmetros que permitam inferir sobre um padrão de informação concernente ao transporte de sedimentos, baseado no depósito residual.

Em geral, estudos são realizados com sedimentos não coesos, que podem ser separados em faixas granulométricas através de peneiramento. Assim, uma amostra é peneirada e cada fração deve ser pesada, com certa precisão, e estes pesos são transformados em percentagens de um sistema fechado. Um problema, freqüentemente abordado, é aquele que se refere aos limites das classes granulométricas e ao número destas classes. Na prática, determina-se uma escala, com intervalos de classes crescentes. O número mínimo de classes deve ser tal que não apresente variabilidade exagerada em cada uma delas. Já o número máximo deve ser aquele que torne o experimento economicamente viável, uma vez que um número muito grande iria onerar as operações de peneiramento, de análise estatística e de representação gráfica.

Neste estudo é feita uma abordagem geral dos métodos de McLaren (1981) e McLaren & Bowles (1985) para a obtenção de informações sobre a direção de transporte de sedimentos. Em seguida, são propostas duas outras alternativas para a análise de dados desta gênese: o método *exato estrutural* e o método das *componentes principais*. Finalizando, introduzimos a noção de dependência espacial através da análise variográfica. Para a análise, foram utilizados dados granulométricos amostrados na parte leste do rio Schelde entre a Bélgica e a Holanda. Os algoritmos de análise numérica dos métodos aqui estudados foram implementados em linguagem S-Plus™.

As linhas estudadas apresentam característica markoviana de primeira ordem. Os resultados obtidos através do *exato estrutural* não são muito similares àqueles obtidos pelo método de "McLaren". Entretanto, utilizando a análise de *componentes principais* , foi determinado um índice possibilitando a obtenção de informações similares àquelas fornecidas pelo método de "McLaren" no que concerne informações sobre a direção do transporte de sedimentos.

Palavras-chave: análise de conglomerados; granulometria; processos markovianos; modelo binomial; Z-escore; variograma; krigagem.

SUMÁRIO

INTRODUÇÃO	**04**
CONGLOMERADOS	**05**
CAMPOS DE MARKOV UNIDIMENSIONAIS DE PRIMEIRA ORDEM	**07**
ANÁLISE ESTRUTURAL	**09**
ANÁLISE CLÁSSICA	**11**
TESTE EXATO	**14**
COMPONENTES PRINCIPAIS	**18**
SÍNTESE	**23**
CONCLUSÕES E PERSPECTIVAS	**25**
AGRADECIMENTOS	**26**
BIBLIOGRAFIA	**26**

1. INTRODUÇÃO

As características texturais de um sedimento podem ser referidas a um conjunto de estatísticas descritivas de sua distribuição granulométrica, associadas a tendências centrais (média) e caudais (desvio-padrão e assimetria). A utilização de características granulométricas de sedimentos na descrição e identificação de ambientes deposicionais tem como pioneiro o trabalho de McLaren (1981) que utiliza mudanças em algumas medidas estatísticas como média (primeiro momento), seleção (segundo momento) e assimetria (terceiro momento), para descrever as distribuições de tamanhos de grãos. Obtêm-se, desta forma e através de comparações pontuais dos parâmetros granulométricos, informações sobre a direção do transporte de sedimentos

A análise de dados em seqüência ou em cadeia que inclui levantamentos mineralógicos ou geoquímicos ao longo de um linha definida por uma propriedade geomorfológica é muito comum em Ciências da Terra. As técnicas usuais para examinar estes dados devem fazer parte dos conceitos fundamentais de Processos Estocásticos.

1.1 Objetivos

Dentro deste contexto, propõe-se neste ensaio: (1) o estudo de alguns parâmetros associados à granulometria ao longo de perfis longitudinais (seqüencialmente selecionados) com vias à identificação de um padrão de transporte de sedimentos. Desta forma, através do enfoque de Cadeias de Markov, sugere-se uma metodologia *estrutural* para a caracterização de depósitos em agregação e/ou erosão; (2) implementar um teste *exato*, sob a suposição de que o processo unidimensional tenha memória (número de vizinhos influentes) de tamanho um; (3) desenvolver um modelo mais simples fazendo-se uso da técnica de análise de *componentes principais* para criar um índice sintético que conjugue as medidas estatísticas utilizadas (média, seleção e assimetria) para obter informações sobre a direção de transporte de sedimentos e (4) analisar a dependência espacial de forma menos restrita, por intermédio de variogramas, na tentativa de interpolar os dados e corroborar a tendência linear.

1.2 Dados

A região sob estudo se estende de Baarland, na Holanda a Rupelmonde, na Bélgica (Figura 1). Concernente à aquisição dos dados, as 867 amostras de sedimentos tratadas e analisadas granulometricamente, nos foram gentilmente cedidas pelo Dr. Patrick McLaren da *GeoSea Consulting* (Canadá), em arquivo ASCII, especialmente para este estudo. Estas amostras foram coletadas com um espaçamento regular de 500 metros ao longo de toda a região, salvo em determinadas áreas de batimetria complexa, onde esta malha foi localmente reduzida a um intervalo de 250 metros.

As linhas selecionadas para este estudo foram determinadas uma ao centro (L2), uma à margem superior (L1) e outra à margem inferior (L3), sem levar em consideração que suas batimetrias pudessem apresentar padrões não similares, para que não tivéssemos que considerar o efeito de suporte proveniente de uma amostragem não regular.

Primeiro, agrupou-se as granulometrias com vias a determinação de estados para o estudo estocástico. Em seguida, determinou-se a ordem estrutural do processo para que o teste exato fosse aplicado. Paralelamente, utilizou-se a metodologia de componentes principais para a determinação de um índice de informação sobre a direção de transporte. Finalmente, analisou-se a estrutura de dependência espacial através da construção de variogramas para a determinação de domínios de influência.

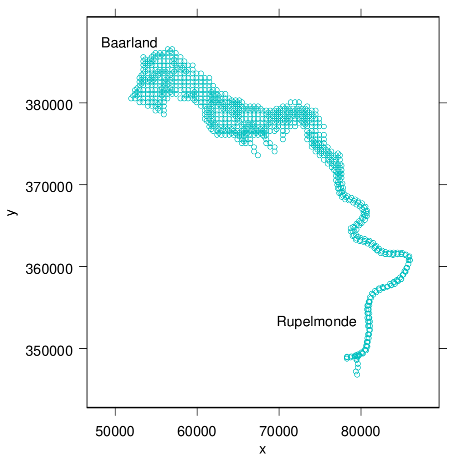

Fig. 1: Área global analisada (867 amostras).

2. CONGLOMERADOS

A análise de conglomerados é freqüentemente denominada classificação automática, onde as similaridades (ou dissimilaridades) entre variáveis são avaliadas com o objetivo de formar agrupamentos com base em características semelhantes através de partições. Propondo-se investigar o conjunto de variáveis granulométricas deste estudo, a partição e categorização foram fundamentadas sob uma discretização nominal utilizando a função de ligação *completa* (Johnson &

Wichern, 1998), que leva em consideração a máxima distância entre grupos, referenciando-se a seus *centróides* – centro de gravidade da massa de dados.

Executando-se o *dendograma* (sob métrica *completa*), realizado para a região, sob um enfoque global, percebe-se a nítida formação de seis conglomerados granulométricos (Figura 2). Note ainda, que detectado através desta técnica, o grau de similaridade, em fase, entre os grupos de variáveis é muito baixo, em torno de 16%.

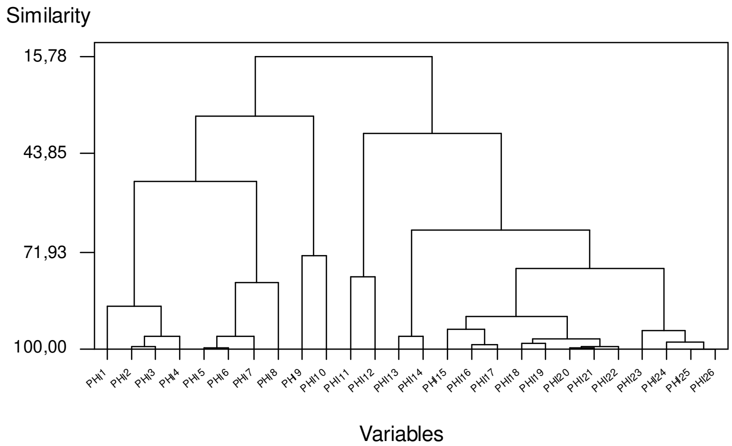

Fig. 2: Dendograma das 26 classificações granulométricas originais concernentes à região global.

Desta forma, estar-se-á estudando, sob um enfoque probabilístico, classes de variáveis granulométricas definidas por meio da determinação de categorias identificadas através de intervalos empíricos por tamanho dos grãos em escala Φ $[\Phi = -\log_2\{\text{Diam(mm)}\}]$. Assim, definiu-se da seguinte forma a classificação, sem perda de generalidades e de informação geológica:

1 – Grânulos e Areia Muito Grossa (GG) $[-2 < \Phi \leq 0]$;

2 – Areia Grossa e Média (GMG) $[0 < \Phi \leq 2]$;

3 – Areia Fina e Muito Fina (GMF) $[2 < \Phi \leq 3{,}5]$;

4 – Silte Grosso (GF) $[3{,}5 < \Phi \leq 5]$;

5 – Silte Fino (S) $[5 < \Phi \leq 7]$;

6 - Argila (A) $[\Phi > 7]$.

A aplicação de modelos matemáticos em geologia geralmente é dificultada pelo pressuposto de casualidade em suposições que invocam eventos totalmente independentes ou, em outro extremo, pela exigência nos modelos determinísticos clássicos de que os eventos estejam condicionados a variáveis específicas, predeterminadas e perfeitamente conhecidas. O modelo de Campos de

Markov, ocupando justamente uma posição intermediária nesse espectro de modelos dinâmicos, presta-se com mais propriedade a ser utilizado na tentativa de explicação de processos geológicos.

3. CAMPOS DE MARKOV UNIDIMENSIONAIS DE PRIMEIRA ORDEM

Um processo estocástico unidimensional $\mathbf{Z} = \{Z(\mathbf{x}), \mathbf{x} \in \Re\}$, é um Campo de Markov de primeira ordem se para qualquer escolha de $\mathbf{x} \in \Re$, tem-se a seguinte relação:

$$P\left(a < Z(\mathbf{x}) \le b \mid Z(\mathbf{x_1}) = z_1, Z(\mathbf{x_2}) = z_2, Z(\mathbf{x_3}) = z_3, ..., Z(\mathbf{x_n}) = z_n\right) = P\left(a < Z(\mathbf{x}) \le b \mid Z(\mathbf{x_n}) = z_n\right),$$

$$0 < \mathbf{x_1} < \mathbf{x}_2 < .... < \mathbf{x}_n < \mathbf{x}.$$

O método dos campos markovianos foi selecionado como ferramenta para análise destes dados de natureza seqüencial, por ser considerado de grande aplicação para o entendimento dos fenômenos geológicos de sedimentação.

Um Campo de Markov com parâmetro espacial discreto $\{\mathbf{Z_n}\}$ é um processo estocástico markoviano cujo espaço de estados é um conjunto finito ou enumerável e para o qual **x** = (0,1,2,...). Seja

$$P_{ij}^{n,n+1} = P\left(Z_{n+1} = j \mid Z_n = i\right).$$

Esta quantidade, $P_{ij}^{n,n+1}$, é dita probabilidade de transição do estado (*i*) na posição (*n*) para o estado (*j*) na posição (*n+1*). Quando $P_{ij}^{n,n+1}$ não depende de *n*, ou seja, não depende da posição, a cadeia possui a propriedade de probabilidade de transição estacionária. Neste caso, sem perda de generalidades, utiliza-se:

$$P_{ij} = P\left(Z_{n+1} = j \mid Z_n = i\right).$$

Os Campos de Markov consideram seqüências ou cadeias de estados em que a transição de um estado para outro, dentro de um intervalo no espaço, depende de estados prévios, de tal forma que as probabilidades associadas às transições entre os estados sejam, necessariamente, estacionárias. Neste sentido, pode-se afirmar que muitos processos geológicos de sedimentação (deposição ou erosão) possuem a propriedade markoviana, que se traduz pela influência de eventos prévios em eventos subseqüentes .

O grau de dependência de um estado em relação a estados prévios nos informa a ordem da cadeia. Desta forma, se existe um padrão de dependência apenas em relação ao estado imediatamente anterior a cadeia é classificada como de *primeira ordem*, se dois estados imediatamente anteriores estão associados a cadeia é dita de *segunda ordem* e envolve uma unidade espacial. No entanto, se os dois estados, que estão associados, não são imediatamente precedentes, a cadeia possui *ordem superior* à segunda e os intervalos espaciais envolvidos podem ser maiores que uma unidade espacial.

Para aplicarmos o método dos Campos de Markov à nossa pesquisa, são necessárias algumas definições prévias, adequando a teoria explicitada acima ao nosso estudo. Assim, doravante, um *processo estocástico* será traduzido como o processo de transporte de sedimentos, caracterizado pela predominância dos materiais granulométricos ao longo de cada uma das linhas amostrais e os *estados da cadeia* serão os seis estágios de predominância granulométrica. O *espaço de estados* será discreto e finito, devendo ser classificado segundo o teor granulométrico. Observe que os Campos de Markov de primeira ordem e dependência simples são os mais utilizados na abordagem de processos sedimentares, tendo em vista sua relativa facilidade de aplicação. No entanto, é comum seqüências sedimentares exibirem superposição de ciclos maiores sobre ciclos menores, situações estas que podem ser representadas por matrizes markovianas de alta ordem, cuja estruturação e abordagem matemática são tanto mais complexas quanto maior o grau de dependência revelado pelo processo markoviano.

As matrizes de probabilidade de transição constituem uma forma simples e prática de apresentação dos Campos de Markov. O cálculo das probabilidades de transição baseia-se na freqüência de ocorrência de transições de um estágio para outro, observadas na representação do processo considerado, isto é, na matriz de registros ou de freqüências de transição. A verificação da propriedade markoviana no processo de transporte de sedimentos é feita geralmente por meio da matriz de transição dos parâmetros granulométricos onde os estágios envolvidos no processo correspondem à granulometria que compõe a seqüência estudada. Desta forma, para se determinar a matriz de registro de transições é necessário que observemos a ordem seqüencial dos estágios granulométricos predominantes.

O diferencial da metodologia estrutural, apresentada neste estudo, em relação à aplicação clássica (McLaren, 1981 e McLaren & Bowles, 1985), é o fato de nosso interesse estar voltado para a região no que tange à influência entre vizinhos mais próximos, e não para uma linha (1D) como um todo. Desta forma, sem perda de informação, caracterizamos a região segundo a análise de dados seqüenciais. É conveniente ressaltarmos que, não estamos interessados em verificar a propriedade markoviana no processo sedimentar, e sim verificar a hipótese do transporte de sedimentos em determinada localidade.

4. ANÁLISE ESTRUTURAL

A partir das matrizes de transições granulométricas, objetiva-se representar os diferentes estágios por meio da sucessão de tamanho de grãos categorizados, que foram depositadas em determinada seqüência sedimentar, buscando caracterizar seu ambiente ou sub-ambiente específico. Assim, as diversas composições ou associações granulométricas de diferentes perfis de um ou diferentes sistemas deposicionais podem ser representadas por diferentes matrizes markovianas. Observe que esta visão simplista não nos permite analisar períodos de não deposição ou mesmo erosivos, pois as freqüências são tabuladas considerado-se unicamente o registro granulométrico local. Problemas similares são evidentes no caso das cadeias, em que intervalos fixos de tamanho da amostra são diretamente relacionados a intervalos discretos na escala espacial de deposição, desprezando-se fatores tais como: diferentes graus de compactação de diferentes locais, hiatos deposicionais ou erosivos, diferentes taxas de sedimentação, entre outros. Tais situações podem ser, em parte, resolvidas pela diminuição dos intervalos dos registros de transições, *i. e.*, mudanças significativas no ambiente sedimentar, não são incluídas numa mesma matriz, mas constituem pontos de separação entre matrizes. Este último caso é de grande importância no que se refere à estacionariedade da cadeia.

A estruturação de uma sucessão de eventos em uma matriz de probabilidades de transições não significa, necessariamente, que o processo físico original seja um processo markoviano. É evidente, pois, que a aplicação do modelo markoviano a seqüências sedimentares prescinde da verificação da propriedade markoviana. Para tanto, as suposições do teste são baseadas nas seguintes condições:

- a matriz de probabilidade de transições deve ser estacionária;
- as matrizes devem ser tabuladas segundo intervalos fixos.

A propriedade markoviana da seqüência deve ser verificada mediante o uso de matrizes tabuladas segundo mudanças do perfil granulométrico. Deve-se enfatizar que a estatística de teste, aqui apresentada, aplica-se somente a matrizes nas quais todos os elementos são positivos e diferentes de zero. As limitações estatísticas impostas por elementos nulos podem ser superadas pela omissão de um grau de liberdade para cada zero. Assim, as hipóteses são definidas como se segue:

H_0: *A probabilidade do (i)-ésimo grupo granulométrico ser seguido pelo (j)-ésimo é a mesma de ser seguido por outra granulometria qualquer, ou seja, na deposição/erosão, as várias granulometrias independem umas das outras.*

H_1: A probabilidade do (i)-ésimo grupo granulométrico ser seguido pelo (j)-ésimo é diferente de ser seguido por uma outra granulometria qualquer, ou seja, no processo de deposição/erosão em determinado local depende daquele imediatamente anterior.

A estatística de teste é assim definida, segundo Anderson & Goodman (1956):

$$-2\log\lambda = 2\sum_{i,j=1}^{m} n_{ij}\log\left[\frac{\hat{p}_{ij}}{\hat{p}_j}\right].$$

Esta estatística de teste para averiguar a propriedade markoviana, segue uma distribuição *qui-quadrada* com $(m-1)^2$ graus de liberdade, onde m é o número de estágios granulométricos, n_{ij} é a freqüência de transições na (ij)-ésima cela, $\hat{p}_{ij}$ é a probabilidade de transições na (ij)-ésima cela e $\hat{p}_j$ é a probabilidade total para a (j)-ésima coluna. Desta forma, a um nível de significância α preestabelecido, rejeita-se a hipótese de que os eventos sejam independentes se o valor calculado para $-2\log\lambda$ exceder ao valor tabelado da χ^2 para $(m-1)^2$ graus de liberdade.

Em nosso estudo, o valor tabelado é obtido a partir da *qui-quadrado* com 25 graus de liberdade e nível de significância de 5%. De acordo com a região de rejeição especificada através do nível de significância, rejeita-se a hipótese nula (H_0) se nos baseamos nos dados apresentados (Tabela 1).

Embora os processos markovianos possam abranger desde os relativamente simples e regulares Campos de Markov de primeira ordem até modelos complexos com memórias que se estendem por dois ou mais intervalos no espaço, o *teste exato* aqui abordado encontra aplicação apenas para o caso de verificação da propriedade markoviana com memória de apenas um passo.

	χ^2_{Calc}	χ^2_{Tab}	χ^2_{Tab}
L1	69.645	37.653 (25gl)	26.29 (16gl)
L2	58.245	37.653 (25gl)	19.65 (11gl)
L3	63.177	37.653 (25gl)	26.29 (16gl)

Tab. 1: Teste estatístico para a verificação da propriedade de dependência de primeira ordem.

Anderson & Goodman (1956) apresentam outros testes que são aplicáveis às cadeias de ordem mais elevadas. Assim, foram também implementados em S-Plus™ testes onde investigou-se a hipótese de que as observações, em pontos sucessivos no espaço unidimensional, segundo uma certa orientação determinada sob um linha, fossem provenientes de um Campo de Markov de *segunda ordem*. As hipóteses são definidas como se segue:

H_0: A probabilidade do (i)-ésimo grupo granulométrico ser seguido pelo (j)-ésimo e pelo (j+1)-ésimo grupo é a mesma de ser seguido por duas outras granulometrias quaisquer, ou seja, no processo de deposição/erosão em determinado local independe daqueles dois imediatamente anteriores.

H_1: A probabilidade do (i)-ésimo grupo granulométrico ser seguido pelo (j)-ésimo e pelo (j+1)-ésimo é diferente de ser seguido por duas outras granulometrias quaisquer, ou seja, no processo de deposição/erosão em determinado local depende dos dois imediatamente anteriores.

Resultaram-se deste teste, *p-valores* muito altos, rejeitando-se, desta forma, a hipótese de que o processo são seja somente de gênese de primeira ordem.

5. ANÁLISE CLÁSSICA

McLaren & Bowles (1985) propõem um modelo que demonstra como as distribuições granulométricas dos depósitos sedimentares mudam na direção de transporte utilizando parâmetros de granulometria. Nos métodos clássicos (McLaren, 1981; McLaren e Bowles, 1985), não é levado em consideração o grau de influência ou a interação entre vizinhos próximos.

Em 1981, McLaren propôs um método dedutivo que utiliza as mudanças espaciais dos parâmetros de granulometria com o objetivo de identificar fontes de sedimentação, dinâmica de transporte e ambientes de deposição, segundo algumas suposições físicas, chegando às seguintes conclusões:

- os sedimentos em transporte devem ser mais finos, com menor variância e mais negativamente assimétrico que os sedimentos na fonte {FB-};
- um depósito residual deve ser mais grosso, com menor variância e mais positivamente assimétrico que a fonte {CB+}.

McLaren e Bowles (1985) utilizam uma aproximação estatística, baseada no *Z-score* ($Z = \frac{X - \mu_X}{\sigma_X} \approx N(0,1)$), para determinar a direção de transporte.

Gao e Collins (1991) confirmaram as tendências anteriores de McLaren & Bowles (1985) e sugeriram mais dois indicativos de transporte: {FB+} ou {CB-}, o que significa grãos mais finos, melhor selecionado e positivamente assimétricos ou grãos mais grossos, melhor selecionado e negativamente assimétricos, respectivamente com relação a cada um dos perfis acima mencionados.

No âmbito de verificar tais informações, implementamos algoritmos em linguagem S-Plus™ com vias ao cálculo destas medidas sugeridas, para estudar a dinâmica granulométrica na determinação do transporte de sedimentos, ao nível de significância de $\alpha = 0.05$ ($Z_{tab} = 1.645$). De acordo com os resultados obtidos, pode-se observar que há indicativo do grão se tornar mais grosso, melhor selecionado e positivamente assimétrico {CB+} na direção norte para as linhas L1, L2 e L3 e na direção sul para as linhas L1 e L2 (Tabela 2a, Tabela 2b e Tabela 2c).

Linha 1		Casos			
N=117		*1*	*2*	*3*	*4*
Direção		F B -	C B +	F B +	C B -
Sul	N	6786	6786	6786	6786
	x	891	1228	203	423
	Z	1.57	**13.94***	-23.68	-15.61
Norte	N	6786	6786	6786	6786
	x	560	2650	334	399
	Z	-10.58	**66.13***	-18.87	-16.49
* Caso mais significativo ao nível de confiança de 5%					

Tab.2a: Resultados da análise do transporte de areia da Linha L1

Linha 2		Casos			
N=118		*1*	*2*	*3*	*4*
Direção		F B -	C B +	F B +	C B -
Sul	N	6903	6903	6903	6903
	x	429	1176	267	433
	Z	-15.79	**11.39***	-21.68	-15.64
Norte	N	6903	6903	6903	6903
	x	783	2207	812	692
	Z	-2.91	**48.92***	-1.85	-6.22
* Caso mais significativo ao nível de confiança de 5%					

Tab.2b: Resultados da análise do transporte de areia da Linha L2

Linha 3		Casos			
N=119		*1*	*2*	*3*	*4*
Direção		F B -	C B +	F B +	C B -
Sul	N	7021	7021	7021	7021
	x	822	790	344	300
	Z	-2.01	-3.16	-19.26	-20.84
Norte	N	7021	7021	7021	7021
	x	585	2965	218	885
	Z	-10.56	**75.32***	-23.80	0.26
* Caso mais significativo ao nível de confiança de 5%					

Tab.2c: Resultados da análise do transporte de areia da Linha L3

6. TESTE EXATO

Propõe-se, nesta etapa, levar em consideração a natureza estrutural do processo, isto é, estudá-lo sob a condição de que o fenômeno se determina localmente como um Campo de Markov unidimensional de primeira ordem. Assim, introduz-se a seguinte metodologia:

- *n* amostras,
- $N = (n\text{-}1)$ vetores que podem exibir uma tendência sugerindo transporte numa direção preestabelecida. Estes são obtidos comparando-se duas amostras vizinhas, na seqüência, a cada vez. Portanto, *N* é o número total de pares unidirecionais possíveis.

São obtidos vinte e sete casos possíveis, dos quais identifica-se oito favoráveis indicativos de tendências equiprováveis, mas apenas quatro destas nos fornecem informações verossímeis com relação ao transporte de sedimentos (Gao & Collins, 1991).

Um teste exato para proporções é aplicado, ao nível de significância α (Dudewicz e Mishra, 1988), para as seguintes hipóteses:

H_0: $p_i = p_0 = 0.125$ *(não existe direção preferencial concernente ao caso "i"),*

i=1{FB-},2{CB+},3{FB+} ou 4{CB-};

H_1: $p_i > p_0 = 0.125$ *(o transporte está ocorrendo em pelo menos uma direção preferencial).*

Desta forma, determina-se um teste onde rejeita-se H_0, ao nível de significância α, se e somente se $P(X_i \geq j) \leq \alpha$ para um determinado "*j*". Neste caso, a variável aleatória X_i é o número observado de pares representando um caso particular "*i*" (apenas uma das quatro configurações acima nos determinariam um padrão de informação condicional) em uma das direções preestabelecidas. Assim, $X_i \sim B(n, p_0)$ sob H_0, uma vez que as restrições de dicotomia, probabilidade constante e independência são facilmente verificadas, através da própria gênese da variável direcional. Assim, se *n* for o número de amostras na seqüência e "*j*" for o ponto de corte, a probabilidade de se rejeitar H_0 é descrita por: $P_p[X_1 + \ldots + X_n \geq j] = \sum_{k=j}^{n} \binom{n}{k} p^k (1-p)^{n-k}$.

No âmbito de abordar numericamente tais informações, implementamos, neste contexto, algoritmos em linguagem S-Plus™, ao nível de significância de $\alpha = 0.05$ com a obtenção do valor crítico "*j*" (ponto de corte) igual a 21 tanto para *n* igual a 117 quanto para 118 e 119. Através dos resultados observou-se que nas três linhas houve o mesmo indicativo de transporte tanto na direção norte como na sul, pois existe uma tendência do grão se tornar mais grosso, melhor selecionado e positivamente assimétrico, sugerindo um ambiente granulométrico {CB+} para todas as linhas em ambas as direções - indicativo de um ambiente caótico de deposição, (Tabela 3a, Tabela 3b e Tabela 3c).

Linha1	*j*=21	Casos			
Direção		F B -	C B +	F B +	C B -
sul	n	117	117	117	117
	x	11	**27***	13	12
norte	n	117	117	117	117
	x	10	**29***	14	11
* Caso mais significativo ao nível de confiança de 5%					

Tab. 3a: *Método Exato Estrutural* para a análise do transporte de sedimentos da linha L1.

Linha36	*j*=21	Casos			
Direção		F B -	C B +	F B +	C B -
sul	n	118	118	118	118
	x	13	**23***	11	12
norte	n	118	118	118	118
	x	13	**24***	14	12
* Caso mais significativo ao nível de confiança de 5%					

Tab. 3b: *Método Exato Estrutural* para a análise do transporte de sedimentos da linha L2.

Linha52	*j*=21	Casos			
Direção		F B -	C B +	F B +	C B -
sul	n	119	119	119	119
	x	10	**23**	13	14
norte	n	119	119	119	119
	x	10	**25***	12	15
* Caso mais significativo ao nível de confiança de 5%					

Tab. 3c: *Método Exato Estrutural* para a análise do transporte de sedimentos da linha L3.

Assim, obtém-se um *teste exato estrutural* robusto e eficiente (tempo de cálculo aproximadamente dez vezes menor) que nos fornece resultados análogos ao teste clássico proposto por McLaren e Bowles (1985).

6.1 Análise utilizando a informação de dependência espacial entre as amostras

A partir da modelagem dos variogramas foi possível determinar o grau de dependência espacial entre as amostras. Os modelos que melhor se ajustaram aos dados analisados, referentes às três linhas foram o modelo com dependência espacial esférica e exponencial. Conforme Figura 3a, Figura 3b e Figura 3c, pode-se afirmar que há uma dependência espacial entre as unidades amostrais que estão a uma distância inferior a 4 *lags* (vizinhos consecutivos) nas linhas L1, L2 e L3, para cada um dos parâmetros.

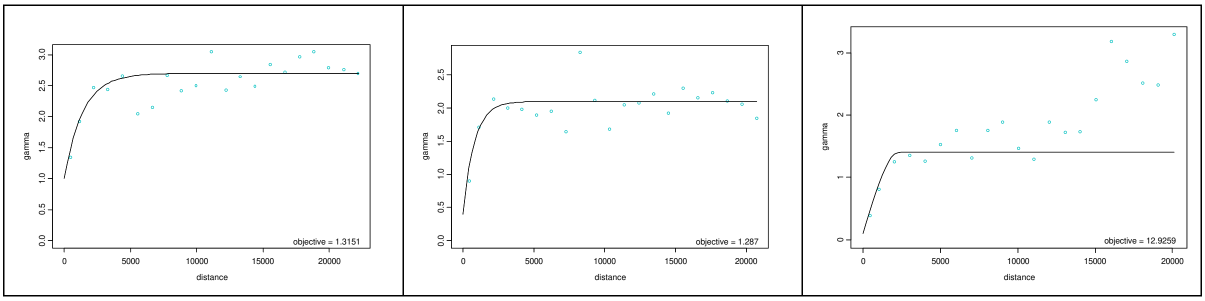

Fig. 3a: *Variogramas* da Média para as linhas L1, L2 e L3.

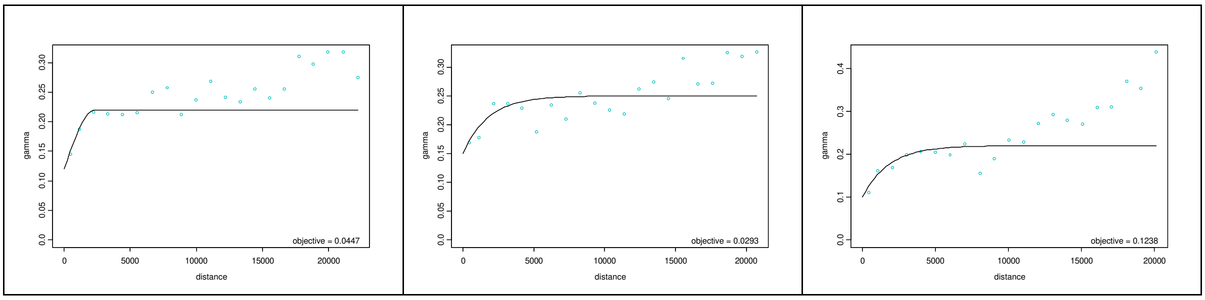

Fig. 3b: *Variogramas* da Seleção para as linhas L1, L2 e L3.

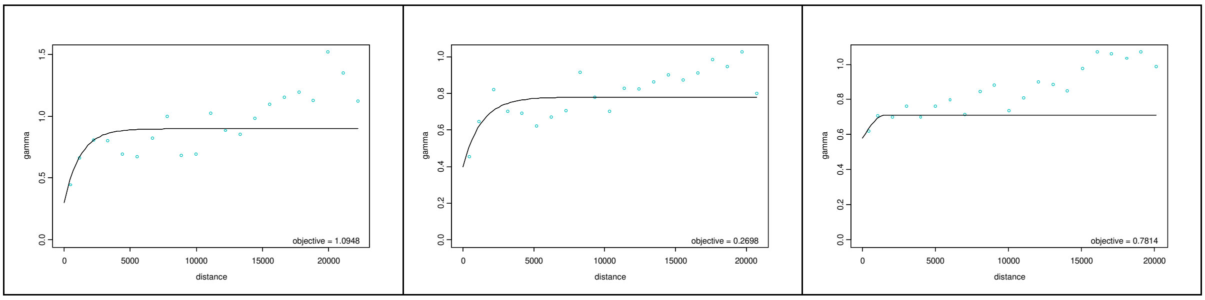

Fig. 3c: *Variogramas* da Assimetria para as linhas L1, L2 e L3.

Uma vez realizada a *krigagem* unidimensional (Tabela 4a, Tabela 4b e Tabela 4c), corrobora-se com veemência, os resultados do *teste exato estrutural* de primeira ordem. Este resultado já era esperado, uma vez que pela definição de *zona de influência*, se o grau de dependência espacial linear é k, então todos os vizinhos de ordem $l \leq k$ contribuem (de forma construtiva, *i. e.*, o sinal do ponderador é positivo) na determinação do operador interpolador.

Linha1	*j*=39	Casos			
Direção		F B -	C B +	F B +	C B -
sul	N	233	233	233	233
	X	14	**75***	8	20
norte	N	233	233	233	233
	X	14	**75***	8	18
* Caso mais significativo ao nível de confiança de 5%					

Tab. 4a: *Método Exato Estrutural Krigado* para a análise do transporte de areia da linha L1.

Linha2	*j*=40	Casos			
Direção		F B -	C B +	F B +	C B -
sul	N	235	235	235	235
	X	10	**62***	16	28
norte	N	235	235	235	235
	X	17	**62***	17	22
* Caso mais significativo ao nível de confiança de 5%					

Tab. 4b: *Método Exato Estrutura Krigadol* para a análise do transporte de areia da linha L2.

Linha3	j=40	Casos			
Direção		F B -	C B +	F B +	C B -
sul	N	237	237	237	237
	X	20	**66***	12	15
norte	N	119	119	119	119
	X	24	**70***	14	15
* Caso mais significativo ao nível de confiança de 5%					

Tab. 4c: *Método Exato Estrutural Krigado* para a análise do transporte de areia da linha L3

7. COMPONENTES PRINCIPAIS

Afim de determinar um índice que associe as medidas (média, seleção e assimetria) envolvidas na informação sobre a direção do transporte de sedimentos, foi utilizada a técnica de estatística multivariada denominada *análise de componentes principais*. Esta metodologia tem como objetivo explicar a estrutura de variância e covariância de um vetor aleatório, composto de p variáveis aleatórias, através da construção de $k \leq p$ combinações lineares independentes das variáveis originais. As componentes principais construídas a partir da matriz de covariância são determinadas pelas variáveis de maior variância, sendo portanto, de pouca utilidade nos casos em que existe uma grande diferença entre as variâncias das variáveis. Este problema é eliminado pela construção das componentes principais utilizando-se a matriz de correlação, que pode ser vista como a matriz de covariâncias das variáveis padronizadas pelas respectivas médias e desvios padrão (Johnson & Wichern, 1998).

Observando a variância explicada, ou seja o quanto os parâmetros granulométricos estão sendo representados através da componente principal, foi selecionada a componente (índice) que obteve o maior valor em cada linha. Não obstante, foram calculados os escores, valores obtidos através da componente principal, para todas as linhas selecionadas. Em seguida, utilizando o método de *McLaren* para determinar as quatro possíveis tendências segundo Gao & Collins (1991), foram identificados o número da amostra, o escore, a média, a variância e a assimetria para cada par de amostra onde havia uma possível tendência. A seguir, foi calculada a razão entre os escores do par, seguindo-se os critérios descritos a seguir:

- Se os escores aumentavam de uma amostra para outra e se os sinais dos mesmos eram iguais, era utilizada a seguinte expressão para o cálculo da razão:

$$Razão = \frac{máx(abs(escore1), abs(escore2))}{mín(abs(escore1), abs(escore2))};$$

- Se os escores diminuíam de uma amostra para outra e se os sinais dos mesmos eram iguais, era utilizada a seguinte expressão para o cálculo da razão:

$$Razão = -\left[\frac{máx(abs(escore1), abs(escore2))}{mín(abs(escore1), abs(escore2))}\right];$$

- Se os escores aumentavam de uma amostra para outra e se os sinais dos mesmos eram diferentes, era utilizada a seguinte expressão para o cálculo da razão:

$$Razão = \frac{máx(escore1, escore2) - mín(escore1, escore2)}{mín(abs(escore1), abs(escore2))};$$

- Se os escores diminuíam de uma amostra para outra e se os sinais dos mesmos eram diferentes, era utilizada a seguinte expressão para o cálculo da razão:

$$Razão = -\left[\frac{máx(escore1, escore2) - mín(escore1, escore2)}{mín(abs(escore1), abs(escore2))}\right].$$

Calculada a razão dos escores através de um algoritmo implementado em S-Plus™, foram construídos intervalos de confiança para a média da razão para cada tendência. Observando aqueles casos que obtiveram maior número de pares, foram analisados seus respectivos intervalos e, aqueles que obtiveram, em média, uma razão dos escores alta, foi considerado como uma das tendências da linha analisada. Para confirmação do resultado obtido pode-se observar a variância explicada através da componente principal da linha estudada, se for alta o resultado tem maior confiabilidade. Assim:

- para a linha L1: n=117, $CP1 = -0.596\overline{X} - 0.545S + 0.590S_K$. O subconjunto contendo as variáveis <u>seleção</u> e <u>assimetria</u> é o mais importante, pois juntas obtiveram R^2=96,1%. Observou-se que os intervalos mais concentrados são referentes ao caso {CB+} tanto na direção norte (1_) quanto na sul (2_). Há indicativo de transporte sedimentar nesses dois casos. (R^2=81,8%) – Figura 4a;

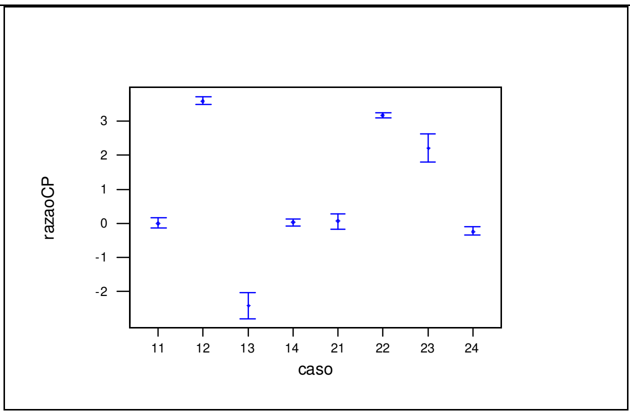

Fig. 4a: Intervalo de Confiança para a linha L1 – *Componentes Principais*.

- para a linha L2: n=118 , $CP1 = -0.586\overline{X} - 0.548S + 0.598S_K$. O subconjunto contendo as variáveis seleção e média é o mais importante, pois juntas obtiveram R^2=92,6%. Observou-se que os intervalos mais concentrados são referentes ao caso {C B +} tanto na direção norte (1_) quanto na sul (2_). Há indicativo de transporte sedimentar nesses dois casos. (R^2=74,0%) – Figura 4b;

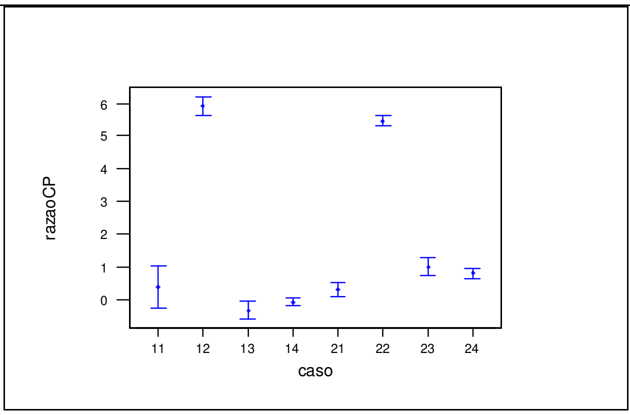

Fig. 4b: Intervalo de Confiança para a linha L2 – *Componentes Principais*.

- para a linha L3: n=119, $CP1 = -0.606\overline{X} - 0.550S + 0.576S_K$. O subconjunto contendo as variáveis seleção e assimetria é o mais importante, pois juntas obtiveram R^2=94,3%. Observou-se que os intervalos mais concentrados são referentes ao caso {CB+} tanto na direção norte (1_) quanto na sul (2_) e o caso {CB-} na direção norte, (1_) sendo que o caso {CB+} na direção norte(1_) possui média mais alta em relação aos outros casos citados. Há indicativo de transporte sedimentar no caso {CB+} na direção norte (1_). (R^2=77,0%) – Figura 4c.

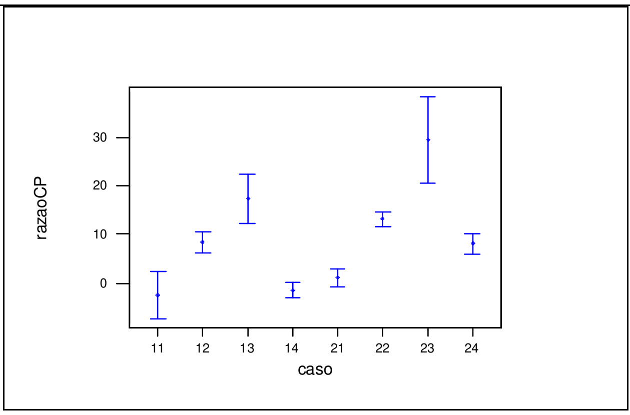

Fig. 4c: Intervalo de Confiança para a linha L3 – *Componentes Principais*.

7.1 Análise utilizando a informação de dependência espacial entre as amostras

A partir da modelagem dos variogramas, foi possível determinar o grau de dependência espacial entre as amostras. Os modelos que melhor se ajustaram aos dados analisados, referentes às três linhas, foram os modelos com dependência espacial esférica e exponencial. A partir da modelagem dos variogramas foi possível determinar o grau de dependência espacial entre as amostras. Assim, pode-se afirmar que há uma dependência espacial entre as unidades amostrais que estão a uma distância inferior a 6 *lags* (vizinhos consecutivos) na linha L1 – Figura 5a, menor que 5 *lags* na linha L2 – Figura5b e menor que 4 *lags* na linha L3 – Figura 5c.

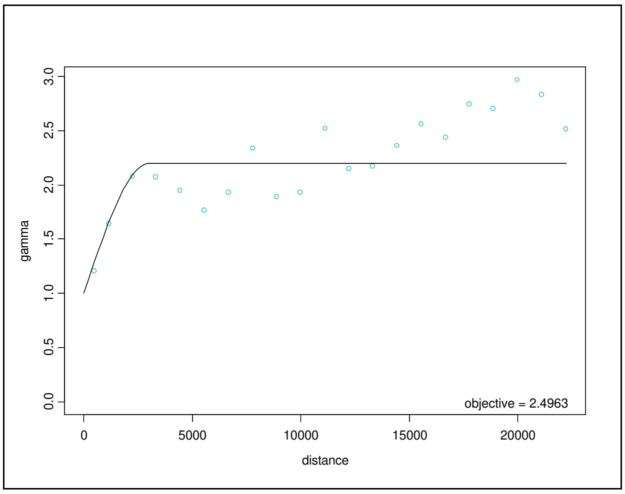

Fig. 5ª: Variograma da CP para a linha L1

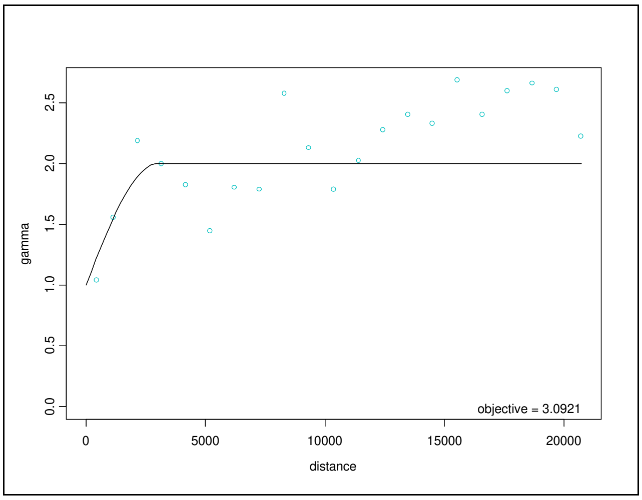

Fig. 5b : Variograma da CP para a linha L2

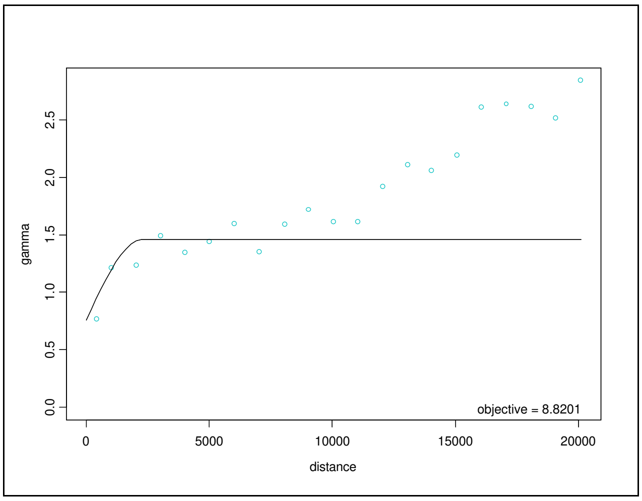

Fig. 5c : Variograma da CP para a linha L3

Uma vez realizada a *krigagem* unidimensional da CP, observa-se, através do gráfico da Figura 6a, que os intervalos mais concentrados em torno da maior média da linha L1 são aquelas do caso {CB+} tanto na direção norte (1_) quanto na sul (2_). Logo pode-se dizer, com 81,8% (R^2) de confiança, que há indicativo de depósito residual. Através do gráfico da Figura 6b, pode-se observar que os intervalos mais concentrados em torno de uma alta média da linha L2 são aqueles do caso {CB+} tanto na direção norte (1_) quanto na sul (2_), logo pode-se dizer com 74% (R^2) de confiança que há indicativo de depósito residual nesta linha. Pela Figura 6c, pode-se observar que os intervalos com médias maiores são aqueles do caso {CB+}, tanto na direção norte (1_) quanto na sul (2_), logo pode-se dizer com 77% (R^2) de confiança que há indicativo de deposição residual.

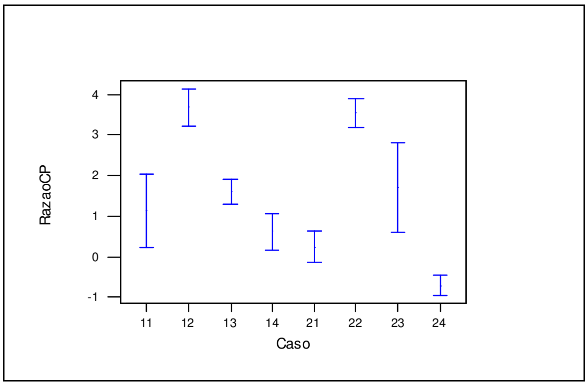

Fig. 6ª: Intervalo de Confiança da CP Krigada para a linha L1.

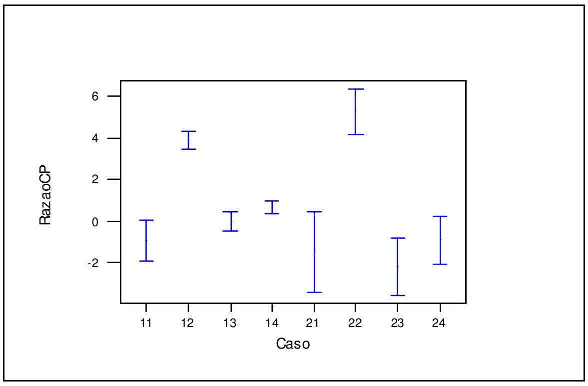

Fig. 6b: Intervalo de Confiança da CP Krigada para a linha L2.

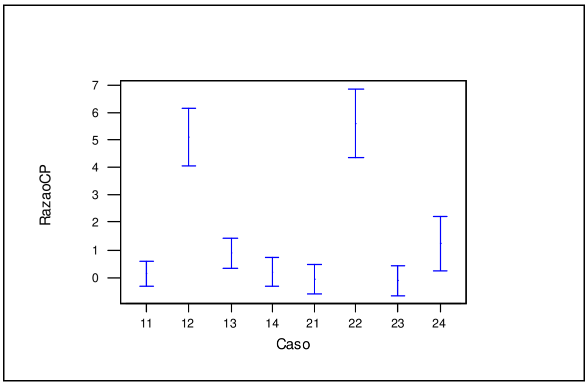

Fig. 6c: Intervalo de Confiança da CP Krigada para a linha L3

8. SÍNTESE

Observando a Tabela 5a, referentes aos resultados do método de *McLaren*, do *teste exato* e *componentes principais* há indicativo de transporte de areia tanto na direção norte como na sul, pois existe uma tendência do grão se tornar mais grosso, melhor selecionado e positivamente assimétrico em ambas as direções. Pode-se observar também que no método de McLaren há indicativo de transporte de areia na direção sul do grão ficar mais grosso, melhor selecionado e negativamente assimétrico.

McLaren		Teste Estrutural		Componentes Principais	
Tendência	Direção	Tendência	Direção	Tendência	Direção
CB+	Sul	CB+	Sul	CB+	Sul
CB+	Norte	CB+	Norte	CB+	Norte
CB-	Sul				

Tab.5a: Síntese dos resultados da análise do transporte de areia da linha L1.

Observando a Tabela 5b, referentes aos resultados do método de *McLaren*, do *teste exato* e *componentes principais* há indicativo de transporte de areia tanto na direção norte como na sul, pois existe uma tendência do grão se tornar mais grosso, melhor selecionado e positivamente assimétrico em ambas as direções. Pode-se observar também que no método de McLaren há indicativo de transporte de areia na direção sul do grão ficar mais grosso, melhor selecionado e negativamente assimétrico.

McLaren		Teste Estrutural		Componentes Principais	
Tendência	Direção	Tendência	Direção	Tendência	Direção
CB-	Sul	FB-	Sul	CB-	Sul
FB-	Norte	FB-	Norte	FB-	Norte
		FB+	Norte		

Tab.5b: Síntese dos resultados da análise do transporte de areia da linha L2.

Observando a Tabela 5c, referentes aos resultados do método de *McLaren* e *componentes principais* há indicativo de transporte de areia na direção sul, pois existe uma tendência do grão se tornar mais grosso, melhor selecionado e negativamente assimétrico. Observa-se também que há uma tendência na direção norte do grão ficar mais fino, melhor selecionado e negativamente assimétrico. E, através do *teste exato* há indicativo de transporte de areia tanto na direção norte como na sul, pois existe uma tendência do grão se tornar mais fino, melhor selecionado e negativamente assimétrico em ambas e também se tornar mais fino, melhor selecionado e positivamente assimétrico na direção norte.

McLaren		Teste Estrutural		Componentes Principais	
Tendência	Direção	Tendência	Direção	Tendência	Direção
CB+	Sul	CB-	Norte	CB+	Sul
CB-	Sul			CB-	Sul
FB+	Norte			FB+	Norte
CB-	Norte			CB-	Norte

Tab.5c: Síntese dos resultados da análise do transporte de areia da linha L3.

Em resumo, nota-se através do método de *McLaren* e por *componentes principais*, que existe um indicativo de transporte de areia na direção sul, pois há uma tendência do grão se tornar mais grosso, melhor selecionado e positivamente assimétrico e mais grosso, melhor selecionado e negativamente assimétrico. Pode-se notar também que há uma tendência na direção norte do grão ficar mais fino, melhor selecionado e negativamente assimétrico. E, através do *teste exato* observa-se que há indicativo de transporte de areia na direção norte pois existe uma tendência do grão se tornar mais grosso, melhor selecionado e negativamente assimétrico.

9. CONCLUSÕES E PERSPECTIVAS

De acordo com os resultados apresentados, podemos concluir que as linhas de seqüências granulométricas estudadas apresentam a característica markoviana de dependência de primeira ordem, dando consistência às suposições do modelo estrutural. Observe que os resultados obtidos podem ter sido influenciados pela classificação granulométrica determinada. Esta deve levar em conta o conhecimento geológico do fenômeno, assim como satisfazer determinadas suposições teóricas estatísticas (como a suposição de normalidade).

Quanto ao teste exato que foi implementado, sabe-se que abordagens devem ser realizadas no que concerne à análise de seu tamanho e poder, para que se possa efetivamente comparar o método clássico ao estrutural. Esta análise, que pode ser realizada lançando-se mão da técnica de *bootstrap*.

Supondo que a memória nas linhas seja de tamanho unitário e observando a validade de todas as suposições de uma distribuição binomial, utilizou-se um *teste exato estrutural* para determinar se uma tendência do transporte de sedimentos ao longo de uma linha amostral era significativa. Os resultados obtidos através desta metodologia não foram muito similares ao método de *McLaren*, pois em algumas linhas foram encontrados menos casos possíveis e em outras foram encontrados diferentes casos, o que já era de se esperar, uma vez que a zona de influência determinada adotada em nossa análise é de memória *um* e não de tamanho *n* (número de elementos da linha).

Através da análise de *componentes principais* foi determinado um índice possibilitando a obtenção de informações similares àquelas fornecidas pelo método de *McLaren* sobre a direção de transporte de sedimentos. Foram encontradas, na maioria das linhas, as mesmas tendências possíveis e, em alguns casos encontradas menos tendências indicativas de transporte, o que nos sugere uma maior precisão do método proposto em relação ao método clássico. Nossa perspectiva é estudar o fenômeno aqui descrito, porém sob sua gênese bidimensional, implementando variantes do método de Gao & Collins (1991). Estas modificações consistem em aplicar as técnicas de *krigagem* pontual para interpolar os parâmetros granulométricos indicadores de transporte, individualmente, e utilizar a metodologia de análise de componentes principais para associar estes parâmetros e determinar a tendência do transporte do sedimento, como uma extensão para o caso bidimensional (2D).

10. AGRADECIMENTOS

Agradecemos à GeoSea Consulting (Canadá), que através do Dr. Patrick McLaren, nos forneceu o arquivo de dados aqui analisados.

11. BIBLIOGRAFIA

- Anderson, T. W. & Goodman, L. A. (1956). Statistical Inference About Markov Chains, *Annals Mathematical Statistics*, Vol. 18, p. 89 – 110.

- Davis, J. C. (1986). *Statistics and Data Analysis in Geology*. 2nd.ed. John Wiley & Sons.

- Dudewicz, E. J. & Mishra, S. N. (1988). *Modern Mathematical Statistics.* John Willey & Sons, New York.

- Gao, S. & Collins, M. (1991). A critique of the ‘ McLaren method’ for defining sediment transport paths – discussion, *Journal of Sedimentary Petrology*, Vol. 61, p. 143 – 146.

- Johnson, R. A. & Wichern, D. W. (1998). *Applied multivariate statistical analysis*. 2nd.ed, Prentice Hall, New Jersey.

- McLaren, P. (1981). An Interpretation of Trends in Grain Size Measures, *Journal of Sedimentary Petrology*, Vol. 51, p. 611 – 624.

- McLaren, P. & Bowles, D. (1985). The Effects of Sediment Transport on Grain-Size Distributions, *Journal of Sedimentary Petrology*, Vol. 55, p. 457 – 470.