ANÁLISE ESTATÍSTICA ESPACIAL APLICADA A CONSTRUÇÃO DE SUPERFÍCIES BATIMÉTRICAS

Thyago de Almeida SILVEIRA¹, Jose Luiz PORTUGAL¹, Lucilene Antunes Correia Marques de SÁ¹, Saulo Roberto de Oliveira VITAL¹

(1) Universidade Federal de Pernambuco. Centro de Tecnologia e Geociências. Departamento de Engenharia Cartográfica - DECart. Av. Acadêmico Hélio Ramos, s/n - 2º andar, Cidade Universitária. CEP: 50740-530. Recife, Pernambuco. Endereço eletrônico: thyago.silveira@gmail.com.

> Introdução Localização da área de estudo Revisão teórica Análise espacial de superfície Materiais e métodos Aquisição das amostras batimétricas Implementação dos interpoladores Análises estatísticas dos resultados Criação da superfície tridimensional Resultados e discussões Aquisição das amostras batimétricas Implementação das interpolações Análises estatísticas dos resultados Criação da superfície tridimensional Conclusões Referências bibliográficas

RESUMO - Nas ultimas décadas os Sistemas de Informações Geográficas (SIG) tem sido incorporados no gerenciamento de zonas costeiras e oceânicas. Esses sistemas permitem a criar mapas e cenas de visualização tridimensional, de forma a obter uma produção mais realista da morfologia do fundo do mar. Face às inúmeras possibilidades de aplicação de modelos de interpoladores em amostras pontuais resultantes de levantamentos batimétricos, torna-se difícil escolher qual interpolador irá gerar a superfície que mais se aproxime da superfície real. Nesse trabalho, é apresentado um estudo estatístico que procurou estabelecer qual interpolador reproduz mais fidedignamente à morfologia da plataforma continental interna, adjacente aos municípios de Paulista, Olinda e Recife. Foram implementadas as interpolações Inverso da Distância Ponderada, Polinomial Local, Funções de Base Radial, Polinomial Global e Krigagem. Os resultados das análises estatísticas indicaram que não existem diferenças significativas entre o Polinomial Local, Funções de Base Radial e Krigagem. Portanto, qualquer um desses três métodos pode ser recomendado. Entretanto, por ser o único interpolador capaz de minimizar tendências e espacialisar a distribuição da variância dos pontos interpolados sobre uma superfície, optou-se por selecionar a Krigagem como o interpolador mais indicado na representação tridimensional da área em estudo.

Palavras-Chave: Sistemas de Informação Geográfica (SIG), Levantamentos Batimétricos, Modelos de Interpoladores, Análises Espaciais de Superfícies.

ABSTRACT - In the last decades, the Geographic Information Systems (GIS) have been incorporated in the management of coastal and oceanic areas. Such GIS allow one to create maps and three-dimensional preview scenes, so it is possible to have a more realistic scenario of the seafloor morphology. Due to the diversity of possibilities concerning the application of interpolation models in specific samples which result from bathymetric surveys, it is hard to choose the interpolator which will create the most realistic surface. This study presents a statistic research which aims at identifying which interpolator makes possible the best recreation of the internal continental platform, by the towns of Paulista, Olinda and Recife, in Brazil. In order to do so, it were implemented the following interpolations: Inverse Distance Weighted, Polynomial Local, Radial Basis Functions, Polynomial Global and Kriging. The results showed that there are no meaningful differences among the Polynomial Local, Radial Basis Functions and Kriging. Therefore, the three methods may be recommended. However, as the only interpolators capable of minimizing trends and spatialize the distribution of errors on a surface, the Kriging is the most appropriate interpolator to represent a three-dimensional scenario of the studied area.

Key words: Geographic Information System (GIS), Bathymetric Surveys, Interpolations Models, Surface Spatial Analysis.

INTRODUÇÃO

O planejamento de ações que envolvam o ambiente costeiro é um desafio encontrado em todo mundo. De acordo com Dias et al. (2007) a procura por habitações próximas ao litoral, a ampliação e diversificação das áreas industriais, o crescimento do turismo litorâneo, e a modificação climática em curso, converteram a faixa do litoral em uma zona de grande complexidade cuja gestão harmônica é muito difícil, necessitando do desenvolvimento de ações de Gerenciamento Costeiro que integrem a tentativa de compatibilização de todos os fatores aludidos, de modo que a exploração e utilização destas áreas seja feita de forma equilibrada e sustentável, visando preservar suas potencialidades para as gerações futuras.

Segundo Absher et al. (2002) zona costeira corresponde à faixa de transição entre o domínio continental e o domínio marinho. No Brasil, a zona costeira apresenta 7.367km de extensão, e é considerada Patrimônio Nacional pela Constituição Federal (BRASIL, 1988) em seu artigo 225, §4º. A utilização deve ser feita na forma da lei, dentro de condições que assegurem a preservação do meio ambiente, inclusive quando ao uso dos recursos naturais.

O extenso litoral brasileiro é composto por diversos ecossistemas, tais como manguezais, restingas, campos de dunas, estuários, recifes de coral, marismas, praias, falésias, costões rochosos, entre outros (Serafini, 2010), e utilizado para finalidades diversas, como lazer, pesca e transportes.

Nesse sentido, conhecer o comportamento morfodinâmico e hidrodinâmico desses ambientes é fundamental e, necessário quando leva-se em consideração a segurança dos banhistas, o manejo de dunas, os critérios para exploração de areia, reconstituição de praias, delimitação submersa de áreas protegidas, entre outras ações que estão intimamente ligados as características geomorfológica deste meio (Nicolodi & Toldo Jr, 2003).

Para tanto, é necessário ter dados confiáveis que possam revelar o comportamento da zona costeira, a fim de proporcionar aos usuários destes recursos, segurança durante suas utilizações. Contudo, devido à própria dinâmica natural dos ecossistemas marinhos, aliado aos processos provocados por intervenções antrópicas, tais quais como a variação das marés ou uma construção portuária (Krug & Noemberg, 2005), a morfologia do fundo oceânico pode sofrer alterações significativas ao longo dos anos.

Tais mudanças podem ser detectadas por meio dos Levantamentos Batimétricos, que associam a posição da embarcação na superfície da água, e que são obtidas a partir da utilização do *Global Position System* (GPS); com a profundidade atingida naquele exato momento, obtida usando o ecobatímetro de mono-feixe (Ayres & Neto, 2004).

O produto resultante é uma malha de pontos tridimensionais que, por si só, não é capaz de gerar diretamente a superfície do fundo do mar imageado. Para construir a superfície que representa tal morfologia, é necessário empregar técnicas de interpolação.

Os interpoladores são funções matemáticas usadas na construção de superfícies contínuas a partir de um conjunto de pontos coletados (Burrough & Mcdonnell, 1998), e são utilizados para a densificação de uma amostra que não cobre todo o domínio de interesse. Atualmente conhecidos diversos modelos são de interpoladores, cada com um suas particularidades e características, e diante de tantas opções para interpolar dados pontuais, tais como dados batimétricos, torna-se difícil escolher qual interpolador irá gerar a superfície que mais se aproxime da superfície real.

Nesse contexto, o objetivo deste trabalho foi estabelecer com base em indicadores estatísticos, o modelo de interpolador que reproduziu mais fidedignamente a morfologia da plataforma continental interna, adjacente aos municípios de Paulista, Olinda e Recife, a partir de um levantamento batimétrico.

LOCALIZAÇÃO DA ÁREA DE ESTUDO

A área piloto para este estudo compreende os municípios de Paulista, Olinda e Recife (Figura 1). Essa região tem suportado um grande crescimento, sendo a área que apresenta maior concentração de atividades econômicas, industriais, de recreação e turismo, e conseqüentemente maior agrupamento dos problemas delas decorrentes (MAI, 2004).



Figura 1. Paulista, Olinda e Recife, municípios pertencentes à área de estudo deste trabalho.

Com o avanço da erosão costeira, vários projetos de contenção foram executados, entretanto, a maioria deles não respeitou os limites requeridos pela dinâmica natural da linha de costa, fato que intensificou ainda mais o processo erosivo na área. Para entender essa problemática, foi criado o Projeto MAI, cujo objetivo foi o levantamento de informações sobre a geologia, geofísica, morfologia costeira, entre outros processos ocorridos ao longo da zona costeira da Região Metropolitana de

Análise espacial de superfície

As análises espaciais de superfícies compreendem superfícies criadas a partir de um conjunto de amostras de pontos espacialmente distribuídas, que representam o comportamento de, por exemplo, dados geológicos, altimétricos e batimétricos (Bailey & Gatrell, 1995). Recife (RMR). Dentro dessa ótica, a obtenção das informações sobre a batimetria tornou-se pertinente, uma vez que a caracterização morfológica deste meio pode ser apresentada por uma superfície interpolada, entretanto, tal representação deve ser construída da forma a modelar a superfície de forma mais real possível e, esse procedimento pode ser realizado através da avaliação das características dos interpoladores com а aplicação de testes estatísticos.

REVISÃO TEÓRICA

A conversão de dados que se encontram na forma de pontos em uma representação na forma de grade regular, é feita com o uso de interpoladores, conforme pode ser visualizado na Figura 2.



Figura 2. Processo de Criação de superfícies usando Interpoladores. FONTE: Adaptado de Câmara, et al. (2004).

As análises espaciais ocorrem após a aquisição dos dados amostrais, e são compostas

das seguintes etapas: Análise da Normalidade dos Dados; Implementação dos Interpoladores;

Análises Estatística dos Resultados; e Criação da Superfície Tridimensional.

Segundo Field (2009), a análise acerca da normalidade dos dados é realizada a através da aplicação do teste de Kolmogorov-Smirnov (Teste de K-S), que compara escores de uma amostra, a escores de uma distribuição normal modelo de mesma média e variância dos valores encontrados na amostra.

De acordo com Sokal & Rohlf (1997), a estatística apropriada do teste é baseada na maior diferença absoluta entre a função de distribuição normal acumulada [$\hat{F}(z_i)$], que corresponde a proporção dos valores esperados menores ou iguais a x; e a freqüência relativa observada acumulada e ajustada [$F_{0,5}$], correspondendo a proporção dos valores observados menores ou iguais a x; em que D_{cal} é módulo do desvio máximo observado, conforme a Equação (1):

$$D_{cal} = \left| \widehat{F}_{(z_i)} - \widehat{F}_{(0,5)} \right| \quad (1)$$

sendo:

 $\widehat{F}_{(0,5)} = \frac{(i-0,5)}{n}$ onde: $\widehat{F}_{(-)} \rightarrow \infty$

 $\hat{F}_{(z_i)} \rightarrow$ refere-se a função de distribuição normal acumulada;

 $\hat{F}_{(0,5)} \rightarrow$ refere-se a freqüência relativa observada acumulada e ajustada;

 $i \rightarrow$ refere-se ao número da amostra;

 $n \rightarrow$ refere-se ao tamanho da amostra;

Em seguida compara-se o D_{cal} , com o D_{tab} , que é o desvio máximo tabelado, para um determinado intervalo de confiança. Quando o valor D_{cal} for maior que o valor crítico tabelado Dt ($D_{cal} > D_{tab}$), conclui-se que a característica em estudo da população não segue a distribuição normal; caso contrário ($D_{cal} < D_{tab}$), concluí-se que a amostra é normalmente distribuída.

O teste de Kolmogorov-Smirnov também pode ser interpretado pelo uso do *p-value*, que corresponde à significância do teste (Field, 2009). Se o teste é não significativo (*p-value* > 0,05) a amostra é normalmente distribuída, caso contrário, a amostra não é normalmente distribuída. Segundo Field (2009) quando o tamanho da amostra é grande, é comum o teste de Kolmogorov-Smirnov apresentar resultado significativo (*p*-value < 0,05), o que nem sempre é valido. Para sanar essa dúvida sugerese, a realização de análises gráficas da distribuição dos dados.

Caso se confirme a não normalidade da amostra, pode-se tentar normalizá-la empregando-se transformações do tipo (Field, 2009):

- Logarítmica (log (Xi)): ao extrair o logaritmo de um conjunto de números reduz-se a assimetria positiva da distribuição;
- Por radiciação (√X_i): ao tomar a raiz quadrada de um conjunto de valores, reduz-se os valores grandes aproximandoos do centro da distribuição;
- Recíproca (1/Xi): ao dividir por 1 está-se diminuindo o impacto dos grandes valores, de forma que eles ficaram próximos de zero.

A transformação logarítmica não pode ser aplicada a valores negativos ou zero, já a transformação por radiciação não pode ser aplicada para valores negativos. Para corrigir esses problemas, uma constante pode ser adicionada aos dados para torná-los maiores que zero.

A análise sobre a normalidade dos dados é importante no momento da decisão acerca da utilização dos testes estatísticos. Caso os dados sigam uma distribuição normal, é possível a aplicação de testes paramétricos, caso contrário, se os dados não forem normais, e não conseguirem ser normalizados pode ser aplicado testes não-paramétricos, ou ainda utilizar o erro, que segundo Field (2009), tende a apresentar um comportamento normal.

A implementação dos Interpoladores são os procedimentos usados na predição de valores de atributos em locais não conhecidos (incertos), a partir de medidas realizadas em pontos com locais conhecidos, para uma mesma área (Burrough & Mcdonnell, 1998). A utilização de interpoladores é necessária quando os dados disponíveis não cobrem o domínio de interesse completamente. A Figura 3 mostra os pontos utilizados em uma interpolação, e a superfície resultante sobrepostas. Pode-se identificar que a superfície infere a construção dos valores inexistentes entre os dois perfis.



Figura 3. Processo de Interpolação.

Burrough (1987) especifica que os interpoladores se dividem em três tipos: Modelos de Interpoladores Locais, Modelos de Interpoladores Globais, e Modelos de Interpoladores Geoestatísticos.

Para os Modelos de Interpoladores Locais cada ponto da superfície é estimado apenas a partir da interpolação das amostras mais próximas, onde predominam efeitos locais. Eles se subdividem em Inverso da Distância Ponderada (IDP), Polinomial Local e Funções de Base Radial (FBR)

O interpolador Inverso da Distância Ponderada é um interpolador local, e tem a característica de exato. O modelo permite a manipulação dos parâmetros de dimensões do raio de busca, o número de vizinhos a serem processados no cálculo e a potência a ser empregada na ponderação da distância (Estrada & Safriet, 2006).

Segundo Bajjali (2002), sua formulação matemática pode ser expressa pela Equação (2):

$$\widehat{Z}_o = \frac{\sum_{i=1}^n Z_i \frac{1}{d_i^p}}{\sum_{i=1}^n \frac{1}{d_i^p}},$$
 (2)

onde:

 $\hat{Z}_{o} \rightarrow \hat{e}$ o valor predito no ponto 0; $Z_{i} \rightarrow \hat{sao}$ os valores dos pontos amostrais vizinhos ao ponto desconhecido Z_{o} ;

 $d_i^{h} \rightarrow \acute{e}$ a distância que separa cada ponto amostral i ao ponto desconhecido 0;

 $n \rightarrow \text{é o número de vizinhos;}$

 $p \rightarrow e a pontencia adotada.$

O interpolador Polinomial Local possui formulação matemática de acordo com a Equação (3).

$$f \{(x, y)\} = \sum_{r+s \le p} (b_{rs}. x^r. y^s), \quad (3)$$

onde:

 $f \{(x, y)\} \rightarrow e o valor de elevação (Z), nas coordenadas (X, Y);$

 $b \rightarrow$ são os coeficientes do polinômio de grau *p*;

 $p \rightarrow \acute{e}$ ordem da superfície de tendência; $r \rightarrow \acute{e}$ o valor correspondente ao grau da primeira variável do polinômio;

 $s \rightarrow \acute{e}$ o valor correspondente ao grau do segunda variável do polinômio.

Neste modelo, são adaptadas "n" superfícies na região de trabalho, sendo estas definidas pelas dimensões do raio de busca (Burrough & Mcdonnell, 1998). O modelo permite a manipulação do grau do polinômio, do número de vizinhos a serem processados no cálculo, e dos pesos a serem atribuídos a cada ponderação alcançada na definição do grau do polinômio (Estrada e Safriet, 2006).

As Funções de Base Radial correspondem a um grupo de interpoladores chamados *Splines* que produzem superfícies suaves (Chin-Shung, et al. 2004). O princípio das *Splines* é minimizar a curvatura total da superfície, garantindo-se que a mesma contenha os pontos amostrais, configurando-se como um interpolador exato (Jakob & Young, 2006).

Nos Modelos de Interpoladores Globais na caracterização do fenômeno em estudo, predomina a variação em larga escala onde a variabilidade local não é relevante. 0 interpolador polinomial global é o representante dos Interpoladores Globais, e possui o mesmo modelo matemático do interpolador polinomial local. Segundo Michael e Triveloni (2006), o interpolador polinomial global é uma técnica que consiste no ajustamento de equações que representam à variação espacial de valores superfícies através de matemáticas (polinômios). Dessa forma, este interpolador tem como resultado uma superfície gradual, que muda de acordo com a definição do grau polinomial usado na interpolação dos dados.

JAKOB e YOUNG (2006) explicitam que interpolador polinomial global ajusta um polinômio à uma superfície toda, diferentemente do interpolador polinomial local que pode ajustar muitos polinômios tendo em vista particularidades de vizinhança.

Os Modelos de Interpoladores Geoestatísticos estabelecem estimativas probabilísticas da qualidade da interpolação. A Krigagem é o interpolador dessa classe, e é considerado como um estimador exato que considera tanto efeitos Locais como Globais em sua predição, ou seja, é função dos dados e de covariância espacial (Chaplot et al., 2006). Segundo Bailey & Gatrell (1995), este interpolador tem por objetivo identificar a correlação espacial existente entre os valores de um conjunto de amostras com seus valores interpolados, e o erro associado ao valor estimado.

De acordo com Burrough & Mcdonnell (1998) a Krigagem pode ser expressa pela Equação (4):

$$Z(s) = \mu(s) + \varepsilon'(s) + \varepsilon, \qquad (4)$$

onde:

 $Z(s) \rightarrow e o valor da função aleatória numa posição s;$

 $\mu(s) \rightarrow \text{é}$ uma função determinística que descreve a componente estrutural Z numa posição s;

 $\varepsilon'(s) \rightarrow \epsilon'$ é um termo estocástico

correlacionado com variação local;

 $\varepsilon^{"} \rightarrow \epsilon$ um ruído aleatório não

correlacionado, normalmente distribuído.

Tal formulação tem interpretação geométrica mostrada na Figura 4, que reproduz o comportamento das amostras correlatas, de acordo com a variação espacial existente entre o conjunto de valores de entrada e o conjunto dos dados interpolados (Burrough & Mcdonnell, 1998).



Figura 4. Variação espacial da variável regionalizada. FONTE: Adaptado de Burrough & Mcdonnell (1998).

Na Figura 4(a) visualiza-se o comportamento amostras que possuem valores da $\mu(s)$ constante, por sua vez, na Figura 4(b), as amostras representadas revelam o $\mu(s)$ com um comportamento tendêncioso.

Quando a função determinística [$\mu(s)$] é constante na região em estudo, os efeitos globais da amostra se tornam quase que inexistentes. Desse modo, a Equação 5 fica reduzida a determinação do termo

estocástico $\varepsilon'(s)$, que é obtido em função de um Variograma (Burrough & Mcdonnell, 1998).

Mello et al., (2005) definem o Variograma como sendo uma técnica de suporte a permite Krigagem, que representar quantitativamente a variação de um fenômeno regionalizado no espaço, e é definido pela Equação (5).



onde:

Figura 5. Variograma. FONTE: Adaptado de Burrough & Mcdonnell (1998).

Segundo Schaffrath et. al., (2007) no variograma são identificados os seguintes parâmetros:

- Alcance ou Range (A): materializa a distância onde as amostras apresentam-se correlacionadas espacialmente;
- Patamar Sill (C₁): valor 011 da semivariância correspondente ao Alcance (A). O Patamar (C_1) indica o ponto que deixa de existir a dependência espacial, dado que a variação da diferença entre pares de amostras torna-se aproximadamente constante;
- Efeito Pepita ou *Nugget* (C_0): representa a interseção da curva com o eixo y. Sua construção aponta uma descontinuidade, que pode ocorrer por se considerar distâncias menores que a menor distância entre as amostras, por erros de medição ou pelo acaso.

Um caso particular da Krigagem ocorre quando a função determinística que descreve a componente estrutural Z numa posição s [$\mu(s)$], é constante na região em estudo, assim a Krigagem passa a ser do tipo Ordinária. Assim, o processo a ser estimado tem uma média desconhecida, mas constante, cujo valor é igual à média dos valores observados nas amostras. Portanto, os pesos são escolhidos de maneira que o valor médio estimado é restringido pelo valor da média das amostras, sendo para isso necessário que a soma dos pesos seja igual a 1 (BARROS FILHO, 2007). Dessa forma, o peso reflete a distância entre as amostras e o ponto a estimar, de forma que quanto mais próximas estiverem as amostras do ponto a estimar, maior será o seu peso no estimador (SOARES, 2000).

A Krigagem Ordinária possui a capacidade de avaliar o grau de incerteza dos parâmetros ajustados modelos teóricos aos de semivariogramas. Tal incerteza é o erro da

 $N(h) \rightarrow e o número de amostras separadas por$ uma distância h;

 $Z_{(x_i)} e^{Z_{(x_i+h)}} \rightarrow são os valores amostrais$ nas posições $x_i e x_i + h$, separados pela distância h;

 $\hat{\gamma}(h) \rightarrow \hat{t}$ é a semivariância de todos os pares de amostras $Z_{(x_i)} e Z_{(x_i+h)}$

A representação gráfica do variograma é

estimativa, que pode ser obtido mediante o procedimento chamado validação do modelo que envolve a re-estimação dos valores conhecidos por meio dos parâmetros ajustados ao modelo do semivariograma (Camargo, 1997).

Barros Filho (2007) afirma que a Krigagem Ordinária minimiza a variância dos erros, espacializa os erros, permitindo a geração de uma superfície de erros.

A superfície gerada pela Krigagem Ordinária corresponde a uma grade, cujos pontos são calculados em função da variação local da amostra, conforme definido na Equação (6) (Burrough & Mcdonnell, 1998).

$$\widehat{Z}(x_0) = \sum_{i=1}^n \lambda_i Z(x_i) , \qquad (6)$$

onde:

 $\overline{Z}(x_0) \rightarrow e$ o valor calculado dos pontos da grade;

 $Z(x_i) \rightarrow i$ é o valor amostral;

 $\lambda_i \rightarrow i$ é o peso determinado pelo variograma,

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

Os pesos (λ_i) são escolhidos quando a estimação de $[\widehat{Z}(x_0)]$ é não tendenciosa, e quando a estimação da variância (σ_e^2) é menor do que qualquer outra combinação linear dos valores observados.

O erro associado ao ponto estimado, conhecido como Variância da Krigagem (σ_{e}^{2}), pode ser interpretado como padrão de confiança para os valores interpolados. A variância da Krigagem pode ser calculada de acordo com a Equação (7) seguinte forma:

$$\sigma_{\sigma}^{2} = \sum_{i=1}^{n} \lambda_{i} \gamma(x_{i}, x_{0}) + \phi, \qquad (7)$$

e obtido quando

$$\sum_{i=1}^{n} \lambda_{i} \gamma(x_{i}, x_{j}) + \phi = \gamma(x_{j}, x_{o})$$
 para todo j.

Nesse caso:

 $\gamma(x_i, x_j) \rightarrow \acute{e}$ a semivariância de \hat{z} entre os pontos amostrais $x_i \in x_j$;

 $\gamma(x_j, x_0) \rightarrow i$ é a semivariância de \hat{z} entre os pontos amostrais x_i e os desconhecidos x_0 ;

 $\phi \rightarrow \dot{e}$ o fator associado que garantem que a soma dos pesos seja 1.

Após a implementação dos interpoladores, é necessários realizar a análise estatística dos resultados advindos de cada interpolador. Essa etapa corresponde ao conjunto de procedimentos que servem para descrever a amostra, bem como são indicadores para a tomada de decisão na avaliação comparativa entre variáveis estudadas.

A análise estatística dos resultados das interpolações é realizada através da Média e Desvio Padrão, da Tendência Média, do Intervalo de confiança, e do Teste t dependente (teste t de amostras emparelhadas).

A Média e o Desvio Padrão correspondem ao cálculo da média aritmética e do desvio padrão da amostra, podendo ser expressos nas Equações (9) (a) e (b).

$$\overline{Z} = \frac{\sum_{i=1}^{n} Z_{i}}{N}, \quad [9(a)]$$
$$S = \sqrt{\frac{\sum_{i=1}^{n} (Z_{i} - \overline{Z})^{2}}{N}}, \quad [9(b)]$$

Onde:

 $Z_i \rightarrow$ refere-se ao dado amostral *i*;

 $\overline{Z} \rightarrow$ refere-se a média amostral;

 $S \rightarrow$ refere-se ao desvio padrão;

 $N \rightarrow$ refere-se ao tamanho da amostra.

A Tendência Média identifica o desvio médio de uma amostra (a média dos erros), e é definida pela Equação (10):

$$E_{m\acute{e}dio} = \frac{\sum_{i=1}^{n} \widehat{Z}_{i} - Z_{i}}{N}, \qquad (10)$$

Onde:

 $\hat{Z}i \rightarrow$ refere-se ao valor amostral predito;

 $Z_i \rightarrow$ refere-se ao valor amostral medido;

 $N \rightarrow$ refere-se ao tamanho das amostras;

O Intervalo de confiança corresponde aos limites entre os quais acredita-se que o valor da media verdadeira estará, dentre um determinado nível de probabilidade. Pode-se construir um intervalo de confiança de acordo com a Equação (11). $IC_{\alpha} = \overline{Z} \mp S_{\overline{z}}$, (11)

Para:

$$S_{\overline{Z}} = \frac{S}{\sqrt{N-1}}$$

Onde:

 $S_{\overline{z}} \rightarrow$ refere-se ao erro padrão da amostras;

 $S \rightarrow$ refere-se ao desvio padrão;

 $N \rightarrow$ refere-se ao tamanho da amostra;

O Teste t dependente (teste t de amostras emparelhadas) compara a diferença média entre duas amostras, com a diferença que espera-se

A metodologia empregada no trabalho foi dividida em quatro passos.

Aquisição das amostras batimétricas

O conjunto de amostras batimétricas, que delimitam a área piloto, foi adquirido junto à coordenação do Projeto de Monitoramento Ambiental Integrado - MAI. Para identificar a natureza do conjunto de amostras, foi realizada a análise exploratória dos dados a partir do teste de normalidade de Kolmogorov-Smirnov (teste K-S) e da análise gráfica.

Implementação dos interpoladores

Os interpoladores empregados foram: Inverso da Distância Ponderada, Polinômio Local, Funções de Base Radial (*Completely Regularized Spline*), Polinômio Global e Krigagem.

Durante a fase de interpolação, foi efetuada a validação cruzada, que consiste em omitir a

encontrar entre as médias das amostras (zero), levando em conta o erro padrão das amostras. O teste t de amostras emparelhadas pode ser descrito de acordo com a Equação (12).

$$t = \frac{\bar{D} - \mu_D}{S_{\overline{Z}}}, \qquad (12)$$

Onde:

 $\overline{D} \rightarrow$ refere-se a diferença média entre duas amostras;

 $\mu_D \rightarrow$ refere-se a diferença esperada entre as médias das amostras;

 $S_{\overline{z}} \rightarrow$ refere-se ao erro padrão entre as amostras;

MATERIAIS E MÉTODOS

posição de cada ponto amostral " P_i " e recalcular o ponto para a mesma posição, considerando os demais pontos. Este procedimento é repetido para os "n" pontos, possibilitando identificar o erro entre os valores preditos e medidos (Bailey & Gatrell, 1995).

Os parâmetros de cada interpolador foram ajustados individualmente de forma que, cada um deles apresentasse o menor erro.

Análises estatísticas dos resultados

As análises foram baseadas na estatística descritiva dos dados, levando em consideração os indicadores Média e Desvio Padrão, Tendência Média, Intervalo de Confiança, e o Teste t de amostras emparelhadas.

Criação da superfície tridimensional

A criação da Superfície foi efetuada pelo interpolador que apresentou os melhores resultados estatísticos.

RESULTADOS E DISCUSSÕES

Aquisição das amostras batimétricas

As amostras batimétricas coletadas na área piloto seguem o trecho compreendido entre os

municípios de Paulista, Olinda, e parte de Recife, conforme mostrado na Figura 6.



Figura 6. Área Piloto.

A amostra é composta de 65.041 pontos, dispostos em 137 perfis verticais separados a uma distância de 200 metros, e dois perfis horizontais espaçados a 2.000 metros um do outro. Todos os pontos foram coletados e referenciados ao Datum vertical SAD 69, e projetados no Sistema de Coordenadas UTM. A análise exploratória foi realizada ao efetuar o teste de K-S, cujos resultados constam na Tabela 1. Verificou-se que a amostra batimétrica tem distribuição diferente da normal (*p*-value < 0,05), em um nível de significância de 95%.

Tab <u>ela</u>	1.	Parâmetros	do	Teste	de	Kolmogorov-Smirnov

Teste K-S						
Variável	D _{cal}	GL	p-value			
Batimetria Medida	.056	65299	.000			



comprovando-se não normalidade da distribuição.



Para tentar normalizar a amostra, foi adicionada uma constante de valor 50 metros aos valores das batimetrias medidas, resultando na conversão de todos os valores negativos em positivos, permitindo a aplicação das transformações logarítmica, por radiciação, e por transformação recíproca. Após as transformações o Teste de K-S foi repetido, e nenhuma delas teve sucesso para normalização, conforme identifica-se na Tabela 2.

Teste de K-S para as Profundidades						
Variáveis	Dcal	GL	p-value			
Nova Prof (Prof+50)	.056	65041	.000			
LN (Nova Prof)	.052	65041	.000			
Raiz (Nova Prof)	.053	65041	.000			
1 / (Nova Prof)	.056	65041	.000			

Tabela 2. Teste de K-S para as profundidades normalizadas	
---	--

Os histogramas das amostras transformadas são mostrados nas Figuras 8, 9 e 10, novamente

comprovando os resultados obtidos pelo teste de K-S.











A não normalidade na distribuição das amostras batimétricas impossibilitou o emprego de testes paramétricos. Como condições de contorno foram utilizados os erros obtidos entre cada ponto batimétrico medido e seu dual calculado por validação cruzada. Para a interpolação pelo método IDP os pontos tiveram que estar contidos dentro de uma elipse com semi-eixo maior de 26.190 metros, e semi-eixo menor de 8.990 metros. O seu semi-eixo menor teve uma direção de 20°, paralela a direção dos perfis, conforme mostrado na Figura 11.

<figure>

Figura 11. Elipse de busca usada na interpolação por IDP.

Para o cálculo de um novo ponto, a quantidade de pontos selecionados foi de no

máximo 15, e de no mínimo 10 por quadrante da elipse. Como conseqüência esses parâmetros

Implementação das interpolações

tiveram por propriedade a captura de pontos em no mínimo 3 perfis consecutivos.

A interpolação utilizando o método Polinômio Local seguiu os mesmos parâmetros definidos para o IDP. Já a interpolação usando a Função de Base Radial foi realizada usando o método *Completely Regularized Spline*, e os parâmetros adotados para essa interpolação foram semelhantes aos do IDP. Para o método Polinomial Global foi escolhido um polinômio de grau 1.

O último interpolador implementado foi a Krigagem. Entretanto, antes de iniciar o processo da Krigagem verificou-se se a amostra apresentava tendência. Isso foi feito através do Gráfico de Tendência, conforme mostrado na Figura 12.



Figura 12. Gráfico de Tendência das amostras batimétricas em estudo.

Na Figura 12 identificam-se tendências decrescentes nos sentidos Leste-Oeste (Eixo Y) e Norte-Sul (Eixo X), conforme observado nas linhas azul e amarelo respectivamente. Essas tendências foram removidas usando um polinômio de primeira ordem.

O método Geoestatístico da Krigagem foi realizado usando a Krigagem Ordinária aplicada aos resíduos provenientes da remoção da tendência da amostra. Os pontos residuais foram ajustados no variograma do tipo esférico, tendo o Alcance (A) de 237,0653 metros, o Patamar (C₁) de 1,9502 metros, e o Efeito Pepita (C₀) de 0,2468 metros (Figura 13).



Figura 13. Variograma do tipo esférico resultante da interpolação por Krigagem Ordinária.

Os parâmetros adotados para a seleção dos pontos vizinhos na construção da interpolação tiveram os mesmos valores usados pelos IDP.

Análises estatísticas dos resultados

Após a implementação dos Interpoladores, o Teste de K-S foi aplicado para os valores dos erros dos resíduos interpolados. O resultado é mostrado na Tabela 3, identificando-se a não normalidade das distribuições (p-value < 0,05).

Teste K-S para os erros Interpolados							
Variáveis	Dcal	GL	p-value				
IDP	.130	65041	.000				
Polinômio Local	.136	65041	.000				
FBR	.130	65041	.000				
Polinômio Global	.040	65041	.000				
Krigagem	.135	65041	.000				

Tabela 3.	Teste da normalidade	para os err	os interpolados.
I abela of	reste du normanadae	pulu 05 cm	ob merpolucios.

Como a amostra em estudo é muito grande (n= 65.041), analisa-se graficamente a distribuição em cada variável interpolada (Figuras 13, 14, 15, 16 e 17), concluindo-se sobre a normalidade de todos os erros.



Figura 14. Histograma com a curva normal dos erros do Polinômio Local.







Figura 17. Histograma com a curva normal dos erros da Krigagem.

A estatística descritiva para a os Erros, é mostrada Tabela 4.

Tabela 4.	Estatística	descritiva	para	dos	Erros	das	Interp	olac	ões.
Labera H	Lotutioticu	ueserni vu	puru	aob	L1105	uub	merp	oraş	000.

Interpoladores	Média dos	Desvio	Erro Padrão	Intervalo para média com 95% de Confiança		
-	erros	Paarao		Limite -	Limite +	
IDP	.0056804	.6979112	.0027365	.0003167	.0110441	
Polinomial Local	.0028979	.7002147	.0027456	0024834	.0082793	
FBR	.0024291	.6922131	.0027142	0028907	.0077490	
Polinomial Global	.0000399	3.576930	.0140254	0274499	.0275298	
Krigagem	.0032213	.6830434	.0026782	0020280	.0084707	

O interpolador IDP apresentou intervalo de confiança para a média dos erros, com 95% de confiança, variando entre 0.0003167 e 0.0110441. Como esse intervalo não contém o zero, que é a média esperada para os erros, esse interpolador foi rejeitado.

O Interpolador Polinomial Global apresentou intervalo de confiança para a média dos erros, com 95% de confiança, variando entre -0.0274499 e 0.0275298. Apesar de esse intervalo conter o zero, o interpolador é rejeitado porque o seu erro padrão (0.0140254) é sete vezes maior que o erro padrão dos demais interpoladores, e também porque é impossível que a superfície real seja um plano inclinado. Por esses motivos, este interpolador é rejeitado.

O Interpoladores RBF, Polinomial Local e Krigagem, apresentaram intervalos de confiança para a média dos erros, com 95% de confiança, contendo o zero, e seus erros padrão possuem valores muito próximos. Assim, qualquer um desses interpoladores pode ser aplicado. Restava verificar se, do ponto de vista estatístico, existia diferença significativa entre eles. Isso foi feito usando o teste *t* de Amostras Emparelhadas.

A Tabela 15 apresenta os resultados deste teste, mostrando que dentro de um nível de confiança de 95%, os três interpoladores não diferem (Sig > 0,05 para todos os pares).

Tabela 5. Estatística descritiva para o teste t de amostras emparelhadas.									
Pares de Interpoladore	Média	Desvio Padrão	Erro Padrão	Intervalo para média com 95% de Confiança		t	GL	Sig	
8				Limite -	Limite +				
Local - FBR	.0004688	.1714870	.0006724	0012632	.0022009	.697	65040	.486	
Local - <i>Krigagem</i>	0003233	.1942125	.0007615	0022850	.0016382	425	65040	.671	
FBR - Krigagem	0007922	.1925926	.0007551	0027374	.0011530	- 1.049	65040	.294	

Dessa forma, pode-se afirmar que qualquer um desses três métodos é indicado para ser usado como interpolador para a área de estudo. Entretanto, em virtude da Krigagem ser o único interpolador capaz de remover tendências e espacialisar a variância dos pontos interpolados sobre uma superfície, seleciona-se este como o mais indicado.

Criação da superfície tridimensional

A superfície batimétrica da área piloto e a superfície da variância dos valores estimados geradas pela Krigagem são apresentadas na Figura 18 (a) e (b), respectivamente.



Figura 18. Superfície batimétrica obtida pela Krigagem (a), e sua Superfície da Variância da Krigagem (b).

Na Figura 18 (b), observa-se que as menores variações estão espacializados mais próximos dos pontos amostrais, e na vizinhança desses pontos, localizados justamente sobre os perfis de dados coletados.

A visualização do modelo tridimensional consiste na transferência do modelo gerado para um plano de visualização 3D, permitindo que o observador perceba a sensação tridimensional original da Plataforma Interna Continental adjacente aos Municípios de Paulista, Olinda e Recife.

A representação tridimensional da superfície gerada pela Krigagem pode ser visualizada em diversos ângulos e perspectivas na Figura 19.



Figura 19. Representação Batimétrica Tridimensional gerada.

Essas duas superfícies sobrepostas permitiram identificar a coordenada batimétrica

predita de qualquer ponto, e o valor da variância associada. De acordo com a Figura

20, ao sobrepor os pontos batimétricos, e as duas superfícies geradas pela Krigagem, encontram-se para os Pontos 01, 02 e 03, as seguintes coordenadas (em metros) mostradas na Tabela 6.



Figura 20. Consulta espacial a superfície interpolada pela Krigagem, e a sua variância associada.

Tabela 6	Coordenadas	dos Pontos	01	02 e 03
Tabcia U.	Coordenadas	uos i ontos	UI.	, 02 0 05.

	1 45 614	of ecolutination acts	1 ontos o1, o = e oc.	
Ponto	Coordenada E	Coordenada N	Batimetria (m)	Variância (m)
01	299033,630	9127074,799	0,17	1,11
02	300072,162	9117021,802	11,40	1,57
03	300122,565	9109288,905	16,84	1,52

CONCLUSÕES

Este trabalho teve como princípio a avaliação de diferentes modelos de interpoladores com o objetivo de encontrar o modelo que representasse de forma mais real um conjunto de amostras batimétricas pertencente a uma área piloto em estudo.

De modo geral, as etapas definidas pela metodologia desenvolvida foram ordenadamente às seguintes: aquisição e análise dos dados batimétricos; implementação das interpolações; análise dos resultados; e criação da superfície tridimensional.

A aquisição dos dados da área piloto ocorreu junto à coordenação do Projeto MAI. A amostra consta de 65.041 pontos batimétricos, espacializados em 2 perfis na direção lesteoeste, e 137 perfis na direção norte-sul, ao longo de uma superfície de aproximadamente 121km², compreendida entre os Municípios Paulista, Olinda, e Recife.

A análise descritiva dos dados batimétricos identificou a impossibilidade da aplicação de testes paramétricos diretamente sobre os pontos batimétricos. Isso aconteceu porque а distribuição desses pontos não era normal, e também por não encontrar um modelo de normalização. A estatística paramétrica não se aplicou ao conjunto de amostras batimétricas da área piloto em estudo. Como condição de contorno, optou-se por fazer uma análise dos erros, entendidos como a diferença entre o valor medido e o valor predito, obtido por validação cruzada, para cada interpolador estudado.

A implementação das interpolações foi caracterizada pela aplicação dos seguintes modelos: Inverso da Distância Ponderada (IDP), Polinomial Local, Função de Base Radial (FBR), Polinomial Global e Krigagem. E resultou na obtenção de valores calculados para os pontos conhecidos. A análise dos resultados dos erros baseou-se nos seguintes indicadores paramétricos: normalidade da distribuição, média, desvio padrão, intervalo de confiança para a média e, diferença média entre duas amostras. O resultado da análise rejeitou os interpoladores IDP e Polinomial Global. O IDP obteve um intervalo de confiança com 95% que não continha o zero. Já o interpolador Polinomial Global apresentou um erro padrão muito alto quando comparados com os demais.

Os demais (Polinomial Local, FBR e Krigagem) foram aceitos, não tendo sido encontrado diferença significativa entre eles em um nível de confiança de 95%, parâmetro comprovado pelo teste t de amostras emparelhadas.

Mesmo assim, dentre os três, selecionou-se a Krigagem como o mais indicado, por possui funções particulares quando comparada aos outros modelos.

Por ser um interpolador Geoestatístico a Krigagem conseguiu remover tendências identificadas, melhorando significativamente os resultados obtidos no momento do procedimento de interpolação de dados. Outra característica identificada é que, em todos os casos, o único caminho para determinar fidedignamente a predição de pontos, seria a estimativa computada por uma série externa de pontos que não foram usados na interpolação original. Particularmente a Krigagem é o único interpolador capaz de produzir a superfície de variação dos valores correspondente aos pontos interpolados, diferenciando-se dos demais interpoladores implementados no trabalho.

A superfície gerada pela Krigagem permite de forma mais fidedigna, a identificação de feições geomorfológicas presentes no relevo submerso da área em estudo, como bancos de areia e recifes de corais, bem como, auxilia na elaboração e no planejamento de ações costeiras que levem em consideração a identificação destas áreas.

Enfatiza-se ainda que, no que diz respeito da metodologia elaborada nesse trabalho, ao interpolar dados advindos de quaisquer que sejam os levantamentos pontuais, é necessário analisar o comportamento do conjunto de amostras de entrada e, utilizar análises estatísticas espaciais para identificar disparidades semelhancas ou nos dados interpoladores, de forma a assegurar a veracidade da representação das amostras em uma superfície contínua.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ABSHER, T. M.; BORZONE, C. A.; BRANDINI, F. P.; LANA, P. da C.; SPACH, H. L.; MARONE, E. **Subsídios para Estudos Ambientais em Áreas Costeiras**. Trabalho técnico – Centro de Estudos do Mar da Universidade Federal do Paraná, 2002.

2. AYRES, A. & NETO, J. A. B. Métodos Diretos e Indiretos de Investigação do Fundo Oceânico. In NETO, J. A. B.; PONZI, V. R. A.; SICHEL. S. E. **Introdução à geologia marinha**. Rio de Janeiro: Interciência, 2004. cap. 6.

3. BAILEY, T. C. & GATRELL, A. C. Interactive Spatial Data Analysis. Prentice Hall; Pap/Dsk edition. 1995. 432p.

4. BAJJALI, W. Model the Effect of Four Artificial Recharge Dams on the Quality of Groundwater using Geostatistical Methods in GIS Environment, Oman. 2002. Journal of Spatial Hydrology Fall, Vol. 5, N°. 2. Department of Biology and Earth Sciences, University of Wisconsin – Superior, Superior, WI 54880.

5. BARROS FILHO, M. N. M. **Krigagem Ordinária** aplicada à análise espacial da habitabilidade intra-urbana. O caso da Cidade do Recife (Brasil). In: VII Congresso Internacional da Sociedade Latino-Americana de Estudos Imobiliários, 2007, São Paulo, SP. VII Congresso Internacional da Sociedade Latino-Americana de Estudos Imobiliários, 2007.

6. BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**. Artigo 225, §4°. Brasília, DF, Senado, 1988.

7. BURROUGH, P. A. & McDONNELL, R. A.

Principles of Geographical Information Systems: Spatial Information Systems. Oxford University Press. Oxford, 1998.

8. BURROUGH, P. A. **Principles of Geographical Information Systems for Land Resources Assessment.** Clarendon Press. Oxford, 1986.

CÂMARA, G.; MONTEIRO, A. M.; MEDEIROS, J.
Introdução à Ciência da Geoinformação. São José dos Campos, INPE, 2004.

10. CAMARGO, E. C. G. Desenvolvimento, implementação e teste de procedimentos geoestatísticos (Krigeagem) no Sistema de Processamento de Informações Georreferenciadas (SPRING). 1997. Dissertação (Mestrado em Sensoriamento Remoto) - Divisão de Processamento de Imagens, Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 1997.

11. CHAPLOT, V.; DARBOUX, F.; BOURENNANE, H. Accuracy of interpolation techniques for the derivation of digital elevation models in relation to landform types and data density. Geomorphology – Vol 77, Issues 1 – 2, Pages 126 – 141, 15 July 2006.

12. CHIN-SHUNG, Y.; SZU-PYNG, K.; FEN-BIN, L.; PEN-SHAN, H. P. **Twelve Different Interpolation Methods:** A Case Study of Surfer 8.0. XXth International Society for Photogrammetry and Remote Sensing (ISPRS) Congress. Istanbul, 2004.

13. DIAS, J. A.; POLETTE, M.; CARMO, J. A. **O Desafio da Gestão Costeira Integrada** (Prefácio). Revista de Gestão Costeira Integrada Vol. 7(1): pág 3-4, 2007. 14.ESTRADA,L.& SAFRIET,C.AdvancedTechniques for Geographic Information Systems. 2006. SPAComputerLab.Disponívelem<www.spa.ucla.edu/up206b/interpolation_methods.htm>,acesso em 25/11/2006.

15. FIELD, A. P. (2009). **Descobrindo a estatística usando o SPSS**. Porto Alegre, Brasil, Artmed, 2a edição.

16. FONSECA, L.; MAYER, L.; PATON, M. ArcView Objects in the Fledermaus Interactive 3-D Visualization System: An Exemple from the STRATAFORM GIS. Undersea with GIS. Chapter 1, page 1-21. ESRI, 2002.

17. JAKOB, A. A. E. & YOUNG, A. F. Uso de Métodos de Interpolação Espacial de Dados nas Análises Sociodemográficas. XV Encontro Nacional de Estudos Populacionais, Caxambú: ABEP, 2006.

 MAI. Monitoramento Ambiental Integrado – Avaliação dos Processos de Erosão Costeira nos Municípios de Paulista, Olinda, Recife e Jaboatão dos Guararapes. Projeto executivo. UFPE - Universidade Federal de Pernambuco, 2004.

19. MELLO, J. M.; BATISTA, J. L. F.; RIBEIRO JR, P. J.; OLIVEIRA, M. S. Ajuste e seleção de modelos espaciais de semivariograma visando à estimativa volumétrica de Eucalyptus grandis. Revista SCIENTIA FORESTALIS n. 69, p.25-37, dez. 2005, Instituto de Pesquisas e Estudos Florestais - IPEF.

20. MICHAEL, R. & TRIVELONI, C. A. P. Uso de Inferência Estatística e Análise de Superfície de Tendência para Avaliação em Massa de Imóveis. UFSC – Departamento de Engenharia Civil. Florianópolis/SC. COBRAC 2006 -Congresso Brasileiro de Cadastro Técnico Multifinalitário -UFSC - Florianópolis - 15 a 19 de Outubro 2006.

21. NICOLODI, J. L. & TOLDO JR, E. E. Morfodinâmica de praias: uma ferramenta para gestores de

ambientes costeiros - o caso da Praia de Fora, Parque Estadual de Itapuã, RS. Natureza e Conservação . pp. 22-32 Artigos Científicos. Vol 1 . nº2 . Out 2003.

22. KRUG, L. A. & NOEMBERG, M. A. Extração de batimetria por sensoriamento remoto de áreas rasas dos sistemas estuarinos do Estado do Paraná - Brasil. Anais XII Simpósio Brasileiro de Sensoriamento Remoto, Goiânia, Brasil, 16-21 abril 2005, INPE, p. 3077-3084.

23. SCHAFFRATH, V. R.; TORMENA, C. A.; GONÇALVES, A. C. A.; OLIVEIRA JUNIOR, R. S. Variabilidade Espacial de Plantas Daninhas em Dois Sistemas de Manejo de Solo. Revista Brasileira de Engenharia Agrícola Ambiental, Campina Grande, v. 11, n. 1, 2007. Disponível em:

http://www.scielo.br/scielo.php?script=sci_arttext&pid=S141 5-43662007000100007&lng=pt&nrm=iso>. Acesso em: 10 Agosto 2007.

24. SERAFINI, L. Z. O Zoneamento Costeiro e as Competências Municipais: Uma Análise a luz da Legislação Ambiental Brasileira. 2010. Disponível em: <www.ibap.org/teses2004/teses2004d26.doc>. Acesso em 04 de Fevereiro de 2010.

25. SOARES, A. **Geoestatística para as ciências da terra e do ambiente**. Lisboa: Instituto Superior Técnico (IST) Press, 2000.

26. SOKAL, R.R. & ROHLF, J.F. **Biometry: The principles and practice of statistics in biological research**. 3. ed. New York: W. H. Freeman and Company, 1997. 850p.

Manuscrito recebido em: 06 de Abril de 2011 Revisado e Aceito em: 10 de Janeiro de 2012