

# MÓDULOS EM MATLAB PARA INTERPOLAÇÃO ESPACIAL PELO MÉTODO DE KRIGAGEM ORDINÁRIA E DO INVERSO DA DISTÂNCIA

*Matlab modules for spatial interpolation by ordinary kriging and inverse distance*

**Alexandre Cândido Xavier<sup>1</sup>**  
**Roberto Avelino Cecílio<sup>2</sup>**  
**Julião Soares de Souza Lima<sup>1</sup>**

**<sup>1</sup>Universidade Federal do Espírito Santo - UFES**  
**Departamento de Engenharia Rural**  
Alto Universitário, s/nº - Cx Postal 16. 29500-000 Alegre-ES.  
xavier@cca.ufes.br  
limajss@yahoo.com.br

**<sup>2</sup>Universidade Federal do Espírito Santo - UFES**  
**Departamento de Engenharia Florestal**  
Av. Carlos Lindemberg, s/n - Cep: 29550-000 - Jerônimo Monteiro/ES  
racecilio@yahoo.com.br

## RESUMO

Este trabalho tem como objetivo apresentar um conjunto de módulos desenvolvidas em MATLAB para interpolação de dados espaciais por meio da krigagem ordinária e do inverso da distância elevada a uma potência. Os dados de entrada são vetores da posição das amostras ( $x$ ,  $y$ ) e seus respectivos atributos ( $z$ ). É considerada dependência espacial do atributo do tipo isotrópica. Os módulos construídos para krigagem ordinária calculam as semivariâncias experimentais, os parâmetros dos modelos teóricos (esférico, gaussiano, exponencial e linear) e suas respectivas estatísticas da tabulação cruzada. As estatísticas da tabulação cruzada também são calculadas no módulo desenvolvido para o método inverso da distância elevada a uma potencia, considerando as potencias de 1 a 6. Os módulos podem ser utilizados juntos de forma que o processo de escolha do modelo seja automatizada, tornando-se a escolha do interpolador um processo de fácil execução.

**Palavras chaves:** Interpolação, Geoestatística, Inverso da Distância Elevada a uma Potência.

## ABSTRACT

The main objective of this paper is to present a Matlab set modules to interpolation spatial data by means of ordinary kriging and inverse distance interpolation. The data incoming are vectors values of positions ( $x$ ,  $y$ ) and theirs respective attributes ( $z$ ). It is considered that the spatial dependence of the attributes is isotropic. The modules implemented for ordinary kriging compute the experimental semivariance, the parameters of the theoretical models (spherical, Gaussian, exponential, and linear) and those respective statistics of cross tabulation. The cross tabulation statistics are also calculated in the module implemented for inverse distance interpolation with inverse distance weighting with values of 1 to 6. The modules can be used together to search model automatically, becoming the interpolated search an easy process to execute.

**Keywords:** Interpolation, Geostatistic, Inverse Distance Interpolation.

## 1. INSTRUÇÕES GERAIS

A interpolação de dados espaciais é necessária sempre que se necessite estimar uma variável numérica qualquer, em uma determinada posição geográfica, em

que não existe esta informação. Para realizar esta tarefa é necessário que exista na região de interesse pontos amostrais, nos quais a variável foi medida. A partir dos valores medidos nos pontos amostrais é que se pode estimar a variável para o local de interesse não medido,

utilizando alguma função matemática em que as variáveis independentes são as amostras da vizinhança (BURROUGH; MACDONNEL, 2000). Os interpoladores de dados espaciais são utilizados nas mais diversas áreas da ciência, podendo-se citar: na saúde (OPROMOLLA *et al.*, 2005; CASTRO; SINGER, 2007); na meteorologia (ERTEKIN e EVRENDILEK, 2007), na hidrologia (SARANGI *et al.*, 2006); na geociência (PEREIRA; REMACRE, 2003; TARNAVSKY *et al.*, 2008), na agrária (OLIVEIRA *et al.*, 2008) e na ecologia (SOUZA *et al.*, 2005; RODRÍGUEZ-LADO *et al.*, 2007).

Existem várias metodologias para interpolar dados espaciais, podendo-se citar: o vizinho mais próximo, o inverso da distância elevada a uma potência, a krigagem e a superfície de tendência (BURROUGH; MACDONNEL, 2000; LANDIM, 2003). Na escolha do método, geralmente, é utilizado apenas um critério de escolha, aquele com o qual o usuário tem maior afinidade. Até pouco tempo atrás, não se levava em consideração a preocupação com a qualidade da estimativa, o que somente vem ocorrendo mais recentemente (FERNANDES; MENEZES, 2005). A qualidade do estimador pode ser avaliada por meio da avaliação dos dados da tabulação cruzada, isto é, comparação entre os dados reais e suas respectivas estimativas. Este método consiste em, uma vez escolhido o interpolador, cada uma das amostras é retirada dos dados originais e, para a sua posição no espaço, é realizada a sua estimativa pelo interpolador escolhido. Várias estatísticas podem ser calculadas e utilizadas para se avaliar a qualidade do interpolador por meio dos dados da tabulação cruzada, por exemplo, o coeficiente de determinação ( $R^2$ ) entre os dados medidos e estimados.

Diversos programas computacionais na área de geoprocessamento têm ferramentas destinada à realização de interpolação espacial, por exemplo: o GS+ (ROBERTSON, 2000), o Spring (CAMARA *et al.*, 1996); ArcGIS (ESRI, 2008); IDRISI (EASTMAN, 1998). Contudo, a escolha do método de interpolação a ser utilizado que fornecerá melhor resultado passa a ser bastante laboriosa, tendo em vista a escolha dos vários parâmetros que são necessários para cada tipo de interpolador. Por exemplo, se a escolha do interpolador

foi pelo método inverso da distância elevada a uma potência, qual potência deveria ser utilizada na ponderação? Caso a escolha tenha sido por krigagem, qual modelo teórico (gaussiano, esférico, exponencial ou linear) deve-se considerar para a estimativa das semivariâncias, e quais valores dos seus respectivos parâmetros (alcance, patamar e efeito pepita)? Ou seja, são várias as opções a serem analisadas para definição do interpolador, o que torna sua busca laboriosa.

Desta forma, o objetivo deste trabalho foi desenvolver um conjunto de módulos em MATLAB para a interpolação de dados espaciais segundo os métodos de interpolação por krigagem ordinária e pelo inverso da distância elevado a uma potência, de maneira a agilizar a escolha do modelo e seus respectivos parâmetros, baseando-se em estatísticas da tabulação cruzada.

## 2. MATERIAL E MÉTODOS

### 2.1 Dados das amostras teste

Para testar os interpoladores foram utilizadas 71 amostras testes obtidas aleatoriamente de uma grade de 21 x 21 amostras geradas a partir da função:

$$f(x, y) = z = 1 + \sin \sqrt{x^2 + y^2} / \sqrt{x^2 + y^2} \quad (1)$$

em que:  $x$  e  $y$  são as coordenadas no plano com variação de -10, -9, -8,..., 9, 10, e  $z$  o valor do atributo, calculado pela Equação 1. Esta função foi utilizada, pois possui grande variabilidade espacial. Também apresenta variação espacial do tipo isotrópica, pressuposto para os modelos de semivariâncias avaliados neste trabalho. Para variabilidade espacial anisotrópica, outras técnicas podem ser utilizadas (BURROUGH e MACDONNELL, 2000), contudo estas fogem ao escopo inicial deste trabalho. A Figura 1 apresenta a função, a grade e as amostras testes utilizadas neste trabalho. Para este trabalho não serão atribuídas unidades para  $x$ ,  $y$  e  $z$ , contudo, em casos práticos,  $x$  e  $y$  teriam unidades de distância (p.ex.: km, m etc) e  $z$  a unidade referente ao atributo que se esteja querendo especializar.

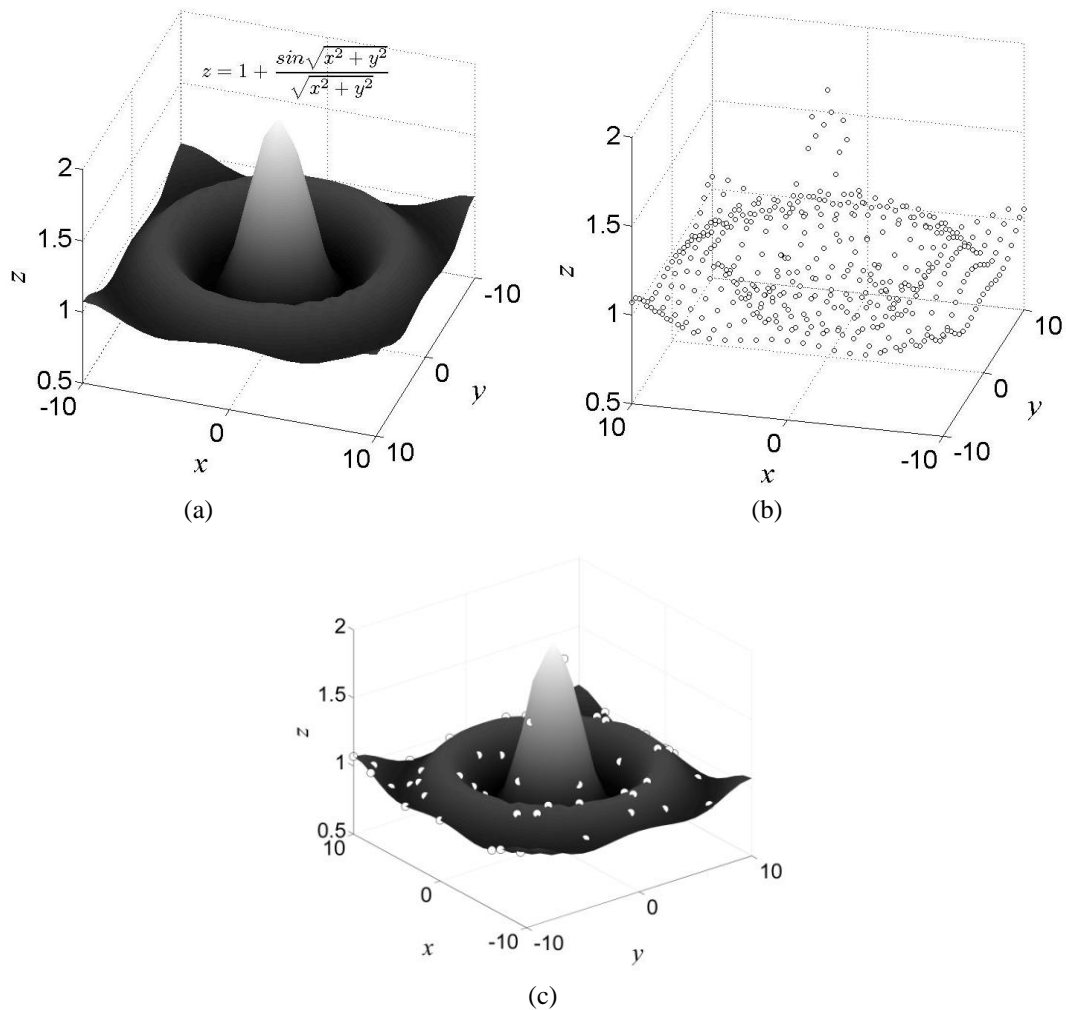


Fig. 1 – Função matemática (a), juntamente com a grade de 21x21 amostras (b) de onde foram obtidas de forma aleatória 71 amostras para teste dos interpoladores (c).

## 2.2 Interpoladores

Foram implementados diversos módulos para interpolação espacial por krigagem ordinária e pelo inverso da distância elevada a uma potência. Os módulos foram escritos em MATLAB 2007b (MATHWORKS, 2007a). Na seqüência apresenta-se uma breve descrição matemática dos interpoladores utilizados no presente trabalho (BURROUGH; MACDONNELL, 1997).

### 2.2.1 Krigagem ordinária

Na interpolação por krigagem ordinária, um atributo  $z$  na posição  $j$ , com coordenadas dado pelo vetor  $\mathbf{x}$ , é estimado por:

$$z(\mathbf{x}_j) = \sum_{i=1}^n \lambda_i \cdot z(\mathbf{x}_i) \quad (2)$$

em que:  $\lambda_i$  são as ponderações a serem aplicadas em cada um dos  $i$ -ésimos valores dos atributos que tenham

dependência espacial com o valor da variável a ser estimado na posição  $\mathbf{x}_j$ , sendo  $\sum_{i=1}^n \lambda_i = 1$ . Os valores  $\lambda_i$  são calculados por:

$$A^{-1}b = \begin{bmatrix} \lambda \\ \phi \end{bmatrix} \quad (3)$$

em que:  $A$  é a matriz de semivariâncias entre os pares de pontos utilizados para a interpolação;  $b$  é o vetor de semivariância entre cada ponto utilizado na interpolação e o ponto a ser estimado ( $\mathbf{x}_j$ ).

As semivariâncias são estimadas por modelos ajustados ao semivariograma experimental ( $\hat{\gamma}$ ) calculado por:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2n} \sum_{i=1}^n \{z(\mathbf{x}_i) - z(\mathbf{x}_i + \mathbf{h})\}^2 \quad (4)$$

em que:  $n$  é o número de pares amostrais do atributo  $z$  separados por uma distância  $\mathbf{h}$ . A plotagem  $\gamma(\mathbf{h})$

versus  $\mathbf{h}$  é o semivariograma experimental. A partir do semivariograma experimental é realizado o ajuste dos modelos teóricos. Os modelos teóricos utilizados foram:

$$\gamma(\mathbf{h}) = c_0 + c_1 \left\{ \frac{3\mathbf{h}}{2a} - 1/2(\mathbf{h}/a)^3 \right\} \quad \text{para } 0 < \mathbf{h} < a \quad (5)$$

$$\gamma(\mathbf{h}) = c_0 + c_1 \quad \text{para } \mathbf{h} \geq a \quad (6)$$

exponencial:

$$\gamma(\mathbf{h}) = c_0 + c_1 \{1 - \exp(-\mathbf{h}/a)\} \quad (7)$$

gaussiano:

$$\gamma(\mathbf{h}) = c_0 + c_1 \{1 - \exp(-\mathbf{h}^2/a^2)\} \quad (8)$$

e o linear:

$$\gamma(\mathbf{h}) = c_0 + \frac{c_1}{a} \mathbf{h} \quad (10)$$

em que:  $a$  é o alcance, isto é, a distância até onde existe dependência espacial;  $c_0$  é o efeito pepita, que mostra a descontinuidade do semivariograma para distâncias menores do que a menor distância entre as amostras;  $c_1 = c - c_0$ , onde  $c$  é o patamar, valor do semivariograma correspondente a seu alcance ( $a$ ), ou seja, valor da semivariância em que não ocorrerá mais dependência espacial.

O próximo passo é ajustar os modelos teóricos as semivariâncias calculadas. Tal ajuste pode ser realizado manualmente. Todavia, tendo como por objetivo automatizar o processo, a obtenção dos parâmetros  $c_0$ ,  $c_1$  e  $a$  foi foram obtidos por meio daqueles que minimizam função:

$$f(a, c_0, c_1) = \sum_{i=1}^n (\hat{\gamma}(\mathbf{h}) - \gamma(\mathbf{h}))^2 \quad (11)$$

Sujeita as seguintes restrições:

$$0 < a < \max(\mathbf{h}) \quad (12)$$

$$c_0 > 0 \quad (13)$$

$$c_0 + c_1 < \max(\gamma(\mathbf{h})) \quad (14)$$

A função utilizada para minimizar  $f$  foi a **fmincon** (MATHWORKS, 2007b). Esta função encontra o mínimo de função não linear sujeita a várias restrições.

### 2.2.2 Inverso da distância

A interpolação pelo inverso da distância considera que o valor da variável a ser estimado na posição  $\mathbf{x}_j$  ( $z(\mathbf{x}_j)$ ) é calculado pelos  $n$  vizinhos mais próximos ponderados pelo inverso da sua distância elevada a uma potência ( $r$ ):

$$z(\mathbf{x}_j) = \sum_{i=1}^n z(\mathbf{x}_i) \cdot d_{ij}^{-r} / \sum_{i=1}^n d_{ij}^{-r}, \quad (15)$$

em que:  $d_{ij}$  é a distância euclidiana entre o  $i$ -ésimo vizinho e a posição  $\mathbf{x}_j$ ;  $i=1, 2, \dots, n$ .

### 2.3 Avaliação do estimador

Para avaliação da precisão dos interpoladores são calculadas várias estatísticas da tabulação cruzada (Anexo 1). Neste trabalho será considerada a raiz quadrada do erro quadrado médio (RMSE) dos valores medidos e estimados pelos interpoladores. Nesta estatística quanto menor RMSE maior precisão:

$$\text{RMSE} = \sqrt{n^{-1} \sum_{i=1}^n (z_j - \hat{z}_j)^2}. \quad (16)$$

## 3. RESULTADOS

Segue abaixo uma descrição dos módulos implementados e os resultados por eles obtidos. Primeiramente será apresentado cada um dos módulos separadamente, e depois, estes funcionando conjuntamente para escolher automaticamente modelo que resulte no menor RMSE.

### 3.1 Descrição dos módulos

Os módulos em MATLAB implementados para interpolação são longos para ser incluída no trabalho, contudo serão apresentadas descrições sucintas dos mesmos. Os módulos são arquivos chamados de M-Files com o nome tipo **módulo.m**, sendo:

- alcance\_minimo** uma vez que no processo para o cálculo das ponderações  $\varphi$  é necessária a existência da matriz de semivariância ( $A$ ), esta função encontra qual deve ser o alcance mínimo para que ela exista em qualquer ponto dentro da área amostral. A entrada desta função são as coordenadas X e Y das amostras medidas;
- semiva** calcula as semivariâncias amostrais ( $\hat{\gamma}$ ); o número de pares utilizados para o respectivo cálculo ( $n$ ); e as distâncias médias entre os pares amostrais ( $\mathbf{h}$ ). Os dados necessários para entrada desta função são: as coordenadas (X, Y) e os respectivos valores dos atributos (Z) das amostras; a medida máxima para a distância entre os pares a ser considerado; o número de

**escomodelo2min** semivariâncias que se quer que compute; o número mínimo de pares para se ter um valor de semivariância; a partir de dados de distância (**h**) e das semivariâncias estimadas, os valores dos parâmetros  $a$ ,  $c_0$  e  $c_1$  são estimados para cada um dos modelos teóricos, esférico, exponencial, gaussiano e linear;

**corre\_kriga** com os dados dos parâmetros dos modelos teóricos, faz-se necessário avaliar, qual destes resultará na melhor precisão na interpolação dos dados. Esta função calcula os valores da tabulação cruzada considerando cada um dos modelos teóricos, podendo apresentar o resultado geral por meio de gráfico, juntamente com as respectivas estatísticas;

**corre\_invdist** função que a partir dos dados X, Y e Z das amostras realiza a tabulação cruzada considerando as potências,  $r = 1, 2, 3, \dots, 6$ , juntamente com estatísticas.

Como resultado da aplicação dos módulos nas amostras testes verificou-se que no módulo **alcance\_minimo**, para garantir a existência da matriz A foi de 11,4, valor que corresponde a 44% da distância

máxima verificada entre dois pares (26,2).

Exemplos de resultados do módulo **semiva** são mostrados na Figura 2, por meio das dispersões das semivariâncias amostrais calculadas com a distância entre os pares (**h**), mantendo-se o mesmo número de semivariâncias a serem calculadas (8) mas com **h** máximo entre os pares de amostras de 44% e 60% da maior distância entre as amostras, Figura 2a e 2b, respectivamente. Para cada semivariância também são apresentados os números pares utilizados para cálculo seu cálculo. Observa-se que o padrão da dispersão das semivariâncias amostrais nestas duas hipóteses são distintas, o que levaria a escolha de diferentes parâmetros para os modelos teóricos da semivariância. Por exemplo, visualmente para o semivariograma da Figura 2a, observa-se que não é verificado de forma clara um patamar ( $c_0 + c_1$ ), de tal forma que o modelo teórico que melhor se ajustaria a estes dados seria o linear. Já para os dados do semivariograma da Figura 2b, ocorreria uma estabilização das semivariâncias com patamar e alcance de aproximadamente 0,06 e 7, respectivamente. Desta forma, com modelos teóricos distintos, diferentes estimativas poderão ser calculadas, resultando em precisões distintas.

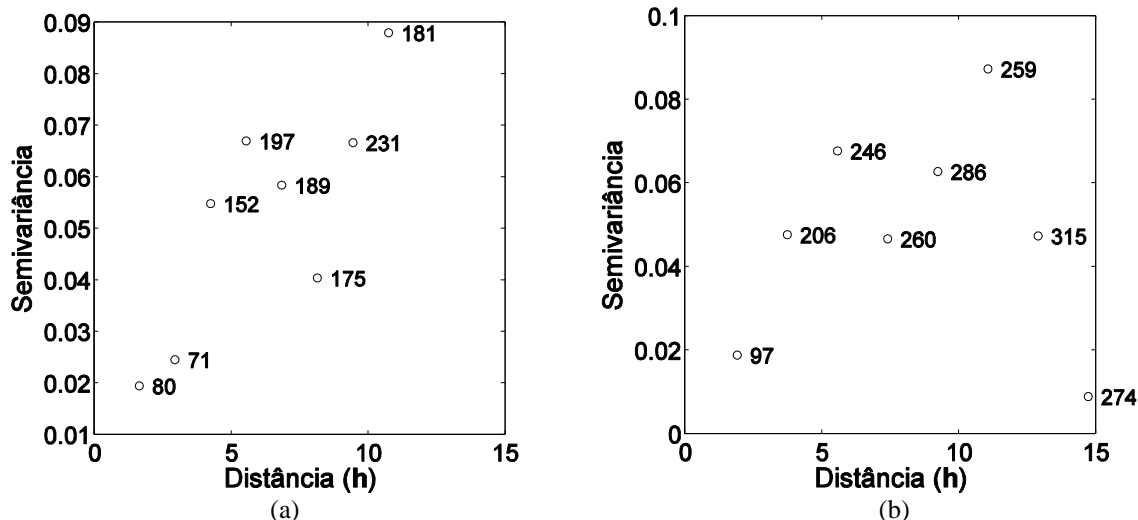


Fig. 2 – Apresentação da semivariância calculada considerando 8 semivariâncias calculadas até uma distância máxima de 44% (a) e 60% (b) da distância máxima entre pares de amostras.

Na Figura 3 são apresentados os resultados da aplicação do módulo **escomodelo2min** para ajustar os parâmetros dos modelos teóricos aos dados de semivariância experimental, com distância máxima entre os pares de 44% e 8 semivariância calculadas. Os parâmetros ( $a$ ,  $c_0$ , e  $c_1$ ), obtidos na minimização da Equação 11, para cada um dos modelos são também

apresentados, sendo os mesmos armazenados para posterior utilização nas estimativas. Neste trabalho não é realizada avaliação do ajuste do modelo teórico aos dados de semivariância, pois o interesse é na precisão da estimativa, que só será observado por meio dos dados da validação cruzada.

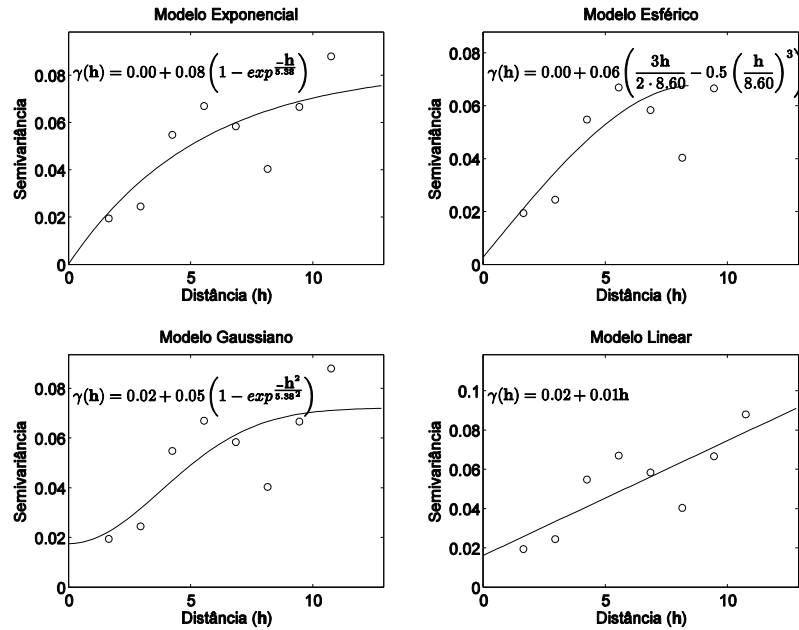


Fig. 3 – Ajuste dos parâmetros ( $a$ ,  $c_0$ , e  $c_1$ ) dos modelos teóricos aos dados de semivariância experimental calculado pelo módulo **escomodelo2min**.

Na Figura 4, a partir dos parâmetros estimados para cada um dos modelos teóricos mostrados na Figura 3, é mostrado o resultado das respectivas tabulações cruzadas pelo método da krigagem ordinária (módulo **corre\_kriga**). Algumas estatísticas para análise da precisão da estimativa são apresentadas para cada um dos modelos teóricos. Pode-

se verificar, considerando o RMSE, que o modelo teórico gaussiano foi o que resultou na maior precisão com RMSE=0,05, enquanto os demais resultaram em valores maiores que 0,10. Verifica-se também que todas as relações foram significativas a 1% de probabilidade (valor-p<0,01)

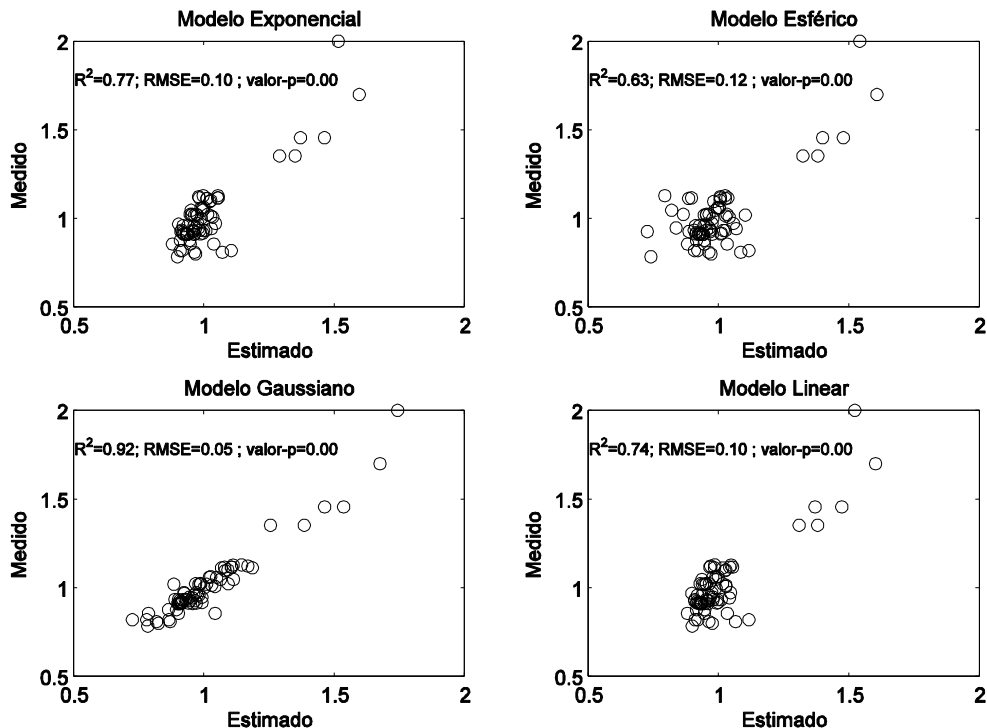


Figura 4 – Tabulação cruzada entre valores medidos e estimados e suas respectivas estatísticas de acordo com o modelo de semivariância.

Para interpolador pelo método inverso da distância elevada a uma potência, o módulo **corre\_invdist** é executado. Na Figura 5 é apresentado um exemplo de resultados da tabulação cruzada, gerados por este módulo, em que foi considerando: as potências 1,2,3,...,6 na ponderação das distâncias considerando; e a utilização dos 6 vizinhos mais próximos. Pode-se observar que todas as tabulações

cruzadas foram significativas. Observando os valores dos RMSE's, o melhor resultado foi para a ponderação do inverso da distância elevada a segunda potência (RMSE=0,148). Excluindo-se a ponderação com o valor de 1, verificou-se que a qualidade da interpolação diminuiu na medida em que a potência da ponderação aumenta.

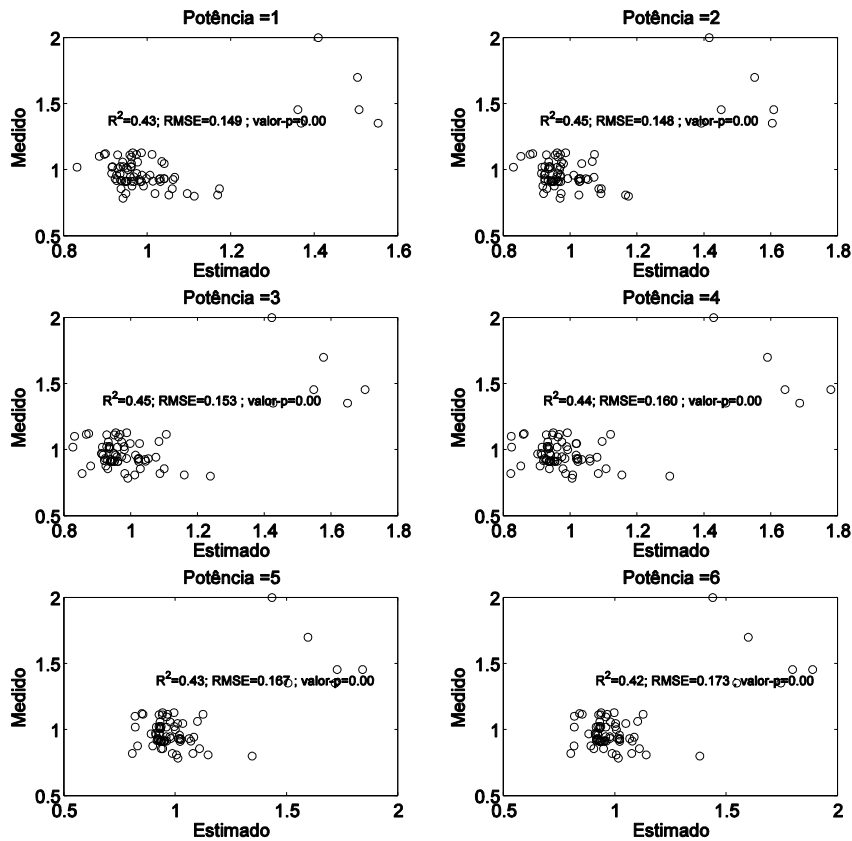


Fig. 5 – Tabulação cruzada com estatísticas pela interpolação pelo inverso da distância elevada a uma potencia.

### 3.2 Procurando modelo

Por meio dos módulos apresentados neste trabalho, pode-se construir um pequeno algoritmo para procurar automaticamente o modelo que apresente melhor resultado, dentre várias possibilidades de semivariogramas teóricos e de número de vizinhos mais próximos para, respectivamente, a interpolação por krigagem ordinária e pelo inverso da distância. Um exemplo de algoritmo que se pode construir é apresentado abaixo. Neste algoritmo são calculadas e armazenadas uma série de estatísticas de tabulações cruzadas para: a) krigagem ordinária considerando uma série de modelos teóricos estimados a partir de diferentes números de semivariâncias, de 8 a 20, e alcance máximo entre os pares amostrais variando de alcance mínimo a 0,8 da distância máxima entre os pares com passo de 0,1; e b) para o inverso da distância ponderações de 1 a 6, e número de amostras mais próximas de 1 até o número de amostras total (71).

```
% ALGORITMO PARA PROCURA DE
% INTERPOLADOR E SEUS RESPECTIVOS
% PARÂMETROS
```

```
dado ← (X,Y,Z)
```

```
% tabulação cruzada para a krigagem ordinária
```

```
alc_min ← alcance_minimo(dado);
```

```
lags ← 8;
```

```
alcance ← alc_min;
```

```
v ← 1;
```

```
% início das iterações para krigagem
```

```
enquanto lags < 20 faça
```

```
    enquanto alcance < .8 faça
```

```
        semiva ← semiva(dado,alcance,lags);
```

```
        betas ← escomodelo2min(semiva);
```

```
        resul ← corre_kriga(betas,dado);
```

```
        resultado_krige(v) ← [resul betas];
```

```
        alcance ← alcance + .1;
```

```

v ← v + 1;
fim-enquanto
lags ← lags + 1;
fim-enquanto

% tabulação cruzada para o inverso da distância

para n_viz ← 1 até 71
resul_invdist(n_viz) ← corre_invdist(dado, n_viz);
fim-para

```

A partir dos dados armazenados nas variáveis

resultado\_krige e resul\_invdist é possível encontrar para cada interpolador e respectivos parâmetros aquele que resultou no melhor resultado segundo alguma estatística da tabulação cruzada. Na Tabela 1 é apresentado o resultado geral segundo o critério de menor RMSE encontrado para cada modelo, além do resultado para os interpolador inverso da distância elevado ponderado por diversas potencias. Todos os resultados foram significativos e aquele que apresentou maior precisão pelo RMSE foi a krigagem ordinária com modelo gaussiano teórico para a semivariância.

TABELA 1 – RESULTADO GERAL DA TABULAÇÃO CRUZADA COM DIVERSOS ÍNDICES ESTATÍSTICOS ENTRE OS VALORES DE MEDIDOS ( $z_i$ ) E ESTIMADOS ( $\hat{z}_i$ ).

Interpolador	Interpolador	R <sup>2</sup>	valor-P	RMSE	d*	c*	EMA*	EMP*
1	Modelo Exponencial	0,77	0,00	0,10	0,90	0,79	0,07	6,99
2	Modelo Esférico	0,75	0,00	0,10	0,90	0,78	0,07	7,30
3	Modelo Gaussiano	0,92	0,00	0,05	0,98	0,94	0,04	3,56
4	Modelo Linear	0,74	0,00	0,10	0,90	0,78	0,07	7,26
5	Inverso distância (r=1)	0,43	0,00	0,15	0,79	0,52	0,11	11,25
6	Inverso distância (r=2)	0,45	0,00	0,15	0,81	0,55	0,11	10,86
7	Inverso distância (r=3)	0,40	0,00	0,15	0,76	0,48	0,11	10,45
8	Inverso distância (r=4)	0,42	0,00	0,15	0,80	0,52	0,11	10,56
9	Inverso distância (r=5)	0,42	0,00	0,16	0,81	0,52	0,11	11,07
10	Inverso distância (r=6)	0,42	0,00	0,17	0,80	0,52	0,12	11,51

\* “d” é Índice de concordância; “c” é índice do confiança; “EMA” é o erro médio absoluto; “EMP” é o erro médio percentual. As fórmulas como são calculados é apresentado em Anexo.

Escolhendo-se o interpolador de maior precisão, é possível, por meio dos seus respectivos parâmetros, realizar a interpolação em qualquer software que possua as ferramentas de interpolação por krigagem ordinária e ou inverso da distância, como o SPRING, SURF, IDRISI, ARC INFO etc. Na Figura 6 é apresentado o resultado da interpolação por krigagem

ordinária, juntamente com as amostras testes utilizadas neste trabalho, segundo os parâmetros do modelo Gaussiano que apresentou o menor RMSE. Verifica-se que a superfície se aproxima de maneira bastante satisfatória aquela que deu sua origem para os pontos amostrais (Figura 1a).

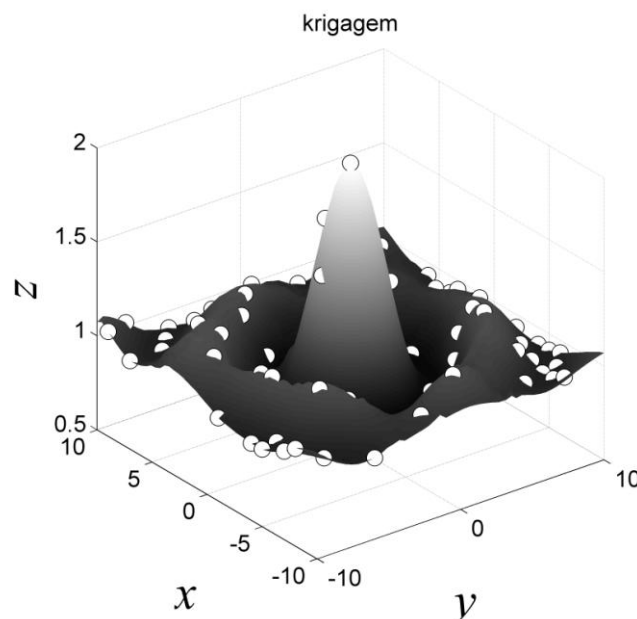


Figura 6 – Resultado da interpolação espacial por krigagem utilizando modelo gaussiano.



### 3. CONCLUSÕES

Neste trabalho é apresentada uma série de módulos em MATLAB para avaliação da interpolação espacial por krigagem ordinária e inverso da distância. A partir dos módulos implementados é possível construir um simples algoritmo para automatizar a escolha dos vários parâmetros que são utilizados para cada uma das técnicas de interpolação. A escolha pode ser realizada por meio diversos índices estatísticos dos dados da tabulação cruzada, isto é, comparação entre os valores reais e os estimados.

### AGRADECIMENTOS

Os autores gostariam de agradecer a Fundação de Apoio à Ciência e Tecnologia do Espírito Santo (FAPES) pelo apoio financeiro para realização deste trabalho (processo número 38939673/2007)

### REFERÊNCIAS BIBLIOGRÁFICAS

BURROUGH, P. A.; MCDONNELL, R. A. **Principles of geographical information systems**. Oxford; New York : Oxford University Press, 2000. 333p.

CAMARA, G.; SOUZA, R.C.M.; FREITAS, U.M.; GARRIDO, J. SPRING: Integrating remote sensing and GIS by object-oriented data modeling. **Computers & Graphics**, v. 20, p. 395-403, 1996.

CASTRO, M. C.; SINGER, B. H. Environment and Health: a methodological approach for spatial assessment of malaria transmission in colonization projects in the Brazilian Amazon. **Revista Brasileira de Estudos de População**, vol. 24, n. 2, p. 247-262, 2007.

EASTMAN, J.R. **IDRISI32 for Windows: User's guide. Version 32.11. Software de sistema de informação geográfica (software)**. Worcester: Clark University, 1998. 192p.

ERTEKIN, C.; EVRENDILEK, F. Spatio-temporal modeling of global solar radiation dynamics as a function of sunshine duration for Turkey. **Agricultural and Forest Meteorology**, v. 145, n. 1-2, p. 36-47, 2007.

ESRI. **Arcgis 9: what is ArcGIS 9.3**. ESRI, 2008. 122p.

FERNADES, M.C.; MENDEZES, P.M. Comparação entre métodos para geração de MDE para a obtenção de observações em superfície real no maciço da Tijuca-RJ. **RBC - Revista Brasileira de Cartografia**.v. 57, n. 2, p. 154-161, 2005.

LANDIM, P. M. B. Análise estatística de dados geológicos. São Paulo: Editora UNESP, 2003.

253p.

MATHWORKS. **Statistics Toolbox™ 6**. User's Guide. The MathWorks, Inc. 2007a.

MATHWORKS. **MATLAB® & Simulink®**. Release Notes for R2008a. The MathWorks, Inc. 2007b.

OLIVEIRA, R. B. DE; LIMA, J. S. DE S.; XAVIER, A. C.; PASSOS, R. R.; SILVA, S. A.; SILVA, A. F. Comparison between soil sampling methods for conilon coffee liming and fertilization recommendation. **Engenharia Agrícola**, vol. 28, n. 1, p. 176-186, 2008.

OPROMOLLA, P. A.; DALBEN, I.; CARDIM, M. Analysis of the spatial distribution of leprosy in The State of São Paulo, 1991-2002. **Revista Brasileira de Epidemiologia**, v.8, n. 4, p. 356-364, 2005.

PEREIRA, W. C.; REMACRE, A. Z. Utilização de variável mineralógica na estimativa de reservas de minério de ferro. **Rem: Revista Escola de Minas**, vol.56, n. 2, p. 123-129, 2003.

ROBERTSON, G.P. **GS+: Geostatistics for the environmental sciences - GS+ User's Guide**. Plainwell: Gamma Desing Software, 2000. 152p.

RODRIGUEZ, R. P.; MARQUES, M. J.; BIENES R. Spatial variability of the soil erodibility parameters and their relation with the soil map at subgroup level, **Science of The Total Environment**. v. 378, n. 1-2, p. 166-173, 2007.

RODRÍGUEZ-LADO, L.; SPAROVEK, G.; VIDAL-TORRADO, P.; DOURADO-NETO, D.; MACÍAS-VÁZQUEZ, F. Modelling air temperature for the state of São Paulo, Brazil. **Scientia Agrícola**, v. 64, n. 5, p. 460-467. 2007.

SARANGI A.; MADRAMOOTOO C. A.; ENRIGHT P. Comparison of Spatial Variability Techniques for Runoff Estimation from a Canadian Watershed **Biosystems Engineering**, v. 95, n. 2, p. 295-308, 2006.

SOUZA, L.A.; SOBREIRA, F.G.; PRADO FILHO, J.F. Cartografia e diagnóstico geoambiental aplicados ao ordenamento territorial do município de Mariana – MG. **RBC - Revista Brasileira de Cartografia**, v. 57, n. 3, p. 189-203, 2005.

TARNAVSKY, E.; GARRIGUES, S.; BROWN, M. E. Multiscale geostatistical analysis of AVHRR, SPOT-VGT, and MODIS global NDVI products, **Remote Sensing of Environment**, v. 112, n. 2, p. 535-549, 2008.

## ANEXO

Seja:  $z_i$  o valor da amostra medida;  $\hat{z}_i$  o valor estimado;  $i = 1, 2, 3, \dots, n$ , sendo  $n$  o número de amostras medidas; e  $\bar{z}$  a média das amostras medidas. As estatísticas calculadas a partir dos dados da tabulação cruzada são:

coeficiente de determinação ( $R^2$ )

$$R^2 = \frac{\sum_{i=1}^n \left( z_i - \frac{\sum z_i}{n} \right) \left( \hat{z}_i - \frac{\sum \hat{z}_i}{n} \right)}{\sqrt{\left[ \sum_{i=1}^n \left( z_i - \frac{\sum z_i}{n} \right)^2 \right] \left[ \sum_{i=1}^n \left( \hat{z}_i - \frac{\sum \hat{z}_i}{n} \right)^2 \right]}}$$

raiz quadrada do erro quadrado médio (RMSE)

$$\text{RMSE} = \left( \frac{\sum_{i=1}^n (z_i - \hat{z}_i)^2}{n} \right)^{1/2}$$

índice de concordância (d)

$$d = 1 - \frac{\sum_{i=1}^n (z_i - \hat{z}_i)}{\sum_{i=1}^n (|\hat{z}_i - \bar{z}| |z_i - \bar{z}|)}$$

índice do confiança (c)

$$c = \sqrt{R^2} \cdot d$$

erro médio absoluto (EMA)

$$\text{EMA} = \frac{\sum_{i=1}^n |z_i - \hat{z}_i|}{n}$$

erro médio percentual (EMP)

$$\text{EMP} = 100 \frac{1}{n} \frac{\sum_{i=1}^n |z_i - \hat{z}_i|}{z_i}$$