

APLICAÇÃO DO DIAGRAMA DE TAYLOR PARA AVALIAÇÃO DE INTERPOLADORES ESPACIAIS EM ATRIBUTOS DE SOLO EM CULTIVO COM EUCALIPTO¹

Daniel Pena Pereira², Julião Soares de Souza Lima³, Alexandre Cândido Xavier³, Renato Ribeiro Passos⁴ e Nilton César Fiedler⁵

RESUMO – As diversas técnicas de interpolação espacial adotadas oferecem desempenhos diferentes, de acordo com as características dos dados iniciais, e é muito comum encontrar avaliadores contando somente com R^2 (coeficiente de determinação múltiplo) e erro residual. Neste trabalho, objetivou-se aplicar o Diagrama de Taylor para analisar 15 métodos diferentes de interpolação espacial, em área com cultivo comercial de eucalipto, para comparação e escolha do melhor método de interpolação de um conjunto de valores da fração granulométrica do solo (argila), na profundidade de 0-0,20 m do solo. De acordo com os resultados, considera-se muito satisfatória a ferramenta do Diagrama de Taylor, pois, além de definir graficamente os melhores métodos de interpolação, essa ferramenta permite fazer escolhas entre eles, dentro de um conjunto menor e mais definido de informações; logo, optou-se pelo método Splines em detrimento da krigagem modelo linear. As estatísticas desses dois métodos estão muito próximas, com pequenas variações, estando o desvio-padrão do Splines mais próximo dos dados observados; logo, é o melhor método de interpolação para argila, na profundidade de 0-0,20 m do solo.

Palavras-chave: Interpoladores gráficos; Distribuição espacial; Textura do solo.

TAYLOR DIAGRAM APPLICATION FOR THE EVALUATION OF SPATIAL INTERPOLATORS IN ATTRIBUTES OF SOIL UNDER CULTIVATION WITH EUCALYPTUS

ABSTRACT – The various spatial interpolation techniques used offer different performances, according to the characteristics of the initial data and it is very common to find evaluators relying only on R^2 (coefficient of multiple determination) and residual error. This study aimed to apply the Taylor diagram to analyze 15 different methods of spatial interpolation in area with commercial cultivation of eucalyptus, to compare and choose the best method of interpolation of a set of values of the size granulometric fraction of the soil (clay) at the depth of 0-0.20 m from the soil. According to the results, the diagram Taylor tool was considered very satisfactory, because in addition to graphically define the best interpolation methods, it allows to make choices among them within a smaller and more defined set of information; therefore, the Splines method was chosen rather than linear kriging model. The statistics of these two methods were very similar, with minor variations, and the standard deviation of the Splines was closer to the observed data, which means that it is the best interpolation method for clay at a soil depth of 0-0.20 m.

Keywords: Graphics interpolators; Spatial distribution; Soil texture.

¹ Recebido em 10.10.2013 aceito para publicação em 28.07.2014.

² IPrograma de Pós-graduação em Produção Vegetal, Universidade Federal do Espírito Santo, UFES, Brasil. E-mail: <daniel.geraes@gmail.com>.

³ Departamento de Engenharia Rural, Universidade Federal do Espírito Santo, UFES, Brasil. E-mail: <limajss@yahoo.com.br> e <alexandre.candido.xavier.ufes@gmail.com>.

⁴ Departamento de Produção Vegetal, Universidade Federal do Espírito Santo, UFES, Brasil. E-mail: <renatoribeiropassos@hotmail.com>.

⁵ Departamento de Engenharia Florestal, Universidade Federal do Espírito Santo, UFES, Brasil. E-mail: <fiedler@pq.cnpq.br>.

1. INTRODUÇÃO

A agricultura de precisão, como ferramenta para melhor gerenciar o processo produtivo, tem como objetivo o aumento da rentabilidade ao produtor e à redução de impactos ambientais, assegurando competitividade e sustentabilidade à atividade rural. Essa técnica vem sendo utilizada nas mais diferentes áreas, como: definição de zonas de manejo (MORAL et al., 2010; AGGELOPOOULOU et al., 2013), controle e aplicação de fertilizantes (SANCHEZ et al., 2012), produtividade de essências florestais (RIGATTO et al., 2005; AMARAL et al., 2013) e qualidade da madeira (LIMA et al., 2006).

Para aplicação de métodos de interpolação, (1) há necessidade de amostragem da variável de interesse; (2) com a finalidade de torná-la contínua, faz-se a interpolação, segundo determinado método; (3) criam-se mapas em cores ou escalas de cinza ou linhas de contorno (LI; HEAP, 2008). O método de interpolação a ser aplicado vai depender de vários fatores, como: as características dos dados, o conhecimento técnico e a disponibilidade de softwares. Quando mais de um interpolador é utilizado, uma ou mais estatísticas dos dados da tabulação cruzada são usadas para escolher o método de interpolação. Podem-se citar como estatísticas utilizadas o coeficiente de determinação (R^2) e o erro médio quadrático.

Os interpoladores normalmente aplicados na espacialização das variáveis de solo são: os interpoladores geoestatísticos (krigagem), inverso da distância ao quadrado (IDQ) e Splines (SOUZA et al., 2010; MORAL et al., 2010; OMRAN, 2012).

Taylor (2001) criou um simples diagrama para comparação de modelos que estimavam variáveis meteorológicas. Esse diagrama fornece visualmente um resumo das estatísticas: desvio-padrão, correlação e raiz quadrada da média do quadrado das diferenças centralizadas dos dados observados e modelados. Dessa forma, evita-se a construção de extensas tabelas.

Objetivou-se desenvolver um modelo para comparação e escolha do melhor método de interpolação de um conjunto de valores de dados relacionados à granulometria do solo (teor de argila), utilizando-se, como ferramenta, o Diagrama de Taylor para analisar 15 métodos diferentes de interpolação espacial, em área com cultivo comercial de eucalipto.

2. MATERIAL E MÉTODOS

2.1. Área de estudo

A área de estudo ($-41,63^\circ$, $-20,94^\circ$ SAD69 Fuso 24S) cobre cerca de 97.580 m^2 , dentro de uma área sob Latossolo Vermelho-Amarelo distrófico no Sul do Estado do Espírito Santo, localizada na parte Norte do Município de São José do Calçado.

2.2. Métodos

Foram coletadas 94 amostras de solo superficiais (0-0,20 m) na área de estudo, segundo um esquema de amostragem em grade regular de $33 \text{ m} \times 33 \text{ m}$, em novembro de 2008, incluindo registro das coordenadas geográficas e altitude. Para determinação dos teores de argila, seguiu-se a metodologia conforme Embrapa (1997).

Precedeu-se à análise exploratória dos dados – que se constituiu na visualização das medidas de dispersão e histogramas de frequência dos dados para checar se existem problemas amostrais com os dados; e análise de *outliers*, por meio de gráficos *boxplot*, com o objetivo de verificar se há ou não tendência dos dados e retirá-los quando constatados, utilizando-se o programa Matlab (MATLAB, 2010). *Outliers* são aqueles valores de um atributo plotado 1,5 vez a diferença interquartílica acima do terceiro quartil e 1,5 vez a diferença interquartílica abaixo do primeiro quartil.

Os interpoladores utilizados foram: a) Inverso do quadrado da distância usando potências de 1 a 6 (BURROUGH; MCDONNELL, 2006); b) Análise geoestatística por krigagem ordinária e considerando para a modelagem da semivariância os modelos linear, esférico, exponencial e gaussiano (VIEIRA et al., 1983); c) Splines (SANDWELL, 1987; AMORIM et al., 2008); d) Vizinho mais próximo (MCROBERTS, 2012); e e) Regressão simples usando os modelos polinomiais linear, quadrático e cúbico e tendo como variáveis independentes as coordenadas x e y em graus Datum SAD69 (BURROUGH; MCDONNELL, 2006). No total, foram 15 métodos de interpolação utilizados.

2.3. O Diagrama de Taylor

Para construção do diagrama de Taylor, foram considerados os dados da tabulação cruzada de cada método, ou seja, os dados observados (r) e os estimados (f).

Uma das estatísticas mais utilizadas para observar a qualidade do modelo é comparando essas duas variáveis pela Raiz Quadrada da Média do Quadrado das Diferenças (E), dado por:

$$E = \left[\frac{1}{N} \sum (f_n - r_n)^2 \right]^{1/2} \quad (1)$$

sendo e dados estimados e observados, respectivamente.

A Raiz Quadrada da Média do Quadrado das Diferenças Centralizadas (E') corresponde à segunda parte da equação 1.

$$E'^2 = s_f^2 + s_r^2 - 2s_f s_r R \quad (2)$$

sendo S_f e S_r os desvios-padrão dos dados estimados e observados, respectivamente; R é o coeficiente de correlação.

Nessa última fórmula, encontram-se quatro estatísticas (E', S_f, S_r, R), que podem ser utilizadas para estudar o padrão da relação entre f e r , de forma a orientar se o modelo é um bom estimador. Taylor (2001) propôs a construção de um diagrama que contenha essas quatro estatísticas. Parte-se da similaridade da equação 2 com a equação da lei dos cossenos, que relaciona um ângulo interno de um triângulo com seus lados ($a^2 = b^2 + c^2 - 2bc \cos \theta$). O resultado geométrico da aplicação da lei dos cossenos na equação 2 está representado na Figura 1.

Fonte: Adaptado de TAYLOR, 2001.
Source: Adapted from TAYLOR, 2001.

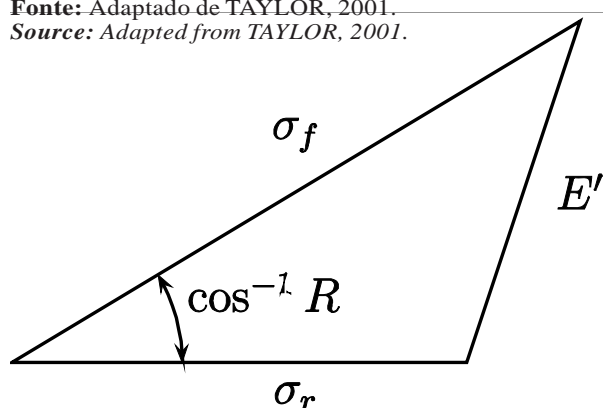


Figura 1 – Representação geométrica da relação das estatísticas E', σ_f, σ_r e R .

Figure 1 – Geometric representation of the relation of statistics E', σ_f, σ_r e R .



A construção do diagrama de Taylor se dá pela representação de $\frac{1}{2}$ ou $\frac{1}{4}$ de círculo, sendo este último o mais comum. Os eixos x e y têm as medidas do desvio-padrão, sendo sobre o eixo x colocado o valor de (S_r). A distância radial da origem à posição representando o modelo é o S_f . O azimute da origem ao modelo é proporcional a R (Figura 2).

A distância entre o ponto de referência no diagrama ao ponto do modelo é o valor de E' . Circunferências com a origem em ($S_r, 0$) e diferentes raios (E') podem ser construídas para apresentar a magnitude de E' (Figura 2).

3. RESULTADOS

Foram utilizados 15 métodos de interpolação, mas como compará-los? Qual método produz o resultado mais consistente? Para uma primeira visualização dos resultados das interpolações, ver Figura 3.

N Figura 3, cada método é representado por um símbolo, em que a distância radial do ponto de origem é proporcional ao desvio-padrão dos dados interpolados e, em azimute do ponto, é o valor de R dos dados observados e estimados pelo método. A correlação do ponto de referência com ele próprio é igual a 1 e, por conseguinte, verifica-se no eixo x .

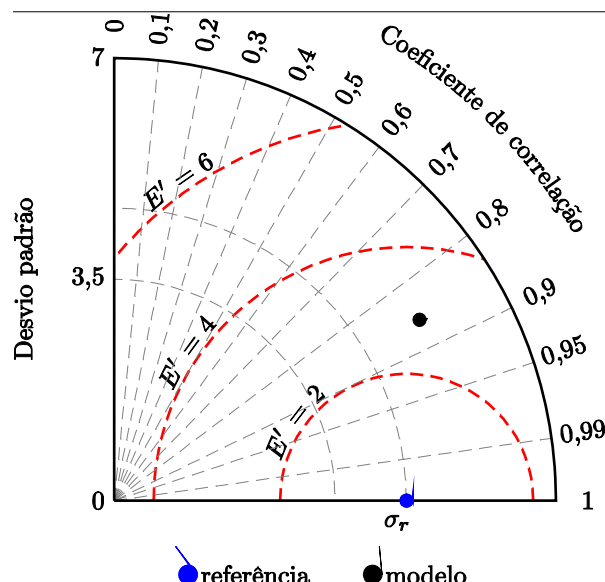


Figura 2 – Exemplo de construção do Diagrama de Taylor.
Figure 2 – Example of Taylor Diagram construction.

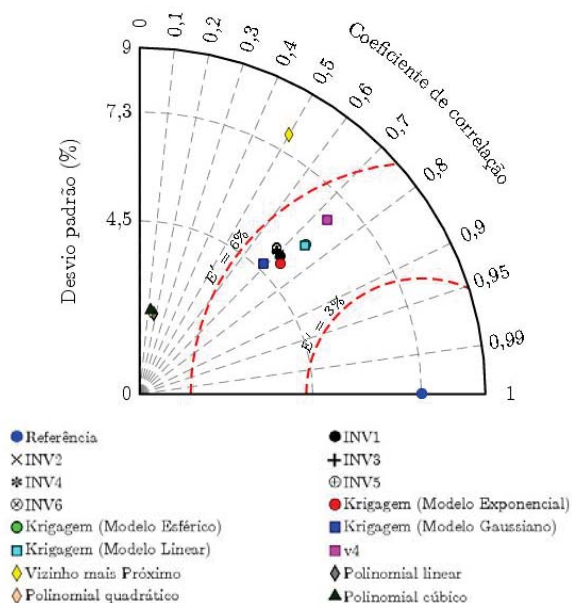


Figura 3 – Aplicação do Diagrama de Taylor representando graficamente as estatísticas e a tabulação cruzada, referentes aos interpoladores utilizados. A saber: Inverso do quadrado da distância usando potências de 1 a 6 (INV 1-6), Análise geoestatística (modelos linear, esférico, exponencial e gaussiano), Splines (v4), Vizinho mais próximo e Regressão simples usando os modelos polinomiais linear, quadrático e cúbico.

Figure 3 – Application of Taylor diagram representing graphically the statistics and of the cross tabulation, concerning interpolators used, namely: Inverse of squared distance using powers of 1 to 6 (INV 1-6), Geostatistical analysis (linear, spherical, exponential, Gaussian models), Splines (v4), Nearest-neighbor and Simple regression models using polynomial linear, quadratic and cubic models.

4. DISCUSSÃO

O diagrama descrito aqui começou a ser visto em uso em muitos estudos, mesmo antes da publicação de sua descrição (RÄISÄNEN, 1997; GATES et al., 1998), inclusive com referências bem anteriores (TAYLOR; PENNER, 1994).

Esse diagrama explora a equivalência demonstrada na Figura 1, para fornecer a maneira conveniente de comparar diversas estatísticas entre dados observados e estimados, no exemplo deste trabalho, dados de argila interpolados. Quanto mais próximo o ponto de um método à referência, melhor o método de interpolação.

Em relação aos métodos testados, notaram-se os métodos do Vizinho Mais Próximo e Polinomiais como os menos adequados para o objetivo proposto. O primeiro mostrou desvio-padrão praticamente igual ao da referência, entretanto o seu valor de E' foi mais distante da referência, apresentando baixa correlação ($< 0,45$). Os métodos polinomiais estão situados na parte inferior do diagrama e à esquerda do gráfico, indicando baixo valor de R e desvios-padrão dos dados interpolados muito menores que aqueles observados nos dados de referência ($S_r = 7,32\%$).

Os modelos do inverso da distância ao quadrado (IDQ) nas potências de 1 a 6 e os modelos geoestatísticos definidos pelo semivariograma exponencial e gaussiano apresentaram padrão semelhante em relação a todas as estatísticas, com R em torno de 0,7 e S_f de aproximadamente 4,7%, menor do que “s”.

A krigagem a partir do semivariograma linear foi o que apresentou o mais alto valor de R e o menor valor de E' . Seu desvio-padrão foi menor do que o da referência. O método Splines merece destaque, além de os valores de R e E' serem próximos aos da krigagem com o semivariograma linear, e o seu desvio-padrão é mais próximo do S_r .

A análise do diagrama de Taylor permite a flexibilidade na escolha do método que será utilizado. Pode por exemplo, a despeito de o modelo Splines não apresentar, nas estatísticas analisadas, todos os melhores resultados, ser o método a utilizar para espacialização da porcentagem de argila. Uma vez que, como apresentado anteriormente, além de o seu valor de R e E' ser próximo dos melhores valores encontrados, o seu desvio-padrão está mais próximo do .

No diagrama de Taylor, o *bias* (que é a diferença entre a média dos valores estimados (f) e a dos observados (r); $bias = f - r$) não é possível ser observado. Todavia, pode-se construir outro diagrama, em que o eixo-x é referente ao valor do *bias* e o eixo-y é a estatística E' , ou visualizar em tabelas. A distância da origem desse sistema ao ponto de um modelo é a estatística E (equação 1). Os valores de *bias* são menores que 0,5%, e o maior *bias* é referente à interpolação por krigagem com o modelo de semivariância gaussiano.

A demonstração e comparação de resultados em formas de tabelas e figuras complexas têm sido a apresentação mais comum entre muitos autores (BABAK; DEUTSCH, 2009; OMRAN, 2012; PAVÃO et al., 2012).

Também, analisou-se o desempenho de interpoladores usando basicamente, desde somente o resíduo médio (MAZZINI; SCHETTINI, 2009) e o coeficiente de correlação (PEZZOPANE et al., 2004; MAILLARD; COSTA-PEREIRA, 2011), o grau de dependência espacial e o coeficiente de determinação (R^2), no caso de análise geoestatística nos trabalhos de Lima et al. (2008) e Sanchez et al. (2012) até análises usando a média, o desvio-padrão, a raiz quadrada do erro médio quadrático (RMSE) e o coeficiente de correlação (AMORIM et al., 2008; MCROBERTS, 2012).

O Diagrama de Taylor aumenta a qualidade da discussão sobre o desempenho e a escolha de determinado modelo, pois é possível analisar simultaneamente uma série de estatísticas dos dados observados e estimados. Neste trabalho, optou-se pela escolha do método Splines para interpolação. Na Tabela 1, mostram-se as estatísticas derivadas da tabulação cruzada por esse método, em termos das médias (\bar{x}), coeficientes de correlação (R), desvios-padrão de referência (S_r) e estimativas (S_p) e médias da raiz quadrada do quadrado das diferenças centralizadas (E'), *bias* ($f-r$).

Na Figura 4 é apresentado o resultado da espacialização da porcentagem de argila pelo método Splines.

Nota-se, pela Figura 4, que a maior parte da área possui teor médio de argila, à exceção de duas pequenas áreas (1) mais claras, apresentando teor mais elevado de argila e localizada em parte convexa, com maior retirada de areia do sistema; e (2) mais escuras, com teor menor, sendo a parte mais baixa na topografia local e com maior acúmulo de areia ao longo do tempo.

4. CONCLUSÕES

1.O Diagrama de Taylor mostra-se poderosa ferramenta para analisar graficamente diferentes resultados de interpolação da argila; pode ser aplicado a outras variáveis de solo e constitui referência para trabalhos em agricultura de precisão.

Tabela 1 – Resumo dos resultados da interpolação da argila pelo método Splines.

Table 1 – Summary of the results of clay interpolation by Splines method.

Método de interpolação	\bar{x}	s	E	R	<i>bias</i>
Teor de argila original (%)	50,38	7,32	0,00	1,00	0,00
Teor de argila interpolado pelo método Splines (%)	50,30	6,63	5,14	0,73	-0,08

Nota: \bar{x} = média; s = desvio-padrão; E = raiz quadrada da média do quadrado das diferenças centralizadas; e R = correlação.
Note: \bar{x} = average; s = standard deviation; E = square root of the average of the square of the centralized differences; and R = correlation.

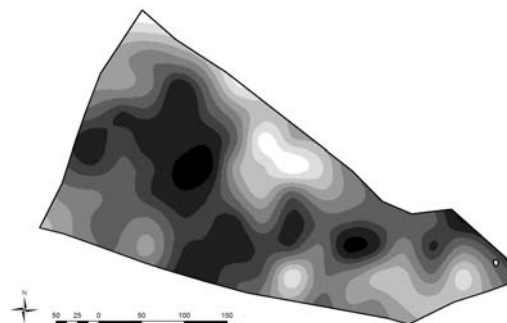


Figura 4 – Mapa da distribuição espacial da argila, interpolado pelo método Spline.

Figure 4 – Map of the spatial distribution of clay interpolated by Splines method.

2.O melhor método encontrado é o Spline, considerando que o seu desvio-padrão se aproxima mais dos dados da amostra, sua correlação e o *bias* são muito semelhantes ao método da krigagem modelo linear, apesar de o E' ser um pouco maior do que esse método.

3.Os resultados indicaram que as estatísticas obtidas na tabulação cruzada devem ser sempre adotadas como critério de definição do melhor método de interpolação espacial, visualizadas facilmente através de diagramas circulares adaptados e que contenham as estimativas da correlação (R), desvio-padrão dos valores observados e estimados e o RMS centralizado (E').

5. REFERÊNCIAS

AGGELOPOULOU, K. et al. Delineation of management zones in an apple orchard in Greece using a multivariate approach. **Computers and Electronics in Agriculture**, v.90, n.1, p.119-130, 2013.

AMARAL, L. P. et al. Influência da floresta alterada na distribuição espacial de três espécies da Floresta Ombrófila Mista avaliada pela geoestatística. **Revista Árvore**, v.37, n.3, p.491-501, 2013.

AMORIM, R. C. F. et al. Avaliação de desempenho de dois métodos de espacialização da precipitação pluvial para o Estado de Alagoas. **Acta Scientiarum Technology**, v.30, n.1, p.87-91, 2008.

BABAK, O.; DEUTSCH, C. V. Statistical approach to inverse distance interpolation. **Stochastic Environmental Research and Risk Assessment**, v.23, n.5, p.543-553, 2009.

BURROUGH, P. A.; MCDONNELL, R. A. **Principles of geographical information systems**. New York: Oxford University Press, 2006. 333p.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA - EMBRAPA. Centro Nacional de Pesquisa de Solos. **Manual de métodos de análise de solo**. 2.ed. Rio de Janeiro: 1997. 212p. (EMBRAPA CNPS. Documentos, 1). Disponível em: <http://www.agencia.cnptia.embrapa.br/Repositorio/Manual+de+Metodos_000fzvhotqk02wx5ok0q43a0ram31wtr.pdf>. Acesso em: 27 maio 2013.

GATES, W. L. et al. **An overview of the results of the atmospheric model intercomparison project (AMIP)**. The Program for Climate Model Diagnosis and Intercomparison: PCMDI Report n° 45. Livermore: 1998. 47 p. (UCRL-JC-129928). Disponível em: <<http://ccpp.llnl.gov/publications/pdf/45.pdf>>. Acesso em: 13 dez. 2012.

LI, J.; HEAP, A. D. **A review of spatial interpolation methods for environmental scientists**. Geoscience Australia. Record, 2008. 137p.

LIMA, J. S. S. et al. Métodos geoestatísticos no estudo da resistência do solo à penetração em trilha de tráfego de tratores na colheita de madeira. **Revista Árvore**, v.32, n.5, p.931-938, 2008.

LIMA, J. S. S. et al. Estudo da viabilidade de métodos geoestatísticos na mensuração da variabilidade espacial da dureza da madeira de paraju (*Manilkara* sp.). Viçosa-MG: **Revista Árvore**, v.30, n.4, p.651-657, 2006.

MAILLARD, P.; COSTA-PEREIRA, P. S. Estimação da idade da regeneração da vegetação de cerrado a partir de imagens Landsat. **Revista Geográfica Acadêmica**, v.5, n.1, p.34-47, 2011.

MATLAB: The Language of Technical Computing. The MathWorks, Matlab Version, R2010a. 2010.

MAZZINI, P. L. F.; SCHETTINI, C. A. F. Avaliação de metodologias de interpolação espacial aplicadas a dados hidrográficos costeiros quase sinóticos. **Brazilian Journal of Aquatic Science and Technology**, v.13, n.1, p.53-64, 2009.

MCROBERTS, R. E. Estimating forest attribute parameters for small areas using nearest neighbors techniques. **Forest Ecology and Management**, v.272, n.1, p.3-12, 2012.

MORAL, F. J.; TERRÓN, J. M.; SILVA, J. R. M. Delineation of management zones using mobile measurements of soil apparent electrical conductivity and multivariate geostatistical techniques. **Soil & Tillage Research**, v.106, n.2, p.335-343, 2010.

OMRAN, E. L. E. Improving the prediction accuracy of soil mapping through geostatistics. **International Journal of Geosciences**, v.3, p.574-590, 2012.

PAVÃO, C. G. et al. Spatial interpolation applied a crustal thickness in Brazil. **Journal of Geographic Information System**, v.4, n.2, p.142-152, 2012.

PEZZOPANE, J.E.M.; SANTOS, E.A.; ELEUTÉRIO, M.M.; REIS, E.F. dos; SANTOS, A.R. dos. Espacialização da temperatura do ar no Estado do Espírito Santo. Santa Maria-RS: **Revista Brasileira de Agrometeorologia**, v.12, n.1, p.151-158, 2004.

RÄISÄNEN, J. Objective comparison of patterns of CO₂ induced climate change in coupled GCM experiments. **Climate Dynamics**, v.13, p.197-211, 1997.

RIGATTO, P. A.; DEDECEK, R. A.; MATTOS, J. L. M. Influência dos atributos do solo sobre a produtividade de *Pinus taeda*. **Revista Árvore**, v.29, n.5, p.701-709, 2005.

SANCHEZ, B. R. et al. Mapeamento das formas do relevo para estimativa de custos de fertilização em cana-de-açúcar. **Engenharia Agrícola**, v.32, n.2, p.280-292, 2012.

SANDWELL, D. T. Biharmonic spline interpolation of GEOS-3 and SEASAT altimeter data. **Geophysical Research Letters**, v.2, p.139-142, 1987.

SOUZA, G. S. et al. Krigagem ordinária e inverso do quadrado da distância aplicados na espacialização de atributos químicos de um argissolo. **Scientia Agraria**, v.11, n.1, p.73-81, 2010.

TAYLOR, K. E. Summarizing multiple aspects of model performance in a single diagram. **Journal of Geophysical Research**, v.106, n.7, p.7183-7192, 2001.

TAYLOR, K. E.; PENNER, J. E. Response of the climate system to atmospheric aerosols and greenhouse gases. **Nature**, v.369, p.734-737, 1994.

VIEIRA, S. R. et al. Geostatistical theory and application to variability of some agronomical properties. **Hilgardia**, v.51, n.3, p.1-75, 1983.