



# **Cálculo de superfícies de tendência, por regressão polinomial, pelo SURFER**

***PAULO M. BARBOSA LANDIM***

**Professor Voluntário do Depto. Geologia Aplicada  
UNESP/Rio Claro**

***ALESSANDRA CRISTINA CORSI***

**Doutoranda em "Geociências e Meio Ambiente"  
UNESP/Rio Claro**

**UNESP/campus de Rio Claro  
Departamento de Geologia Aplicada - IGCE  
— Laboratório de Geomatématica —  
Texto Didático 05  
2001**

**Reprodução autorizada desde que citada a fonte**

**Norma 6023-2000/ABNT ( <http://www.abnt.org.br>):**

**LANDIM, P.M.B. & CORSI, A.C. Cálculo de superfícies de tendência, por regressão polinomial, pelo SURFER 6. DGA,IGCE,UNESP/Rio Claro, Lab. Geomatématica, Texto Didático 05, 11 pp. 2001. Disponível em <<http://www.rc.unesp.br/igce/aplicada/textodi.html>>.**

**Acesso em:....**

## **DÚVIDAS**

Dúvidas, questões, sugestões, etc. sobre o texto deverão ser encaminhadas para o endereço [plandim@rc.unesp.br](mailto:plandim@rc.unesp.br), as quais serão sempre bem recebidas

## CÁLCULO DE SUPERFÍCIES DE TENDÊNCIA, POR REGRESSÃO POLINOMIAL, PELO “SURFER®”

### Regressão polinomial

- método pelo qual uma superfície contínua é ajustada, por critérios de regressão por mínimos quadrados, aos valores de  $Z_i$  como uma função linear das coordenadas X-Y dos pontos amostrados e irregularmente distribuídos
- a equação matemática utilizada para o ajuste da superfície baseia-se nos polinômios não-ortogonais
- o ajuste é incrementado pela adição de termos adicionais (ordens) à equação polinomial
- após o ajuste da superfície aos dados amostrados, segundo o grau desejado, os valores de  $Z^*_i$  para os nós da grade são calculados
- após a solução das equações, com a determinação dos coeficientes, as mesmas são utilizadas para o cálculo de  $Z^*$  para qualquer valor de X-Y, no caso a localização dos nós das células da grade
- se necessário, é possível calcular os resíduos entre a superfície gerada e os valores originais.
- podem ser obtidos contornos muito suaves em superfícies de alto grau e as isolinhas podem não ser fiéis aos dados originais
- técnica é adequada para "remoção" de tendências e destaque de resíduos.
- alguns artefatos indesejados podem ser gerados nas bordas e no interior do mapa quando houver áreas sem dados amostrados e com a utilização de ordens polinomiais elevadas.
- cuidados devem ser tomados quando da aplicação da análise de tendência:
  - procurar tecer considerações apenas em relação à área coberta pelos pontos evitando as extremidades dos mapas;
  - o número de pontos deve ser maior que o número de coeficientes do polinômio a ser calculado;
  - o arranjo dos pontos, ainda que irregular, deve ser casual e razoavelmente bem distribuído, evitando agrupamentos;
  - quando da inversão da matriz, por programas em microcomputador, podem ocorrer problemas com os resultados obtidos para superfícies de mais alto grau; isso porque em sistemas com valores de diversos dígitos, tipo UTM, a precisão computacional se deteriora exigindo *palavras* de dupla precisão; mesmo assim pode ocorrer limitações e, então, a solução é a transformação das coordenadas  $x_i$  e  $y_i$  conforme as equações, que fornecem valores para as coordenadas entre 0 e 1 e não modifica a forma das superfícies:

$$x^* = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad y^* = \frac{y_i - y_{\min}}{y_{\max} - y_{\min}}$$

- **Equações polinomiais**

Superfície de primeira ordem:  $Z^* = a + bX + cY$

Superfície de segunda ordem:  $Z^* = a + bX + cY + dX^2 + eXY + fY^2$

Superfície de terceira ordem:  $Z^* = a + bX + cY + dX^2 + eXY + fY^2 + gX^3 + hX^2Y + iXY^2 + jY^3$

Superfície de quarta ordem:  $Z^* = a + bX + cY + dX^2 + eXY + fY^2 + gX^3 + hX^2Y + iXY^2 + jY^3 + kX^4 + lX^3Y + mX^2Y^2 + nXY^3 + oY^4$

onde:

$Z^*$  = valor estimado de  $Z^*_i$  para o nó da célula (variável dependente)

X e Y = coordenadas  $X_i$  e  $Y_i$  (variáveis independentes)

a...o = coeficientes que proporcionam o melhor ajuste aos dados amostrados.

- **Ajuste das superfícies polinomiais aos dados observados**

Computadas a soma de quadrados da variável dependente, a soma de quadrados devido à superfície polinomial e a soma de quadrados dos resíduos, pode-se obter uma indicação da validade da superfície calculada utilizando a análise de variância:

variação total dos dados:  $SQT = \sum Z_i^2 - [\sum Z_i]^2/n$

variação dos dados devido à superfície calculada:

$SQP = \sum Z_i^{*2} - [\sum Z_i^*]^2/n$

variação devido aos resíduos:  $SQR = SQT - SQP$

porcentagem de ajuste da superfície = coeficiente de determinação =  $R^2$

$R^2 = ( SQP / SQT ) 100\%$

Fontes de variação	SQ	g.l.	MQ	F
Regressão polinomial	$SQP$	$m$	$MSP$	$\frac{MSP}{MSR}$
Resíduos	$SQR$	$n - m - 1$	$MSR$	
Total	$SQT$	$n - 1$		

$m$  = número de coeficientes da equação polinomial, não contando o termo  $a_0$ .

$n$  = número de observações

$H_0$  = não ocorre ajuste significativo da superfície aos dados.

$H_1$  = ocorre ajuste significativo da superfície calculada aos dados.

<b><i>Vantagens</i></b>	<b><i>Desvantagens</i></b>
<ul style="list-style-type: none"> <li>• uma única superfície é gerada.</li> <li>• fácil definição de parâmetros.</li> <li>• a mesma superfície é gerada mesmo com mudança na orientação da grade.</li> <li>• tempo para cálculo de superfícies de baixa ordem é baixo.</li> <li>• contempla tanto as tendências regionais quanto anomalias locais.</li> <li>• estima valores acima e abaixo dos amostrados.</li> </ul>	<ul style="list-style-type: none"> <li>• extrapola valores de Z para além dos limites da área amostrada.</li> <li>• anomalias locais não são vistas em mapas de superfícies de baixa ordem, porém podem ser destacadas em mapas para os resíduos.</li> <li>• utilização torna-se facilmente abusiva: a tentativa de especificar um ajuste de uma superfície de alta ordem pode ser maior do que o bom senso quanto ao resultado</li> <li>• a quantidade de RAM necessária aumenta exponencialmente com o aumento da ordem do polinômio.</li> </ul>
<b><i>Quando usar</i></b>	<b><i>Quando não usar</i></b>
<ul style="list-style-type: none"> <li>• número adequado de pontos amostrados estiver disponível, sempre maior que o número de coeficientes da equação: <ul style="list-style-type: none"> <li>• ordem=1, coeficientes=2, pontos <math>\geq 3</math></li> <li>• Ordem=2, coeficientes=5, pontos <math>\geq 5</math></li> </ul> </li> <li>• dados forem regularmente distribuídos.</li> <li>• como um "pré-processamento", para remover a tendência regional antes de usar o método da krigagem ou estimar por IQD.</li> <li>• "gerar" novos dados em áreas com dados esparsos.</li> </ul>	<ul style="list-style-type: none"> <li>• poucos dados, com distribuição irregular ou para uma superfície real com alta variabilidade local</li> <li>• pontos amostrados em <i>clusters</i> e valores de Z altamente variáveis.</li> <li>• superfície for descontínua por falhas ou inconformidades.</li> <li>• amplitude da superfície variar drasticamente ou erráticamente: anomalias locais de grande variação.</li> </ul>

### **Análise de superfícies de tendência**

O comportamento espacial de variáveis mapeáveis pode ser mostrado com os valores distribuindo-se segundo curvas de mesmo valor, também conhecidas como *isopletas*. Tais mapas, como os topográficos ou os de isópacas, com linhas de mesma espessura de camadas, fornecem importantes informações, porém, em algumas situações os padrões de variação não se mostram muito claros devido a flutuações locais ou a valores anômalos. É comum nessas circunstâncias falar-se em tendências regionais que são mascaradas por anomalias locais. O método da análise de superfícies de tendência pode, então, ser utilizado para evidenciar tal situação, pois, segundo

esse procedimento define-se, além das grandes e sistemáticas mudanças existentes na área, aquelas pequenas, aparentemente não ordenadas flutuações, que se impõem aos padrões mais gerais.

A análise de superfícies de tendência é uma técnica relativamente simples e muito útil quando os mapas de tendência e os respectivos resíduos podem ser interpretados a partir de um ponto de vista espacial ou então quando o número de observações é limitado de modo que a interpolação possa ser baseada nesses poucos dados.

No modelo linear simples, se possuído um conjunto de dados nos quais foram medidas duas variáveis  $x$  e  $y$ , cuja correlação entre si indica um comportamento linear, pode-se ajustar uma reta que melhor se encaixe a esses pares de valores pelo método dos mínimos quadrados. Esse processo permite a construção de uma única reta em relação a qual a somatória das diferenças ao quadrado entre os valores observados menos os correspondentes computados é mínima.

Partindo desse caso bidimensional, para o modelo linear geral, a analogia é óbvia com o caso tridimensional, onde se deseja correlacionar a distribuição de uma variável dependente  $z$  em função das coordenadas  $x$ , no sentido leste-oeste, e  $y$ , no sentido norte sul. Nessas circunstâncias, deve-se calcular, em vez de uma reta, uma superfície que melhor se adapte ao conjunto de observações por meio de técnicas matemáticas que fornecerão a melhor superfície mapeável e objetiva. Uma dessas técnicas é a análise de superfícies de tendência. Com a aplicação dessa análise consegue-se separar dados mapeáveis em duas componentes: uma de natureza regional, representada pela própria superfície, e outra que revela as flutuações locais, representadas pelos valores residuais.

Se as coordenadas forem determinadas a partir de uma grade regular em que os intervalos são iguais segundo cada uma das duas direções e se existe a possibilidade da variação de  $z_i$  ocorrer segundo um padrão cíclico, o modelo da análise das séries de Fourier pode ser aplicado. Se as observações, porém, não obedecem a uma periodicidade e são coletadas segundo uma grade regular é possível efetuar uma análise de tendência a partir de polinômios ortogonais.

A coleta tendo sido feita, porém, de modo irregular, o que normalmente acontece em Geologia, o recurso a ser usado é o do método da regressão polinomial, tentando encaixar a preliminarmente uma superfície linear aos dados, em seguida uma quadrática, uma cúbica e assim por diante. O método usual para o ajustamento aos dados é o da regressão pelos mínimos quadrados. Em seguida essas superfícies, e os respectivos desvios, são examinadas para que se verifique a sua implicação geológica. Em alguns casos, como em problemas de suavização, o interesse é pelo melhor ajuste aos dados e assim procura-se pela superfície de mais alto grau possível. Em outros, como na detecção de anomalias, o que interessa são os resíduos e calculam-se, então, superfícies de baixo grau com os respectivos mapas de resíduos positivos e negativos.

### **EXEMPLO**

Os dados para este exemplo foram retirados do arquivo *example.dat*, que acompanha o programa GeoEas, para a variável “cadmio”. Para os cálculos foi usado o SURFER na versão 6.

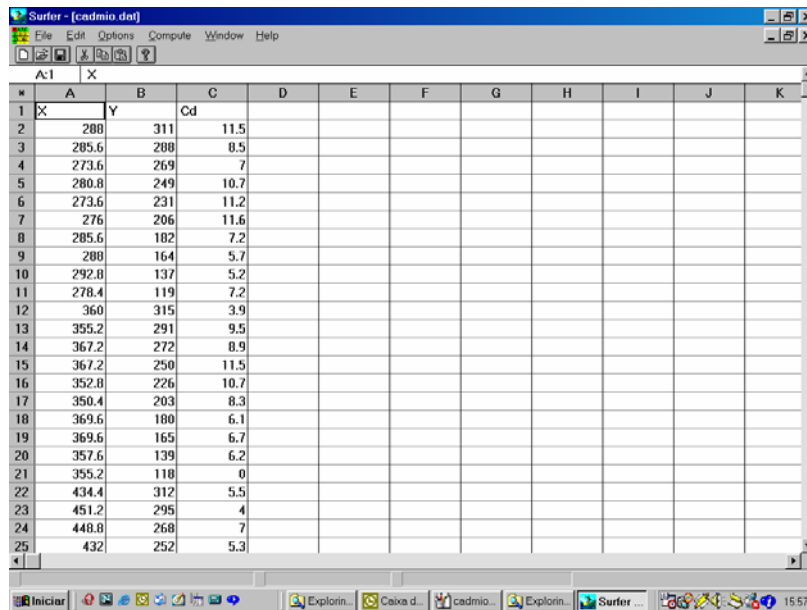
X	Y	Cd
288	311	11.5
285.6	288	8.5
273.6	269	7
280.8	249	10.7
273.6	231	11.2
276	206	11.6
285.6	182	7.2
288	164	5.7
292.8	137	5.2
278.4	119	7.2
360	315	3.9
355.2	291	9.5
367.2	272	8.9
367.2	250	11.5
352.8	226	10.7
350.4	203	8.3
369.6	180	6.1
369.6	165	6.7
357.6	139	6.2
355.2	118	0
434.4	312	5.5
451.2	295	4
448.8	268	7
432	252	5.3
441.6	228	11.6
441.6	204	9
444	182	14.5
441.6	160	12.1
432	140	0.9
444	119	0
254.4	172	3.2
254.4	128	1.2
254.4	299	1.7
333.6	301	1.2

333.6	271	7.6
333.6	194	11.6
333.6	163	8.7
412.8	285	5.8
254.4	257	3.8
412.8	172	10.4
412.8	150	10
492	282	7.1
492	249	4.4
492	315	10.4
492	150	1.6
444	190	15
436.8	240	3.4
360	195	6.8
345.6	210	10.8
254.4	216	14.9
280.8	216	9.9
307.2	216	11.6
333.6	216	6.5
360	216	10.1
386.4	216	11.8
412.8	216	11
439.2	216	16.7
465.6	216	11.6
492	216	6.9
345.6	216	9.9

- **Gravação de dados:**

Entrar em “File/New/Worksheet” para iniciar a gravação da planilha de dados.

Por definição os valores da coordenada X devem ser gravados na primeira coluna e os valores da coordenada Y na segunda. A variável Z (ou variáveis  $Z_1, Z_2 \dots$ ) serão gravadas a partir da terceira coluna, lembrando que o método somente pode ser aplicado à uma variável por vez. O arquivo será do tipo ASCII Files ( \*.dat)

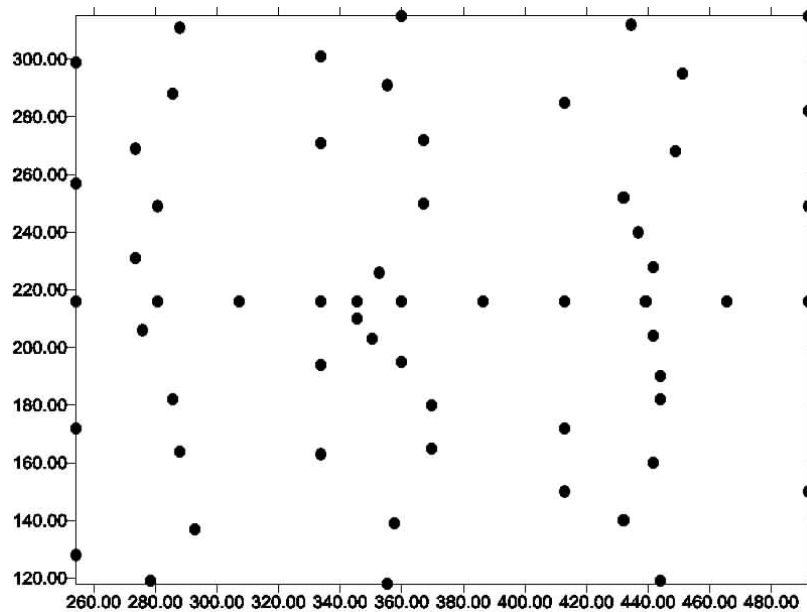


The screenshot shows the Surfer software window titled "Surfer - [cadmio.dat]". The window contains a spreadsheet with the following data:

	A	B	C	D	E	F	G	H	I	J	K
1	X	Y	Cd								
2		208	311	11.5							
3		205.6	208	8.5							
4		273.6	269	7							
5		280.8	249	10.7							
6		273.6	231	11.2							
7		276	206	11.6							
8		205.6	182	7.2							
9		208	164	5.7							
10		292.8	137	5.2							
11		278.4	119	7.2							
12		360	315	3.9							
13		355.2	291	9.5							
14		367.2	272	8.9							
15		367.2	250	11.5							
16		352.8	226	10.7							
17		350.4	203	8.3							
18		369.6	180	6.1							
19		369.6	165	6.7							
20		357.6	139	6.2							
21		355.2	118	0							
22		434.4	312	5.5							
23		451.2	295	4							
24		448.8	268	7							
25		432	252	5.3							

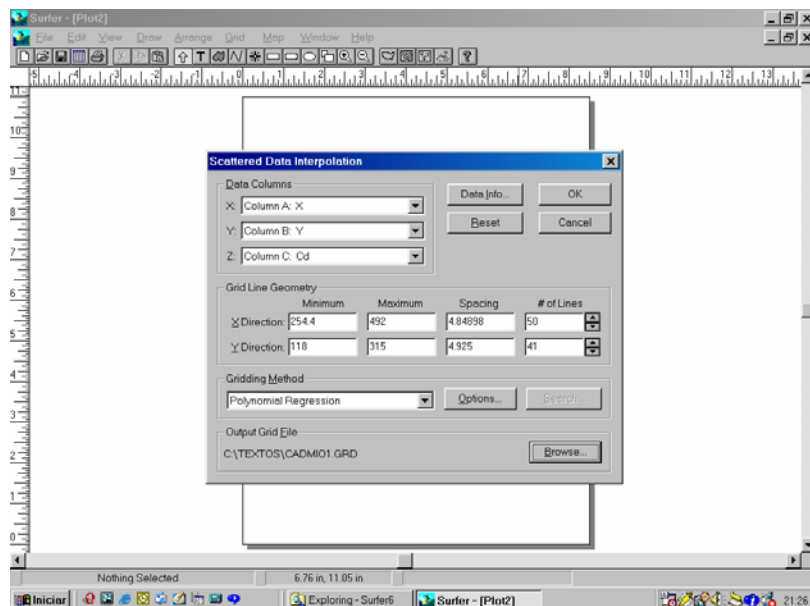
- **Distribuição dos pontos**

Para verificar a distribuição dos pontos entrar em “Map/Post” ou “Map/Classed Post” e escolher o arquivo \*.dat. Os pontos poderão vir acompanhados dos respectivos valores (*Label*) ou não.



- **Cálculo da superfície**

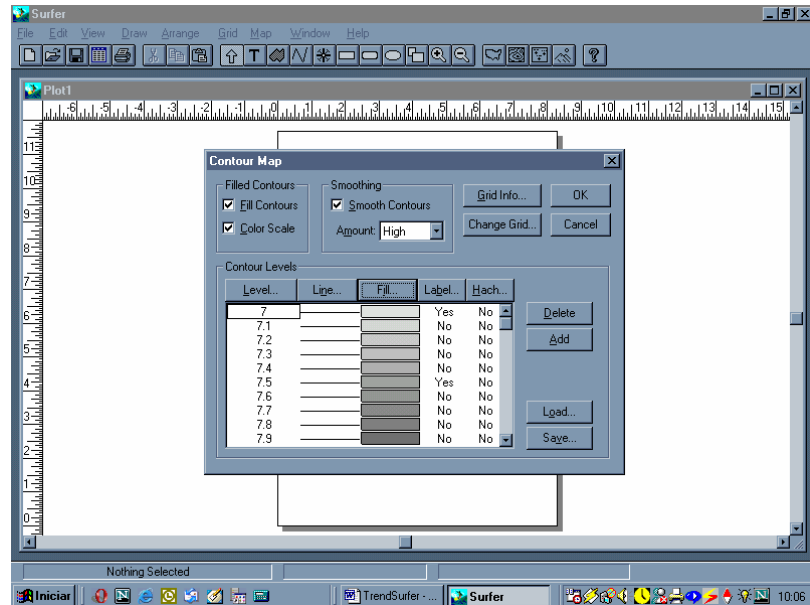
Para o cálculo da superfície entrar em “Grid/Data” e, em seguida, escolher a opção “Gridding Method: Polynomial Regression” e em “Options” o grau da superfície desejada. Caso queira gravar o arquivo com os valores do reticulado com um nome especial, faça-lo usando “Browse”



- **Impressão da superfície**

Entrar em “Map/Contour” e escolher o arquivo \*.grd. Em seguida, na janela “Contour Map” decidir sobre “Fill Contours”, “Color Scale”, “Smoothing”, “Level”, “Fill” e “Label”.

O resultado será um arquivo \*.srf, que poderá ser gravado e impresso



- **Cálculo e impressão do Mapa de resíduos**

Entrar em “Grid/Residuals” e escolher o arquivo “\*.grid” e, em seguida, o correspondente arquivo “\*. dat”. Na janela “Grid Residuals” aparecerá a informação “Store residuals in column [ ]”. Abrir a janela “Worksheet” e regravar o arquivo “\*.dat”, agora com os valores residuais calculados. Abrir a janela “Plot”, e em seguida “Grid”.

Na janela “Gridding method” escolher em “Data Columns” como varival Z, a coluna com os valores residuais e em “Gridding Method” um algoritmo para a confecção do mapa de valores residuais. Neste exemplo foi escolhido “Minimum Curvature”.

Para a impressão do mapa de resíduos seguir os mesmos passos anteriormente descritos em “Impressão da superfície”.

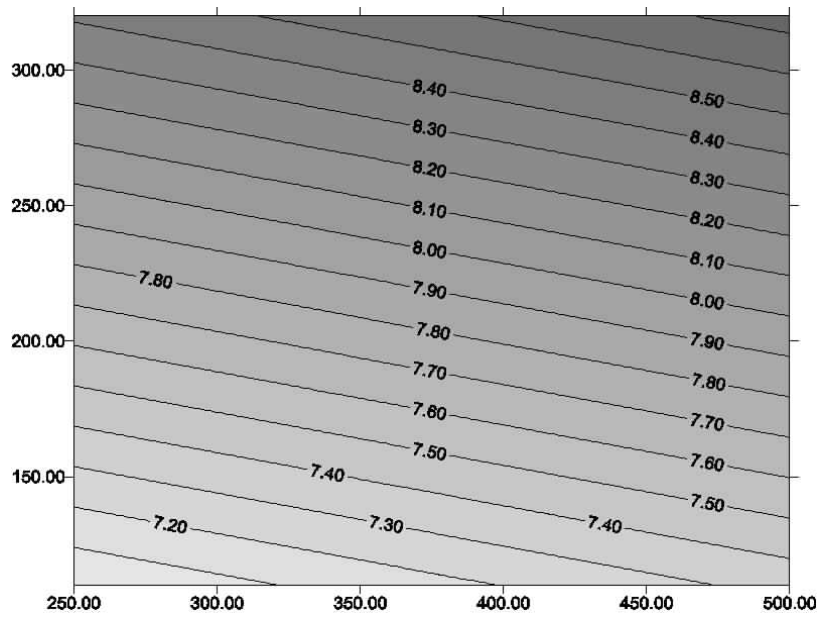
## **BIBLIOGRAFIA RECOMENDADA**

Krajewski, S.A. & Gibbs, B.L. (1966) – Understanding Contouring: A practical Guide to Spatial Estimation and Contouring Using a Computer and Basics of Using Variograms: Gibbs Associates

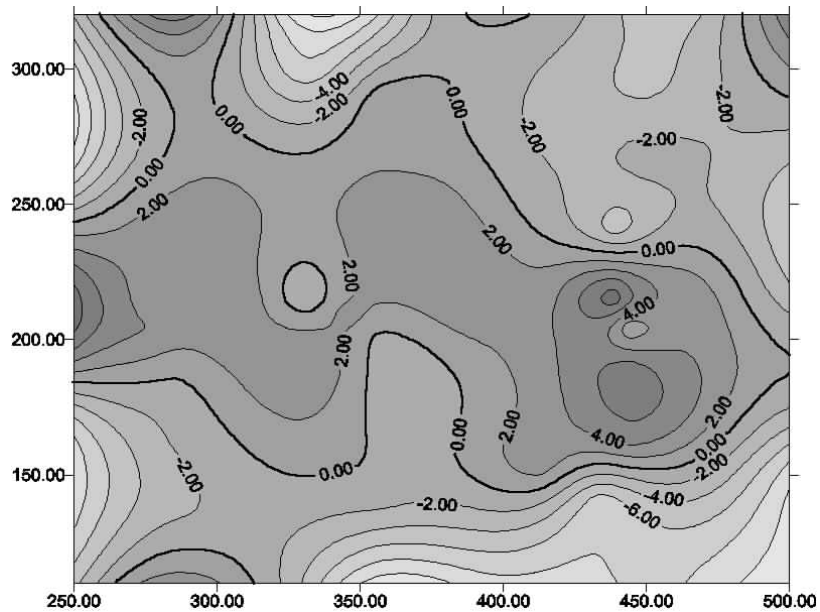
Landim, P.M.B. (1998) – Análise Estatística de Dados Geológicos: Editora Unesp

Surfer 6.0 (1995) – User's guide. Contouring and 3D Surface Mapping for Scientists and Engineers: Golden Software, Inc.

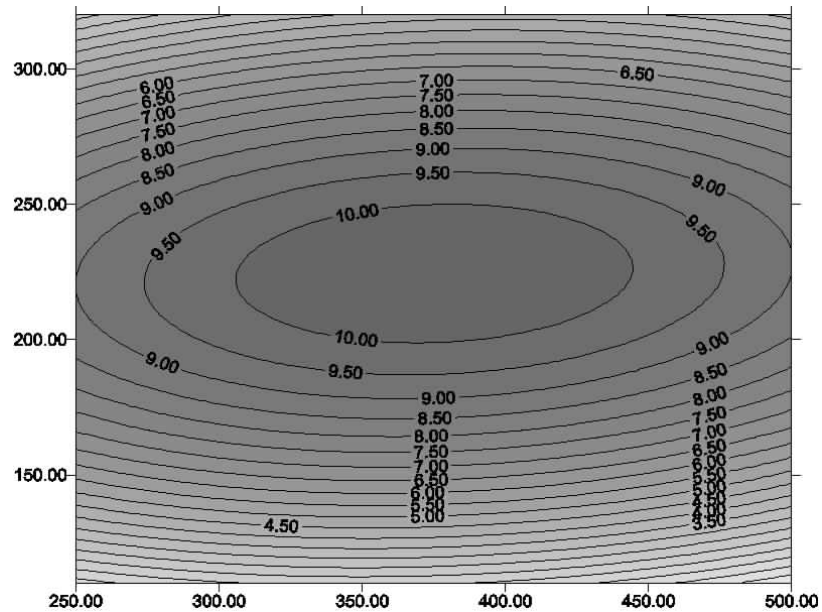
USEPA – United States Environmental Protection Agency. GEO-EAS, version 1.2.1. [ s.l.]: USEPA, 1991. Conjunto de programas, disponível em < <http://www.sph.umich.edu/~aelon/geoeas/>> em 06 fev. 2002.



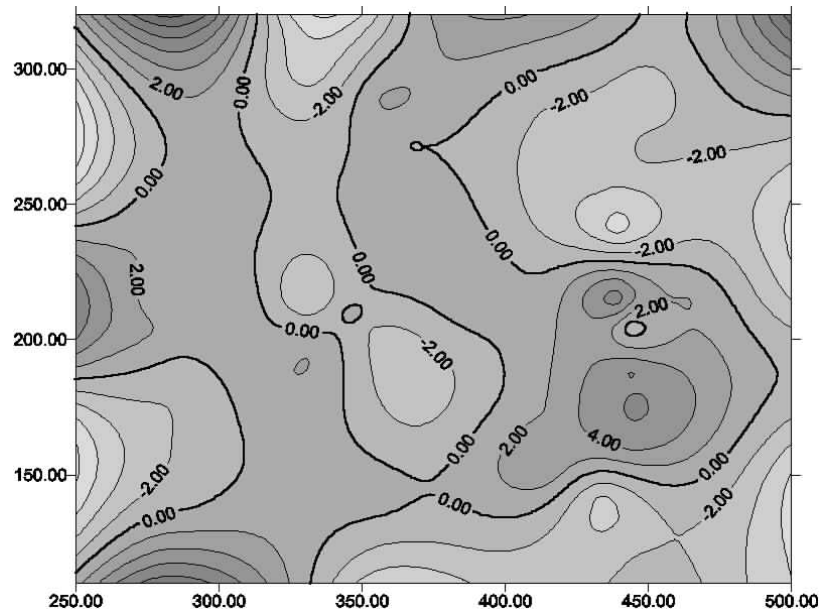
- Superfície de 1º grau, calculada pelo algoritmo “regressão polinomial”
- Equação da superfície, gravada no “bloco de notas”:
  - $z(x,y) = 5.94059 + 0.00671035y + 0.00131154x$
- $R^2 = 0,009737$  (calculado pela função “RQUAD” existente no Excel/estatística)



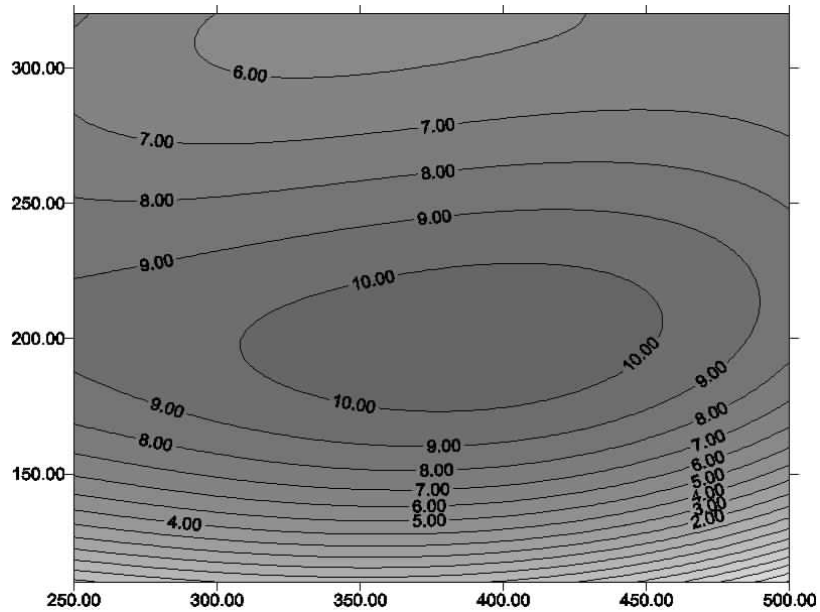
- Mapa de resíduos referentes à superfície de grau 1, calculado pelo algoritmo “curvatura mínima”



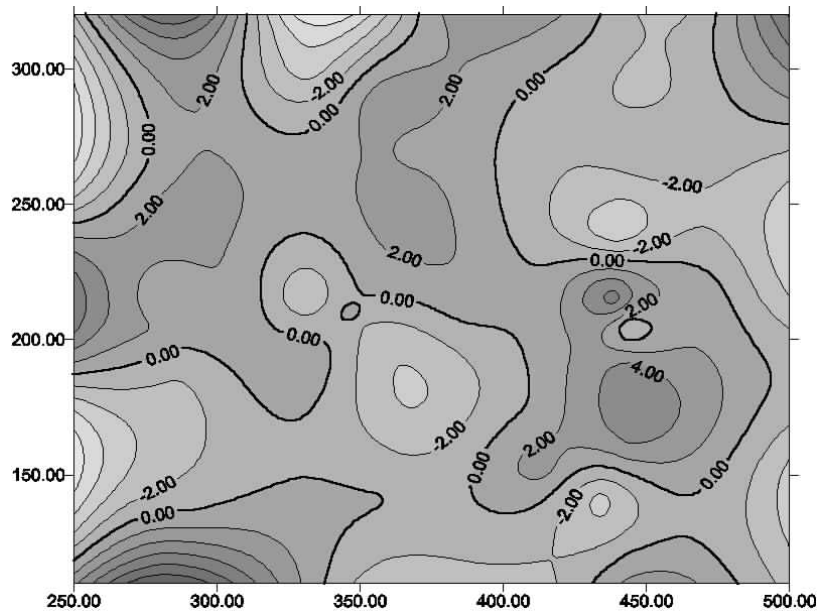
- Superfície de 2º grau, calculada pelo algoritmo “regressão polinomial”
- Equação da superfície, gravada no “bloco de notas”:
  - $z(x,y) = -32.9993 + 0.287639y + -0.000677274y^2 + 0.0595063x + 4.3638e-005xy + -9.23639e-005x^2$
- $R^2 = 0,333646$  (calculado pela função “RQUAD” existente no Excel/estatística)



- Mapa de resíduos referentes à superfície de grau 2, calculado pelo algoritmo “curvatura mínima”



- Superfície de 3º grau, calculada pelo algoritmo “regressão polinomial”
- Equação da superfície, gravada no “bloco de notas”:
  - $z(x,y) = -85.943 + 1.08613y + -0.00393305y^2 + 5.49438e-006y^3 + 0.0630072x + -0.000699097xy + -8.66794e-007xy^2 + 0.000123355x^2 + 1.52758e-006x^2y + -4.88373e-007x^3$
- $R^2 = 0,390644$  (calculado pela função “RQUAD” existente no Excel/estatística)



- Mapa de resíduos referentes à superfície de grau 3, calculado pelo algoritmo “curvatura mínima”