

METODOLOGIA GEOESTATÍSTICA PARA DADOS COM  
TENDÊNCIA REGIONALIZADA

Luiz Alberto Amaral Nardi

Dissertação apresentada à Universidade Estadual Paulista “Júlio de Mesquita Filho” para a obtenção do título de Mestre em Biometria.

BOTUCATU  
São Paulo - Brasil  
Fevereiro – 2015

**METODOLOGIA GEOESTATÍSTICA PARA DADOS COM  
TENDÊNCIA REGIONALIZADA**

**Luiz Alberto Amaral Nardi**

Orientador: Prof. Dr. **Paulo Milton Barbosa Landim**

Dissertação apresentada à Universidade Estadual Paulista “Júlio de Mesquita Filho” para a obtenção do título de Mestre em Biometria.

**BOTUCATU**  
São Paulo - Brasil  
Fevereiro – 2015

## Ficha Catalográfica

Nardi, Luiz Alberto Amaral.

Metodologia geoestatística para dados com tendência regionalizada / Luiz Alberto Amaral Nardi. - Botucatu, 2015

Dissertação (mestrado) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências de Botucatu

Orientador: Paulo Milton Barbosa Landim

Capes: 10202005

1. Geologia - Métodos estatísticos. 2. Krigagem. 3. Análise espacial (Estatística).

Palavras-chave: Geoestatística; Krigagem Ordinária; Krigagem Universal; Krigagem dos Resíduos.

## Dedicatória

Dedico esta dissertação aos meus pais: Francisco José Nardi (*in memoriam*) e Ernestina Maria do Amaral Nardi e ao meu irmão: Francisco José Nardi Filho. Também aos meus avós maternos: Sebastião Félix do Amaral (*in memoriam*) e Conceição Fonseca do Amaral e aos avós paternos: Francisco Nardi (*in memoriam*) e Maria Aparecida Rossetti Nardi (*in memoriam*). Tudo o que aprendi e todo amor sincero que tive nesta vida sempre vieram de vocês. E também a você, Daniela Herreira Carvalho, que me apoia e me dá alento e carinho nas horas em que mais preciso.

## Agradecimentos

Agradeço a Deus fazendo o que nos diz o salmista: “Vinde, manifestemos nossa alegria ao Senhor, aclamemos o rochedo de nossa salvação; apresentemo-nos diante Dele com louvores e cantemos-Lhe alegres cânticos!” (Salmo 94:1-2).

Agradeço a todas as pessoas que fizeram parte da minha formação intelectual e pessoal.

Agradeço aos colegas, funcionários e professores do Departamento de Bioestatística pelo convívio sempre amistoso tanto nas penúrias do dia-a-dia quanto nas confraternizações e nos (incontáveis) “cafés-com-prosa” na copa.

Agradeço ao professor Paulo Milton Barbosa Landim que muito me honrou aceitando me orientar e possibilitando, com isso, que eu pudesse desfrutar dos seus conhecimentos e vivências científicas ímpares.

Finalmente, agradeço à CAPES, pelo recurso financeiro destinado à elaboração deste trabalho.

## Sumário

	Página
<b>LISTA DE FIGURAS</b>	<b>vii</b>
<b>LISTA DE TABELAS</b>	<b>x</b>
<b>RESUMO</b>	<b>xi</b>
<b>SUMMARY</b>	<b>xiii</b>
<b>1 INTRODUÇÃO</b>	<b>1</b>
<b>2 VARIÁVEIS REGIONALIZADAS E FUNÇÕES ALEATÓRIAS</b>	<b>4</b>
2.1 Variáveis regionalizadas . . . . .	4
2.2 Processos Estocásticos . . . . .	6
<b>3 HIPÓTESES DE ESTACIONARIEDADE</b>	<b>8</b>
3.1 Momentos de uma função aleatória . . . . .	8
3.1.1 Momento de Primeira Ordem . . . . .	8
3.1.2 Momentos de Segunda Ordem . . . . .	8
3.2 Hipótese de Estacionariedade . . . . .	8
3.3 Hipótese Intrínseca . . . . .	11
3.4 Ergodicidade . . . . .	14
<b>4 SEMIVARIOGRAMAS</b>	<b>15</b>
4.1 Método dos momentos . . . . .	15
4.1.1 Semivariograma Empírico . . . . .	16

	vi
4.1.2 Amostragem . . . . .	16
4.1.3 Cálculo de semivariogramas empíricos . . . . .	17
4.2 Modelos para Representar os Semivariogramas Empíricos: Semivariogramas Teóricos . . . . .	20
4.2.1 Método dos Mínimos Quadrados Ponderados . . . . .	21
4.2.2 Ajuste à sentimento ou “visualmente” . . . . .	21
4.2.3 Procedimentos para seleção de modelos . . . . .	22
4.3 Efeito Pepita: valor do semivariograma teórico na origem . . . . .	28
4.4 Fenômenos Anisotrópicos . . . . .	29
<b>5 KRIGAGEM ORDINÁRIA E A KRIGAGEM UNIVERSAL</b>	<b>32</b>
5.1 Krigagem Ordinária . . . . .	32
5.2 Krigagem Universal . . . . .	36
5.2.1 Equações da Krigagem sob suposição da Hipótese Intrínseca . . . . .	38
5.2.2 Krigagem Local e Global . . . . .	45
5.2.3 Vantagens e Desvantagem da Krigagem Universal . . . . .	45
5.2.4 Método da Krigagem dos Resíduos: Método Direto . . . . .	47
5.2.5 Método da Krigagem dos Resíduos: Iterativo . . . . .	48
<b>6 APLICAÇÕES A DOIS CASOS DA LITERATURA</b>	<b>50</b>
6.1 Material . . . . .	50
6.2 Método . . . . .	51
6.3 Resultados e Discussões . . . . .	51
6.3.1 Dados do Aquífero High Plains . . . . .	51
6.3.2 Dados do Aquífero Wolfcamp . . . . .	72
<b>7 CONCLUSÕES</b>	<b>86</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>94</b>

## Lista de Figuras

	Página
1 Componentes de uma variável regionalizada. . . . .	5
2 $\alpha_1$ : direção da continuidade; $a_1$ : distância na qual a correlação espacial vai se esvanecendo; $C( \mathbf{h} , \alpha_1) + \gamma( \mathbf{h} , \alpha_1)$ =valor do patamar= $C(0)$ . Fonte: CALVETE & RAMÍREZ (1996) . . . . .	13
3 A) Malha quadrada B) Malha retangular, com indicação das diagonais para o cálculo dos semivariogramas experimentais. Círculo vazio:Ponto não amostrado, Círculo cheio: ponto amostrado. . . . .	18
4 Esquema mostrando a pesquisa de pares para cálculo de semivariogramas no caso da distribuição irregular. A) O dispositivo de pesquisa é centrado no ponto 1 B) O dispositivo de pesquisa se move e é centrado no ponto 2.	18
5 Exemplo de Semivariograma Experimental . . . . .	19
6 Semivariograma Teórico: Efeito Pepita . . . . .	23
7 Semivariograma Teórico: Modelo Esférico . . . . .	24
8 Semivariograma Teórico: Modelo Exponencial . . . . .	25
9 Semivariograma Teórico: Modelo Gaussiano . . . . .	26
10 Anisotropia. . . . .	30
11 Tipos de Anisotropia: A) Geométrica, B)Zonal e C) Mista . . . . .	31
12 Decomposição da variável regionalizada. . . . .	39
13 Localização espacial dos 327 amostras do Aquifero High Plains. . . . .	52
14 Histograma da Variável WTable. . . . .	53
15 Gráfico Quantil-Quantil para Distribuição Normal. . . . .	53

16	Distribuição dos Valores da variável WTable por sua localização geográfica	54
17	Semivariograma para direção $0^\circ$ .	55
18	Semivariograma para direção $45^\circ$ .	55
19	Semivariograma para direção $90^\circ$ .	56
20	Semivariograma para direção $135^\circ$ .	56
21	Modelo 1 sobreposto aos dados do Aquífero High Plains.	59
22	Modelo 2 sobreposto aos dados do Aquífero High Plains.	59
23	Ajuste por um Polinômio Quadrático, da Variável WTable.	60
24	Histograma dos Resíduos obtidos quando se modela WTable por um Polinômio Quadrático	61
25	Gráfico Quantil-Quantil Normal para os Resíduos da Variável WTable, considerando o Modelo 1.	61
26	Semivariograma dos Resíduos da Variável WTable quando considerado o Modelo 1, nas direções $0^\circ$ , $45^\circ$ , $90^\circ$ e $135^\circ$ . A linha tracejada indica a variância de $Z(x)$	62
27	Mapa da Krigagem Ordinária dos Resíduos considerando o Modelo 1.	63
28	Mapa dos Desvios-Padrão considerando o Modelo 1.	63
29	Mapa de Krigagem Universal, considerando o Modelo 1, para Variável WTable.	64
30	Superfície Estimada por Ajuste Polinomial por meio de um Polinômio Quadrático.	65
31	Superfície Estimada por Krigagem Universal considerando o Modelo 1.	65
32	Ajuste considerando um Polinômio Linear, da Variável WTable.	66
33	Histograma dos Resíduos obtidos quando se modela a Variável WTable por um Polinômio Linear.	67
34	Gráfico Quantil-Quantil Normal para os Resíduos da Variável WTable considerando Modelo 2.	67
35	Semivariograma dos Resíduos da Variável WTable nas direções $0^\circ$ , $45^\circ$ , $90^\circ$ e $135^\circ$ . A linha tracejada indica a variância de $Z(x)$	68

36	Mapa da Krigagem Ordinária dos Resíduos, considerando o Modelo 2. . . . .	69
37	Mapa dos Desvio-padrão, considerando o Modelo 2. . . . .	69
38	Mapa de Krigagem Universal para Variável WTable. . . . .	70
39	Superfície Estimada por Ajuste Polinomial. . . . .	71
40	Superfície Estimada por Krigagem Universal. . . . .	71
41	Localização espacial dos 85 amostras do Aquífero Wolfcamp. . . . .	72
42	Histograma da do Nível Piezométrico do Aquífero Wolfcamp. . . . .	73
43	Gráfico Quantil-Quantil para Distribuição Normal. . . . .	74
44	Distribuição dos Valores da variável Nível Piezométrico por sua localiza- ção geográfica . . . . .	74
45	Semivariograma na direção 0° para variável Nível Piezométrico. . . . .	75
46	Semivariograma na direção 45° para variável Nível Piezométrico. . . . .	75
47	Semivariograma na direção 90° para variável Nível Piezométrico. . . . .	76
48	Semivariograma na direção 135° para variável Nível Piezométrico. . . . .	76
49	Modelo 1 sobreposto aos dados do Aquífero Wolfcamp. . . . .	79
50	Modelo 2 sobreposto aos dados do Aquífero Wolfcamp. . . . .	79
51	Modelo 3 sobreposto aos dados do Aquífero Wolfcamp. . . . .	80
52	Nível Piezométrico ajustado pelo Modelo 2. . . . .	80
53	Histograma dos Resíduos Obtidos Quando se Modela o Nível Piezométrico por Um Polinômio Quadrático. . . . .	81
54	Gráfico Quantil-Quantil para Distribuição Normal para os Resíduos. . . . .	81
55	Semivariograma dos Resíduos. As cores correspondem aos ângulos assim como na Figura 26. . . . .	83
56	Krigagem dos Resíduos obtidos. . . . .	84
57	Mapa dos Desvios-Padrão da Krigagem dos Resíduos. . . . .	84
58	Mapa da Krigagem Universal da Variável Nível Piezométrico. . . . .	85
59	Mapa de Superfície da Variável Nível Piezométrico obtido através da Kri- gagem Universal. . . . .	85

## Lista de Tabelas

	Página
1 Modelos teóricos de semivariogramas com efeito pepita . . . . .	28
2 Classificação dos Graus de Aleatoriedade . . . . .	29
3 Estatísticas Descritivas da variável WTable . . . . .	52
4 Valores dos coeficientes do Polinômio Quadrático e seus respectivos valores-p. . . . .	57
5 Valores dos coeficientes do Polinômio Linear e seus respectivos valores-p.	58
6 Comparação, através da ANOVA, dos Modelos 1 e 2. . . . .	58
7 Estatísticas Descritivas dos Resíduos considerando o Modelo 1. . . . .	60
8 Estatísticas Descritivas dos Resíduos considerando o Modelo 2. . . . .	66
9 Estatísticas Descritivas da Variável Nível Piezométrico . . . . .	73
10 Valores dos coeficientes do Polinômio Cúbico e seus respectivos valores-p.	77
11 Valores dos coeficientes do Polinômio Quadrático e seus respectivos valores-p. . . . .	77
12 Valores dos coeficientes do Polinômio Linear e seus respectivos valores-p.	78
13 Comparação, através da ANOVA, dos Modelos 1,2 e 3. . . . .	78
14 Estatísticas Descritivas dos Resíduos Obtidos quando se Modela a Variável Nível Piezométrico Por Um Polinômio de Segundo Grau. . . . .	82

# METODOLOGIA GEOESTATÍSTICA PARA DADOS COM TENDÊNCIA REGIONALIZADA

Autor: LUIZ ALBERTO AMARAL NARDI

Orientador: Prof. Dr. PAULO MILTON BARBOSA LANDIM

## RESUMO

Georges Matheron, baseado na teoria das variáveis regionalizadas, sua criação, desenvolveu a Geoestatística, inicialmente aplicada em mineração, mas que atualmente é de larga aplicação em várias áreas do conhecimento cujos problemas possuem forte ligação com a sua localização espacial. Segundo essa metodologia, o método estimador mais usual é a Krigagem Ordinária, que leva em consideração a estrutura de variâncias e covariâncias entre as amostras que, neste caso, dependem da localização geográfica dos pontos amostrados. Para utilização essa técnica, parte-se da pressuposição que a função aleatória  $Z(\mathbf{x})$ , que descreve o fenômeno estudado, apresente estacionariedade de segunda ordem ou satisfaça a hipótese intrínseca. Isto implica, em ambos os casos, que  $E[Z(\mathbf{x})] = m$ , sendo  $m$  desconhecido, mas constante. Se, porventura, esta condição não for satisfeita e se a função aleatória  $Z(x)$  puder ser escrita como a seguinte soma  $Z(\mathbf{x}) = m(\mathbf{x}) + Y(\mathbf{x})$ , onde  $m(\mathbf{x})$  é o valor da

esperança de  $Z(\mathbf{x})$  e pode ser escrita como um polinômio de grau baixo e  $Y(\mathbf{x})$  constituir a parte aleatória, deve-se usar a metodologia da Krigagem Universal, da qual a Krigagem Ordinária é um caso particular. No entanto, a Krigagem Universal possui uma dificuldade metodológica que é a de se supor conhecido o semivariograma dos Resíduos. Uma das maneiras de contornar este problema é utilizar a Krigagem dos Resíduos ou Krigagem Residual. O objetivo desta dissertação é apresentar, embora de maneira não rigorosa, o formalismo matemático que embasa a teoria da Krigagem, os problemas relativos a Krigagem Universal, a maneira de contorná-los por meio da Krigagem Residual e as críticas que são feitas a este método. Por fim, serão exibidos dois exemplos retirados da literatura, cujo fenômeno em estudo não cumpre a as condições exigidas para a Krigagem Ordinária e nos quais, portanto, deve-se aplicar a Krigagem Universal que, nestes casos, será feita através da Krigagem dos Resíduos.

# GEOSTATISTICAL METHODOLOGY FOR TREND REGIONALIZED DATA

Author: LUIZ ALBERTO AMARAL NARDI

Adviser: Prof. Dr. PAULO MILTON BARBOSA LANDIM

## SUMMARY

George Matheron, based on the regionalized variable theory, his creation, developed Geostatistics, initially applied in mining, but currently applied largely in many fields of knowledge whose problems have a strong connection with their spatial location. According to this methodology, the most common estimator method is the Ordinary Kriging, which takes into consideration the variance and covariance structures between samples that, in this case, depend on the geographic location of sample points. For the use of this technique, it starts with the assumption that the random function  $Z(\mathbf{x})$ , which describes the studied phenomenon, shows second-degree stationarity or satisfies the intrinsic hypothesis. This implies, in both cases, that  $E[Z(\mathbf{x})] = m$ , being  $m$  unknown, but constant. If, perhaps, this condition is not satisfied and the random function  $Z(\mathbf{x})$  can be written as the following sum  $Z(\mathbf{x}) = m(\mathbf{x}) + \epsilon(\mathbf{x})$ , where  $m(\mathbf{x})$  is the expectation value of  $Z(\mathbf{x})$  and can be written

as a low-degree polynomial and  $\epsilon(\mathbf{x})$  composes its random part, the Universal Kriging methodology should be used, in which the Ordinary Kriging is a particular case. However, the Universal Kriging has a methodological difficulty, which is assuming the Residual semivariogram as known. One of the ways to get around this is using the Residual Kriging. The objective of this dissertation is to present, although not in a rigorous way, the mathematical formalism that underlies the Kriging theory, the problems related to the Universal Kriging, the way to get around this using the Residual Kriging, and the criticisms that are made to this method. Lastly, it will be shown two examples taken from the literature, whose studied phenomenon does not meet the requirements for the Ordinary Kriging and which, thus, the Universal Kriging must be applied, in these cases, through the Residual Kriging.

# 1 INTRODUÇÃO

O título de fundador da Geoestatística é do francês Georges Matheron (1930-2000), professor da École de Mines de Fontenebleau onde fundou, em 1968, o “*Centre de Geostatistique et de Morphologie Mathématique*” (AGTERBERG, 2004) e, durante toda a década de 60 do século passado, desenvolveu matematicamente a teoria das variáveis regionalizadas, “fundamento de toda a geoestatística”(JOURNEL, 1989).

Os trabalhos de Matheron tiveram como inspiração inicial o artigo do Engenheiro de Minas, D.G. Krige intitulado: “*A statistical approach of some basic mines valuation problems on the Witwatersran*” no “*Journal of the Chemical, Metallurgical and Mining Society of South Africa*” publicado em 1951 e também os trabalhos do professor da Universidade Técnica de Delft, na Holanda, H.J. De Wijs dos anos de 1951 e 1953 (YAMAMOTO & LANDIM, 2013).

Krige, trabalhando com dados de concentração de ouro, concluiu que somente a informação dada pela variância seria insuficiente para explicar o fenômeno em estudo e que, para tanto, seria necessário levar em consideração a distância entre as observações. Segundo ANDRIOTTI (1988), Krige verificou erros sistemáticos na avaliação dos teores médios de blocos quando se considerava apenas amostras interiores aos mesmos, desprezando as amostras exteriores. O erro que acontecia era a superestimação de blocos ricos e a subestimação de blocos pobres. Com a finalidade de minimizar estes erros, Krige passou a utilizar métodos semelhantes aos das médias móveis. A partir daí surge um dos fundamentos que diferenciam a geoestatística da de outros métodos estatísticos, pois, aquela, leva em consideração a localização geográfica e a dependência espacial entre as amostras (DRUCK et al., 2004). Algumas

das técnicas apresentadas por Krige foram aperfeiçoadas e fundamentadas matematicamente por Matheron que, em homenagem ao engenheiro de minas, nomeou a técnica de "Krigeage", traduzida posteriormente para o inglês como "Kriging" e, em português, Krigagem.

Os primeiros resultados obtidos por Matheron foram publicados no "*Traité de Géostatistique Appliquée: volume 1*" de 1962 e depois no volume 2 de 1963. Matheron foi um cientista modesto que preferiu trabalhar isoladamente, deixando seus estudos armazenados na biblioteca do seu instituto em Fontainebleau (AGTERBERG, 2004).

O termo Geoestatística se deve ao fato de se utilizar uma metodologia estatística aplicada a dados georreferenciados, como são, em geral, os dados provenientes das Ciências da Terra. Entretanto, estas técnicas tem se mostrado promissoras em vários outros campos das ciências, da Agricultura à Zoologia (CRESSIE, 2011).

Embora o desenvolvimento da teoria das variáveis regionalizadas tenha sido desenvolvida por Georges Matheron, a preocupação com a localização dos locais de amostragem e como isso poderia afetar a estrutura de variabilidade dos dados remonta aos trabalhos de Fisher. CRESSIE (2011) relata que Fisher escreveu no seu clássico "*Design of Experiments*" de 1935: "*after choosing the area we usually have no guidance beyond the widely verified fact that patches in close proximity are commonly more alike, as judged by the yield of crops, than those which are further apart*" o que quer dizer que Fisher notou que o rendimento das culturas eram semelhantes em regiões próximas. A variabilidade que Fisher percebeu quando verificou as diferenças entre rendimentos das culturas pode ser explicada, em grande parte, pelas propriedades ambientais e físicas do solo que, em particular, possuem grande dependência espacial.

O método da Krigagem foi descrito pela primeira vez, em detalhes, no "*Traité de Géostatistique Appliquée: volume 2*", chamado "*Le Krigeage*" de 1963. A Krigagem é um interpolador exato, ou seja, o valor da estimativa obtida por ele, nos pontos amostrais, é o verdadeiro valor destes pontos, a menos do chamado Efeito

Pepita. Efeito Pepita é um conceito que será visto no capítulo 4. Além disso, este interpolador possui as propriedades estatísticas de não-enviesamento e eficiência. A Krigagem nada mais é que um interpolador que atribui pesos às amostras que serão usadas para estimar os valores não-amostrados. Isto é feito através de uma matriz de variâncias e covariâncias espacial que deve ser estimada. A Krigagem Ordinária é uma técnica que supõe a estacionariedade de 2ª ordem ou ainda da hipótese intrínseca. Estes conceitos serão vistos com detalhes no capítulo 5. Entre os vários tipos de Krigagem, a mais usual é a Krigagem Ordinária (*krigeage ordinaire*, em francês, traduzida para o português como Krigagem Ordinária). SOARES (2000) nos relata algo interessante com relação à esta tradução dizendo que a nomenclatura mais apropriada seria Krigagem Normal. No entanto, como na maior parte da literatura existente a técnica é chamada de Krigagem Ordinária, manter-se-á, neste trabalho, este termo.

A técnica da Krigagem provocou fortes discussões entre geólogos e engenheiros de minas quando foi apresentada. AGTERBERG (2004) relata que Whitten, no artigo "*The general linear equation in prediction of gold content in Witwatersland rocks*" de 1966, preferia a interpolação polinomial (análise de superfície de tendências) ao que Matheron respondeu num artigo de 1967 intitulado "*Kriging or Polynomial Interpolation Procedures: a contribution to polemics in mathematical geology*".

Já a Krigagem Universal trata dos casos em que há tendência, ou seja, a esperança da função aleatória na região de estudo não é constante e pode ser escrita através de combinações lineares de funções analíticas (OLEA, 1999). O primeiro a se preocupar com os casos de não estacionariedade foi o próprio Matheron provocado por um trabalho de Journel, de 1969, que estava estudava o assoalho do oceano (ARMSTRONG, 1984).

A Krigagem Universal como técnica apareceu em 1970 no artigo "*Random Function and its application in Geology*" apresentado no *Geostatistics colloquium in Lawrence, Kansas, EUA*. Nesta abordagem a concentração de elementos de um

bloco de rocha teve seu valor esperado chamado de “*drift*” que dependia da localização espacial. Pode-se considerar a Krigagem Universal como uma generalização da Krigagem Ordinária no sentido que será visto mais à frente, neste trabalho.

O objetivo deste trabalho é apresentar a metodologia para dados com tendência regionalizada, ou seja, dados que devem ser trabalhados através da técnica da Krigagem Universal. Nos capítulos 2, 3 e 4 serão visto os conceitos de Variáveis Regionalizadas, Funções Aleatórias, Hipóteses de Estacionariedade e Semivariograma que são os conceitos fundamentais que embasam a teoria da Krigagem. Já no capítulo 5 será visto a Krigagem Ordinária que é o primeiro método criado por Matheron e, em seguida, a Krigagem Universal que é o método utilizado para dados que apresentam tendência. No capítulo 6 serão apresentados alguns exemplos de aplicação da Krigagem Universal e no capítulo 7 será apresentada uma Conclusão.

## 2 VARIÁVEIS REGIONALIZADAS E FUNÇÕES ALEATÓRIAS

### 2.1 Variáveis regionalizadas

Segundo MATHERON (1971), quando um fenômeno se propaga no espaço e exibe certa estrutura espacial, diz-se que este fenômeno está regionalizado. Se  $Z(x)$  denota o valor de um ponto  $\mathbf{x} \in \mathbb{R}^n$  (em geral, nas aplicações  $n = 1, 2$  ou  $3$ ) de uma característica  $Z$  do fenômeno,  $Z(x)$  é dita ser uma variável regionalizada.

Ainda segundo DRUCK et al. (2004) a variação de uma variável regionalizada pode ser expressa pela soma de três componentes:

1. uma componente estrutural, associada a um valor médio constante ou a uma tendência constante;

2. uma componente aleatória, espacialmente correlacionada;
3. um ruído aleatório ou erro residual.

Seja  $\mathbf{x} \in \mathbb{R}^2$ . Então  $\mathbf{x} = (x_1, x_2)$  é a localização da variável regionalizada  $Z(\mathbf{x})$ , ou seja, suas coordenadas na região em estudo. Com o que foi dito acima, pode-se escrever:

$$Z(\mathbf{x}) = m(\mathbf{x}) + \epsilon'(\mathbf{x}) + \epsilon''$$

na qual:

$m(\mathbf{x})$ : é uma função determinística associada ao valor médio, que descreve a componente estrutural de  $Z$  em  $\mathbf{x}$ ;

$\epsilon'(\mathbf{x})$ : é um termo estocástico correlacionado, que varia localmente;

$\epsilon''$ : é um ruído aleatório não correlacionado, com Distribuição Normal com média 0 e variância  $\sigma^2$ .

Isto pode ser visto na Figura 1 retirada de (DRUCK et al., 2004).

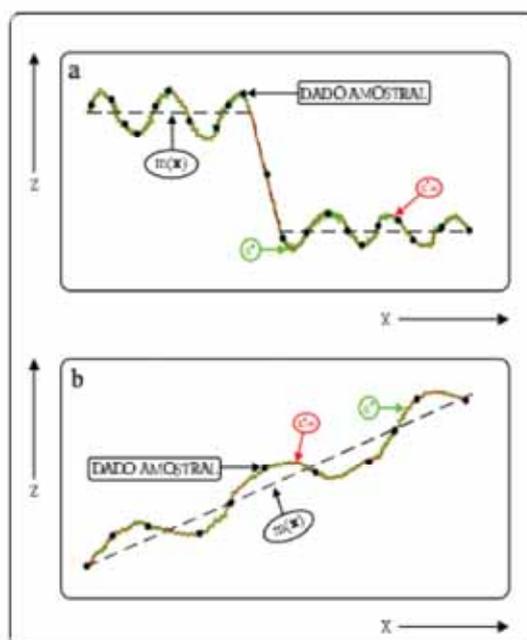


Figura 1 - Componentes de uma variável regionalizada.

Devido às funções regionalizadas apresentarem um comportamento imprevisível e que não podem ser representadas através de funções com expressões matemáticas simples, Matheron propôs que se estudassem os valores resultantes da variável regionalizada como funções aleatórias, ou seja, como um processo estocástico.

## 2.2 Processos Estocásticos

### Definição 2.2.1

Seja  $D$  um domínio espacial num espaço euclidiano  $n$ -dimensional e  $\Omega$  um espaço amostral adequado. Um processo estocástico (função aleatória) é uma coleção de variáveis aleatórias  $\{Z(\mathbf{x}, \omega) : \mathbf{x} \in D, \omega \in \Omega\}$  (OLEA, 1999).

No caso da Geoestatística, usualmente  $\mathbf{x} \in D \subset \mathbb{R}^2$ . Sendo  $Z(\mathbf{x})$  é uma função aleatória, quando varia-se  $\mathbf{x} \in \mathbb{R}^2$ ,  $Z(\mathbf{x})$  gera-se uma família de variáveis aleatórias. Em outras palavras, a cada ponto do espaço associa-se uma variável aleatória.

Da mesma maneira que ao se tomar uma amostra  $\omega$  de uma variável aleatória  $Z$  (de acordo com sua função de distribuição) se obtém um valor numérico  $Z(\omega)$ , ao se tomar uma amostra aleatória de uma função aleatória  $Z(\mathbf{x})$  se obtém uma função  $Z(\mathbf{x}, \omega)$  que coincide com o que se definiu ser uma variável regionalizada, ou seja representa uma possível observação de um campo aleatório (CALVETE & RAMÍREZ, 1996). A função  $Z(\mathbf{x}, \omega)$  se denomina uma realização ou observação da função aleatória  $Z(\mathbf{x})$ .

Aparentemente pode parecer estranha a utilização de uma função aleatórias como modelo para representar a valor desconhecido em um local. Isto é, lidar com múltiplos valores para prever um valor desconhecido. Mas o fato é que os modelos estocásticos são capazes de sintetizar um único e razoável resultado de uma variável aleatória (OLEA, 1999).

### Definição 2.2.2

Considere uma função aleatória  $Z(\mathbf{x})$  definida em  $\mathbb{R}^n$ . Para quaisquer

$k$  pontos  $x_i, i = 1, \dots, k$ , o vetor aleatório  $[Z(\mathbf{x}_1), \dots, Z(\mathbf{x}_k)]$  se caracteriza por sua função de distribuição  $k$ -variável.

$$F_{x_1 \dots x_k} = Prob(Z(\mathbf{x}_1) \leq Z_1, \dots, Z(\mathbf{x}_k) \leq Z_k)$$

O conjunto de todas essas distribuições para todo  $k$  e para qualquer seleção de pontos em  $\mathbb{R}^n$  constituem a lei espacial de probabilidades da função aleatória  $Z(\mathbf{x})$ .

No entanto, há que se responder duas perguntas quando os resultados da variável regionalizada são considerados como um processo estocástico:

A) Que sentido tem considerar um fenômeno natural (que se sabe único) como uma observação ou uma realização de uma função aleatória?

B) É possível fazer inferência estatística (ou seja, determinar os parâmetros estatísticos que definem a lei de probabilidades do fenômeno) a partir de uma só observação?

A pergunta A) pode ser entendida da seguinte maneira. Imagine uma mineração de ouro. Retirada uma amostra de um dado local jamais existirá outra situação semelhante para que se possa inferir sobre a lei de distribuição de probabilidades da variável aleatória “quantidade de ouro” daquele local. E, de certa forma, a pergunta A) conduz à pergunta B): é possível, com somente uma realização, inferir a lei de distribuição de probabilidades? Em geral, a resposta é não. E é por isso que, na Geoestatística, faz-se necessário a suposição de certas hipóteses, como, por exemplo, que a propriedade em estudo varie de maneira igual em toda região do espaço, fazendo com que hajam poucos parâmetros a serem inferidos. Estas hipóteses serão descritas no próximo capítulo.

## 3 HIPÓTESES DE ESTACIONARIEDADE

### 3.1 Momentos de uma função aleatória

#### 3.1.1 Momento de Primeira Ordem

O momento de primeira ordem de uma função aleatória é:

$$E[Z(\mathbf{x})] = m(\mathbf{x}), \forall \mathbf{x}. \quad (1)$$

#### 3.1.2 Momentos de Segunda Ordem

Já os momentos centrados de segunda ordem de uma função aleatória são:

$$\sigma^2 = Var[Z(\mathbf{x})] = E[Z(\mathbf{x}) - m(\mathbf{x})]^2 \quad (2)$$

$$Cov(Z(\mathbf{x}_i), Z(\mathbf{x}_j)) = E\{[Z(\mathbf{x}_i) - m(\mathbf{x}_i)][Z(\mathbf{x}_j) - m(\mathbf{x}_j)]\} \quad (3)$$

$$\gamma[Z(\mathbf{x}_i), Z(\mathbf{x}_j)] = \frac{1}{2}E[Z(\mathbf{x}_i) - Z(\mathbf{x}_j)]^2 \quad (4)$$

### 3.2 Hipótese de Estacionariedade

LLOYD (2011) relata que a estacionariedade pode ser dividida, para os propósitos da Geoestatística, em três classes:

- i) Estacionariedade estrita.
- ii) Estacionariedade de segunda ordem.
- iii) Estacionariedade intrínseca.

Deve-se levar em consideração é que estacionariedade é uma propriedade da função aleatória e não dos dados. Por esta razão é um equívoco igualar o

conceito de estacionariedade com o de homogeneidade. Ou seja, a estacionariedade não pode ser testada através dos dados.

Supor que uma função aleatória é estacionária é equivalente a dizer que esta função se repete no espaço e que esta repetição proporciona a informação equivalente à muitas realizações da mesma função aleatória, possibilitando assim a inferência estatística. (CALVETE & RAMÍREZ, 1996)

### Definição 3.2.1

Uma função aleatória é dita ser estritamente estacionária se a sua função de distribuição for invariante mediante uma translação por um vetor  $\mathbf{h}$  qualquer, ou seja, a distribuição do vetor aleatório:  $(Z(\mathbf{x}_1), \dots, Z(\mathbf{x}_n))$  tem a mesma distribuição de:  $(Z(\mathbf{x}_1 + \mathbf{h}), \dots, Z(\mathbf{x}_n + \mathbf{h}))$ .

Dá-se o nome ao vetor  $\mathbf{h}$  de *lag* (distância e direção) com respeito à localização  $\mathbf{x}_i$ .

### Definição 3.2.2

Uma função aleatória é dita ser estacionária de segunda ordem se:

$$\text{i) } E[Z(\mathbf{x}_i)] = m, \forall \mathbf{x}_i$$

ii) Para todo par  $\{Z(\mathbf{x} + \mathbf{h}), Z(\mathbf{x})\}$  a covariância existe e só depende de  $\mathbf{h}$ . Ou seja, a covariância torna-se uma função do *lag*  $\mathbf{h}$ . Esta função será denotada por  $C(\mathbf{h})$ . A expressão da função de covariância é dada em (9):

$$C(\mathbf{h}) = E\{[Z(\mathbf{x}) - m][Z(\mathbf{x} + \mathbf{h}) - m]\} \quad (5)$$

$$= E\{[Z(\mathbf{x})][Z(\mathbf{x} + \mathbf{h})] - [Z(\mathbf{x})]m - [Z(\mathbf{x} + \mathbf{h})]m + m^2\} \quad (6)$$

$$= E\{[Z(\mathbf{x})][Z(\mathbf{x} + \mathbf{h})]\} - mE\{[Z(\mathbf{x})]\} - mE\{[Z(\mathbf{x} + \mathbf{h})]\} + m^2 \quad (7)$$

$$= E\{[Z(\mathbf{x})][Z(\mathbf{x} + \mathbf{h})]\} - m^2 - m^2 + m^2 \quad (8)$$

$$= E\{[Z(\mathbf{x})][Z(\mathbf{x} + \mathbf{h})]\} - m^2 \quad (9)$$

Se  $\mathbf{h} = \mathbf{0}$ , obtêm-se a seguinte relação:

$$C(\mathbf{0}) = E\{[Z(\mathbf{x})][Z(\mathbf{x})]\} - m^2 \quad (10)$$

$$= E[Z^2(\mathbf{x})] - m^2 \quad (11)$$

$$= E[Z^2(\mathbf{x})] - [E(Z(\mathbf{x}))]^2 \quad (12)$$

$$= Var[Z(\mathbf{x})] \quad (13)$$

### Resultado 3.2.1a

A função  $C(\mathbf{h})$  é uma função par (LANA, 1989).

De fato, tem-se que  $E[Z(\mathbf{x})Z(\mathbf{x} + \mathbf{h})] - m^2 = C(\mathbf{h})$ . Seja  $\mathbf{x}_1 = \mathbf{x}_2 - \mathbf{h}$ , então

$$C(\mathbf{h}) = E[Z(\mathbf{x}_1)Z(\mathbf{x}_1 + \mathbf{h})] - m^2 = E[Z(\mathbf{x}_2 - \mathbf{h})Z(\mathbf{x}_2)] - m^2 = C(-\mathbf{h})$$

Este resultado é importante pois mostra que a função de covariância é a mesma para direções opostas.

### Resultado 3.2.1b

A função  $C(\mathbf{h})$  é limitada pela variância de  $Z(\mathbf{x})$  (LANA, 1989).

De fato, pela desigualdade de Cauchy-Schwarz, tem-se:

$$2|Z(\mathbf{x})Z(\mathbf{x} + \mathbf{h})| \leq |Z(\mathbf{x})|^2 + |Z(\mathbf{x} + \mathbf{h})|^2 \quad (14)$$

$$2|Cov(\mathbf{x}, \mathbf{x} + \mathbf{h})| \leq |VarZ(\mathbf{x})| + |Var[Z(\mathbf{x} + \mathbf{h})]| \quad (15)$$

$$|C(\mathbf{h})| \leq C(\mathbf{0}) = Var[Z(\mathbf{x})]. \quad (16)$$

### Definição 3.2.3

Define-se a função correlograma (ou função de autocorrelação) para cada *lag*  $\mathbf{h}$  como sendo a razão:

$$\rho(\mathbf{h}) = \frac{C(\mathbf{h})}{C(\mathbf{0})}$$

No entanto, em muitos casos de estudo, supor a estacionariedade de segunda ordem é fazer uma suposição demasiadamente forte porque em muitos fenômenos naturais a variância ou dispersão pode ser ilimitada. Por este motivo,

Matheron definiu a chamada hipótese intrínseca (LLOYD, 2011).

### 3.3 Hipótese Intrínseca

#### Definição 3.3.1

Para que uma variável aleatória cumpra a hipótese intrínseca deve-se ter:

1.  $E[Z(\mathbf{x}_i)] = m, \forall \mathbf{x}_i$
2. A variância dos incrementos deve ser finita, ou seja,  $\gamma(\mathbf{h}) = Var[Z(\mathbf{x} + \mathbf{h}) - Z(\mathbf{x})]$  deve ser finita. Chamamos esta variância dos incrementos  $Z(\mathbf{x} + \mathbf{h}) - Z(\mathbf{x}), \forall \mathbf{x}, \mathbf{s} + \mathbf{h} \in D \subset \mathbb{R}^2$ , de semivariograma. Note-se que o semivariograma é uma função de  $\mathbf{h}$ .

Deve-se observar que a estacionariedade de segunda ordem implica a hipótese intrínseca, mas a hipótese intrínseca não implica estacionariedade de segunda ordem. Isto resta claro já que, se o processo é estacionário de segunda ordem, então os momentos de segunda ordem são constantes em toda região de estudo. E no caso da hipótese intrínseca isso valerá apenas para pequenas regiões nas quais a variância é finita. Então, se não for possível supor a estacionariedade de segunda ordem, deve-se utilizar a hipótese intrínseca e, portanto, a função semivariograma.

Existem três caracterizações importantes para a função semivariograma:

- i)  $\gamma(\mathbf{0}) = 0$
- ii)  $\sum_{i=1}^k \sum_{j=1}^k \alpha_i \bar{\alpha}_j 2\gamma(\mathbf{x}_i - \mathbf{x}_j) \leq 0$ . Para todo inteiro positivo inteiro  $k$ , conjunto de localização espaciais  $\{\mathbf{x}_i, i = 1, \dots, k\}$  e conjunto de números complexos  $\{\alpha_i, i = 1, \dots, k\}$  satisfazendo  $\sum_{i=1}^k \alpha_i = 0$ . Esta condição é extremamente importante pois garante que o modelo baseado nas variâncias é não-negativo e, portanto, o semivariograma é não-positivo.

$$\text{iii) } \gamma(|\mathbf{h}|) = \gamma(-|\mathbf{h}|)$$

A condição ii) é conhecida como condicional não-positiva definida (CRESSIE, 2011).

Cabe fazer uma consideração à respeito do termo “semivariograma”. Como  $\gamma(\mathbf{h}) = \frac{1}{2}E[Z(\mathbf{x} + \mathbf{h}) - Z(\mathbf{x})]^2 = \frac{1}{2}Var\{Z(\mathbf{x} + \mathbf{h}) - Z(\mathbf{x})\}$  alguns autores costumam chamar a quantidade  $2\gamma(\mathbf{h})$  de variograma. No entanto, neste trabalho, sempre será usada a expressão semivariograma para designar  $\gamma(\mathbf{h})$ .

Há uma importante relação entre a função semivariograma e a função covariância que é a seguinte:

### Resultado 3.3.2

$$\gamma(\mathbf{h}) = \frac{1}{2}E[Z(\mathbf{x} + \mathbf{h}) - Z(\mathbf{x})]^2 \quad (17)$$

$$= \frac{1}{2}E[Z^2(\mathbf{x} + \mathbf{h}) - 2Z(\mathbf{x} + \mathbf{h})Z(\mathbf{x}) + Z^2(\mathbf{x})] \quad (18)$$

$$= \frac{1}{2}\{E[Z^2(\mathbf{x} + \mathbf{h})] - E[2Z(\mathbf{x} + \mathbf{h})Z(\mathbf{x})] + E[Z^2(\mathbf{x})]\} \quad (19)$$

$$= \frac{1}{2}\{E[Z^2(\mathbf{x})] - 2E[Z(\mathbf{x} + \mathbf{h})Z(\mathbf{x})] + E[Z^2(\mathbf{x})]\} \quad (20)$$

$$= \frac{1}{2}\{2E[Z^2(\mathbf{x})] - 2E[Z(\mathbf{x} + \mathbf{h})Z(\mathbf{x})]\} \quad (21)$$

$$= \{E[Z^2(\mathbf{x})] - E[Z(\mathbf{x} + \mathbf{h})Z(\mathbf{x})]\} \quad (22)$$

$$= Var[Z(\mathbf{x})] + \{E[Z(\mathbf{x})]\}^2 - [C(\mathbf{h}) + m^2] \quad (23)$$

$$= C(\mathbf{0}) + m^2 - C(\mathbf{h}) - m^2 \quad (24)$$

$$= C(\mathbf{0}) - C(\mathbf{h}) \quad (25)$$

Quando a função covariância decresce a função semivariograma cresce para valores crescentes de  $\mathbf{h}$ . Em outras palavras, quando a covariância diminui para valores maiores de  $\mathbf{h}$  a função semivariograma cresce até o limite imposto pela  $Var[Z(\mathbf{x})]$ . Isto pode ser observado na Figura 2.

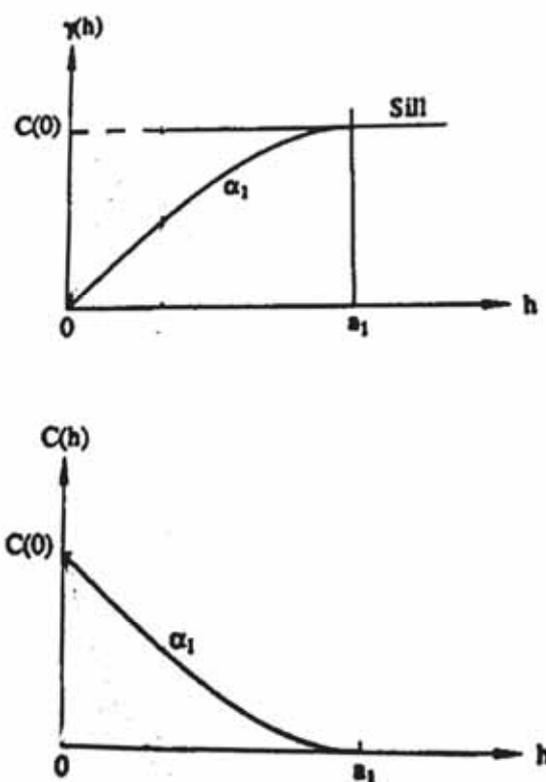


Figura 2 -  $\alpha_1$  : direção da continuidade;  $a_1$ : distância na qual a correlação espacial vai se esvanecendo;  $C(|\mathbf{h}|, \alpha_1) + \gamma(|\mathbf{h}|, \alpha_1) = \text{valor do patamar} = C(0)$ . Fonte: CALVETE & RAMÍREZ (1996)

### Resultado 3.3.3

A função  $\gamma(\mathbf{h})$  é uma função par. De fato, pelo Resultado 3.2.1a temos:

$$\gamma(\mathbf{h}) = C(\mathbf{0}) - C(\mathbf{h}) = C(\mathbf{0}) - C(-\mathbf{h}) = \gamma(-\mathbf{h}).$$

A importância deste resultado é o mesmo que foi apresentado no **Resultado 3.2.1a**, ou seja, o valor do semivariograma é o mesmo para direções opostas.

### Resultado 3.3.4

Este resultado mostra a relação da função semivariograma com a fun-

ção correlograma  $\rho(\mathbf{h})$  :

$$\gamma(\mathbf{h}) = C(\mathbf{0}) - C(\mathbf{h}) \quad (26)$$

$$= \text{Var}[Z(\mathbf{x})][1 - \rho(\mathbf{h})]. \quad (27)$$

### 3.4 Ergodicidade

Será possível, com apenas uma única observação, saber se a função aleatória satisfaz a estacionariedade de segunda ordem, a hipótese intrínseca ou nenhuma das duas? Não existe nenhum teste estatístico para isto (CALVETE & RAMÍREZ, 1996). Pode-se facilitar a inferência estatística considerando-se a hipótese da ergodicidade. Segundo CALVETE & RAMÍREZ (1996) um processo se diz ergódico for possível determinar algumas estatísticas dele com apenas uma única observação. Pode-se relaxar esta hipótese considerando-se, por exemplo, ergocidade na média, ergodicidade na variância, etc.

Para um exemplo do que significa a hipótese de ergodicidade considere uma certa propriedade  $Z$  num ponto  $x_0$  numa região mineradora. Cada realização desta função aleatória deveria ser considerada como regiões distintas. Ou seja, a hipótese de ergodicidade surge como uma maneira de resolver o problema de não haver “múltiplas regiões mineradoras”.

## 4 SEMIVARIOGRAMAS

O semivariograma  $\gamma(\mathbf{h})$  de uma função aleatória é uma função que quantifica a sua estrutura de correlação espacial do fenômeno em estudo. Até agora o que foi visto foram propriedades do **semivariograma teórico**. No entanto, para efeitos práticos é necessário estimá-lo, e com isso encontrar o **semivariograma experimental** ou **empírico**. Matheron fez isso utilizando o método dos momentos que será apresentado na subseção 4.1.

### 4.1 Método dos momentos

Talvez o método mais antigo para se obter estimadores seja o método dos momentos, remontando à época de Karl Pearson no final do século XIX.

Seja:

$$m_r = \frac{1}{n} \sum_{i=1}^n Y_i^r,$$

$r \geq 1$  o  $r$ -ésimo momento amostra de uma amostra aleatória  $Y_1, \dots, Y_n$ . E seja:

$$\mu_r = E[Y^r],$$

$r \geq 1$  o  $r$ -ésimo momento populacional.

O método dos momentos consiste na obtenção de estimadores para  $\boldsymbol{\theta}=(\theta_1, \dots, \theta_k)$ , resolvendo-se as equações:

$$m_r = \mu_r,$$

$r = 1, \dots, k$ .

Este e outros métodos de estimação se encontram em livros clássicos como BOLFARINE (2010) e CASELLA (2010).

#### 4.1.1 Semivariograma Empírico

Seja  $\frac{1}{2}Y^2 = \frac{1}{2}[Z(\mathbf{x} + \mathbf{h}) - Z(\mathbf{x})]^2$ , portanto:

$$E[\frac{1}{2}Y^2] = E[\frac{1}{2}[Z(\mathbf{x} + \mathbf{h}) - Z(\mathbf{x})]^2] = \gamma(\mathbf{h}),$$

Então, pelo método dos momentos apresentados acima, o estimador para o semivariograma é:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(\mathbf{x} + \mathbf{h}) - Z(\mathbf{x})]^2,$$

onde:  $N(h)$  = número de pares de pontos amostrais separados pelo *lag*  $\mathbf{h}$ .

Segundo CRESSIE (1991) este estimador para o semivariograma não é robusto para *outliers*, por isso foram propostos outros estimadores em vários trabalhos probabilísticos como por exemplo CRESSIE & HAWKINS (1980) e GENTON (1998). Entretanto, segundo OLEA (1999), embora hajam outros estimadores propostos que apresentam maior robustez, o mais comumente utilizado ainda é o de momentos.

#### 4.1.2 Amostragem

Segundo YAMAMOTO & LANDIM (2013), existem 3 tipos de amostras a serem consideradas:

**i) Amostragem aleatória simples:** Quando se fala sobre amostragem aleatória simples, quer-se dizer que dentre  $N$  unidades de uma população escolhe-se  $n$ , sem reposição e de maneira aleatória, de forma que qualquer amostra tenha a mesma probabilidade de ser selecionada.

No caso da Geoestatística, os " $n$ " que serão sorteados serão as coordenadas geográficas de uma dada região em estudo.

**ii) Amostragem aleatória estratificada:** A amostragem aleatória estratificada é feita subdividindo-se a região em estudo em células de dimensões fixas nas direções leste-oeste e na direção norte-sul. Dentro de cada células, a coordenada geográficas é escolhida ao acaso e desta maneira

o ponto é selecionado. Ao final deste processo, o número de pontos selecionados será igual ao número de células.

**iii) Amostragem Sistemática:** A amostragem sistemática é feita através de uma malha regular definida através de um ponto escolhido aleatoriamente. No entanto, na prática, a escolha deste ponto é feita pelo pesquisador para otimizar a coleta das unidades dentro da região de estudo.

Estes mesmos autores afirmam que o melhor método de amostragem é a Sistemática, pois, como se espera que haja correlação entre os dados coletados, este tipo de amostragem é o que mais captará tal efeito. Entretanto, nem sempre é possível este tipo de amostragem devido aos acidentes geográficos como por exemplo: rios, lagos, vegetação, etc.

#### 4.1.3 Cálculo de semivariogramas empíricos

Se os dados apresentarem uma distribuição espacial regular, ou seja, se a amostragem for sistemática através de uma malha onde os pontos foram igualmente distribuídos, podemos calcular o semivariograma experimental com facilidade. Se a malha for quadrada teremos as direções Leste-Oeste e Norte-Sul para o vetor  $\mathbf{h}$  (*lag*) além das diagonais do quadrado. Já se a malha for retangular devemos calcular a direção do vetor  $\mathbf{h}$  de acordo com o comprimento dos lados deste retângulo. Para o quadrado, as direções diagonais são de 45 e 315 graus. Já no caso de uma malha retangular, para se determinar as diagonais, devemos usar o valor dos lados deste retângulo. Sejam  $L_1$  e  $L_2$  os comprimentos dos lados de um retângulo. O ângulo entre estes lados é de 90 graus. Portanto, para obtermos o valor do ângulo de inclinação das diagonais, basta calcular:  $\arctg\left(\frac{L_1}{L_2}\right)$  e  $\arctg\left(\frac{L_2}{L_1}\right)$ . O esquema da Figura 3 encontra-se em (YAMAMOTO & LANDIM, 2013, p.37):

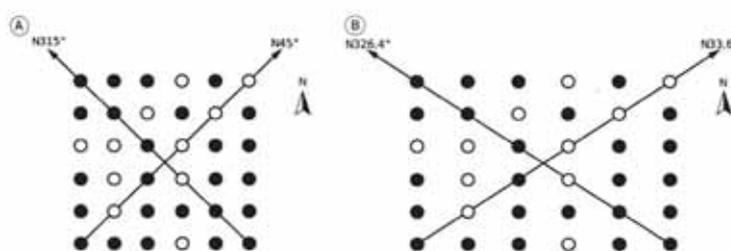


Figura 3 - A) Malha quadrada B) Malha retangular, com indicação das diagonais para o cálculo dos semivariogramas experimentais. Círculo vazio: Ponto não amostrado, Círculo cheio: ponto amostrado.

Já se a malha for irregular, ou seja, se a distribuição dos pontos amostrados for irregular, deve-se adicionar alguns parâmetros para que seja possível ajustar um semivariograma experimental satisfatório.

Os parâmetros definidos são os seguintes: Tolerância do *lag* ou passo, ou seja, do módulo do vetor  $h$  e a tolerância angular.

O esquema da Figura 4 encontra-se em (YAMAMOTO & LANDIM, 2013):

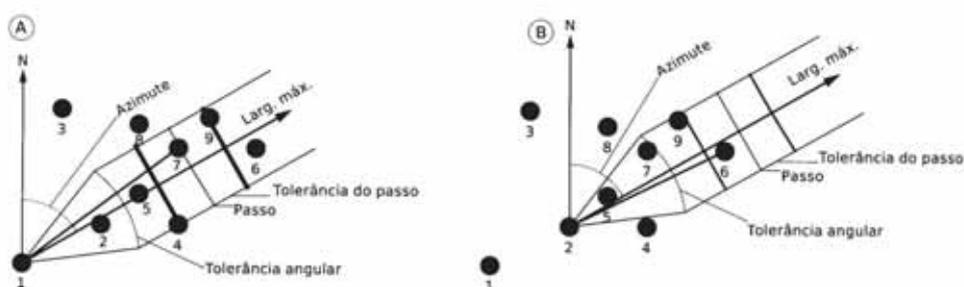


Figura 4 - Esquema mostrando a pesquisa de pares para cálculo de semivariogramas no caso da distribuição irregular. A) O dispositivo de pesquisa é centrado no ponto 1 B) O dispositivo de pesquisa se move e é centrado no ponto 2.

O ângulo no qual será investigado o semivariograma é chamado de

azimute e a direção norte é definida como sendo a origem.

A Figura 4 mostra que fixado o ponto 1, observa-se que o ponto 7 é encontrado dentro da janela formada pela tolerância do ângulo e dos passos. Então a diferença entre os pontos 7 e 1 é considerada no cálculo do semivariograma. Fixando-se agora no ponto 2, o ponto 6 encontra-se dentro da janela, ou seja, ele será considerado no cálculo do semivariograma junto com o ponto 2 e assim sucessivamente.

Feito os ajustes para os cálculos dos semivariogramas, encontra-se um conjunto de pontos que podem ser colocados num sistema cartesiano ortogonal onde o eixo das ordenadas é o tamanho do *lag* e o eixo das ordenadas será o valor do semivariograma  $\hat{\gamma}(\mathbf{h})$  como o da Figura 5:

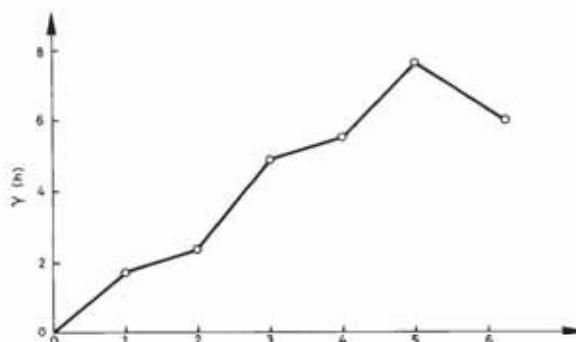


Figura 5 - Exemplo de Semivariograma Experimental

Segundo OLEA (1999) há três boas razões para se empregar os semivariogramas na Geoestatística:

1. O semivariograma é uma estatística que avalia o decrescimento das médias entre duas variáveis aleatórias como a distância entre os incrementos destas variáveis, conduzindo a muitas aplicações na análise exploratória de dados.
2. Está demonstrado que é impossível obtermos a Krigagem sem o conhecimento da covariância ou do semivariograma. Na prática, não são conhecidas nem a

função de covariância e nem a função semivariograma. Por isso devemos estimá-las.

3. Para as equações de Krigagem serão usados os semivariogramas ao invés da função covariância, dado que as vezes a variância pode não ser finita e isto seria um problema caso quisesse se usar a função de covariância. O semivariograma exige suposições mais fracas que a função covariância. Um exemplo clássico de um processo que possui semivariograma e não possui covariância é o processo discreto de Werner-Levy.

Além disso, ISAACS & SRIVASTAVA (1989) e CRESSIE & GRONDONA (1992) enumeram mais duas:

4. O estimador do semivariograma é insensível a adição de uma constante à função aleatória, já o estimador de covariância depende de tal constante.
5. O estimador do semivariograma é menos sensível que o estimador da covariância para a exigência de que a  $E[Z(\mathbf{x})]$  seja constante.

Com a estimação dos semivariogramas, obtemos apenas pontos. E o objetivo é conhecer a variância dos incrementos em regiões, por isso há a necessidade de se ajustar funções matemáticas aos pontos obtidos. A maneira como isto é feito será apresentado na próxima seção.

## 4.2 Modelos para Representar os Semivariogramas Empíricos: Semivariogramas Teóricos

Para se ajustar funções aos pontos obtidos na estimação do semivariograma há, segundo CALVETE & RAMÍREZ (1996) e FUENTES (2011) os seguintes métodos:

1. Método dos mínimos quadrados ordinário;
2. Método dos mínimos quadrados ponderados;

3. Método dos mínimos quadrados não-lineares;
4. Método Jackknife;
5. Método da Verossimilhança;
6. Método da Validação Cruzada;
7. Ajuste à sentimento ou “visualmente”.

Destes métodos, falar-se-á apenas do Método dos Mínimos Quadrados Ponderados e do Método do ajuste à sentimento ou “visuamente”.

#### 4.2.1 Método dos Mínimos Quadrados Ponderados

CRESSIE (1989), num trabalho de 1985 *“Fitting Variogram Models by Weighted Least Squares”*, propõe um método que, em suas palavras, “tem um compromisso entre a eficiência e a simplicidade” e que propõe minimizar a soma de quadrados:

$$\omega(\mathbf{h}) = \frac{\sum_{k=1}^K \left\{ \frac{2\hat{\gamma}(\mathbf{h}(k))}{2\hat{\gamma}(\mathbf{h}(k); \boldsymbol{\theta})} - 1 \right\}^2}{|N(\mathbf{h}(k))|}$$

com relação aos parâmetros  $\boldsymbol{\theta}$  do semivariograma. A sequência  $\mathbf{h}(1), \dots, \mathbf{h}(K)$  denotam os *lags* nos quais o estimador  $\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [Z(\mathbf{x} + \mathbf{h}) - Z(\mathbf{x})]^2$ , foi obtido.

FUENTES (2011) reescreve a expressão acima da seguinte maneira:  $\omega(\mathbf{h}) = \sum_{\mathbf{u} \in U} \frac{N(\mathbf{h}_{\mathbf{u}})}{\gamma(\mathbf{h}_{\mathbf{u}}; \boldsymbol{\theta})^2} [\hat{\gamma}(\mathbf{h}_{\mathbf{u}}) - \gamma(\mathbf{h}_{\mathbf{u}}; \boldsymbol{\theta})]^2$ , na qual toma-se  $U = \{\mathbf{u}; N(\mathbf{h}_{\mathbf{u}}) \geq G_1, |\mathbf{h}_{\mathbf{u}}| \leq G_2\}$ , em que  $G_1 = 30$  e  $G_2 =$  metade do maior *lag* dos dados. Os pesos são pequenos se  $N(\mathbf{h}_{\mathbf{u}})$  é pequeno ou  $\gamma(\mathbf{h}_{\mathbf{u}}; \boldsymbol{\theta})$  é grande. Então a estimação não paramétrica (semivariograma empírico) em *lags* grandes tende a receber pequenos pesos.

#### 4.2.2 Ajuste à sentimento ou “visualmente”

O ajuste a sentimento consiste em selecionar os parâmetros dos semivariogramas levando-se em conta a estrutura do semivariograma empírico e a estrutura do fenômeno físico, o comportamento do semivariograma em pequenas distâncias (o

ajuste deve ser bom nos pontos iniciais). O procedimento é feito utilizando softwares que interagem com o usuário de maneira que este possa ajustar a curva (função) manualmente aos pontos do semivariograma empírico.

### 4.2.3 Procedimentos para seleção de modelos

Segundo FUENTES (2011) pode-se seguir os seguintes procedimentos para selecionar os modelos:

1. Inspeção visual do gráfico do semivariograma empírico.
2. Minimizar os Resíduos ponderados da função de soma de quadrados,  $\omega(\mathbf{h})$
3. Maximizando a função de verossimilhança  $L(\hat{\beta}; \hat{\theta})$
4. Através de um critério que penaliza a função de Verossimilhança, por exemplo, o critério de informação de Akaike:  $AIC = L(\hat{\beta}; \hat{\theta}) - p$ . Ele pode ser estimado através da expressão:  $A = \{n \cdot \log_e \left(\frac{2\pi}{n}\right) - n + 2\} + n \cdot \log_e(RSS) + 2 \cdot p$ , onde  $n$  é o número de *lags* no qual a semivariância é estimada,  $p$  é o número de parâmetros e  $RSS$  é a soma de quadrados dos Resíduos (OLIVER & WEBSTER, 1989).

FUENTES (2011) afirma que o estimador de Mínimos Quadrados Ponderados tem sido o mais escolhido pelos pesquisadores por ser mais fácil implementação computacional que, por exemplo, o Estimador de Máxima Verossimilhança.

Daqui para frente e até o fim desta seção considerar-se-á  $h = |\mathbf{h}|$ . Os principais semivariogramas teóricos que podem ser ajustados e selecionados pelos métodos apresentados anteriormente são:

#### a) Efeito pepita puro:

Não é comum empregá-lo sozinho e sim combinando-o com outro modelo. Ele é indicativo de um fenômeno sem nenhuma autocorrelação.

$$\gamma(\mathbf{h}) = \begin{cases} C_0, & \text{se } h > 0 \\ 0, & \text{se } h = 0. \end{cases}$$

Onde  $C_0$  é o chamado *sill* ou patamar. Seu gráfico é exibido na Figura 6.



Figura 6 - Semivariograma Teórico: Efeito Pepita

Os parâmetros dos próximos modelos teóricos são o *Sill* ou patamar que será representado pela letra  $C$  e que coincide com a variância da variável  $Z$ ; o *range* ou alcance “ $a$ ” que representa a zona de influência em torno de um ponto na qual, depois dela, a autocorrelação é nula e há também a distância integral definida como sendo  $d = \frac{1}{C} \int_0^{+\infty} [C - \gamma(h)] dh$ , ou seja, é o valor de  $h$  tal que a área debaixo da curva do semivariograma é igual à aquela área formada pelo assintota horizontal  $C(0)$  e a curva do semivariograma. Este valor serve para se medir o grau de correlação espacial da variável.

#### b) Modelo Esférico:

$$\gamma(\mathbf{h}) = \begin{cases} \frac{C}{2} \left[ 3 \left( \frac{h}{a} \right) - \left( \frac{h}{a} \right)^3 \right], & \text{se } h \leq a \\ C, & \text{se } h > a. \end{cases}$$

O coeficiente angular na origem é igual a  $1,5 \frac{C}{a}$ . Isto prova-se facilmente:

$$\frac{\partial \gamma(\mathbf{h})}{\partial h} = \frac{C}{2} \left[ \frac{3}{a} - 3 \left( \frac{h}{a} \right)^2 \frac{1}{a} \right]$$

Portanto, na origem ( $h=0$ )

$$\frac{\partial \gamma(h)}{\partial h} \Big|_{h=0} = \frac{3C}{2a}$$

A distância integral  $d$  é:

$$d = \frac{1}{C} \int_0^a C \left[ 1 - \frac{3}{2} \left( \frac{h}{a} \right) + \frac{1}{2} \left( \frac{h}{a} \right)^3 \right] dh = \frac{5}{12} a$$

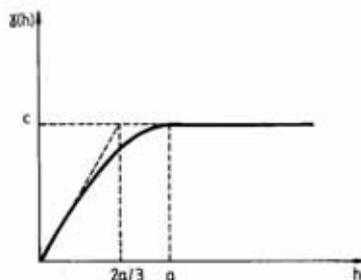


Figura 7 - Semivariograma Teórico: Modelo Esférico

O modelo esférico, exibido na Figura 7, é um dos mais empregados na prática. Ele alcança o patamar para uma distância finita, ou seja,  $h = a$ . Este modelo é indicativo de fenômenos contínuos ou com quantidade enumerável de descontinuidades, ainda que não deriváveis.

### c) Modelo exponencial:

$$\gamma(h) = C \left[ 1 - \exp\left(-\frac{h}{a}\right) \right]$$

O modelo exponencial, dado pela Figura 8 alcança seu patamar assintoticamente:

$$\lim_{h \rightarrow \infty} C \left[ 1 - \exp\left(-\frac{h}{a}\right) \right] = C$$

Para efeitos práticos, sugere-se tomar como alcance efetivo a distância a qual  $\gamma(h) = 0.95C$ , que é aproximadamente igual a  $a' = 3a$ , sendo “a” a distância do semivariograma esférico.

A inclinação da reta tangente à origem é  $\frac{C}{a}$  como se demonstra facilmente:

$$\frac{\partial\lambda(h)}{\partial h} = \frac{C}{a} \left[ \exp\left(-\frac{h}{a}\right) \right]$$

$$\frac{\partial\lambda(h)}{\partial h} \Big|_{h=0} = \frac{C}{a}$$

A distância integral d é obtida através de:

$$d = \frac{1}{C} \int_0^{+\infty} C - C \left[ 1 - \exp\left(-\frac{h}{a}\right) \right] dh = a = \frac{a'}{3}$$

O semivariograma exponencial, exibido na Figura 8, é representativo de fenômenos contínuos salvo em um conjunto de pontos (em uma dimensão), retas (em duas direções) ou planos (em três dimensões).

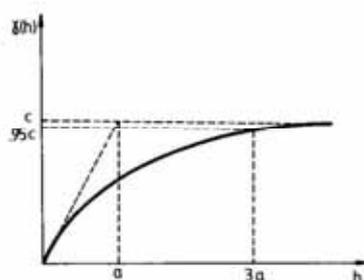


Figura 8 - Semivariograma Teórico: Modelo Exponencial

#### d) Modelo Gaussiano:

$$\gamma(h) = C \left[ 1 - \exp\left(-\frac{h^2}{a^2}\right) \right]$$

O modelo gaussiano, exibido na Figura 9, também atinge seu patamar, assim como o modelo exponencial, assintoticamente:

$$\lim_{h \rightarrow +\infty} C \left[ 1 - \exp\left(-\frac{h^2}{a^2}\right) \right] = C$$

Define-se, para efeitos práticos,  $a' = \sqrt{3}a$ , valor para o qual  $\gamma(h) = 0.95C$

A inclinação da reta tangente na origem é:

$$\frac{\partial\gamma(h)}{\partial h} = \frac{2hC}{a^2} \exp\left(-\frac{h^2}{a^2}\right)$$

$$\frac{\partial\gamma(h)}{\partial h} \Big|_{h=0} = 0$$

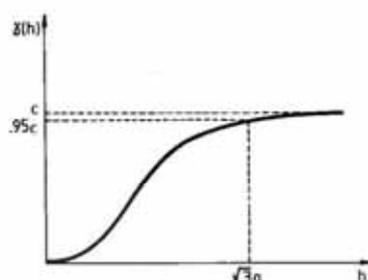


Figura 9 - Semivariograma Teórico: Modelo Gaussiano

Este modelo se emprega para fenômenos suaves, ou seja, contínuo em todos os pontos e deriváveis na sua maioria

Para outros exemplos de modelos teóricos de semivariograma, deve-se consultar CALVETE & RAMÍREZ (1996), OLEA (1999) e YAMAMOTO & LANDIM (2013).

Espera-se que o comportamento do semivariograma teórico seja regido por:

$$\lim_{|h| \rightarrow \infty} \frac{\gamma(\mathbf{h})}{|h|^2} = 0$$

Ou seja, espera-se que função semivariograma cresça mais lentamente que o quadrado da distância. Se isto não ocorrer é de se esperar que a função aleatória não atenda a hipótese intrínseca. Além disso ela deve obedecer às caracterizações dadas na Seção 3.3. A obtenção de novas funções semivariograma parece não ser algo

fácil como se pode notar na tentativa feita por CONCEIÇÃO (2013) de usar algumas funções de distribuição de probabilidades como possíveis funções semivariograma.

MATHERON (1971) observa que, para uma função unidimensional na qual o semivariograma tenha a forma  $\gamma(h) = kh^\theta$ , a variância é proporcional a  $\gamma(h)$  e inversamente proporcional ao número de pares  $N(h)$ , o que implica que as flutuações do semivariograma tendem a diminuir quando se aumentam o número de pares e tendem a aumentar para valores grandes de  $h$ . Algumas maneiras de se diminuir tais flutuações são:

1. Aumentar o número de pares  $N(h)$  em cada intervalo, considerando intervalos maiores ou utilizando-se uma tolerância angular maior.
2. Aumentar o número de dados disponíveis.

As flutuações dos semivariogramas empíricos para valores grandes de  $|h|$  não constituem um problema já que os resultados da Krigagem são pouco afetados por valores do semivariograma a longas distâncias.

Ainda segundo ANDRIOTTI (1988) é de grande importância o comportamento dos semivariogramas na origem, ou seja, para pequenos valores de  $h$ . Acontece que há uma forte relação entre este comportamento e a regularidade espacial da Função Aleatória  $Z(x)$  e portanto da Variável Regionalizada representada por ela.

Existem quatro tipos de semivariograma de acordo com seu comportamento próximo à origem:

1. (Comportamento Parabólico)

$\gamma(h) \sim A|h|^2$  quando  $h \rightarrow 0$ .  $\gamma(h)$  sendo duas vezes derivável e a Função Aleatória derivável, tem-se que há uma grande regularidade espacial na Variável em estudo.

2. (Comportamento Linear)

$\gamma(h) \sim A|h|$  quando  $h \rightarrow 0$ . Neste tipo de comportamento tem-se que a Variável em estudo tem menos regularidade espacial que no caso anterior.

3. (Descontinuidade na Origem)

Este é o chamado Efeito Pepita que será visto em seguida, na secção 4.4.

4. (Efeito Pepita Puro)

Reflete a inexistência de correlação espacial entre as Variáveis Aleatórias  $Z(\mathbf{x})$  e  $Z(\mathbf{x} + \mathbf{h})$ .

Deve-se, neste momento, fazer a consideração de que em alguns fenômenos o semivariograma não apresentará o comportamento de se estabilizar em algum ponto, como acontece no caso do semivariograma esférico, ou assintoticamente, como nos semivariogramas gaussiano e exponencial. Se isto ocorrer, se diz que o fenômeno possui tendência. Nestes casos, deve-se procurar algum método que faça com esta tendência seja retirada para que se possa aplicar a técnica da Krigagem Universal que será vista à frente.

### 4.3 Efeito Pepita: valor do semivariograma teórico na origem

Matematicamente, o valor dos semivariogramas teóricos na origem deveria ser 0, ou seja,  $\gamma(\mathbf{0}) = 0$ . Isto, em geral, não acontece. O valor na origem passa a ser  $C_0 > 0$ . Este valor  $C_0$  é denominado Efeito Pepita. Alguns dos motivos para o aparecimento deste valor são a variabilidade do fenômeno espacial bem como de alterações na escala da amostragem (YAMAMOTO & LANDIM, 2013). No caso do aparecimento do Efeito Pepita, os modelos apresentarão como valor de patamar:  $C + C_0$ . Isto levará à redefinição das expressões dos semivariograma teóricos, (OLEA, 1999), conforme descritos na Tabela 1.

Tabela 1: Modelos teóricos de semivariogramas com efeito pepita

Esférico:	$\gamma(\mathbf{h}) = \begin{cases} C_0 + \frac{C}{2} \left[ 3 \left( \frac{h}{a} \right) - \left( \frac{h}{a} \right)^3 \right], & \text{se } h \leq a \\ C_0 + C, & \text{se } h > a. \end{cases}$
Exponencial:	$\gamma(h) = C_0 + C \left[ 1 - \exp\left(-\frac{h}{a}\right) \right]$
Gaussiano:	$\gamma(h) = C_0 + C \left[ 1 - \exp\left(-\frac{h^2}{a^2}\right) \right]$

Com isso define-se  $E = \frac{C_0}{C}$  chamado grau de aleatoriedade (GUERRA, 1988). Segundo este mesmo autor, este grau de aleatoriedade pode ser classificado em:

Tabela 2: Classificação dos Graus de Aleatoriedade

Grau de aleatoriedade	Componente Aleatória
$E \leq 0,15$	Pequena
$0,15 \leq E \leq 0,30$	Significativa
$E > 0,30$	Muito Significativa

#### 4.4 Fenômenos Anisotrópicos

Um fenômeno espacial é chamado isotrópico se apresenta o mesmo semivariograma para todas as direções do espaço. Se isto não ocorre, o fenômeno é chamado de anisotrópico. Pode-se interpretar geometricamente o conceito de isotropia e a anisotropia imaginando que a isotropia equivaleria à círculos concêntricos nos quais os valores do semivariograma são iguais para cada direção e a anisotropia equivaleria à elipses concêntricas nos quais os valor são mais pronunciando em uma direção como está exibido na Figura 10 retirada de (CALVETE & RAMÍREZ, 1996). A anisotropia pode ser representada matematicamente por

$$\Omega(\theta) = \{A^2 \cos^2(\theta - \phi) + B^2 \sin^2(\theta - \phi)\}^{\frac{1}{2}},$$

na qual  $A$  é o gradiente do semivariograma na direção da máxima variação do parâmetro de distância e  $B$  é o gradiente do semivariograma na direção da menor variação do parâmetro de distância,  $\phi$  é o ângulo da direção de maior variação e  $\theta$  é o azimute (OLIVER & WEBSTER, 1989).

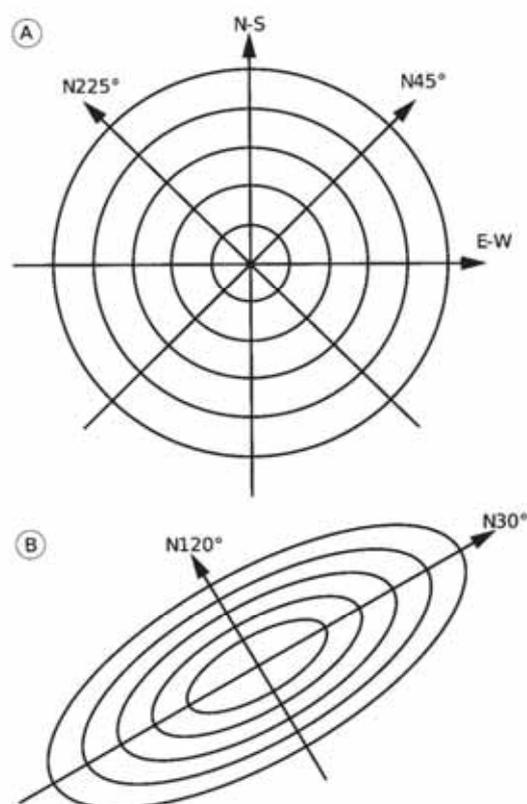


Figura 10 - Anisotropia.

Existem três tipos de anisotropia: anisotropia geométrica, anisotropia zonal, anisotropia mista.

A anisotropia geométrica ocorre quando existe um patamar com dois alcances diferentes. Já a anisotropia zonal ocorre quando existem patamares diferentes para um único alcance. A anisotropia mista ocorre quando o fenômeno apresenta patamares e amplitudes diferentes. A Figura 11 ilustra estes três tipos de anisotropia (YAMAMOTO & LANDIM, 2013):

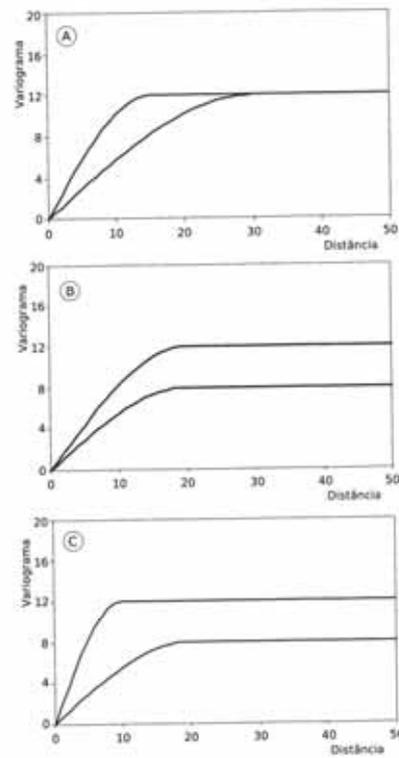


Figura 11 - Tipos de Anisotropia: A) Geométrica, B) Zonal e C) Mista

Quando a presença de anisotropia é detectada, deve-se proceder à correção, pois como aponta o trabalho de ROSSONI & LIMA (2012) as estimativas geoestatísticas de dados que tem a anisotropia levada em consideração apresenta menores erros quadráticos médios. Processos para correção da anisotropia geométrica e zonal são apresentados em YAMAMOTO & LANDIM (2013), no entanto, podem ser resumidos em se aplicar transformações lineares a fim de rotacionar e redimensionar a elipse que anteriormente representava o fenômeno transformando-a numa circunferência.

## 5 KRIGAGEM ORDINÁRIA E A KRIGAGEM UNIVERSAL

A diferença entre a Krigagem e outros métodos de interpolação é a maneira como os pesos são atribuídos às diferentes amostras (DRUCK et al, 2004). Estes pesos serão encontrados quando a expressão para variação da covariância com a distância, obtida através dos semivariogramas, for incorporada às equações da Krigagem. A soma destes pesos deverá ser igual a um.

### 5.1 Krigagem Ordinária

A Krigagem Ordinária pode ser considerado o “algoritmo fundamental da geoestatística”(DEUSTCH & JOURNAL, 1998 apud LLOYD, 2011).

A Krigagem Ordinária é um método local de estimativa e, portanto, a estimativa de um ponto não amostrado resulta das combinações lineares dos pontos na vizinhança (OLEA, 2009).<sup>7</sup>

Sejam:

- $\hat{Z}_{OK}(\mathbf{x}_0)$  o estimador de Krigagem Ordinária de um ponto  $\mathbf{x}_0$  não amostrado,
- $\mathbf{x}_i$  as coordenadas geográficas de um local  $i$ ,
- $k$  o número de observações consideradas,
- $m$  a média de  $Z(\mathbf{x})$ ,  $\forall x$  na região considerada
- $\lambda_i$  os pesos de cada amostra.

Considerar-se-á, neste trabalho, que a função aleatória  $Z$  satisfaz a hipótese intrínseca, que é uma hipótese mais fraca do que a estacionariedade de 2ª ordem. Isto quer dizer:

- $E[Z(\mathbf{x})] = m$

- $Var[Z(\mathbf{x}) - Z(\mathbf{x} + \mathbf{b})] = 2\gamma(\mathbf{h})$ , quando  $\gamma(\mathbf{h})$  é o semivariograma da função aleatória  $Z$ .

Tem-se:

$$\hat{Z}_{OK}(\mathbf{x}_0) = \sum_{i=1}^k \lambda_i Z(\mathbf{x}_i). \quad (28)$$

Deseja-se que este estimador seja não-viciado, ou seja:

$$E[\hat{Z}_{OK}(\mathbf{x}_0) - Z(\mathbf{x}_0)] = 0. \quad (29)$$

Desenvolvendo-se esta expressão:

$$E[\hat{Z}_{OK}(\mathbf{x}_0)] - E[Z(\mathbf{x}_0)] = 0 \quad (30)$$

$$E\left[\sum_{i=1}^k \lambda_i Z(\mathbf{x}_i)\right] - E[Z(\mathbf{x}_0)] = 0 \quad (31)$$

$$\sum_{i=1}^k \lambda_i E[Z(\mathbf{x}_i)] - E[Z(\mathbf{x}_0)] = 0 \quad (32)$$

$$\sum_{i=1}^k \lambda_i m - m = 0 \quad (33)$$

$$m\left(\sum_{i=1}^k \lambda_i - 1\right) = 0 \quad (34)$$

$$\sum_{i=1}^k \lambda_i = 1. \quad (35)$$

Deseja-se que a variância do erro de estimação seja o menor possível, ou seja, deseja-se minimizar a expressão:

$$\sigma_E^2 = Var[\hat{Z}_{OK}(\mathbf{x}_0) - Z(\mathbf{x}_0)]. \quad (36)$$

Desenvolvendo-se esta expressão:

$$\sigma_E^2 = E[(\hat{Z}_{OK}(\mathbf{x}_0) - Z(\mathbf{x}_0))^2] - \{E[\hat{Z}_{OK}(\mathbf{x}_0) - Z(\mathbf{x}_0)]\}^2 = \quad (37)$$

$$E[(\hat{Z}_{OK}(\mathbf{x}_0) - Z(\mathbf{x}_0))^2] = E \left[ \left( \sum_{i=1}^k \lambda_i Z(\mathbf{x}_i) - \sum_{i=1}^k \lambda_i Z(\mathbf{x}_0) \right)^2 \right] = \quad (38)$$

$$E \left[ \left( \sum_{i=1}^k \lambda_i (Z(\mathbf{x}_i) - Z(\mathbf{x}_0)) \right)^2 \right] = \quad (39)$$

$$E \left[ \left( \sum_{i=1}^k \lambda_i (Z(\mathbf{x}_i) - Z(\mathbf{x}_0)) \right) \left( \sum_{j=1}^k \lambda_j (Z(\mathbf{x}_j) - Z(\mathbf{x}_0)) \right) \right] \quad (40)$$

$$\sum_{i=1}^k \sum_{j=1}^k \lambda_i \lambda_j E[(Z(\mathbf{x}_i) - Z(\mathbf{x}_0))(Z(\mathbf{x}_j) - Z(\mathbf{x}_0))] \quad (41)$$

Antes de prosseguir, deve-se obter uma expressão para este semivariograma:

$$\gamma(\mathbf{x}_i - \mathbf{x}_j) = \frac{1}{2} E[(Z(\mathbf{x}_i) - Z(\mathbf{x}_j))^2] = \quad (42)$$

$$\frac{1}{2} E[((Z(\mathbf{x}_i) - Z(\mathbf{x}_0)) - (Z(\mathbf{x}_j) - Z(\mathbf{x}_0)))^2] = \quad (43)$$

$$\frac{1}{2} E[(Z(\mathbf{x}_i) - Z(\mathbf{x}_0))^2] + \frac{1}{2} E[(Z(\mathbf{x}_j) - Z(\mathbf{x}_0))^2] - E[(Z(\mathbf{x}_i) - Z(\mathbf{x}_0))(Z(\mathbf{x}_j) - Z(\mathbf{x}_0))] \quad (44)$$

Substituindo-se (44) na equação (41), obtém-se:

$$E[(Z(\mathbf{x}_i) - Z(\mathbf{x}_0))(Z(\mathbf{x}_j) - Z(\mathbf{x}_0))] = \gamma(\mathbf{x}_i - \mathbf{x}_0) + \gamma(\mathbf{x}_j - \mathbf{x}_0) - \gamma(\mathbf{x}_i - \mathbf{x}_j). \quad (45)$$

Logo:

$$\sigma_E^2 = \sum_{i=1}^k \sum_{j=1}^k \lambda_i \lambda_j [\gamma(\mathbf{x}_i - \mathbf{x}_0) + \gamma(\mathbf{x}_j - \mathbf{x}_0) - \gamma(\mathbf{x}_i - \mathbf{x}_j)]. \quad (46)$$

Tirando-se o fator comum a  $\sum_{i=1}^k \lambda_i$  e  $\sum_{j=1}^k \lambda_j$  que valem um e tendo em conta que:  $\sum_{i=1}^k \lambda_i \gamma(\mathbf{x}_i - \mathbf{x}_0) = \sum_{j=1}^k \lambda_j \gamma(\mathbf{x}_j - \mathbf{x}_0)$ , já que  $i$  e  $j$  são índices mudos, (46) resulta na seguinte expressão:

$$\sigma_E^2 = - \sum_{i=1}^k \sum_{j=1}^k \lambda_i \lambda_j \gamma(\mathbf{x}_i - \mathbf{x}_j) + 2 \sum_{i=1}^k \lambda_i \gamma(\mathbf{x}_i - \mathbf{x}_0) \quad (47)$$

Esta é a expressão a ser minimizada sujeita à restrição  $\sum_{i=1}^k \lambda_i = 1$ . Para isto, deve-se utilizar multiplicadores de Lagrange  $\psi$ . A função objetivo a ser minimizada será:

$$L(\lambda_1, \dots, \lambda_k, \psi) = - \sum_{i=1}^k \sum_{j=1}^k \lambda_i \lambda_j \gamma(\mathbf{x}_i - \mathbf{x}_j) + 2 \sum_{i=1}^k \lambda_i \gamma(\mathbf{x}_i - \mathbf{x}_0) - 2\psi \left( \sum_{i=1}^k \lambda_i - 1 \right) \quad (48)$$

Para isso deve-se fazer:

$$\frac{\partial L}{\partial \lambda_i} = 0, \text{ para } i = 1, \dots, k \text{ e } \frac{\partial L}{\partial \psi} = 0$$

E com isso obtém-se (n+1) equações:

$$\begin{cases} \sum_{j=1}^k \lambda_j \gamma(\mathbf{x}_i - \mathbf{x}_j) + \psi = \gamma(\mathbf{x}_0 - \mathbf{x}_i) \\ \sum_{i=1}^k \lambda_i = 1 \end{cases}$$

Neste caso, a variância da Krigagem Ordinária será:

$$\sigma_{KO}^2 = \sum_{i=1}^k \lambda_i \gamma(\mathbf{x}_0 - \mathbf{x}_i) + \psi \quad (49)$$

O sistema acima pode ser escrito de maneira matricial:

$$\begin{pmatrix} 0 & \gamma_{12} & \gamma_{13} & \cdots & \gamma_{1n} & 1 \\ \gamma_{21} & 0 & \gamma_{23} & \cdots & \gamma_{2n} & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \gamma_{n3} & \cdots & 0 & 1 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \psi \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \\ 1 \end{pmatrix}$$

Onde:  $\gamma_{ij} = \gamma(\mathbf{x}_i - \mathbf{x}_j)$  e  $\gamma_i = \gamma(\mathbf{x}_i - \mathbf{x}_0)$

Com isso encontra-se os valores dos  $\lambda$ 's (pesos) que serão atribuídos às amostras que serão utilizadas para estimar a resposta nos pontos não amostrados.

Em ANDRIOTTI (1988) há várias considerações interessantes sobre a Krigagem dentre as quais destacamos três:

1. A Krigagem é aditiva, ou seja, pode-se fazer um Krigagem Global ou fazer Krigagens Parciais e depois fazer sua composição, o que não acontece com as variâncias de Krigagem.
2. O Sistema de Krigagem e a Variância de Krigagem não dependem dos valores individuais, o que permite prever a qualidade da estimação por Krigagem através da configuração da malha de amostragem a ser implantada, simplesmente.
3. O Sistema de Krigagem e a Variância de Krigagem levam em conta a geometria do domínio a ser estimado, as distâncias entre estimadores e estimados, a geometria da informação e a estrutura do fenômeno a ser estudado (pois utiliza a função semivariograma ou a covariância).

## 5.2 Krigagem Universal

A Krigagem Ordinária é uma metodologia que só pode ser aplicada em fenômenos onde a função aleatória envolvida possui estacionariedade de segunda ordem ou satisfaz a hipótese intrínseca. Em ambos os casos uma das suposições é que a esperança da função aleatória seja constante em toda região de análise. Se isto não ocorrer, ou seja, se a média não puder ser considerada constante dentro de uma “janela móvel” LLOYD (2011) não há garantia de que a Krigagem Ordinária produzirá estimadores não viciados (CLARK & HARPER, 2000). No entanto, em vários campos de estudo as funções não são estacionárias. Pode-se ver isso como por exemplo nos trabalhos de ABOUFIRASSI & MARIÑO (1983) e NEUMAN & JACOBSON (1984) que estudaram características de aquíferos, no trabalho de FERREIRA et al. (2013) que busca estudar o melhoramento da representação cartográfica de superfícies batimétricas e ainda no trabalho de ATKINSON (2001) na área de geografia física. Além disso, outras áreas como a Meteorologia e a Engenharia de Reservatórios lidam com variáveis não estacionárias. Inclusive um dos motivos que levaram Matheron a estudar os fenômenos não estacionários foram estudos do assoalho oceânico feitos por Journel. (ARMSTRONG, 1984).

MATHERON (1971) sugere a Krigagem Universal, chamada posteriormente de Krigagem com Tendência (LLOYD, 2011), Krigagem Não-Estacionária (LEUANGTHONG et al., 2008) ou ainda Krigagem com Média Desconhecida (CHILÈS & DELFINER, 1999) como metodologia para resolver este tipo de problema. Inicialmente, MATHERON (1971) criticou a estimação de regiões realizadas por meio de Superfícies de Tendência Polinomiais dizendo que é pouco natural impor à natureza um ajuste polinomial. Suas críticas foram as seguintes:

1. Os métodos de ajuste por superfícies polinomiais frequentemente apresentam uma confusão entre o conceito e o modo de operação e há aparentemente três interpretações possíveis com relação à tendência (*trend*):

Sendo  $Z(\mathbf{x})$  é a Variável Regionalizada de interesse e  $P(\mathbf{x})$  o polinômio obtido pelo método dos mínimos quadrados para ajustar os valores experimentais nos pontos  $\mathbf{x}_1, \dots, \mathbf{x}_n$ , os valores  $P(\mathbf{x})$  podem ser tomados como um ou outro dos seguintes significados incompatíveis:

- a)  $P(\mathbf{x})$  é uma estimativa da esperança  $E[Z(\mathbf{x})]$  *a priori*. Este é o sentido do “drift” que será adotado na teoria da Krigagem Universal.
- b)  $P(\mathbf{x})$  é uma estimativa do valor de  $Z(\mathbf{x})$  tomada pela variável regionalizada no ponto  $\mathbf{x}$ . Este sentido é o mesmo do dado na Krigagem Pontual ou Ordinária.
- c) Por último,  $P(\mathbf{x})$  é algumas vezes tomada como média móvel.  $P(\mathbf{x}_0)$  será a estimativa do valor médio de  $Z(\mathbf{x})$  em torno de uma área a ser especificada ao redor do ponto  $\mathbf{x}_0$ .

Ainda segundo MATHERON (1971) estas distinções são importantíssimas na teoria e na prática. Isto porque em alguns estudos o objetivo é reconstruir o mecanismo na qual produz o fenômeno está sendo estudado como, por exemplo, em geofísica, onde as noções de anomalias regionais correspondem ao conceito de “drift” que será dado nas equações de Krigagem Universal correspondente ao caso a). Nos problemas de contornos submarinos, o sentido que é tomado é

o do caso b). Nos casos de mineração, o importante é estimar o valor da média do teor de algum mineral procurado em algum local e isto corresponde ao caso c).

2. Já a segunda objeção se relaciona com o método dos mínimos quadrados. Matheron afirma que cada uma das interpretações que podem ser dadas ao polinômio, a), b) ou c) são incompatíveis entre si.
3. A terceira e última objeção é que o método dos mínimos quadrados não apresenta uma maneira de se determinar a magnitude do erro produzido pela estimativa do “drift”. MATHERON (1969) afirma que a variância dos Resíduos não é uma estimativa da variância isto porque, segundo ele, a variância da diferença  $Z(\mathbf{x}_i) - P(\mathbf{x}_i)$  nos pontos aonde os valores  $Z(\mathbf{x}_i)$  são conhecidas em geral são muito mais baixas do que as diferenças aonde os pontos  $Z(\mathbf{x}_i)$  são desconhecidos e portanto a Variância dos Resíduos não seria uma boa estimativa da variância de  $Z(\mathbf{x})$  (Caso b). Além disso não é a estimativa da variância do “drift” (Caso a). Por fim, Matheron afirma que há no máximo um *link* entre a diferença  $Z(\mathbf{x}_i) - P(\mathbf{x}_i)$  considerada como um estimador da esperança *a priori*.

### 5.2.1 Equações da Krigagem sob suposição da Hipótese Intrínseca

Sendo  $\gamma(h) = \frac{1}{2}E[(Z(\mathbf{x} + \mathbf{h}) - Z(\mathbf{x}))^2]$  o semivariograma representativo do fenômeno, e se, por hipótese, é válida a separação  $Z(\mathbf{x}) = m(\mathbf{x}) + Y(\mathbf{x})$ , ou seja, a função aleatória puder ser decomposta em uma componente  $m(\mathbf{x})$  que expressa uma tendência no valor da esperança da função aleatória, mais uma parte aleatória, representada por  $Y(\mathbf{x})$  chega-se à seguinte conclusão:

$$\frac{1}{2}[m(\mathbf{x} + \mathbf{h}) - m(\mathbf{x})]^2 + \frac{1}{2}[Y(\mathbf{x} + \mathbf{h}) - Y(\mathbf{x})]^2 = \frac{1}{2}[m(\mathbf{x} + \mathbf{h}) - m(\mathbf{x})]^2 + \gamma_Y(h),$$

ou seja, que para fenômenos com entornos pequenos,  $|\mathbf{h}| \approx 0$ , temos que  $\gamma_Z(h) \approx \gamma_Y(h)$ , isto quer dizer, o semivariograma da parte aleatória é, aproximadamente, o semivariograma relativo à função aleatória  $Z$ .

**Definição 5.2.1:** Define-se o “drift” de uma função aleatória como sendo o valor:  $E[Z(\mathbf{x})] = m$ .

Pode-se ver na Figura 12, que se encontra em OLEA (1999), o significado do “drift” numa variável regionalizada.

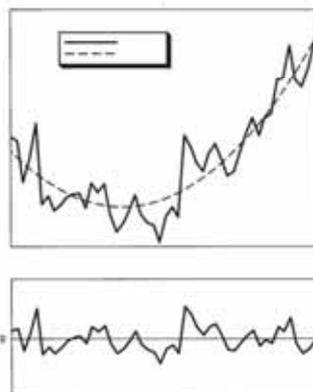


Figura 12 - Decomposição da variável regionalizada.

Na Figura 12 a linha contínua representa a Variável a ser estudada e a linha tracejada o seu “Drift”. Na figura de cima temos a variável em estudo e seu “drift” e na figura de baixo temos o Resíduo depois da subtração do “drift” da Variável em estudo.

**Definição 5.2.2:** Define-se o resíduo  $Y(\mathbf{x})$  de uma função aleatória como:  $Y(\mathbf{x}) = Z(\mathbf{x}) - E[Z(\mathbf{x})]$

Por hipótese,  $Y(\mathbf{x})$  honra a hipótese intrínseca, ou seja:

1.  $E[Y(\mathbf{x})]$  é constante
2.  $Var[Y(\mathbf{x}) - Y(\mathbf{x} + \mathbf{h})] = 2\gamma_Y(\mathbf{h})$

**Resultado 5.2.1:** A esperança de  $Y(\mathbf{x})$  é nula.

$$E[Y(\mathbf{x})] = E\{Z(\mathbf{x}) - E[Z(\mathbf{x})]\} \quad (50)$$

$$= E[Z(\mathbf{x})] - E\{E[Z(\mathbf{x})]\} \quad (51)$$

$$= E[Z(\mathbf{x})] - E[Z(\mathbf{x})] = 0 \quad (52)$$

**Definição 5.2.3:** Definimos o “drift”  $m(\mathbf{x})$  como sendo uma combinação linear de  $n$  funções analíticas, ou seja:

$$m(\mathbf{x}) = \sum_{l=0}^n a_l f_l(\mathbf{x}),$$

com  $f_0(\mathbf{x}) = 1$ .

**Resultado 5.2.2** Sejam  $\hat{Z}(\mathbf{x}) = \sum_{i=1}^k \lambda_i Z(\mathbf{x}_i)$  o estimador de Krigagem Universal e  $m(\mathbf{x})$  o “drift” como definido no Definição 5.2.3. Então:

$$E[\hat{Z}(\mathbf{x})] = \sum_{l=0}^n \sum_{i=1}^k a_l \lambda_i f_l(\mathbf{x}_i)$$

Segue a demonstração:

$$E[\hat{Z}(\mathbf{x})] = E \left[ \sum_{i=1}^k \lambda_i Z(\mathbf{x}_i) \right] \quad (53)$$

$$= \sum_{i=1}^k \lambda_i E[Z(\mathbf{x}_i)] \quad (54)$$

$$= \sum_{i=1}^k \lambda_i \sum_{l=0}^n a_l f_l(\mathbf{x}_i) \quad (55)$$

$$= \sum_{l=0}^n \sum_{i=1}^k a_l \lambda_i f_l(\mathbf{x}_i) \quad (56)$$

**Resultado 5.2.3** Sejam  $\hat{Z}(\mathbf{x})$  e  $m(\mathbf{x})$  o “drift” da Definição 5.2.3. Este estimador será não viciado se, e somente se:

$$\sum_{i=1}^k \lambda_i f_l(\mathbf{x}_i) = f_l(\mathbf{x}_0),$$

para  $l = 0, 1, \dots, n$ .

Para se demonstrar, matematicamente, que um estimador é não viciado, deve-se mostrar que vale a relação:

$$E[\hat{Z}(\mathbf{x}) - Z(\mathbf{x})] = 0$$

( $\Rightarrow$ )

Vamos mostrar a necessidade:

Suponhamos que o estimador é não-viciado. Então:

$$E[\hat{Z}(\mathbf{x}_0)] = Z[(\mathbf{x}_0)]$$

Pelo Resultado 5.2.2, temos:

$$\sum_{l=0}^n \sum_{i=1}^k a_l \lambda_i f_l(\mathbf{x}_i) = \sum_{l=0}^n a_l f_l(\mathbf{x}_0)$$

Isto pode ser escrito como:

$$\sum_{l=0}^n a_l \sum_{i=1}^k \lambda_i f_l(\mathbf{x}_i) = \sum_{l=0}^n a_l f_l(\mathbf{x}_0)$$

O que é verdade se:

$$\sum_{i=1}^k \lambda_i f_l(\mathbf{x}_i) = f_l(\mathbf{x}_0),$$

para  $l = 0, 1, \dots, n$ .

( $\Leftarrow$ )

A demonstração da suficiência segue tomando-se o caminho inverso da demonstração acima.

Note-se que a condição  $\sum_{i=1}^k \lambda_i = 1$  é apenas uma das restrições, isto quando  $l=0$ . Para a Krigagem Universal esta restrição é apenas uma das restrições. Por exemplo, se o “drift” é linear, ou seja, se a função linear que a representa é linear, tem-se,  $l=3$ :  $f_1(\mathbf{x}) = 1$ ,  $f_2(\mathbf{x}) = x$  e  $f_3(\mathbf{x}) = y$ . Ou seja, é equivalente a:  $\sum_{i=1}^k \lambda_i = 1$ ,  $\sum_{i=1}^k \lambda_i x_i = x$  e  $\sum_{i=1}^k \lambda_i y_i = y$ .

**Resultado 5.2.4** Sejam  $\hat{Z}(\mathbf{x}_0)$  o estimador não viciado de Krigagem Universal,  $\lambda_i$  os pesos dados na definição de  $\hat{Z}(\mathbf{x}_0)$ . Se  $Y(\mathbf{x})$  é o resíduo definido na Definição 5.2.2, então vale:

$$\hat{Z}(\mathbf{x}_0) - Z(\mathbf{x}_0) = \sum_{i=1}^k \lambda_i Y(\mathbf{x}_i) - Y(\mathbf{x}_0)$$

Rescrevendo  $\hat{Z}(\mathbf{x}_0) - Z(\mathbf{x}_0)$  de acordo com a definição 5.2.2 e  $\hat{Z}(\mathbf{x}_0) = \sum_{i=1}^k \lambda_i Z(\mathbf{x}_i)$ :

$$\hat{Z}(\mathbf{x}_0) - Z(\mathbf{x}_0) = \sum_{i=1}^k \lambda_i \{Y(\mathbf{x}_i) + E[Z(\mathbf{x}_i)]\} - \{Y(\mathbf{x}_0) + E[Z(\mathbf{x}_0)]\} \quad (57)$$

$$= \sum_{i=1}^k \lambda_i Y(\mathbf{x}_i) - Y(\mathbf{x}_0) + \left\{ E \left[ \sum_{i=1}^k \lambda_i Z(\mathbf{x}_i) \right] - E[Z(\mathbf{x}_0)] \right\} \quad (58)$$

$$= \sum_{i=1}^k \lambda_i Y(\mathbf{x}_i) - Y(\mathbf{x}_0) + E[\hat{Z}(\mathbf{x}_0) - Z(\mathbf{x}_0)] \quad (59)$$

$$= \sum_{i=1}^k \lambda_i Y(\mathbf{x}_i) - Y(\mathbf{x}_0) \quad (60)$$

**Resultado 5.2.5** Sejam  $\hat{Z}(\mathbf{x}_0)$  o estimador não viciado de Krigagem Universal, e seja  $\gamma_Y(\mathbf{h})$  o semivariograma dos Resíduos. Então a estimação da variância  $\sigma^2(\mathbf{x}_0)$  é igual a:

$$2 \sum_{i=1}^k \lambda_i \gamma_Y(\mathbf{x}_0, \mathbf{x}_i) - \sum_{i=1}^k \sum_{j=1}^k \lambda_i \lambda_j \gamma_Y(\mathbf{x}_i, \mathbf{x}_j)$$

Pela definição de  $\sigma^2(\mathbf{x}_0) = Var[\hat{Z}(\mathbf{x}_0) - Z(\mathbf{x}_0)]$ .

Usando o Resultado 5.2.4:

$$\hat{Z}(\mathbf{x}_0) - Z(\mathbf{x}_0) = \sum_{i=1}^k \lambda_i Y(\mathbf{x}_i) - Y(\mathbf{x}_0)$$

Daí, os passos são os mesmos utilizados na seção 5.1 à partir da equação (36).

**Definição 5.2.4** Sejam  $\sigma^2(\mathbf{x}_0)$ ,  $\lambda_i$  os pesos relacionados ao estimador da Krigagem Universal, e  $\mu_l$  os multiplicadores de Lagrange. A função objetivo para a Krigagem Universal é:

$$L(\lambda_i, \mu_l) = \sigma^2(\mathbf{x}_0) + 2\mu_0 \left( \sum_{i=1}^k \lambda_i - 1 \right) + 2 \sum_{l=1}^n \mu_l \left( \sum_{i=1}^k \lambda_i f_l(\mathbf{x}_i) - f_l(\mathbf{x}_0) \right)$$

Fazendo-se as derivadas:

- $\frac{\partial L}{\partial \lambda_i} = 2\lambda_Y(\mathbf{x}_0, \mathbf{x}_i) - 2 \sum_{j=1}^k \lambda_j \gamma_Y(\mathbf{x}_i, \mathbf{x}_j) + 2\mu_0 + 2 \sum_{l=1}^n \mu_l f_l(\mathbf{x}_i) = 0$ , para  $i = 1, \dots, k$   
e

- $\frac{\partial L}{\partial \mu_0} = 2 \left( \sum_{i=1}^k \lambda_i - 1 \right) = 0$
- $\frac{\partial L}{\partial \mu_l} = 2 \left( \sum_{i=1}^k \lambda_i f_l(\mathbf{x}_i) - f_l(\mathbf{x}_0) \right) = 0$

E isso conduz ao seguinte sistema de equações:

$$\left\{ \begin{array}{l} \sum_{i=1}^k \lambda_i \gamma_Y(\mathbf{x}_i, \mathbf{x}_1) - \mu_0 - \sum_{l=1}^n \mu_l f_l(\mathbf{x}_1) = \gamma_Y(\mathbf{x}_0, \mathbf{x}_1) \\ \sum_{i=1}^k \lambda_i \gamma_Y(\mathbf{x}_i, \mathbf{x}_2) - \mu_0 - \sum_{l=1}^n \mu_l f_l(\mathbf{x}_2) = \gamma_Y(\mathbf{x}_0, \mathbf{x}_2) \\ \dots\dots\dots \\ \sum_{i=1}^k \lambda_i \gamma_Y(\mathbf{x}_i, \mathbf{x}_k) - \mu_0 - \sum_{l=1}^n \mu_l f_l(\mathbf{x}_k) = \gamma_Y(\mathbf{x}_0, \mathbf{x}_k) \\ \sum_{i=1}^k \lambda_i = 1 \\ \sum_{i=1}^k \lambda_i f_1(\mathbf{x}_i) = f_1(\mathbf{x}_0) \\ \sum_{i=1}^k \lambda_i f_2(\mathbf{x}_i) = f_2(\mathbf{x}_0) \\ \dots\dots\dots \\ \sum_{i=1}^k \lambda_i f_n(\mathbf{x}_i) = f_n(\mathbf{x}_0) \end{array} \right.$$

Resolvendo-se o sistema acima e encontra-se os valores de  $\lambda_i$  e  $\mu_l$  e a expressão da variância da Krigagem Universal é dada por:

$$\sigma^2(\mathbf{x}_0) = \sum_{i=1}^k \lambda_i \gamma_Y(\mathbf{x}_0, \mathbf{x}_i) - \mu_0 - \sum_{l=1}^n \mu_l f_l(\mathbf{x}_0)$$

Deve-se levar em conta que, embora esteja considerando no sistema  $k$  pontos da região em estudo, em geral toma-se apenas  $k' \leq k$  que são o entorno do ponto  $\mathbf{x}_0$ .

Podemos escrever o sistema acima matricialmente como:

$$A = \begin{pmatrix} \gamma_Y(\mathbf{x}_1, \mathbf{x}_1) & \gamma_Y(\mathbf{x}_2, \mathbf{x}_1) & \cdots & \gamma_Y(\mathbf{x}_k, \mathbf{x}_1) & 1 & f_1(\mathbf{x}_1) & f_2(\mathbf{x}_1) & \cdots & f_n(\mathbf{x}_1) \\ \gamma_Y(\mathbf{x}_1, \mathbf{x}_2) & \gamma_Y(\mathbf{x}_2, \mathbf{x}_2) & \cdots & \gamma_Y(\mathbf{x}_k, \mathbf{x}_2) & 1 & f_1(\mathbf{x}_2) & f_2(\mathbf{x}_2) & \cdots & f_n(\mathbf{x}_2) \\ \cdots & \cdots \\ \gamma_Y(\mathbf{x}_1, \mathbf{x}_k) & \gamma_Y(\mathbf{x}_2, \mathbf{x}_k) & \cdots & \gamma_Y(\mathbf{x}_k, \mathbf{x}_k) & 1 & f_1(\mathbf{x}_k) & f_2(\mathbf{x}_k) & \cdots & f_n(\mathbf{x}_k) \\ 1 & 1 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 \\ f_1(\mathbf{x}_1) & f_1(\mathbf{x}_2) & \cdots & f_1(\mathbf{x}_k) & 0 & 0 & 0 & \cdots & 0 \\ f_2(\mathbf{x}_1) & f_2(\mathbf{x}_2) & \cdots & f_2(\mathbf{x}_k) & 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots \\ f_n(\mathbf{x}_1) & f_n(\mathbf{x}_2) & \cdots & f_n(\mathbf{x}_k) & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$X = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \\ \mu_0 \\ \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} \text{ e } \mathbf{a} = \begin{pmatrix} \gamma_Y(\mathbf{x}_0, \mathbf{x}_1) \\ \gamma_Y(\mathbf{x}_0, \mathbf{x}_2) \\ \vdots \\ \gamma_Y(\mathbf{x}_0, \mathbf{x}_k) \\ 1 \\ f_1(\mathbf{x}_0) \\ f_2(\mathbf{x}_0) \\ \vdots \\ f_n(\mathbf{x}_0) \end{pmatrix}. \text{ E daí deve-se resolver } \mathbf{A}\mathbf{X} = \mathbf{a},$$

ou seja,  $\mathbf{X} = \mathbf{A}^{-1}\mathbf{a}$ .

Este sistema é simétrico. A existência e a unicidade é garantida pela regularidade da matriz. Para isto basta assegurar-se que as funções  $f_i$  sejam linearmente independentes no conjunto de dados. Isto se satisfaz normalmente já que o número de amostras é maior que o número de funções que ajustam o “drift”. Além disso, em geral, a malha de amostragem é irregular e portanto não há a possibilidade que as funções sejam linearmente dependentes. Problemas podem surgir se a malha for regular, situação que antes se desejava anteriormente, na Krigagem Ordinária.

### 5.2.2 Krigagem Local e Global

Quando se diz que foi feita uma Krigagem local, quer se dizer que a estimação para cada ponto não amostrado foi feita utilizando-se pontos do entorno deste ponto.

Já a estimação global é feita utilizando-se todos os pontos para estimar cada ponto não amostrado.

### 5.2.3 Vantagens e Desvantagem da Krigagem Universal

Deve-se notar, antes de mais nada, que a Krigagem Ordinária é um caso particular da Krigagem Universal, bastando apenas supor que  $Z(\mathbf{x})$  é uma variável não-intrínseca com  $k = 1$ ,  $a_l = 1$ ,  $f_l(\mathbf{x}) = 1$ . Por isso, Krigagem Universal compartilha com a Krigagem Ordinária as suas boas propriedades: é um estimador exato, não enviesado e que produz o menor erro quadrático. O método é ainda bastante robusto para parâmetros mal especificados (OLEA, 1999). Pode-se notar que embora o “drift” deva ser especificado isto é feito automaticamente através das equações da Krigagem Universal.

Os maiores problemas da Krigagem Universal são:

- Especificar o semivariograma dos Resíduos que não é o mesmo que os dos dados originais
- O problema da seleção do “drift”. Neste caso, a decisão cabe ao usuário e é mais relevante em extrapolações para além da envoltória convexa do espaço amostral. Não há uma estatística direta para guiar a busca do drift. Em alguns casos a escolha do “drift” é clara, no entanto, em outras situações, a experiência do profissional de Geoestatística será fundamental.

No caso do primeiro problema, há algumas maneiras de se estimar semivariogramas que sirva para propósitos acima (CALVETE & RAMÍREZ, 1996). Estes autores sugerem que se faça:

1. Supor que  $\gamma_Y \approx \gamma_Z$ . Isto está justificado para pequenos entornos. Um extensão disto é empregar um semivariograma linear com tangente passando pela origem igual a do semivariograma experimental de  $Z(\mathbf{x})$ . Este tipo de enfoque é um dos mais usados na prática.
2. Calcular o semivariograma na direção em que  $Z(\mathbf{x})$  seja estacionária. Se uma direção não tem tendência, pode-se usá-la para todas as direções e supondo-se que o fenómeno é isotrópico. Este enfoque está limitado às suposições de que existe uma direção em que não haja tendência e que existam dados suficientes nesta direção para que seja possível estimar o semivariograma. Por outro lado, a hipótese da isotropia do semivariograma e da sua estacionariedade é difícil de comprovar.
3. Calcular o semivariograma na região em que  $Z(\mathbf{x})$  é intrínseca em uma sub-região. Daí calcula-se o semivariograma somente em todos os dados nos pontos pertencentes ao local aonde não existe tendência. Esta situação se dá, frequentemente, nas regiões do contorno da região estudada.
4. Calcular o semivariograma por validação cruzada.

Pode-se realizar a Krigagem Universal através de um método indireto conhecido como Krigagem dos Resíduos. Ajusta-se um polinômio aos dados e obtêm-se Resíduos. Posteriormente irá se aplicar a Krigagem Ordinária à estes Resíduos. Deve-se enfatizar que a Krigagem Universal não é equivalente a este processo.

Esta metodologia é criticada por ARMSTRONG (1984) no trabalho intitulado "*Problems with Universal Kriging*". Ela relata dois problemas fundamentais:

- O semivariograma dos Resíduos é estimado com viés (uma sobrestimação) do verdadeiro semivariograma da variável regionalizada.
- Em segundo lugar é extremamente difícil deduzir o grau do polinômio que aproxima o "drift" ou o tipo do semivariograma da variável regionalizada.

Diante dessas dificuldades ela conclui que embora a teoria da Krigagem Universal seja matematicamente correta, a dificuldade em estimar o “drift” e o semivariograma ao mesmo tempo torna este método de difícil aplicação na prática.

Discordando desta autora, CALVETE & RAMÍREZ (1996) argumentam que isto seria mais uma insatisfação de expectativas já que, segundo estes autores, estas questões são coerentes com a hipótese do método.

Além disto, dizem que se for aumentado o número de pontos no entorno do ponto que será estimado ou utilizando-se, diretamente, a Krigagem global, isto poderá melhorar a estimação.

A Krigagem Universal não é tão popular como a Krigagem Ordinária. Alguns motivos são:

- Para a predição os resultados são similares com os Krigagem Ordinária para vizinhanças locais (pequenos entornos).
- A estimação pode ser perigosa para extrapolação.
- É praticamente impossível checar a qualidade do ajuste da tendência se modelos locais forem empregados.

#### 5.2.4 Método da Krigagem dos Resíduos: Método Direto

Para se realizar a Krigagem dos Resíduos deve-se supor conhecida o “drift”  $m(\mathbf{x})$ . Neste caso será feita uma estimação de  $m(x)$  ou seja, deve-se encontrar  $\hat{m}(\mathbf{x})$ . Os Resíduos se calculam segundo:

$$R(\mathbf{x}) = Z(\mathbf{x}) - \hat{m}(\mathbf{x}), \quad (61)$$

que é uma aproximação da componente estocástica apresentada no início da seção 5.2.1 e cujo semivariograma é  $\gamma_Y(h)$ . Nesta metodologia, será calculado o semivariograma  $\gamma_R(h)$  e depois será aplicado a Krigagem Ordinária utilizando-se este semivariograma.

O método da Krigagem dos Resíduos pelo método direto é feito da seguinte maneira:

1. Supõe-se conhecida a forma da tendência com base em conhecimentos do pesquisador sobre o fenômeno físico.
2. Ajusta-se o “drift”  $\hat{m}(x)$  por mínimo quadrados ordinários.
3. Calcula-se os Resíduos mediante a diferença  $\mathbf{R}(\mathbf{x}) = \mathbf{Z}(\mathbf{x}) - \hat{\mathbf{m}}(\mathbf{x})$
4. Calcula-se o semivariograma experimental ajustando os Resíduos obtidos no passo anterior.
5. Realiza-se a Krigagem Ordinária dos Resíduos utilizando o semivariograma  $\gamma_R(\mathbf{h})$ .
6. Obter o valor da variável  $Z^*(\mathbf{x}_0)$  em um ponto amostrado  $\mathbf{x}_0$ ) como a soma do valor do “drift”  $\hat{m}(x)$  calculada em 2 mais o valor da Krigagem dos Resíduos  $R^*(\mathbf{x}_0)$  obtida no passo 5.

Embora este método seja um dos mais utilizados na prática nota-se uma inconsistência interna já que a estimação do “drift” é feita através de mínimos quadrados ordinários que tem, como pressuposição metodológica, a independência entre as observações, o que não é caso.

### 5.2.5 Método da Krigagem dos Resíduos: Iterativo

Este método consiste basicamente em ajustar o “drift” por mínimos quadrados, calcular o semivariograma dos Resíduos e repetir estes passos iterativamente. Ao se chegar no fim do processo, aplicar a Krigagem Ordinária.

Para se ajustar o “drift” primeiro deve-se indicar o grau do polinômio. Para isto ajusta-se os polinômios aos dados em grau crescente até obter um semivariograma dos Resíduos aproximadamente isótropo e estacionário.

Sejam  $\mathbf{Z}$ ,  $\mathbf{m}$  e  $\boldsymbol{\epsilon}$  os verdadeiros valores da variável em estudo, do “drift” e da componente estocástica respectivamente. Sejam  $\hat{\mathbf{Z}}$ ,  $\hat{\mathbf{m}}$  e  $\mathbf{R}$  os valores estimados utilizando-se mínimos quadrados. Sejam ainda  $\mathbf{V}_R$  e  $\mathbf{V}_\epsilon$  as matrizes de covariância de  $\mathbf{R}$  e  $\epsilon$  respectivamente. Os valores de  $\mathbf{V}_{R_{ij}}$  e  $\mathbf{V}_{\epsilon_{ij}}$  são dados por:

$$\mathbf{V}_{R_{ij}} = \sigma_R^2 - \gamma(\mathbf{x}_i - \mathbf{x}_j) \quad (62)$$

$$\mathbf{V}_{\epsilon_{ij}} = \sigma_\epsilon^2 - \gamma(\mathbf{x}_i - \mathbf{x}_j) \quad (63)$$

aonde  $\sigma_R^2$  e  $\sigma_\epsilon^2$  são as variâncias de  $R(\mathbf{x})$  e  $\epsilon(\mathbf{x})$  respectivamente e  $\gamma_R(\mathbf{h})$  e  $\gamma_\epsilon(\mathbf{h})$  seus semivariogramas. Sendo o drift como expresso na seção 5.2.3 e  $\mathbf{X}_{ij} = f_i(x_j)$  e denotando por  $\mathbf{a}$  o vetor de incógnitas  $a_i$ , podemos escrever  $m(\mathbf{x})$  como:

$$\mathbf{Z} = \mathbf{X}\mathbf{a} + \boldsymbol{\epsilon} \quad (64)$$

Se  $\mathbf{V}_\epsilon$  for conhecida, para encontrar um estimador não viciado de  $\hat{\mathbf{a}}$ , podemos minimizar a expressão usando o critério de mínimos quadrados:

$$(\mathbf{Z} - \mathbf{X}\mathbf{a})'\mathbf{V}_\epsilon^{-1}(\mathbf{Z} - \mathbf{X}\mathbf{a}) \quad (65)$$

O estimador de  $\mathbf{a}$ ,  $\hat{\mathbf{a}}$  resultante é dado por:

$$\hat{\mathbf{a}} = \mathbf{V}_a\mathbf{X}'\mathbf{V}_\epsilon^{-1}\mathbf{Z} \quad (66)$$

em que  $\mathbf{V}_a = (\mathbf{X}'\mathbf{V}_\epsilon^{-1}\mathbf{X})^{-1}$ . A partir desta equação obtém-se  $\mathbf{R}$ , pois  $\hat{\mathbf{m}} = \mathbf{X}\hat{\mathbf{a}}$  e  $\hat{\mathbf{R}} = \mathbf{Z} - \hat{\mathbf{m}}$ . A matriz de covariância dos Resíduos estimados  $\hat{\mathbf{R}}$  é dada por (NEUMAN & JACOBSON, 1984 apud CALVETE & RAMIREZ, 1996):

$$\mathbf{V} = (\mathbf{I} - \mathbf{P})\mathbf{V}_\epsilon \quad (67)$$

em que  $\mathbf{I}$  é a matriz identidade e  $\mathbf{P}$  é igual a  $\mathbf{P} = \mathbf{X}\mathbf{V}_a\mathbf{X}'\mathbf{V}_\epsilon^{-1}$

Note-se que  $\mathbf{V}_\epsilon - \mathbf{V}_R = \mathbf{P}\mathbf{V}_\epsilon = \mathbf{X}\mathbf{V}_a\mathbf{X}'$ , ou seja, a diferença entre as matrizes de covariância é tanto maior quanto maior o número de dados com relação aos parâmetros, portanto reduz a incerteza de  $\hat{\mathbf{a}}$  e por consequência  $\mathbf{V}_a$ .

Tem-se que  $\mathbf{V}_\epsilon \geq \mathbf{V}_R$  porque a diferença entre essas matrizes é definida positiva. O patamar de  $\gamma_\epsilon$  será um sempre um pouco maior que  $\gamma_R$ .

Como nesta fase desconhece-se  $V_\epsilon$  utiliza-se a matriz identidade para se começar a iteração.

A segunda fase começa calculando-se  $V_R$ . Com esta matriz de covariância ajusta-se um polinômio de grau definido na fase 1. Isto dará lugar a novos Resíduos e a um novo semivariograma  $\gamma_R$ . Com este último torna-se a calcular  $V_R$  até que o processo convirja e sejam calculados  $\hat{m}(x)$  e o semivariograma. Em geral, segundo CALVETE & RAMÍREZ (1996), são necessárias poucas iterações, cerca de três ou quatro, para que este processo convirja. Uma vez conhecido o “drift” e o semivariograma, aplica-se a Krigagem Ordinária e obtém o valor de  $Z^*(\mathbf{x}_0)$  em cada ponto somando-se o valor do drift  $\hat{m}(\mathbf{x}_0)$  mais o valor da Krigagem dos Resíduos como é feito no método direto.

## 6 APLICAÇÕES A DOIS CASOS DA LITERATURA

### 6.1 Material

Serão analisados dois conjuntos de dados. O primeiro conjunto (ANEXO 1), que se encontra em OLEA (1999), são relativos ao Aquífero High Plains, que se localiza no Estado do Kansas, EUA, datando do período “Terciário-Quaternário, constituído por areias aluviais e eólicas associadas a depósitos de preenchimento de vales de uma drenagem que prossegue para Leste a partir das Montanhas Rochosas” (LANDIM et al., 2002). A variável a ser analisada será a WTable que é a cota do nível hidrostático do aquífero. As coordenadas (X,Y) dos pontos amostrais são dadas em milhas.

Já o segundo conjunto de dados (ANEXO 2), que se encontra em

CLARK & HARPER (2000) é proveniente do Aquífero Wolfcamp, no estado do Texas, EUA. O Departamento de Energia dos EUA estava interessado em estudar altos níveis de resíduos nucleares em sal, basalto e tufo. Um dos lugares estudados foi a divisa entre o nordeste do Texas, o Novo México e Oklahoma. O Aquífero Wolfcamp, localizado nesta região, apresentava-se como um possível depósito de radionuclídeos e isto poderia possibilitar que eles se espalhassem contaminando outros lugares (HARPER & CLARK, 2010).

A variável a ser analisada neste conjunto de dados será o Nível Piezométrico. As coordenadas (X,Y) dos pontos amostrais são dadas em milhas.

Utilizar-se-á, para as análises estatísticas e geoestatísticas, o Software Surfer 8.0 e o Software R 3.1.1.

## **6.2 Método**

Para analisar os dados far-se-á uma análise descritiva dos dados e posteriormente utilizar-se-á a Krigagem Universal por meio da Krigagem dos Resíduos (Método Direto) para criar uma Mapa de valores estimados para a cota do nível hidrostático (WTable) no caso dos dados provenientes do Aquífero High Plains e do Nível Piezométrico no caso dos dados provenientes do Aquífero Wolfcamp.

## **6.3 Resultados e Discussões**

### **6.3.1 Dados do Aquífero High Plains**

Os primeiros dados provenientes do Aquífero High Plains, possui o seguinte Mapa de pontos amostrados exibido na Figura 13:

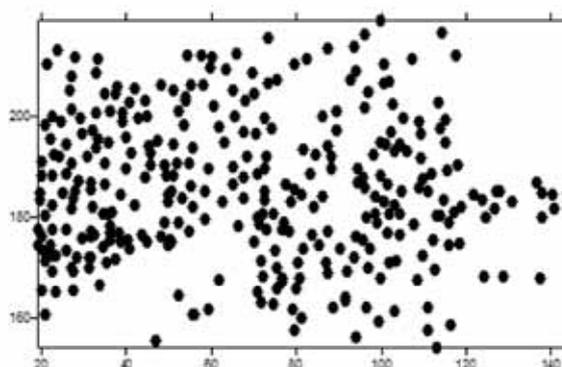


Figura 13 - Localização espacial dos 327 amostras do Aquífero High Plains.

Por meio do Software R 3.1.1, obteve-se as estatísticas descritivas da variável WTable resumidas na Tabela 3:

Tabela 3: Estatísticas Descritivas da variável WTable

Número de dados	327
Mínimo	2.202,84
Máximo	3856,71
1º Quartil	2.698,78
3º Quartil	3419,75
Média	3062,86
Mediana	3050,14
Desvio-padrão	416,17
Coef. de Assimetria	-0.04
Curtose	-1,11

Encontrou-se o Histograma dos Dados e Gráfico Quantil-Quantil para Distribuição Normal e o que estão exibidos nas Figuras 14 e 15:

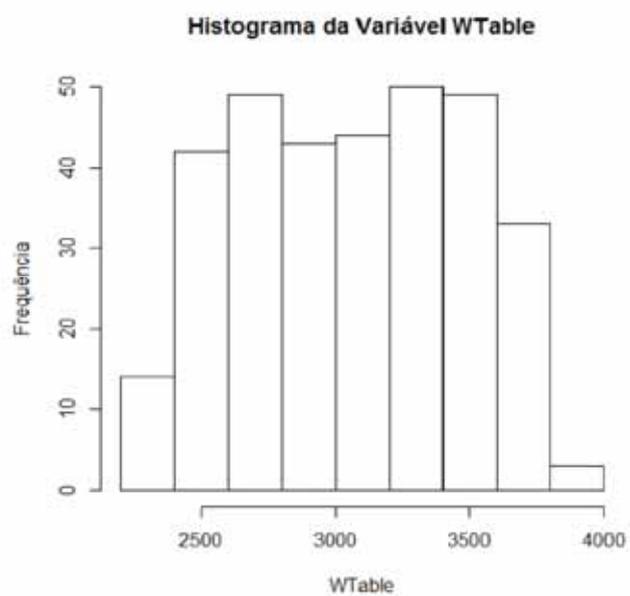


Figura 14 - Histograma da Variável WTable.

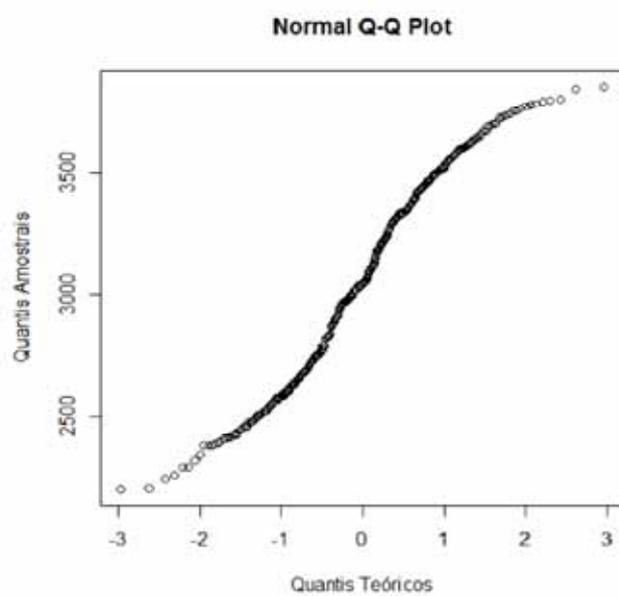


Figura 15 - Gráfico Quantil-Quantil para Distribuição Normal.

Através do teste de Shapiro-Wilks obteve-se  $W = 0,9679$  ( $p \leq 0,01$ ). Com isso concluiu-se que os dados não apresentam normalidade (5 % de significância) no entanto apresentam boa simetria (Coef. Simetria=-0,04). Em geral, deseja-se que os dados apresentem normalidade ou que se faça alguma transformação a fim de se obter esta propriedade, no entanto, a Krigagem é um método robusto para dados que não atendam esta condição como se pode ler em BREGT et al. (1991) e (YOST et al., 1982).

Foi obtida uma visualização da distribuição dos valores da Variável WTable, com relação à sua localização geográfica, através do Software Surfer 8, exibida na Figura 16, a fim de se visualizar a presença ou não de características que indiquem tendência nos dados.

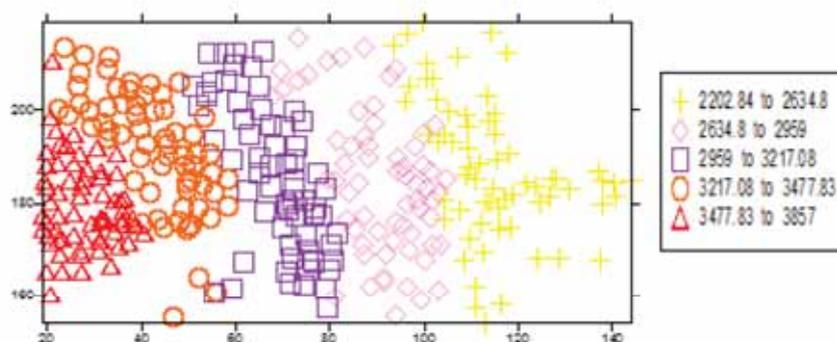


Figura 16 - Distribuição dos Valores da variável WTable por sua localização geográfica

Pode-se observar que há um aumento dos valores seguindo uma direção que vai da esquerda para direita no Mapa. Isto é uma provável indicação de tendência nos dados.

Ao se calcular o semivariograma dos dados nas direções  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  e  $135^\circ$  com tolerância angular igual a  $22,5^\circ$ , utilizando-se o software Surfer 8.0, obteve-se os seguintes gráficos exibidos nas Figuras 17, 18, 19 e 20, respectivamente:

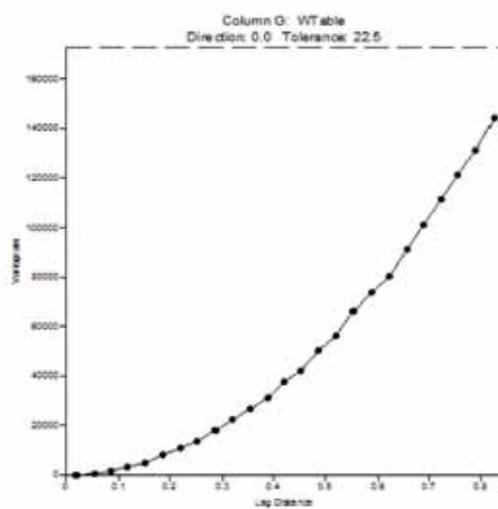


Figura 17 - Semivariograma para direção 0°.

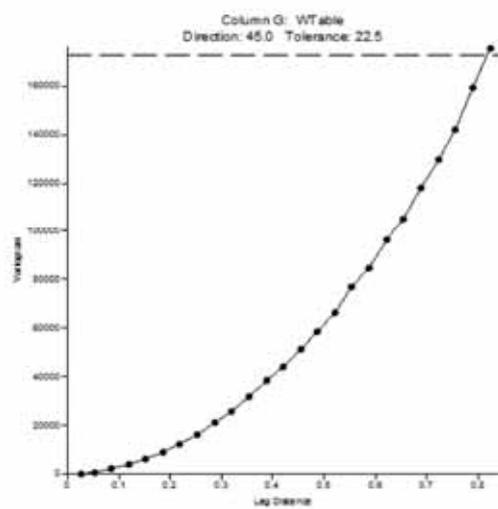


Figura 18 - Semivariograma para direção 45°.

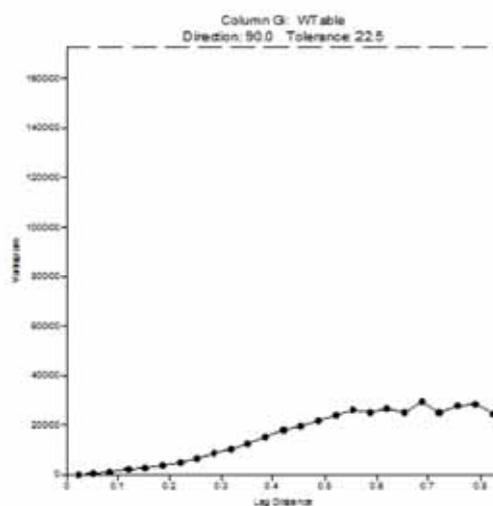


Figura 19 - Semivariograma para direção 90°.

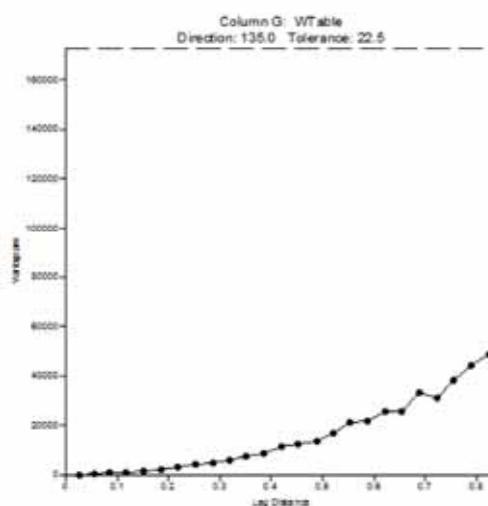


Figura 20 - Semivariograma para direção 135°.

Observa-se a presença de tendência para todas as direções, exceto a direção 90°, apresentada na Figura 19. No entanto, para esta direção, o semivariograma se estabiliza muito abaixo do valor da variância dos dados, indicada pela linha tracejada.

Constatando-se que os semivariogramas não se estabilizavam em nenhum valor, ou seja, que o fenômeno estudado possui tendência, optou-se por realizar a Krigagem Universal por meio da Krigagem dos Resíduos utilizando-se o método direto. Para tanto, fez-se necessário realizar um ajuste polinomial aos dados como exige esta técnica.

CLARK & HARPER (2000) sugerem como método de investigação para se obter o grau do polinômio que melhor se ajustará aos dados, a análise de modelos, através da Análise de Variância (ANOVA). Estes autores, no entanto, frisam que este método é apenas para auxiliar na escolha de qual polinômio utilizar, já que uma pressuposição da ANOVA, a não-correlação entre os dados, não é atendida e, além disso, os dados podem não ter Distribuição Normal.

Para a comparação entre os modelos utilizou-se o Software R. O primeiro modelo proposto (Modelo 1) foi o Polinomial Quadrático, ou seja,  $m_1(x, y) = aX^2 + bY^2 + cX + dY + e$  e o segundo modelo proposto (Modelo 2) foi o Polinomial Linear, isto é:  $m_2(x, y) = fX + gY + h$ . Os valores obtidos para cada modelo estão exibidos nas Tabelas 4 e 5.

Tabela 4: Valores dos coeficientes do Polinômio Quadrático e seus respectivos valores-p.

Coefficientes	Estimado	Erro-padrão	Valor de T	Valor-p
Intercepto	2325,00	307,5	7,559	$4,9 \times 10^{-13}$
X	-21,56	0,901	-23,929	$\leq 2 \times 10^{-16}$
Y	-24,74	3,224	7,673	$2,04 \times 10^{-13}$
X <sup>2</sup>	0,00959	0,001963	4,886	$1,63 \times 10^{-6}$
Y <sup>2</sup>	-0,08537	0,008536	-10,0	$\leq 2 \times 10^{-16}$
XY	0,04155	0,004576	9,081	$\leq 2 \times 10^{-16}$
			<i>R</i> <sup>2</sup>	0,9903

Tabela 5: Valores dos coeficientes do Polinômio Linear e seus respectivos valores-p.

Coeficientes	Estimado	Erro-padrão	Valor de T	Valor-p
Intercepto	4714,648	34,07	138,38	$\leq 2 \times 10^{-16}$
X	-12,561	0,077	-162,22	$\leq 2 \times 10^{-16}$
Y	-4,259	0,179	-23,78	$\leq 2 \times 10^{-16}$
			$R^2$	0,9879

Constatou-se que todos os coeficientes do modelo linear e quadrático foram significativos.

Procedeu-se à comparação entre os modelos, cujo resultado se encontra exibido na Tabela 6:

Tabela 6: Comparação, através da ANOVA, dos Modelos 1 e 2.

Modelo	GL Resíduo	SQR	GL.	SQ	F	Valor-p
Modelo 1	321	394914				
Modelo 2	324	681615	-3	-286701	77,68	$\leq 2 \times 10^{-16}$

Obtém-se, da comparação do Modelo 2 com o Modelo 1 um valor- $p \leq 0,05$  que indica que o Modelo 1 (Polinômio Quadrático) é preferível ao Modelo 2 (Polinômio Linear).

É de grande valia, para a tomada de decisão, observar as superfícies obtidas com o ajuste dos Polinômios, sobrepostas aos dados separados exibidos na Figura 16.

Visualmente, a sensação é de que realmente o Modelo 1 se encaixa melhor aos dados que o Modelo 2, como está exibido nas Figuras 21 e 22.

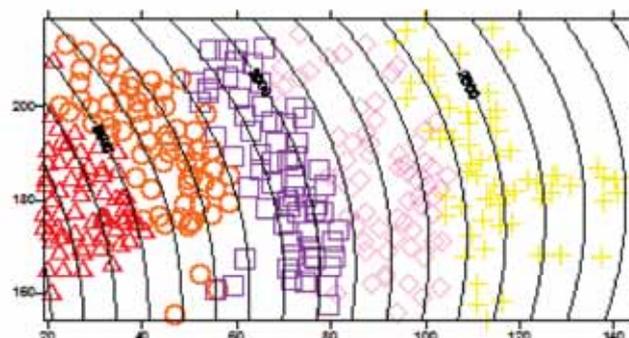


Figura 21 - Modelo 1 sobreposto aos dados do Aquífero High Plains.

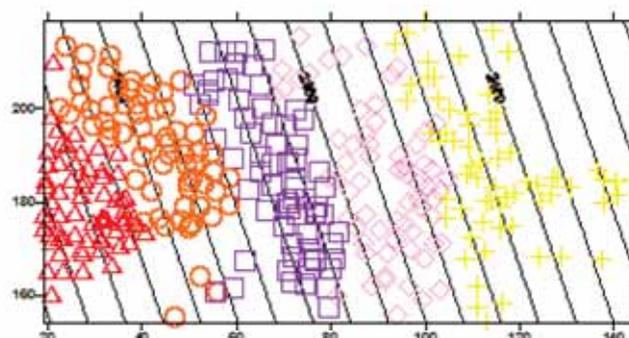


Figura 22 - Modelo 2 sobreposto aos dados do Aquífero High Plains.

Embora já tenha sido feita as observações à respeito da pressuposições para aplicação da técnica de ANOVA, considerar-se-á os  $R^2$  como tendo expressividade, ou seja, como se fosse numa ANOVA onde as pressuposições fossem atendidas. Dessa consideração, obtém-se que o coeficiente de determinação do Modelo 1,  $R_1^2 = 99,03$ , ou seja, explica 99,03% da variabilidade dos dados enquanto o coeficiente de determinação do Modelo 2,  $R_2^2 = 98,79$ , ou seja, explica 98,79% da variabilidade dos dados. Como a diferença é muito pequena optou-se por fazer duas Krigagens, uma para cada modelo.

Considerando o Modelo 1, para o Polinômio:  $m_1(X, Y) = 2325 - 21,56X - 24,74Y + 0,04255XY + 0,00959X^2 - 0,08537Y^2$  obteve-se o Mapa dos valores na Figura 23:

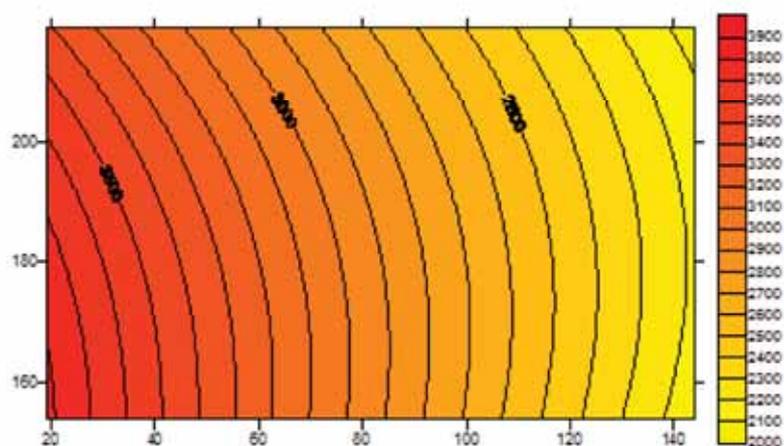


Figura 23 - Ajuste por um Polinômio Quadrático, da Variável WTable.

Através deste ajuste polinomial, obteve-se os Resíduos. As estatísticas descritivas dos mesmos estão descritas na Tabela 7:

Tabela 7: Estatísticas Descritivas dos Resíduos considerando o Modelo 1.

Número de dados	327
Mínimo	-190,114
Máximo	93,0
1º Quartil	-16,412
3º Quartil	19,16
Média	0,00
Mediana	5,165
Desvio-padrão	34,805
Coef. de Assimetria	-1.023
Curtose	3,351

Obteve-se, posteriormente, o Histograma dos Resíduos, exibido na Figura 24 considerando o Modelo 1, na Figura 23 e o Gráfico Quantil-Quantil da Figura 25.

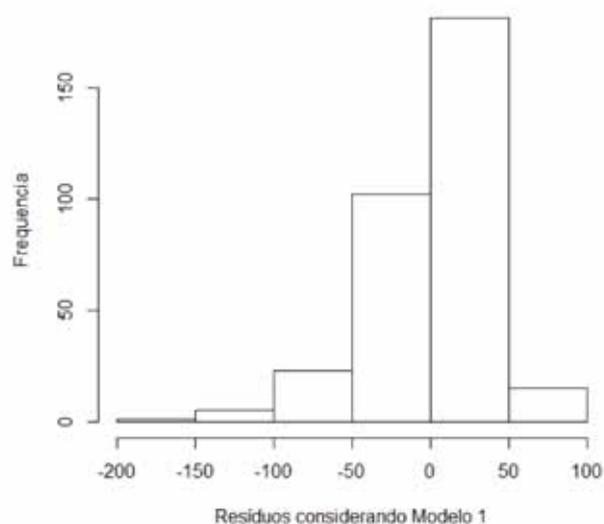


Figura 24 - Histograma dos Resíduos obtidos quando se modela WTable por um Polinômio Quadrático

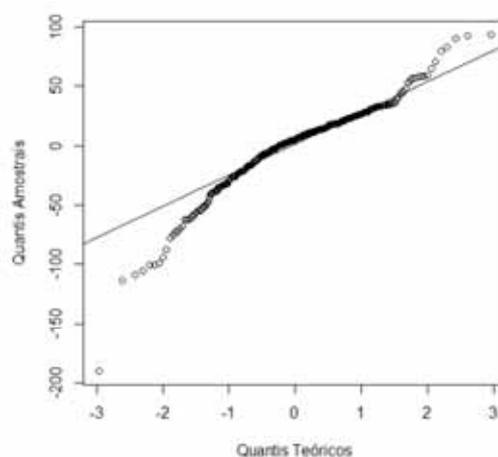


Figura 25 - Gráfico Quantil-Quantil Normal para os Resíduos da Variável WTable, considerando o Modelo 1.

Efetuiu-se o teste de Shapiro-Wilks para o Histograma dos Resíduos do Modelo 1 e obteve-se  $W=0.9338$  ( $p \leq 0,001$ ), portanto estes Resíduos não possuem normalidade. O gráfico apresenta assimetria à direita como se pode observar na Figura 23, com Coeficiente de Assimetria=  $-1,023$ .

Encontrados e analisados os Resíduos, obteve-se o semivariograma dos mesmos e investigou-se a presença de anisotropia. Constatou-se a presença de anisotropia zonal como se pode observar na Figura 26:

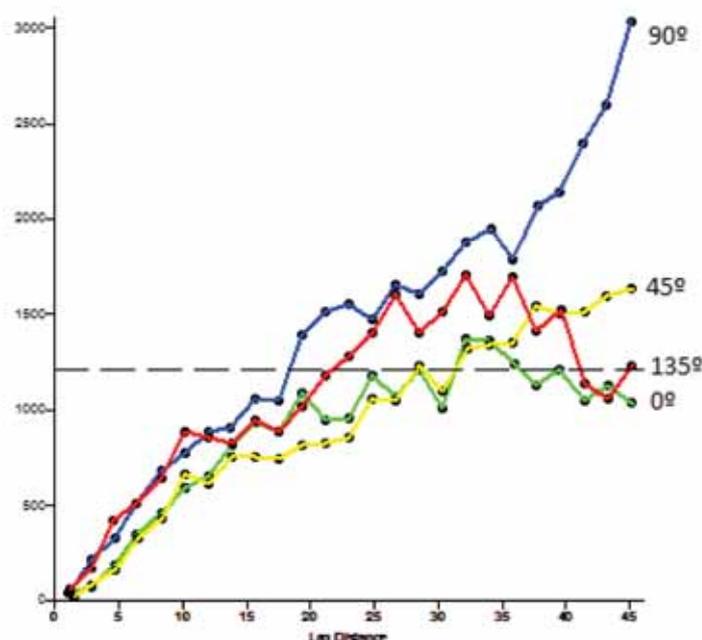


Figura 26 - Semivariograma dos Resíduos da Variável WTable quando considerado o Modelo 1, nas direções  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  e  $135^\circ$ . A linha tracejada indica a variância de  $Z(x)$

O modelo de semivariograma que foi ajustado foi o Exponencial com Efeito Pepita  $C_0 = 0$ , Sill  $C = 2255$ , Alcance  $a = 48,99$  milhas, com raio de anisotropia  $R = 2$  e ângulo de anisotropia de  $4,42^\circ$ . A investigação foi realizada com tolerância angular de  $22,5^\circ$ . Com o semivariograma em mãos, realizou-se a Krigagem Ordinária dos Resíduos e obteve-se o Mapa da Figura 27 e o Mapa dos

Desvios-padrão da Krigagem dos Resíduos dado na Figura 28:

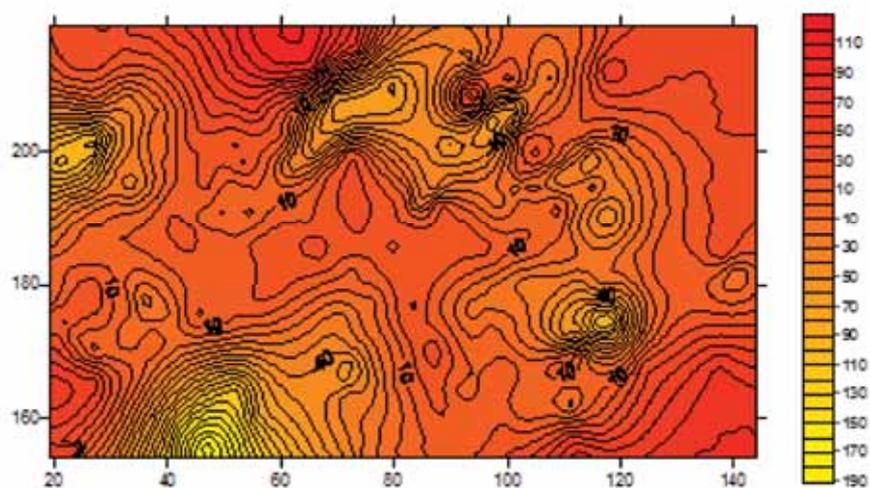


Figura 27 - Mapa da Krigagem Ordinária dos Resíduos considerando o Modelo 1.

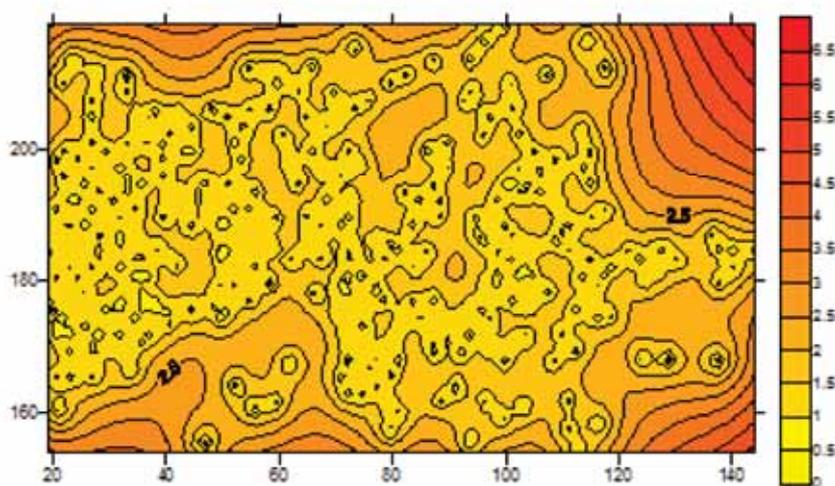


Figura 28 - Mapa dos Desvios-Padrão considerando o Modelo 1.

Através da expressão (61), se conclui que:  $Z(\mathbf{x}) = R(\mathbf{x}) + \hat{m}(\mathbf{x})$ . Portanto, somando-se o Mapa da Krigagem do Resíduos com o Mapa dos Valores dos Dados Estimados através do Modelo 2, obtém-se a Krigagem Universal da Variável

WTable que se encontra na Figura 29:

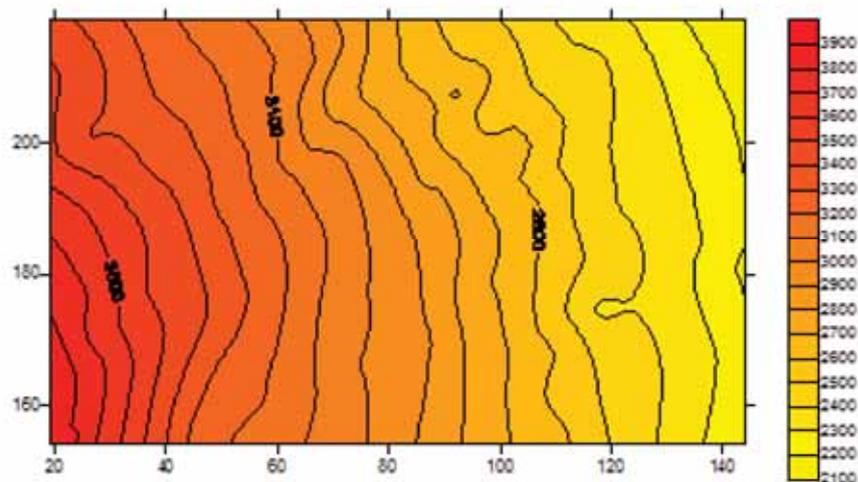


Figura 29 - Mapa de Krigagem Universal, considerando o Modelo 1, para Variável WTable.

Pode-se comparar a superfície obtida através do ajuste polinomial na Figura 30 com a superfície gerada pela Krigagem Universal na Figura 31 e notar que a metodologia da Krigagem Universal permitiu a percepção de ondulações nos valores que são mais condizentes com a realidade física do que a Modelagem por meio de Superfícies de Tendência Polinomiais.

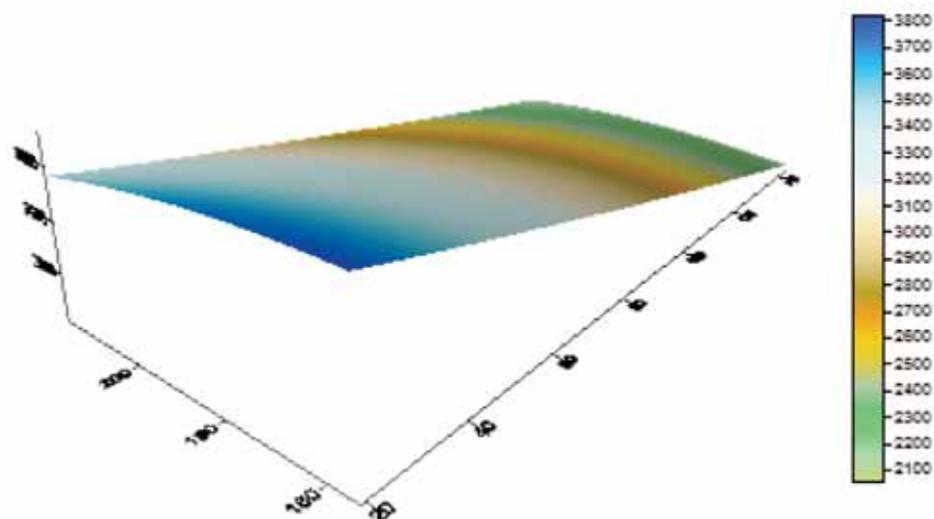


Figura 30 - Superfície Estimada por Ajuste Polinomial por meio de um Polinômio Quadrático.

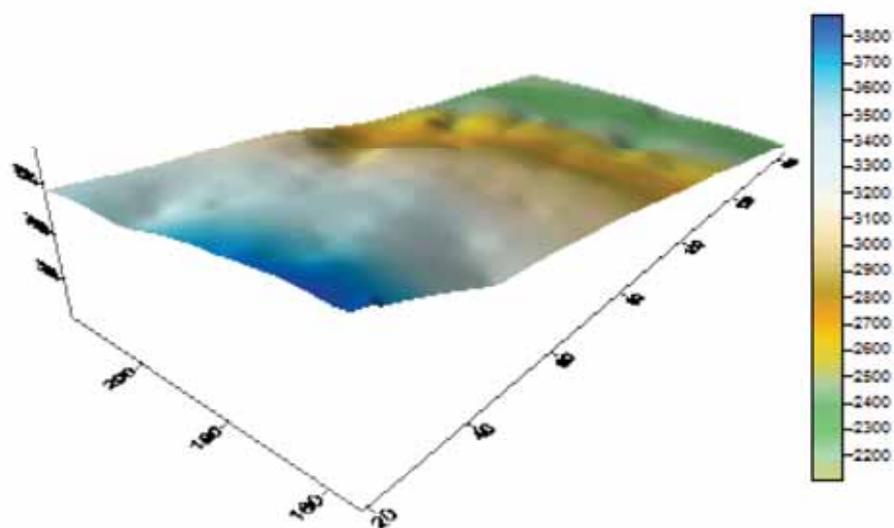


Figura 31 - Superfície Estimada por Krigagem Universal considerando o Modelo 1.

Considerando, agora, o Modelo 2, para o polinômio:  
 $m_2(X, Y) = 4714,65 - 4,26Y - 12,56X$  obteve-se o Mapa dos valores na

Figura 32:

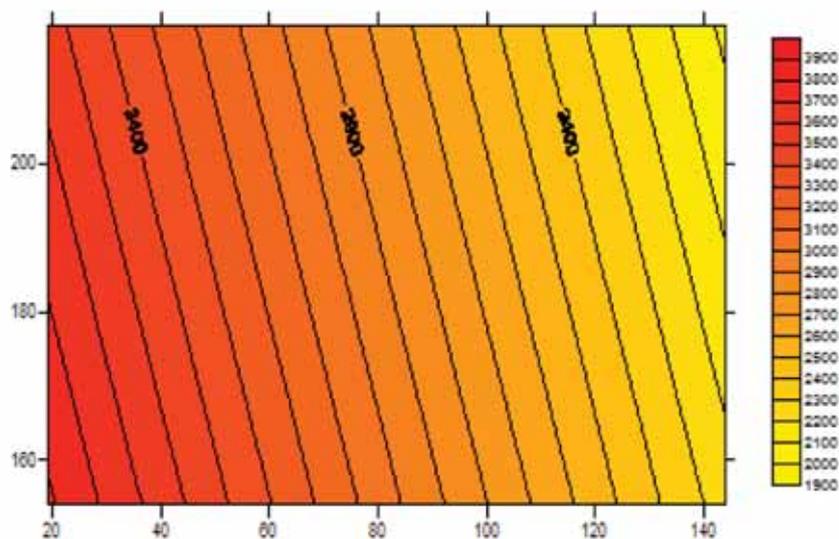


Figura 32 - Ajuste considerando um Polinômio Linear, da Variável WTable.

Através deste ajuste polinomial, obteve-se os Resíduos. As estatísticas descritivas estão descritas na Tabela 8:

Tabela 8: Estatísticas Descritivas dos Resíduos considerando o Modelo 2.

Número de dados	327
Mínimo	-230,864
Máximo	98,992
1º Quartil	-24,999
3º Quartil	31,068
Média	0,00
Mediana	9,535
Desvio-padrão	46,07
Coef. de Assimetria	-0,948
Curtose	1,565

Obteve-se, posteriormente, o Histograma dos Resíduos, considerando o Modelo 2, exibido na Figura 33 e o Gráfico Quantil-Quantil dado na Figura 34 :

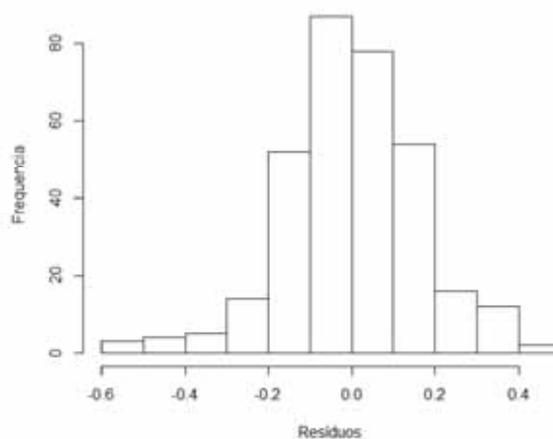


Figura 33 - Histograma dos Resíduos obtidos quando se modela a Variável WTable por um Polinômio Linear.

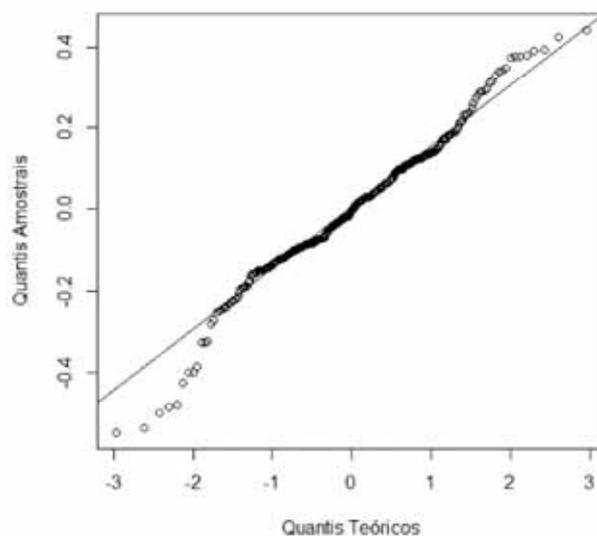


Figura 34 - Gráfico Quantil-Quantil Normal para os Resíduos da Variável WTable considerando Modelo 2.

Efetou-se o teste de Shapiro-Wilks e obteve-se  $W=0.9804$  ( $p \leq 0,001$ ), portanto os Resíduos também não possuem normalidade. No entanto, assim como a variável WTable, apresenta boa simetria (Coef. de Simetria=-0,168).

Encontrados os Resíduos, obteve-se o semivariograma dos mesmos e investigou-se a presença de anisotropia. Constatou-se a presença de anisotropia zonal como se pode observar na Figura 35:

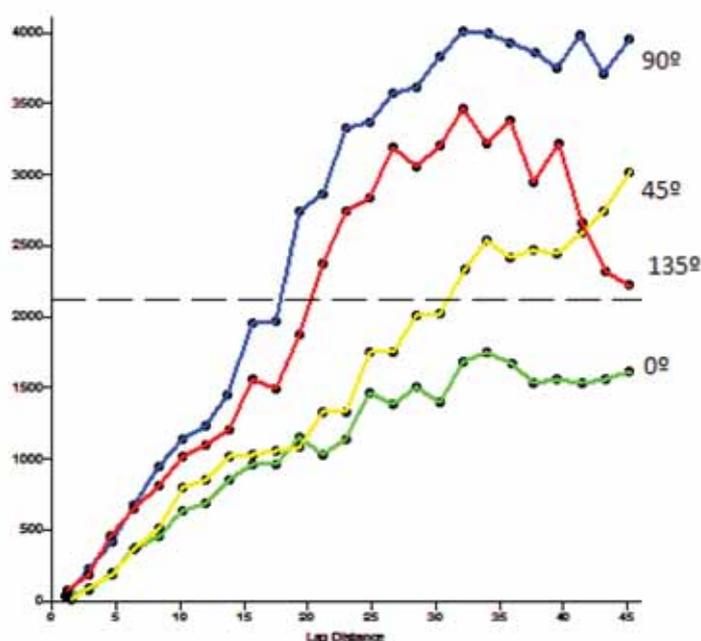


Figura 35 - Semivariograma dos Resíduos da Variável WTable nas direções 0°, 45°, 90° e 135°. A linha tracejada indica a variância de  $Z(x)$

O modelo de semivariograma que foi ajustado foi o Exponencial com Efeito Pepita  $C_0 = 0$ , Sill  $C = 8546$ , Alcance  $a = 141,6$  milhas, com raio de anisotropia  $R = 2$  e ângulo de anisotropia de  $7,64^\circ$ . A investigação foi realizada com tolerância angular de  $22,5^\circ$ . Com o semivariograma em mãos, realizou-se a Krigagem Ordinária dos Resíduos e obteve-se o Mapa da Figura 36:

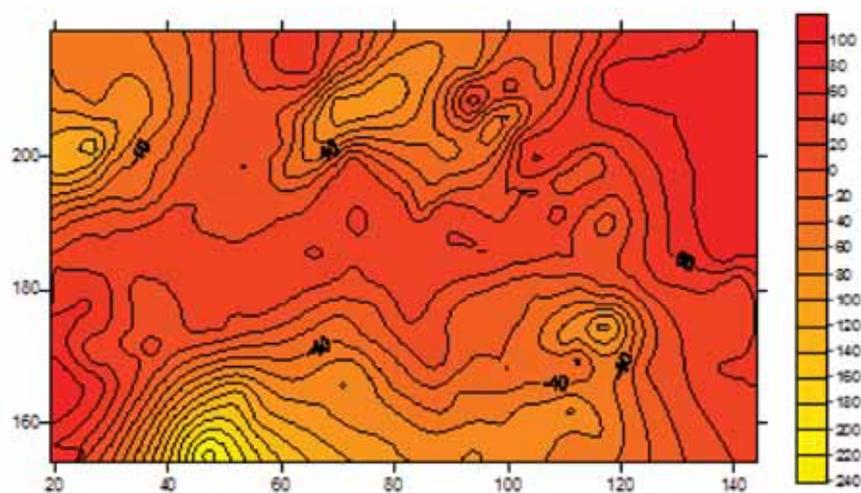


Figura 36 - Mapa da Krigagem Ordinária dos Resíduos, considerando o Modelo 2.

O Mapa dos Desvios-padrão da Krigagem dos Resíduos se encontra dado na Figura 37:

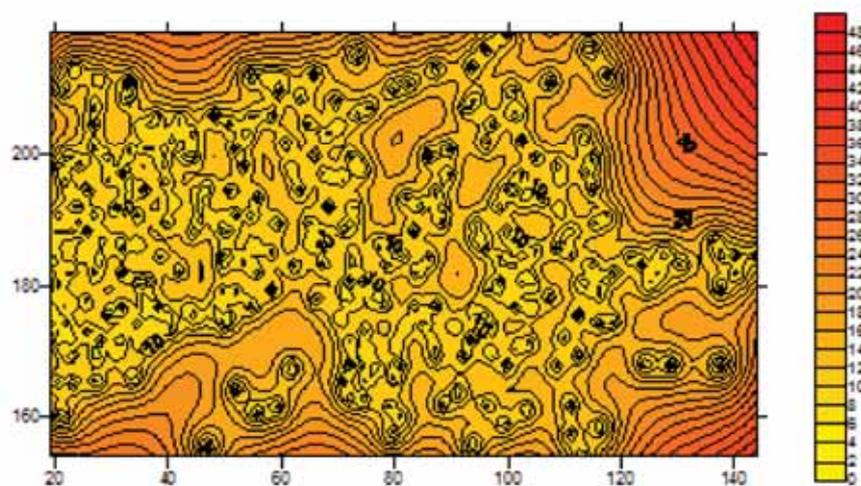


Figura 37 - Mapa dos Desvio-padrão, considerando o Modelo 2.

Através da expressão (61), se conclui que:  $Z(\mathbf{x}) = R(\mathbf{x}) + \hat{m}(\mathbf{x})$ . Portanto, somando-se o Mapa da Krigagem do Resíduos com o Mapa dos Valores dos

Dados Estimados através do Polinômio de Primeiro Grau obtém-se a Krigagem Universal da Variável WTable que se encontra na Figura 38:

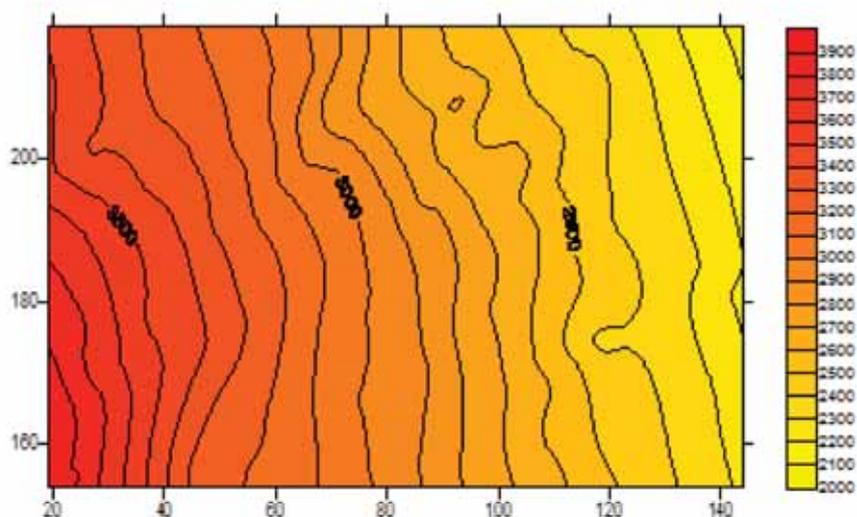


Figura 38 - Mapa de Krigagem Universal para Variável WTable.

Pode-se comparar a superfície obtida através do ajuste polinomial dada na Figura 39 com a superfície gerada pela Krigagem Universal exibida na Figura 40 e notar que a metodologia da Krigagem Universal permitiu a percepção de ondulações que são mais condizentes com a realidade do que a modelagem por meio de Superfície de Tendências Polinomiais.

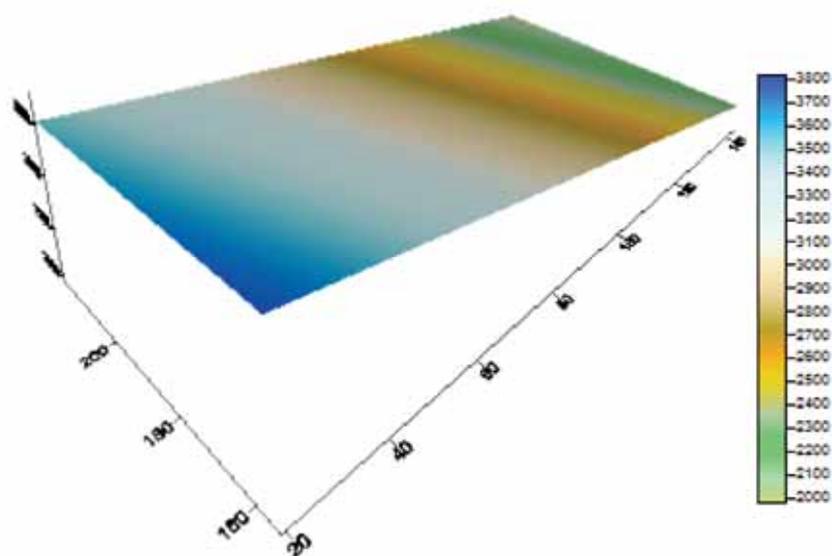


Figura 39 - Superfície Estimada por Ajuste Polinomial.

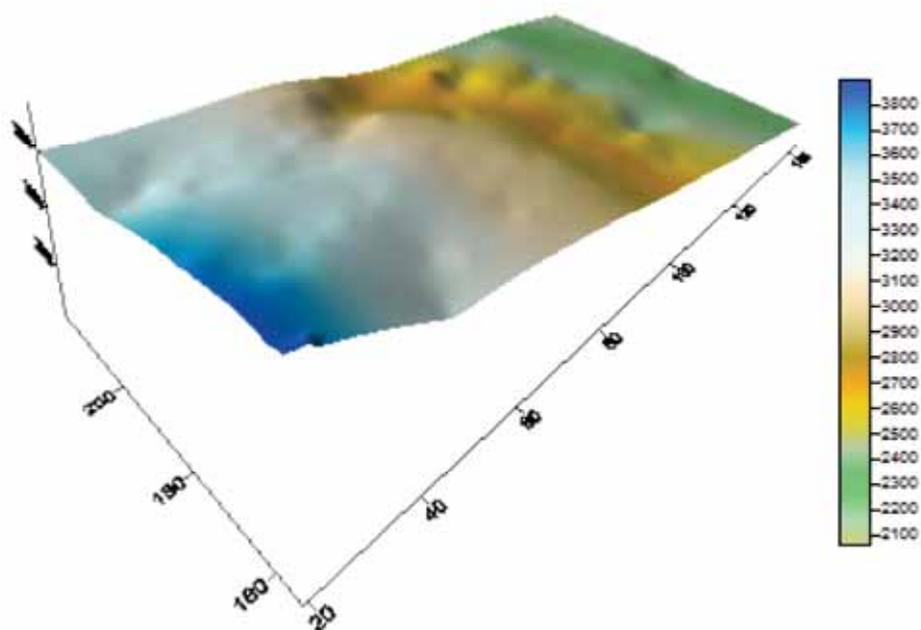


Figura 40 - Superfície Estimada por Krigagem Universal.

Há poucas diferenças entre a Figura 31, que é o Mapa da Krigagem Universal quando se considera o “drift” como sendo um Polinômio Quadrático e a Figura 40, que é o Mapa da Krigagem Universal quando se considera o “drift” como sendo um Polinômio Linear, cabendo ao pesquisador da área avaliar qual a melhor modelagem.

### 6.3.2 Dados do Aquífero Wolfcamp

Já o seguinte conjunto de dados possui o Mapa de amostragens exibidos na Figura 41:

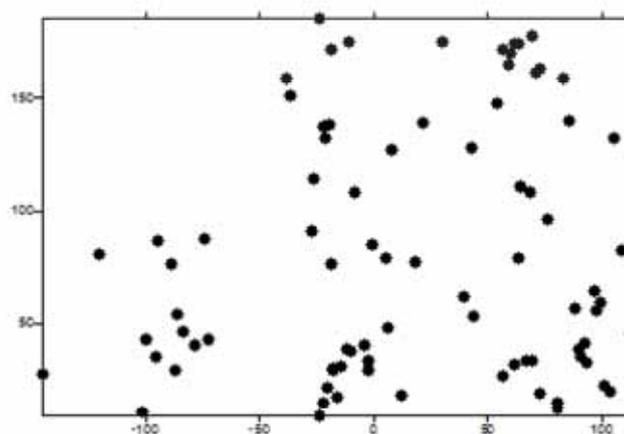


Figura 41 - Localização espacial dos 85 amostras do Aquífero Wolfcamp.

As estatísticas descritivas da Variável Nível Piezométrico, obtidas através do Software R 3.1.1, são as apresentadas na Tabela 9:

Tabela 9: Estatísticas Descritivas da Variável Nível Piezométrico

Número de dados	85
Mínimo	1024,00
Máximo	3571,00
1º Quartil	1548,00
3º Quartil	2540,00
Média	2002,28
Mediana	1797,00
Desvio-padrão	611,34
Coef. de Assimetria	0,493
Curtose	-0,527

O Histograma da Variável Nível Piezométrico está dado na Figura 42:

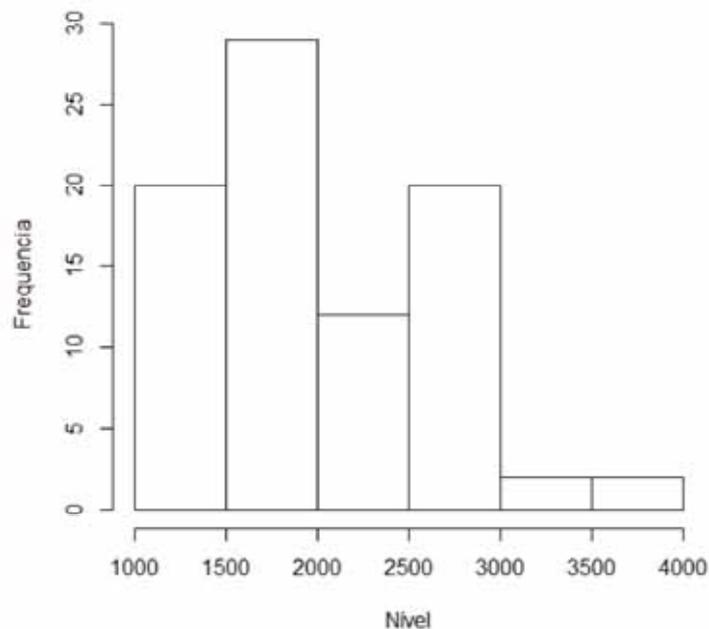


Figura 42 - Histograma da do Nível Piezométrico do Aquífero Wolfcamp.

E o Gráfico Quantil-Quantil para Distribuição Normal na Figura 43:

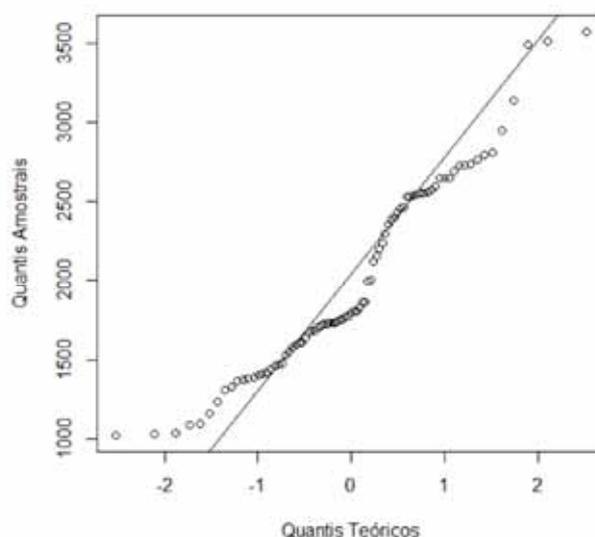


Figura 43 - Gráfico Quantil-Quantil para Distribuição Normal.

Obviamente os dados não apresentam normalidade. Realizado um teste de Shapiro-Wilks obtve-se:  $W=0.947$  ( $p=0,001586$ ). Estes dados também não tem boa simetria, tendo sido obtido o coeficiente de assimetria=0,493

Obteve-se uma visualização da distribuição dos valores da Variável Nível Piezométrico, com relação à sua localização geográfica, através do Software Surfer 8 dada na Figura 44 a fim de se visualizar a presença ou não de características que indiquem tendência nos dados.

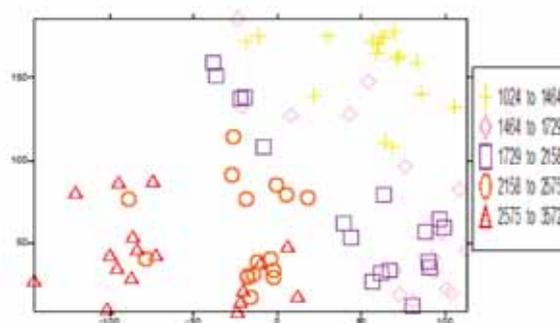


Figura 44 - Distribuição dos Valores da variável Nível Piezométrico por sua localização geográfica

Pode-se observar que há um aumento dos valores seguindo uma direção que vai da diagonal superior direita até a diagonal inferior esquerda. Isto é um indicio de tendência nos dados.

O semivariograma obtido dos pontos, para as direções  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  e  $135^\circ$  com tolerância angular de  $22,5^\circ$  são exibidos nas Figuras 45, 46, 47 e 48, respectivamente.

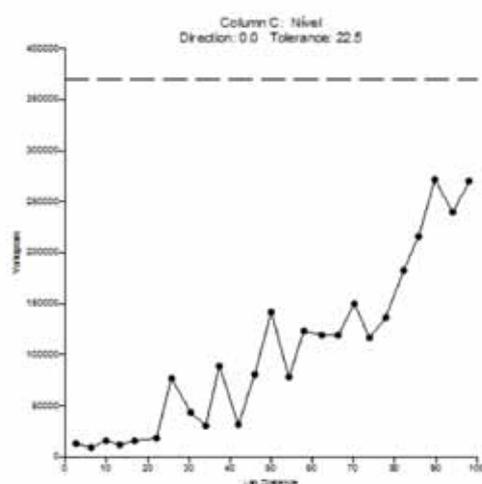


Figura 45 - Semivariograma na direção  $0^\circ$  para variável Nível Piezométrico.

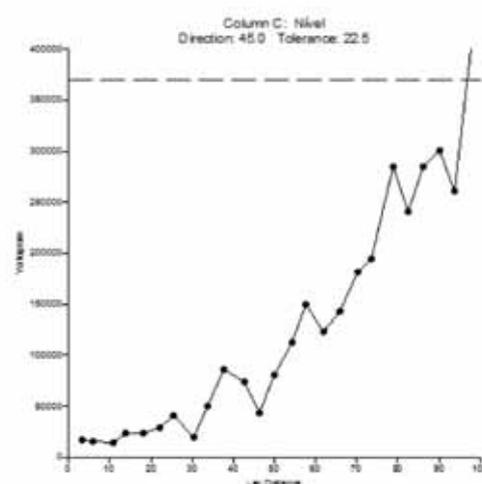


Figura 46 - Semivariograma na direção  $45^\circ$  para variável Nível Piezométrico.

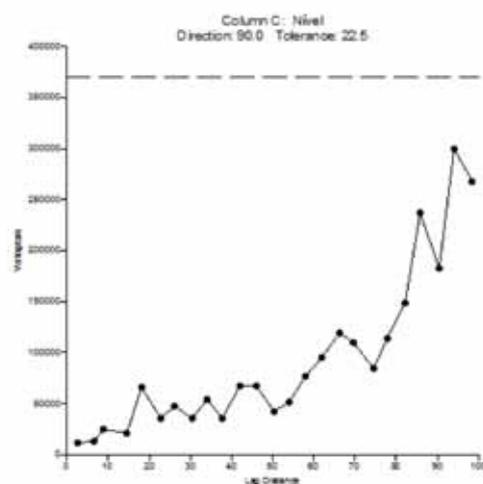


Figura 47 - Semivariograma na direção 90° para variável Nível Piezométrico.

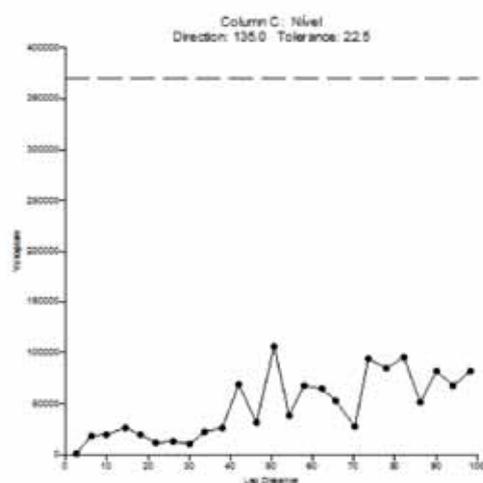


Figura 48 - Semivariograma na direção 135° para variável Nível Piezométrico.

Como os semivariogramas apresentaram tendência, com exceção do semivariograma na direção de 135°, embora ele se estabilize muito abaixo da variância dos dados, dada pela linha tracejada, é necessário se aplicar o método da Krigagem Universal por meio da Krigagem dos Resíduos pelo método direto. Para isso, deve-se ajustar os dados a uma superfície. Utilizou-se o Software R 3.1.1 para se escolher a

melhor superfície de ajuste.

O primeiro modelo proposto (Modelo 1) foi o Polinomial Cúbico, ou seja,  $m_1(x, y) = aX^2 + bY^2 + cX^2Y + dY^2X + eX + fY + g$  e o segundo modelo proposto (Modelo 2) foi o Polinomial Quadrático, isto é:  $m_2(x, y) = hX^2 + iY^2 + jX + kY + l$  e ainda o (Modelo 3), o Polinomial Linear:  $m_3(x, y) = mX + nY + o$ . Os valores obtidos para cada modelo estão nas Tabelas 10 e 11 e 12.

Tabela 10: Valores dos coeficientes do Polinômio Cúbico e seus respectivos valores-p.

Coeficientes	Estimado	Erro-padrão	Valor de T	Valor-p
Intercepto	2522,00	53,38	47,250	$\leq 2 \times 10^{-16}$
X	-7,423	0,876	-8,470	$\leq 1,15 \times 10^{-12}$
Y	-4,651	1,150	-4,042	$1,23 \times 10^{-4}$
$X^2Y$	0,0001318	0,00008451	1,559	0,122994
$Y^2X$	0,0001045	0,0000496	2,108	0,0382378
$X^3$	-0,000003294	0,00007804	-0,042	0,966441
$Y^3$	-0,00006949	0,00003456	-2,011	0,047832
			$R^2$	0,9135

Tabela 11: Valores dos coeficientes do Polinômio Quadrático e seus respectivos valores-p.

Coeficientes	Estimado	Erro-padrão	Valor de T	Valor-p
Intercepto	2481,00	68,11	35,428	$\leq 2 \times 10^{-16}$
X	-8,374	0,552	-15,156	$\leq 2 \times 10^{-16}$
Y	-2,039	1,764	-1,156	0,251099
$X^2$	0,001437	0,004988	0,288	0,773980
$Y^2$	-0,02467	0,009299	-2,653	0,009651
$XY$	0,0268	0,007412	3,616	0,000525
			$R^2$	0,913

Tabela 12: Valores dos coeficientes do Polinômio Linear e seus respectivos valores-p.

Coeficientes	Estimado	Erro-padrão	Valor de T	Valor-p
Intercepto	2591,3758	38,9697	66,50	$\leq 2 \times 10^{-16}$
X	-6,7514	0,3439	-19,63	$\leq 2 \times 10^{-16}$
Y	-5,9870	0,4067	-14,72	$\leq 2 \times 10^{-16}$
			$R^2$	0,8920

Procedeu-se à comparação entre os modelos, cujo resultado se encontra exibido na Tabela 13:

Tabela 13: Comparação, através da ANOVA, dos Modelos 1, 2 e 3.

Modelo	GL Resíduo	SQR	GL.	SQ	F	Valor-p
Modelo 1	78	2714133				
Modelo 2	79	2729843	-1	-15709	0,4515	0,5036248
Modelo 3	82	3390164	-3	-660322	6,3255	0,0006767

Obtém-se, da comparação do Modelo 2 com o Modelo 1 um valor- $p \geq 0,05$  que indica que o Modelo 2 (Polinômio Quadrático) é preferível ao Modelo 1 (Polinômio Cúbico). Já da comparação do Modelo 3 (Polinômio Linear) com o Modelo 2 (Polinomial Linear) obtém-se um valor- $p \leq 0,05$  o que mostra que o modelo 2 é preferível ao Modelo 3.

É de grande valia, para a tomada de decisão, observar as superfícies obtidas com o ajuste dos polinômios, sobrepostas aos dados separados exibido na Figura 44.

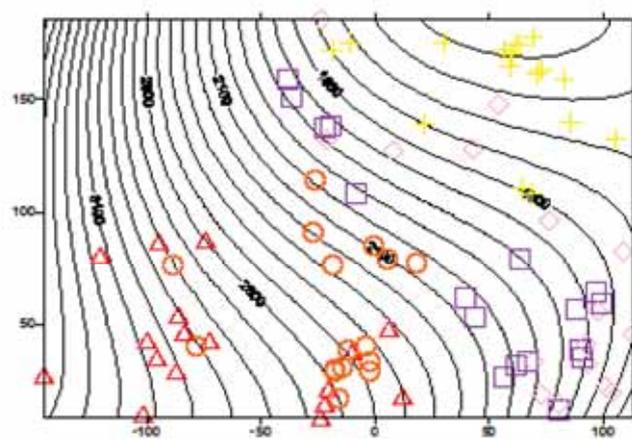


Figura 49 - Modelo 1 sobreposto aos dados do Aquífero Wolfcamp.

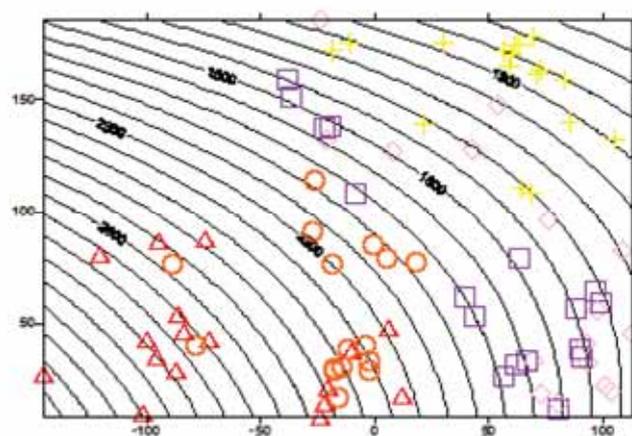


Figura 50 - Modelo 2 sobreposto aos dados do Aquífero Wolfcamp.

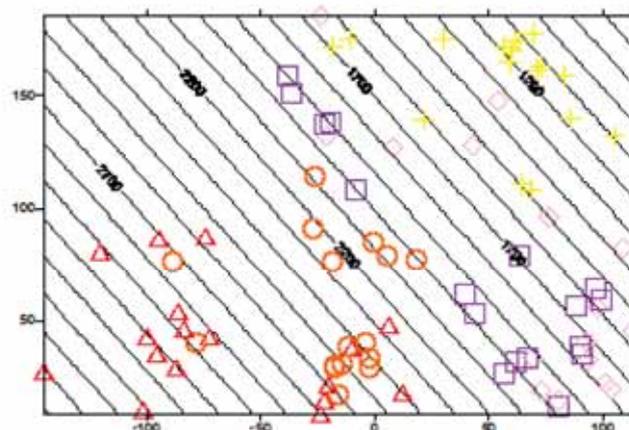


Figura 51 - Modelo 3 sobreposto aos dados do Aquífero Wolfcamp.

Através dos resultados obtidos pela ANOVA e da observação dos Mapas sobrepostos, optou-se por modelar os dados do Nível Piezométrico por um Polinômio Quadrático, dado pelo Modelo 2 cuja equação é:

$$m_2(X, Y) = 2481 - 8,374X - 0,02467Y^2 + 0,0268XY.$$

Através deste ajuste, obteve-se a seguinte Mapa dado pela Figura 52:

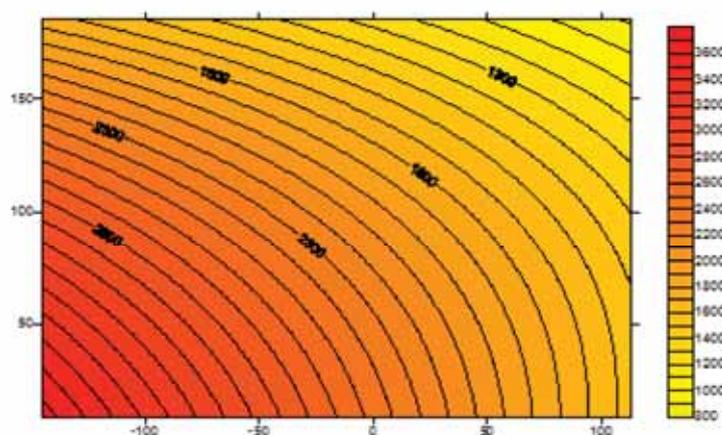


Figura 52 - Nível Piezométrico ajustado pelo Modelo 2.

Feito isto, obteve-se os Resíduos que apresentam o seguinte Histograma dado pela Figura 53 e o Gráfico Quantil-Quantil para Distribuição Normal dada pela Figura 54:

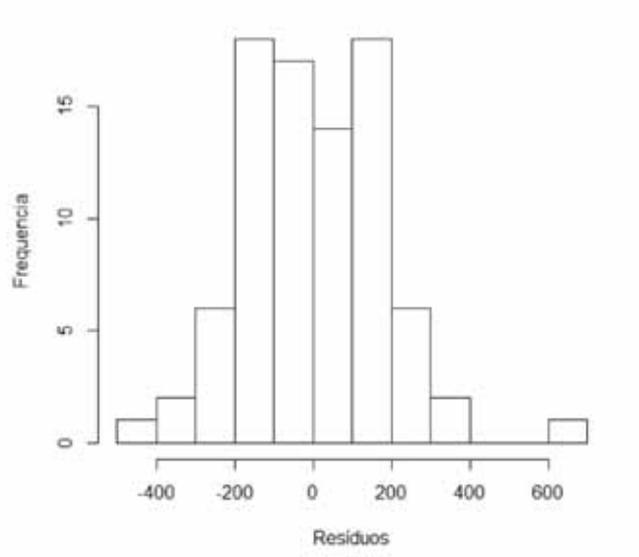


Figura 53 - Histograma dos Resíduos Obtidos Quando se Modela o Nível Piezométrico por Um Polinômio Quadrático.

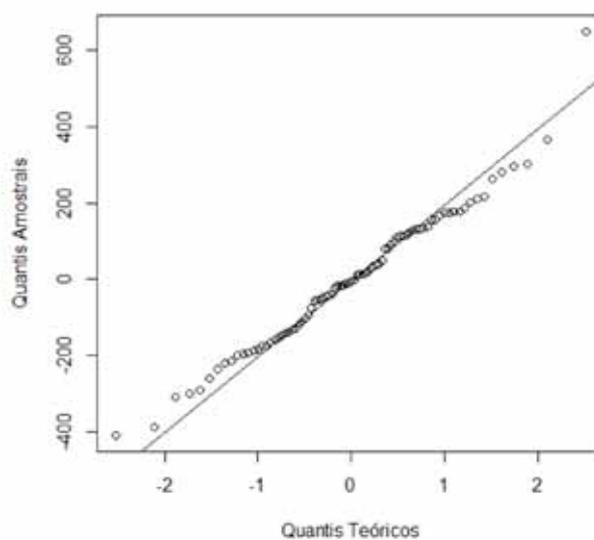


Figura 54 - Gráfico Quantil-Quantil para Distribuição Normal para os Resíduos.

As estatísticas descritivas destes Resíduos encontram-se na Tabela 14:

Tabela 14: Estatísticas Descritivas dos Resíduos Obtidos quando se Modela a Variável Nível Piezométrico Por Um Polinômio de Segundo Grau.

Número de dados	85
Mínimo	-407,5317
Máximo	647,9659
1º Quartil	-138,2426
3º Quartil	129,1848
Média	0,021
Mediana	-5,7565
Desvio-padrão	180,27
Coef. de Assimetria	0,3457
Curtose	0,7439

Observa-se que os Resíduos não apresentam boa simetria (Coef. de Assimetria= 0,3457) e não apresentam normalidade como pode-se observar na Figura 54 e através do teste de normalidade de Shapiro-Wilks que resultou em  $W = 0,9815$  com valor- $p = 0,2636$ ).

Para realizar a Krigagem dos Resíduos, estudou-se o semivariograma nas direções  $0^\circ, 45^\circ, 90^\circ$  e  $135^\circ$ . Observou-se que os semivariogramas apresentam isotropia, como pode ser observado na Figura 55:

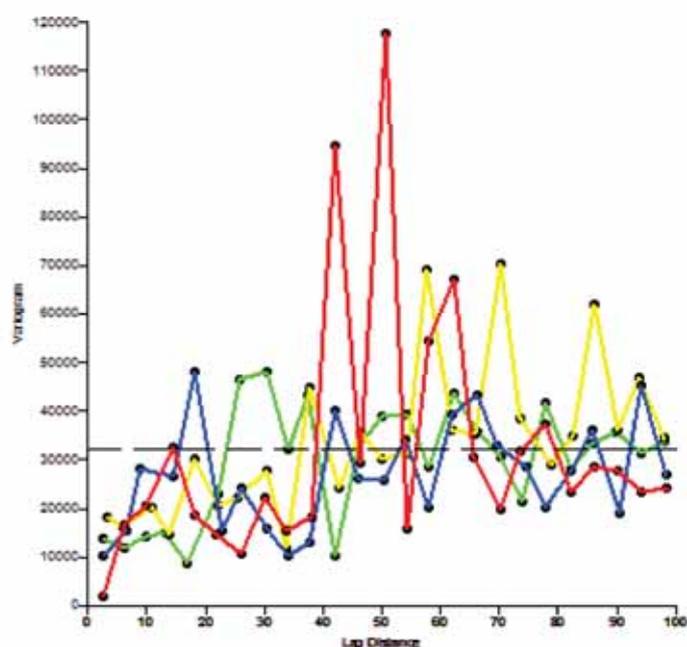


Figura 55 - Semivariograma dos Resíduos. As cores correspondem aos ângulos assim como na Figura 26.

Considerando-se este último resultado, ajustou-se ao semivariograma considerando-o omnidirecional, ou seja, com o mesmo valor em todas as direções. Utilizando-se o método do ajuste à sentinela obteve-se o Modelo Exponencial com Efeito Pepita  $C_0 = 5000$ , Patamar  $C = 25000$  e Alcance  $a = 11$  milhas.

$$\gamma(h) = 5000 + 25000 \left( 1 - \exp\left(-\frac{h}{11}\right) \right)$$

Feito isto, procedeu-se a Krigagem dos Resíduos. Com isto obteve-se o seguinte Mapa dado pela Figura 56 e o Mapa dos Desvios-Padrão dados na Figura 57:

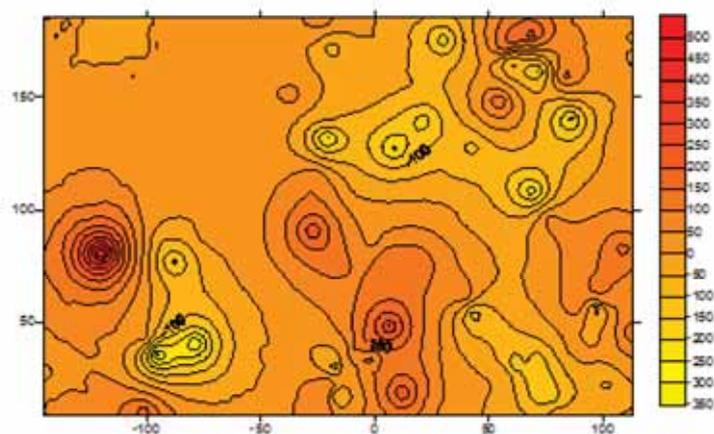


Figura 56 - Krigagem dos Resíduos obtidos.

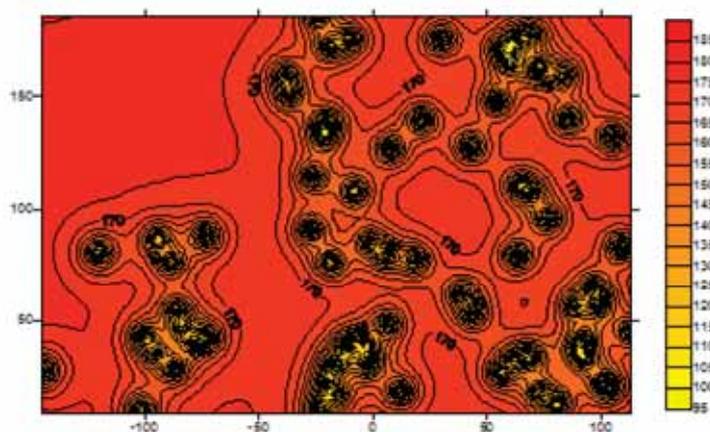


Figura 57 - Mapa dos Desvios-Padrão da Krigagem dos Resíduos.

Através da expressão (61), se conclui que:  $Z(\mathbf{x}) = R(\mathbf{x}) + \hat{m}(\mathbf{x})$ . Portanto, somando-se o Mapa da Krigagem do Resíduos com o Mapa dos Valores dos Dados Estimados através do Polinômio de Segundo Grau obtém-se a Krigagem Universal da Variável Nível Piezométrico que se encontra na Figura 58. Um Mapa da Superfície da Variável Nível Piezométrico obtido através da Krigagem Universal é apresentado na Figura 59:

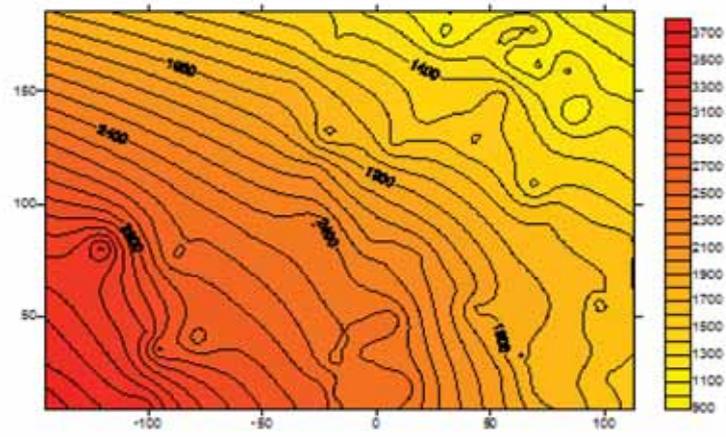


Figura 58 - Mapa da Krigagem Universal da Variável Nível Piezométrico.

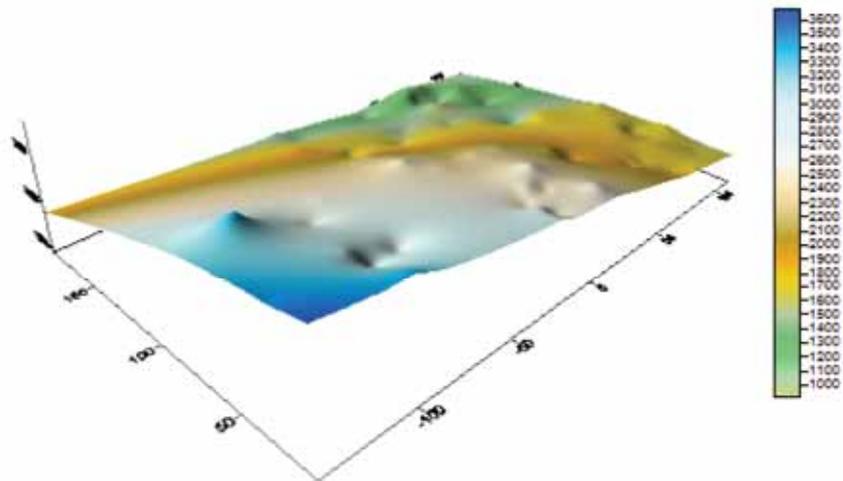


Figura 59 - Mapa de Superfície da Variável Nível Piezométrico obtido através da Krigagem Universal.

## 7 CONCLUSÕES

A Geoestatística, embora tenha encontrado hoje muitas aplicações nos mais variados campos que exigem a estimativa de valores em pontos não amostrados, logo no início foi alvo de muitas críticas, como as que foram feitas no artigo: “*Matheronian Geostatistics: Quo Vadis?*” por PHILIP & WATSON (1986) em que vários pressupostos básicos da teoria elaborada por Matheron foram postos à prova. No entanto, estas críticas receberam respostas de diversos autores como, por exemplo, SRIVASTAVA (1986), no artigo “*Philip and Watson: Quo Vadunt?*” e também de JOURNEL (1986) em: “*Geoestatistics: Models and Tools for the Earth Sciences*”. O artigo de Srivastava aborda mais questões epistemológicas enquanto o de Journel enfoca questões teóricas e técnicas.

Malgrado estas críticas, a metodologia de Matheron prosperou e continuou sendo cada vez mais estudada e aperfeiçoada pois o que prevaleceu foi a sua aplicabilidade. Apesar disso, a metodologia Geoestatística não é um ramo da Estatística Aplicada perfeitamente resolvida em todos os seus detalhes, restando ainda muito por fazer. O que foi apresentado neste texto é um exemplo disto.

O enfoque proposto por Matheron para resolver problemas em que a tendência fosse constatada levava em consideração o pressuposto de se conhecer o semivariograma dos Resíduos do problema a ser modelado. Utilizando a Krigagem Universal como estimador, pretende-se calcular, simultaneamente, num ponto  $\mathbf{x}_0$  os valores devidos à não estacionariedade (em geral coeficientes do polinômio que constituem a tendência) e os pesos que serão atribuídos aos pontos amostrais que estimarão  $\mathbf{x}_0$ . Mas isto leva à seguinte questão: se não houve ainda um ajuste através de um modelo, como conhecer os Resíduos e, mais ainda, o Semivariograma dos mesmos? O problema é que a função que estima a tendência tem que ser conhecida *a priori* o que nem sempre é possível. Para contornar estes problemas foram sugeridas algumas técnicas como a Krigagem dos Resíduos, apresentada neste trabalho, a técnica das Covariâncias Generalizadas e as Funções Aleatórias de Ordem-K, que é

uma generalização da hipótese intrínseca aonde a ideia é a mesma utilizada nas séries temporais em que são construídos acréscimos de ordem sucessivamente crescente até que para certa ordem os mesmos sejam estacionários(PEREIRA, 1983). Além disso o conceito de estacionariedade é um conceito controverso. Este assunto é amplamente explorado em (MYERS, 1989).

Pretendeu-se, através deste trabalho de caráter metodológico, mostrar a técnica da Krigagem Universal ciente de que, quando na presença de tendência regionalizada dos dados, o modelo a ser encontrado, como regra geral, é bastante dependente da escolha pessoal do modelo de tendência. Em outras palavras, como o método não é do tipo “caixa-preta”, o usuário deve ter uma familiaridade com os dados sob tratamento e, daí, a necessidade de uma análise preliminar dos mesmos.



## ANEXO 1

ID	X(milhas)	Y(milhas)	WTable				
1	33,0251	211,1625	3338,5000	50	100,5709	210,1548	2587,4500
2	27,6039	211,7596	3419,0000	51	94,1408	208,8262	2697,7900
3	23,4755	213,1902	3460,5000	52	113,2792	202,7570	2484,0300
4	48,1021	206,0002	3239,2400	53	101,9658	206,7629	2526,1500
5	50,9160	205,1546	3203,7800	54	102,4742	202,2735	2606,3400
6	53,7079	203,6900	3179,6000	55	100,6002	206,5484	2524,3000
7	37,9040	205,8745	3308,1600	56	96,9864	204,8952	2562,4200
8	41,9804	205,4776	3285,0000	57	95,9115	201,8117	2606,0800
9	37,3412	204,2778	3335,8000	58	92,7338	207,2467	2705,9400
10	32,6712	208,6855	3337,5000	59	115,0442	199,3409	2423,7600
11	34,8629	204,4937	3343,8500	60	113,7663	197,5125	2446,8200
12	27,0648	207,7953	3435,6600	61	108,9708	198,7492	2542,8300
13	26,6960	204,9456	3420,5000	62	109,1563	196,3813	2508,5500
14	21,3555	210,1772	3495,9000	63	105,1406	199,4675	2607,7800
15	52,1157	201,0107	3204,1500	64	102,0822	196,6898	2634,8000
16	53,3729	198,2317	3222,9000	65	89,4544	200,8569	2681,7100
17	44,2512	203,1559	3293,6500	66	114,9330	194,8630	2479,4100
18	42,8572	199,4767	3329,4000	67	117,7779	190,4224	2413,4300
19	44,6167	199,9101	3305,1500	68	115,4126	189,3713	2456,1300
20	40,5064	202,7949	3319,2000	69	109,1567	191,6532	2576,3300
21	39,0560	200,7323	3341,6000	70	101,1560	194,4774	2624,0000
22	36,2145	201,0841	3373,1500	71	104,1400	194,5200	2589,1000
23	38,9741	198,7455	3366,5800	72	105,9665	193,1031	2588,5200
24	32,9463	200,4574	3394,8500	73	102,8561	192,9382	2632,3500
25	26,9149	201,4477	3387,2200	74	99,5477	194,6382	2656,8600
26	28,9313	199,6150	3428,9300	75	98,1160	192,1842	2670,2400
27	24,5832	198,8092	3453,2600	76	116,2962	158,4376	2519,8600
28	22,2745	199,9092	3469,2000	77	110,9863	157,4683	2602,7200
29	51,1156	194,5814	3244,9000	78	112,9605	154,0635	2598,7700
30	47,0309	195,1001	3312,5000	79	99,2812	159,1609	2730,4600
31	45,1226	194,0491	3339,4000	80	93,8213	155,9822	2791,7100
32	45,6825	192,5378	3342,3000	81	143,8855	184,8444	2202,8400
33	39,7259	195,8586	3377,3600	82	137,6752	184,9081	2291,4800
34	35,6988	194,7582	3413,9200	83	140,2777	184,5072	2245,6800
35	40,9655	192,7061	3400,8700	84	136,7204	186,7946	2292,0900
36	31,9330	197,1324	3445,1300	85	127,9815	184,9885	2379,8200
37	33,2345	195,4673	3423,8000	86	126,9868	185,0072	2386,9400
38	32,7874	193,7404	3453,8200	87	130,5502	183,0754	2345,4500
39	29,5374	196,4736	3467,1200	88	121,7319	184,4792	2432,6900
40	25,6308	194,5269	3516,6500	89	123,8276	183,3199	2413,2100
41	20,7089	198,2343	3497,1400	90	140,5912	181,6449	2208,1700
42	21,8360	195,5665	3533,4500	91	137,5697	179,8106	2256,1000
43	22,6936	192,4133	3587,2100	92	126,7775	181,5243	2392,5200
44	114,3558	216,4205	2381,4700	93	124,6209	180,0682	2408,7800
45	99,5748	218,7589	2507,3800	94	137,3235	167,8766	2320,3500
46	96,2094	216,2334	2547,9500	95	128,7085	168,0449	2397,0800
47	93,6695	213,8183	2574,9200	96	124,2019	168,2432	2442,6400
48	117,5811	212,0854	2423,7700	97	81,2487	159,9197	2947,3300
49	107,2019	211,3543	2457,0600	98	79,2976	157,3827	2970,0700
				99	55,7476	160,7370	3217,5000

100	55,3681	160,6289	3213,4000	150	118,9385	182,0445	2460,5900
101	73,2018	215,3713	2877,9900	151	117,0708	180,8507	2479,7800
102	87,2346	213,2528	2644,7100	152	113,7428	180,2001	2501,7000
103	82,5099	211,2607	2699,7700	153	115,9152	178,7647	2479,5500
104	79,5468	210,2206	2719,6500	154	112,8774	180,2223	2510,1400
105	65,6701	212,5068	3025,4100	155	103,1477	182,2233	2647,3700
106	63,4432	209,1013	3033,0200	156	101,9917	180,6354	2652,9100
107	60,0682	211,8351	3095,4500	157	104,3425	180,5706	2618,0500
108	57,6035	211,9045	3120,6300	158	107,3827	178,4934	2585,6600
109	54,2807	212,1245	3150,0200	159	100,5786	183,0373	2686,2300
110	59,4990	209,7372	3086,3000	160	98,4286	180,7229	2710,2200
111	75,3935	207,3525	2792,5200	161	98,8703	178,8439	2691,8700
112	73,1361	206,4219	2827,1400	162	96,5590	179,7737	2734,6900
113	69,4924	208,3998	2866,1400	163	118,3942	174,7027	2381,6000
114	70,0903	204,2760	2885,1200	164	115,9295	174,3992	2411,0400
115	67,8211	203,1043	2926,4700	165	112,2600	175,4992	2493,6300
116	64,9159	205,0675	2972,5300	166	109,6980	172,3250	2534,1400
117	58,0142	206,2972	3110,5500	167	104,2352	176,4602	2630,0000
118	55,0352	206,0114	3164,1100	168	101,2798	176,9150	2677,2000
119	86,0997	199,8323	2748,0100	169	97,1361	173,7835	2745,5600
120	89,2171	196,9995	2736,3100	170	93,8586	176,2295	2764,0000
121	72,5264	199,6022	2982,3700	171	95,6994	175,3152	2752,1500
122	74,1933	197,5669	2969,1100	172	90,3101	173,8292	2834,2000
123	67,6048	196,8915	3040,3100	173	112,4779	169,5090	2567,6900
124	60,4560	201,8681	3097,2500	174	108,3335	167,6246	2621,4600
125	65,6075	199,9406	2985,6400	175	103,6043	171,2430	2650,0100
126	61,5563	197,8468	3065,0200	176	101,6173	170,9246	2682,6700
127	53,9271	202,9370	3180,0100	177	96,9059	169,8062	2754,5700
128	87,1644	194,1953	2774,2600	178	99,9321	167,7303	2726,6500
129	84,3980	192,2889	2827,6400	179	93,4916	171,8850	2790,0000
130	88,0863	192,0487	2827,9400	180	92,0500	169,1846	2821,5000
131	81,7169	193,2432	2889,2500	181	110,9167	161,9446	2548,2000
132	72,7623	192,1373	3021,9700	182	103,2174	161,1585	2662,1000
133	70,5556	196,4274	3020,9600	183	96,4484	161,9768	2762,2600
134	68,1806	192,1478	3071,9000	184	91,4141	164,1020	2825,1500
135	62,9415	194,8103	3099,4200	185	91,3942	163,3566	2826,6800
136	55,6806	193,5597	3199,9900	186	51,7339	190,7071	3260,2000
137	111,6057	188,4828	2525,1000	187	48,8932	190,1853	3294,5000
138	109,6054	187,6597	2567,2400	188	51,0258	188,2435	3287,1000
139	108,6944	185,9378	2579,1900	189	49,1813	188,1829	3315,1000
140	110,8969	185,0093	2556,5600	190	45,4478	190,1810	3345,7500
141	104,8345	185,1682	2635,2200	191	44,2528	187,9809	3392,2500
142	99,6046	188,0425	2691,9100	192	39,7315	189,6335	3437,4900
143	101,5313	186,6284	2678,7100	193	34,7158	190,5556	3480,5800
144	98,3917	183,9612	2723,4700	194	31,3849	186,9377	3553,2000
145	94,3439	189,5293	2757,1700	195	34,7757	186,3197	3513,3700
146	95,5401	187,7616	2748,3300	196	30,6251	191,8224	3517,0800
147	94,2733	186,9189	2774,7700	197	27,0822	190,7316	3569,0300
148	95,8626	185,8886	2759,7700	198	25,6080	188,4281	3611,7000
149	114,5573	183,2942	2504,2400	199	28,5482	187,3037	3578,8900

200	24,0170	192,1113	3579,3800	250	37,2449	171,6596	3544,5500
201	20,0650	190,7845	3632,1400	251	31,1982	172,1315	3617,3300
202	22,4684	188,3127	3644,1900	252	31,4448	172,1219	3599,0300
203	19,8193	188,0519	3671,3000	253	35,3410	170,8570	3561,0200
204	50,5329	185,0250	3317,2300	254	31,0964	169,8944	3625,6100
205	49,5249	184,4368	3333,8000	255	25,8040	173,2173	3693,9300
206	52,9500	183,8255	3288,8000	256	28,2189	171,8745	3662,9700
207	49,8355	182,9327	3335,5100	257	27,0232	170,0540	3671,3100
208	38,6522	184,5640	3473,9600	258	27,3486	169,0449	3688,5100
209	41,9350	182,4556	3440,6200	259	20,1760	173,3274	3787,5700
210	36,5035	180,9038	3503,0100	260	22,2709	172,3668	3756,4700
211	31,5658	185,5616	3555,9500	261	23,1434	172,2053	3734,4700
212	31,7627	181,8203	3569,6800	262	20,5958	171,4400	3798,8000
213	34,5335	180,7249	3524,3300	263	22,6361	169,2354	3776,5200
214	35,7574	180,6813	3510,4900	264	52,2229	164,2074	3228,2000
215	24,8281	184,8474	3645,2600	265	33,6661	166,4359	3607,3800
216	27,9438	184,9635	3610,0300	266	27,3086	165,4343	3741,0700
217	27,5099	183,6113	3618,0800	267	19,9513	165,3593	3856,7100
218	27,1665	181,5080	3647,7000	268	23,2146	165,2236	3804,6500
219	19,6768	184,8165	3703,3800	269	88,0328	189,6884	2836,9200
220	19,7418	183,4426	3726,9400	270	86,2767	184,1468	2885,0600
221	22,4515	182,2054	3699,2900	271	83,3324	188,4692	2900,0100
222	52,0251	178,8821	3318,4000	272	79,5451	185,3516	2952,6200
223	49,0621	178,9744	3365,3300	273	81,2418	184,5525	2938,1700
224	50,1567	175,3316	3355,2800	274	72,4617	189,4082	3039,6500
225	50,6424	175,0683	3351,6400	275	77,2298	186,4168	2982,7000
226	49,6177	174,2279	3370,9300	276	72,5650	184,9323	3050,1400
227	43,6757	176,5397	3440,8000	277	70,8359	188,5857	3054,7700
228	47,9655	176,2704	3389,7900	278	67,4943	187,9392	3103,2000
229	44,7205	175,0090	3441,7000	279	64,9632	189,7653	3114,9000
230	37,9554	176,7410	3488,6000	280	64,9694	186,4048	3145,6000
231	40,0335	175,5458	3467,7000	281	54,7976	190,6004	3217,0800
232	38,5134	175,1015	3490,4000	282	59,1199	189,4709	3189,7200
233	36,0202	178,1812	3500,8500	283	55,0144	186,2439	3262,5400
234	31,6705	177,0931	3595,3500	284	58,2145	185,1430	3232,4200
235	31,3852	176,2322	3598,9200	285	78,8485	183,0084	2977,4700
236	31,3740	175,9840	3601,4300	286	83,9998	182,1043	2917,0000
237	32,2509	176,3251	3595,2000	287	72,6327	183,1903	3037,2600
238	36,2031	176,6796	3506,5600	288	74,6690	180,7681	3024,3000
239	36,0331	175,5648	3518,5000	289	76,9707	178,9576	2994,6600
240	34,2423	174,0109	3558,3600	290	67,3448	183,5905	3124,8900
241	25,6261	177,5859	3699,9000	291	72,0467	180,4738	3052,2400
242	29,3902	175,8118	3638,9000	292	70,7943	179,8872	3067,4500
243	20,6026	180,1650	3730,7000	293	62,3618	182,8856	3185,4300
244	22,2441	177,4778	3736,9500	294	55,3642	182,1297	3271,8300
245	19,3015	177,3677	3757,8100	295	58,2550	179,4164	3252,0500
246	19,9205	176,0804	3779,7400	296	87,3088	177,1509	2872,6900
247	19,2008	174,5002	3795,3300	297	85,2476	174,2296	2911,2800
248	21,8617	174,5023	3765,3000	298	78,4026	177,2962	2982,9000
249	40,4530	173,5373	3477,8300	299	83,4603	176,5234	2943,3500

300	81,6348	173,7151	2962,4000
301	76,4262	177,6049	2995,2000
302	72,8179	177,4682	3038,7200
303	74,9134	173,2984	3018,5000
304	71,3614	178,3805	3055,2800
305	66,1671	178,0281	3128,2000
306	70,1365	174,9350	3057,6700
307	54,8070	177,0526	3299,6600
308	86,7648	171,0738	2910,8000
309	87,3239	168,9427	2905,1800
310	80,1779	171,0207	2966,2000
311	80,8238	167,8911	2959,0000
312	75,2981	170,0515	3005,8400
313	76,5813	167,7688	2994,5800
314	72,1362	168,0447	3014,7000
315	76,5641	167,2716	2992,2000
316	71,7469	171,2886	3043,6200
317	61,6779	167,3454	3174,0000
318	88,3738	162,0726	2880,2800
319	79,8799	165,8026	2970,6500
320	78,9856	161,5968	2970,1000
321	75,1538	165,8270	2996,5000
322	71,7231	162,9643	3056,2500
323	74,5462	162,4894	3014,9800
324	70,8059	165,2248	3039,5000
325	59,0965	161,5782	3187,2400
326	46,7660	155,5366	3231,9000
327	20,9567	160,4509	3845,5600

## ANEXO 2

ID	X	Y	Nivel Piez.				
1	42,80	127,60	1464,00	50	11,70	18,70	2646,00
2	-27,40	90,80	2553,00	51	61,70	32,50	1739,00
3	-1,20	84,90	2158,00	52	69,60	33,80	1674,00
4	-18,60	76,50	2455,00	53	66,70	33,90	1868,00
5	96,50	64,60	1756,00	54	-36,70	150,90	1865,00
6	108,60	82,90	1702,00	55	-19,60	137,80	1777,00
7	88,40	56,50	1805,00	56	-21,30	131,80	1579,00
8	90,00	39,30	1797,00	57	-22,40	137,10	1771,00
9	93,20	33,10	1714,00	58	21,10	139,30	1408,00
10	97,60	56,30	1466,00	59	7,70	126,80	1527,00
11	90,60	35,10	1729,00	60	-8,30	107,80	2003,00
12	92,60	41,80	1638,00	61	56,70	171,30	1386,00
13	99,50	59,20	1736,00	62	59,00	164,50	1089,00
14	-24,10	184,80	1476,00	63	69,00	177,20	1384,00
15	-26,10	114,10	2200,00	64	70,90	161,40	1030,00
16	56,30	26,80	1999,00	65	73,00	163,00	1092,00
17	73,00	18,90	1680,00	66	59,70	170,10	1161,00
18	80,30	12,60	1806,00	67	61,90	174,30	1415,00
19	80,20	14,60	1682,00	68	63,70	173,90	1231,00
20	68,80	107,80	1306,00	69	5,60	79,10	2300,00
21	76,40	96,00	1722,00	70	18,20	77,40	2238,00
22	64,50	110,40	1437,00	71	85,70	139,80	1038,00
23	43,40	53,60	1828,00	72	105,10	132,00	1332,00
24	39,10	62,00	2118,00	73	-101,60	10,70	3510,00
25	112,80	45,50	1725,00	74	-145,20	28,00	3490,00
26	54,30	147,80	1606,00	75	-74,00	88,00	2594,00
27	6,10	48,30	2648,00	76	-94,50	86,60	2650,00
28	-3,80	40,40	2560,00	77	-88,80	76,70	2533,00
29	-2,20	29,90	2544,00	78	-120,30	80,80	3571,00
30	-2,40	33,80	2386,00	79	-86,00	54,40	2811,00
31	-2,20	33,70	2400,00	80	-72,80	43,10	2728,00
32	63,20	79,50	1757,00	81	-100,20	42,90	3136,00
33	-10,80	175,10	1402,00	82	-78,80	40,80	2553,00
34	-19,00	171,90	1364,00	83	-83,70	46,50	2796,00
35	-38,60	158,50	1735,00	84	-95,60	35,80	2691,00
36	83,10	159,10	1376,00	85	-87,60	29,40	2946,00
37	-21,80	15,00	2729,00				
38	-23,60	9,40	2766,00				
39	-20,10	22,10	2736,00				
40	-16,60	17,30	2432,00				
41	29,90	175,10	1024,00				
42	100,90	23,00	1611,00				
43	101,30	23,00	1548,00				
44	103,30	20,30	1591,00				
45	-14,30	31,30	2540,00				
46	-18,10	30,20	2352,00				
47	-18,10	29,50	2528,00				
48	-9,90	38,10	2575,00				
49	-12,20	39,10	2468,00				

## REFERÊNCIAS BIBLIOGRÁFICAS

- ABOUFIRASSI, M.; MARIÑO, M. Kriging of Water Levels in the Souss Aquifer, Marroco. **Mathematical Geology**, v.15, n.4, p.537–551, 1983.
- AGTERBERG, F. Georges Matheron: Founder of Spatial Statistics. **Earth Science History**, p.1–6, 2004.
- ANDRIOTTI, J. Introdução à Geoestatística. **Acta Geologica Leopoldensia**, v.11, n.27, p.5–81, 1988.
- ARMSTRONG, M. Problems with Universal Kriging. **Mathematical Geology**, v.16, n.1, p.101–108, 1984.
- ATKINSON, P. Geographical Information in Science: GeoComputation and nonstationarity. **Progress in Physical Geography**, v.25, n.1, p.111–122, 2001.
- BOLFARINE, H. **Introdução à Inferência Estatística**. Rio de Janeiro/RJ: Sociedade Brasileira de Matemática - SBM, 2010. 159p.
- BREGT, A.; MCBRATNEY, A.; WOPEREIS, M. Construction of Isolinear Maps of Soil Attributes with Empirical Confidence Limits. **Soil Science Society of America Journal**, v.55, n.1, p.14–19, 1991.
- CALVETE, J.; RAMÍREZ, J. **Geoestadística: Aplicaciones a la hidrología subterránea**. Barcelona-España: Centro Internacional de Métodos Numéricos en Ingeniería, 1996. 484p.
- CASELLA, G. **Inferência Estatística**. São Paulo: Cengage Learning, 2010. 588p.

- CHILÈS, J.; DELFINER, P. **Geostatistics: Modeling Spatial Uncertainty**. New Jersey: John Wiley and Sons, 1999. 695p.
- CLARK, I.; HARPER, W. **Practical Geostatistics**. Columbus/Ohio: Escosse North America Llc, 2000. 215p.
- CONCEIÇÃO, S. Discussão sobre a obtenção de funções semivariograma a partir de distribuições de probabilidades. Brasília, 2013. 63p. Dissertação (Mestrado) - Universidade de Brasília.
- CRESSIE, N. Geostatistics. **The American Statistician**, v.43, n.4, p.197–202, 1989.
- CRESSIE, N. **Statistical for Spatial Data**. New York: John Wiley & Sons, 1991.
- CRESSIE, N. **Statistical for Spatio-Temporal Data**. New Jersey: John Wiley & Sons, 2011. 215p.
- CRESSIE, N.; GRONDONA, M. A comparison of variogram estimation with covariogram estimation. In: MARDIA, K. (Ed.). **The Art of the Statistical Science**. New York: John Wiley and Sons, 1992. p.191–208.
- CRESSIE, N.; HAWKINS. Robust Estimation of the Variogram: I. **Mathematical Geology**, v.12, n.1, p.115–121, 1980.
- DRUCK, S.; CARVALHO, M.; CÂMARA, G.; MONTEIRO, A. **Análise Espacial de Dados Geográficos**. Brasília: EMBRAPA, 2004.
- FERREIRA, I.; SANTOS, G.; RODRIGUES, D. Estudo Sobre a Utilização Adequada da Krigagem na Representação Computacional de Superfícies Batimétricas. **Revista Brasileira de Cartografia**, v.65, n.5, p.831–842, 2013.
- FUENTES, M. **Geostatistic: Lectures Notes**. Madrid: Disponível em: <http://www.stat.ncsu.edu/people/fuentes/courses/madrid/lectures/geostatistics.pdf>, 2011.

GENTON, M. Highly Robust Variogram Estimation. **Mathematical Geology**, v.30, n.2, p.213–221, 1998.

GUERRA, P. **Geoestatística Operacional**. Departamento Nacional de Produção Mineral, 1988.

HARPER, W.; CLARK, I. Amarillo by morning: Data Visualization in Geostatistics. In: ICOTSS, 2010. **Data and Context in Statistics Education: Towards an evidence-based society. Proceedings of the Eighty International Conference On Teaching Statistics**; resumos. Ljubljana, Slovenia: In C. Reading, 2010. 6.

ISAAKS, E.; SRIVASTAVA, R. **An Introduction to Applied Geostatistics**. New York: Oxford University Press, Inc, 1989.

JOURNEL, A. Geostatistics: Models and Tools for the Earth Sciences. **Mathematical Geology**, v.18, n.1, p.119–141, 1986.

JOURNEL, A. **Fundamentals of Geostatistics in Five Lessons**. American Geophysical Union, 1989. 40p.

LANA, M. O. A. Método Multiquádrico como alternativa para predição de processos estocásticos. Campinas, 1989. 74p. Dissertação (Mestrado) - Instituto de Matemática da Universidade Estadual de Campinas-UNICAMP.

LANDIM, P.; STURARO, J.; MONTEIRO, C. **Krigagem Ordinária para Situações com Tendência Regionalizada**. Rio Claro: Disponível em: <http://www.rc.unesp.br/igce/aplicada/DIDATICOS/LANDIM/tkrigagem.pdf>, 2002.

LEUANGTHONG, O.; KHAN, D.; DEUTSCH, C. **Solved Problems in Geostatistics**. New Jersey: John Wiley and Sons, 2008. 207p.

LLOYD, C. **Local Models for Spatial Statistics Analysis**. Boca Raton, FL: CRC Press, Taylor and Francis Group, 2011. 215p.

MATHERON, G. **Krigeage Universel. Fascicule 1.** Fontainebleau: École Supérieure des Mines de Paris, Centre de Geostatistique, 1969. 84p.

MATHERON, G. **The Theory of Regionalized Variables and its applications.** Fontainebleau: École Supérieure des Mines de Paris, Centre de Geostatistique, 1971. 215p.

MYERS, D. To be or not to be...Stationary? That is the question. **Mathematical Geology**, v.21, n.3, p.347–362, 1989.

NEUMAN, S.; JACOBSON, E. Analysis of Nonintrinsic Spatial Variability by Residual Kriging with Applications to Regional Groundwater Levels. **Mathematical Geology**, v.16, n.5, p.499–521, 1984.

OLEA, R. **Geostatistics for Engineers and Earth Scientists.** Norwell, Massachusetts: Kluwer Academic Publishers, 1999. 303p.

OLEA, R. **A Practical Primer on Geostatistics.** Virginia: U.S. Geological Survey, 2009. 215p.

OLIVER, M.; WEBSTER, R. Geostatistics in Physical Geography. Part I: Theory. **Transactions of the Institute of British Geographers**, v.14, n.3, p.259–269, 1989.

PEREIRA, H. Estimaco de Funoes Aleatrias No Estacionrias. In: DE INVESTIGAO CIENTFICA, I. N. (Ed.). **Encontros sobre Mtodos Quantitativos Aplicados s Variveis Regionalizadas.** Lisboa: Instituto Nacional de Investigao Cientfica, 1983. p.115–131.

PHILIP, G.; WATSON, D. Matheronian Geostatistics: Quo Vadis? **Mathematical Geology**, v.16, n.1, p.93–117, 1986.

ROSSONI, D.; LIMA, R. A importncia da correo da Anisotropia em Anlises Geoestatsticas. **Revista da Estatstica da UFOP**, v.2, p.217–221, 2012.

SOARES, A. **Geoestatística para Ciências da Terra e do Ambiente**. Lisboa: IST Press, 2000. 206p.

SRIVASTAVA, R. Philip and Watson: Quo Vadunt? **Mathematical Geology**, v.18, n.1, p.141–146, 1986.

YAMAMOTO, J.; LANDIM, P. **Geoestatística: Conceitos e Aplicações**. São Paulo: Oficina de Textos, 2013. 215p.

YOST, R.; UEHARA, G.; FOX, R. Geostatistical Analysis of soil chemical properties of large land areas. II. Kriging. **Soil Science Society of America Journal**, v.46, p.1033–1037, 1982.