

numist

# MODELAÇÃO E ESTIMAÇÃO

Uma introdução à geoestatística

PEDRO CORREIA



2010

# ÍNDICE

RESUMO INTRODUTÓRIO.....	4
CARACTERIZAÇÃO DA DISPERSÃO ESPACIAL .....	4
MÉTODO COMPUTACIONAL DE CÁLCULO .....	6
VARIOGRAFIA .....	8
MODELOS DE VARIOGRAMAS .....	12
ESTIMAÇÃO POR KRIGAGEM.....	15
INPUT DE DADOS NUM SOFTWARE DE GEOESTATÍSTICA.....	17
EXERCÍCIO DE KRIGAGEM .....	21
INTRODUÇÃO AO S-GEMS .....	41
INTERFACE DO S-GEMS.....	41
PAINEL DE ALGORITMOS .....	42
PAINEL DE VISUALIZAÇÃO .....	45
COMANDOS BÁSICOS NO S-GEMS .....	48
INSERIR DADOS NO S-GEMS .....	48
FERRAMENTAS DO S-GEMS .....	51
UTILIZAÇÃO DO SOFTWARE S-GEMS NA ESTIMAÇÃO DO TEOR DE CONTAMINANTES NUMA HORTA .....	54
DADOS DA HORTA.....	54
PREPARAÇÃO DOS DADOS .....	55
ESTATÍSTICOS DESCRITIVOS .....	58
ESTUDO DE VARIOGRAFIA .....	60
KRIGAGEM.....	62
ANÁLISE DOS RESULTADOS .....	65
ESTIMAÇÃO DA IMPEDÂNCIA NUM RESERVATÓRIO PETROLÍFERO .....	67
DADOS DO RESERVATÓRIO.....	67
ESTATÍSTICOS DESCRITIVOS.....	67
ESTUDO DE VARIOGRAFIA .....	70
KRIGAGEM DA IMPEDÂNCIA.....	74
EXERCÍCIO 1 .....	77
EXERCÍCIO 2 .....	77
EXERCÍCIO 3 .....	78

EXERCÍCIO 4 .....	78
EXERCÍCIO 5 .....	79
BIBLIOGRAFIA.....	80

# RESUMO INTRODUTÓRIO

Vamos começar com uma breve introdução sobre estimação, nomeadamente com krigagem. Mais à frente poderemos ver todo o processo ao pormenor.

## CARACTERIZAÇÃO DA DISPERSÃO ESPACIAL

A interpolação espacial de valores amostrados pode ser feita com métodos de duas grandes classes: os determinísticos e o geostatísticos que mais à frente falaremos. Começemos por um exemplo prático no qual precisamos de caracterizar a dispersão espacial do fósforo numa horta com base em seis amostras (Figura 1).

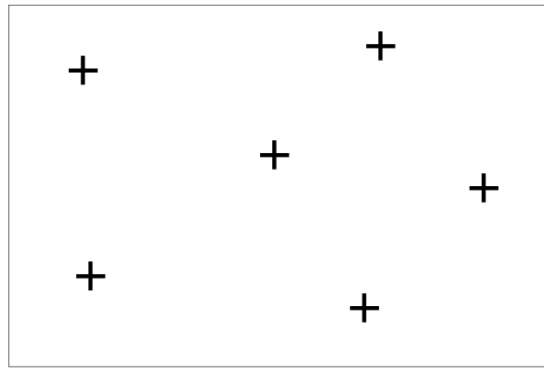


Figura 1 - Horta de onde se retiraram seis amostras para análise da quantidade de fósforo no solo.

Só existem seis locais que sabemos com toda a certeza a quantidade exacta de fósforo, tudo o resto permanece desconhecido. É necessário inferir qual o valor de fósforo que cada uma das zonas terá consoante a amostragem que foi feita. De entre os métodos determinísticos podemos utilizar, por exemplo, diagramas de Voronoi (método dos polígonos de influência) no qual cada ponto da horta toma um valor igual à amostra que se encontra mais próxima (Figura 2).

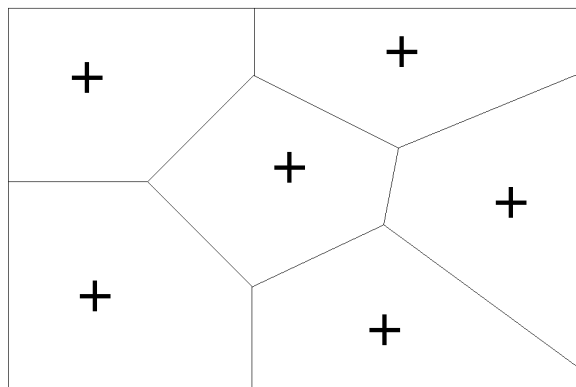


Figura 2 - Método dos polígonos de influência aplicado ao caso da horta.

Cada uma das arestas dos polígonos de influência representa o limite que define se um ponto se encontra mais próximo de uma amostra, ou de outra. Outro método bastante conhecido é o

de triangulação no qual cada local toma o valor, ponderado ou constante, da média feita pelos vértices de cada triângulo representativos das amostras.

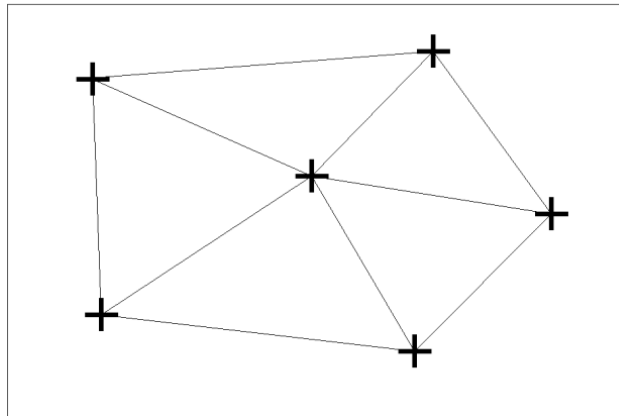


Figura 3 - Método da triangulação aplicado ao caso da horta.

Por último e, possivelmente, o caso mais usado de entre os métodos determinísticos é o do inverso da potência da distância (mais vulgarmente o do inverso do quadrado da distância por se usar a potência de 2 na estimação, Figura 4).

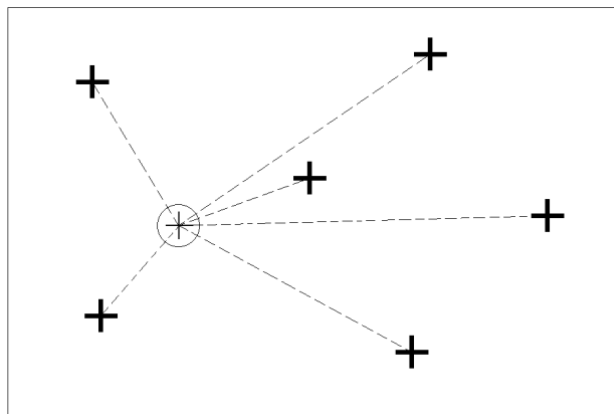


Figura 4 - Inverso da potência da distância aplicado ao caso da horta.

Cada ponto do espaço é calculado utilizando a seguinte fórmula ( $p = \text{potência}$ ):

$$z = \frac{\sum_{i=1}^N \frac{1}{d_i^p} z_i}{\sum_{i=1}^N \frac{1}{d_i^p}}$$

A aplicação destes métodos dá-nos um mapeamento da quantidade de fósforo na horta como podemos ver na Figura 5 onde foi aplicado o método dos polígonos de influência.

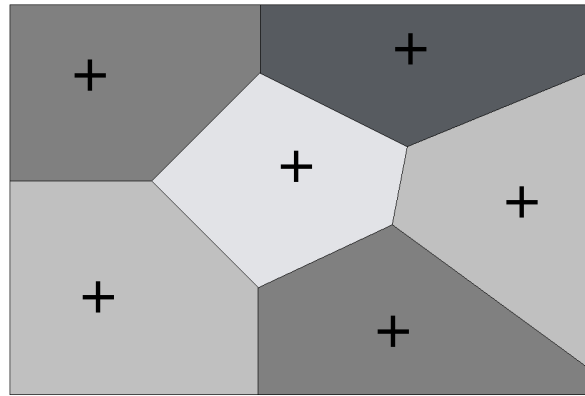


Figura 5 - Mapeamento da quantidade de fósforo na horta com base no método dos polígonos de influência.

Cada uma das cores representa uma dada quantidade de fósforo. Agora temos uma noção da dispersão espacial do fósforo na horta mas poderá a mesma ser melhorada? Métodos determinísticos de inferência espacial, como são os acima explicados, têm por base critérios estritamente geométricos quase arbitrários dado que a dispersão de um elemento numa dada área ou volume corresponde geralmente a um ou mais fenómenos específicos com continuidade dependente da direcção. Por esse e outros motivos se recorre, em muitos casos e especialmente nas ciências da terra e do ambiente, a métodos geostatísticos. Antes de explicar estes métodos e dado que os mesmos são, geralmente, efectuados em recurso a um computador, é importante explicar as limitações que o mesmo tem e que estruturas necessita para efectuar cálculos que, no fim, darão origem a um mapeamento de um qualquer fenómeno como é o nosso caso do fósforo na horta.

## MÉTODO COMPUTACIONAL DE CÁLCULO

Um computador é uma ferramenta com limitações e não consegue “compreender” o que é um espectro contínuo. Imaginemos que temos uma linha com uma amostra real no princípio e precisamos calcular valores ao longo da mesma. Podemos fazê-lo à distância que acharmos melhor, quantas vezes for necessário, sem necessidade de quebrar a continuidade da linha. O computador não consegue efectuar o mesmo tipo de raciocínio (Figura 6).

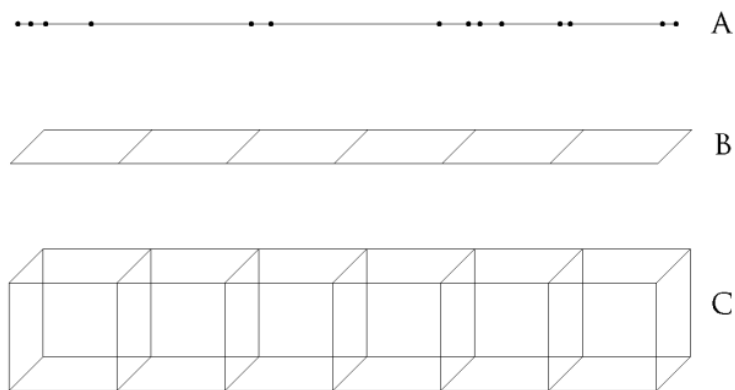


Figura 6 - Método computacional para discretizar uma entidade contínua.

Para colmatar esta deficiência o computador discretiza transformando uma linha (A) num aglomerado de intervalos (B) que, num mapa de inferência espacial, são reconhecidos como blocos (C). Se aplicarmos a mesma lógica para o caso da horta, que se entende bidimensional (ainda não foi considerado a profundidade), vamos ficar com uma discretização por blocos no espaço 2D (Figura 7).

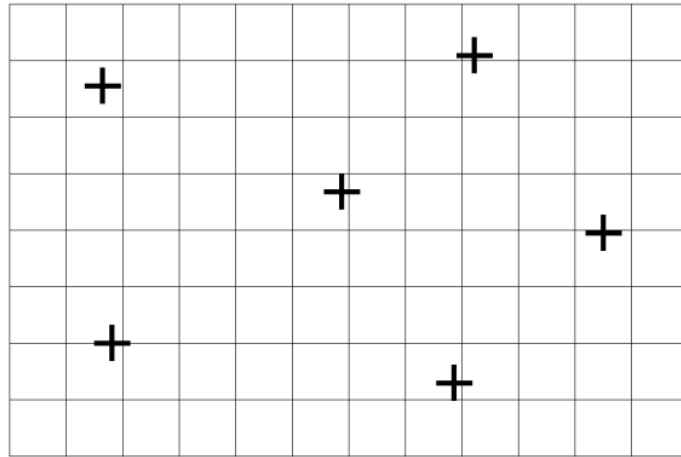


Figura 7 - Discretização em blocos da horta.

A partir daqui o computador irá efectuar um cálculo para cada bloco da discretização da horta. Se formos aplicar o método dos polígonos de influência ao caso da horta através do computador iremos ter o resultado da Figura 8.

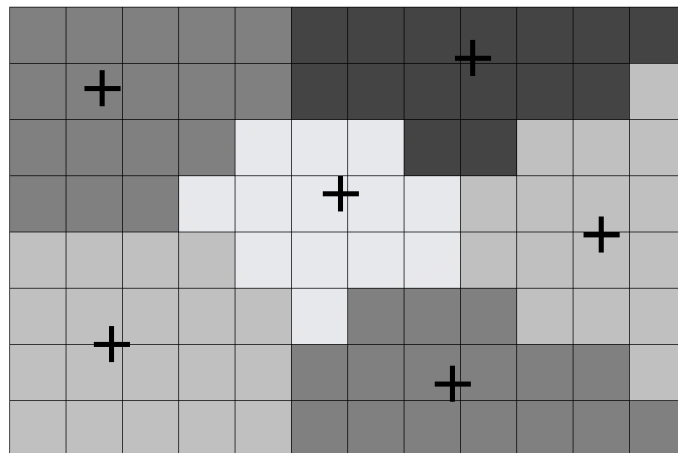


Figura 8 - Aplicação do método dos polígonos de influência num computador.

O tamanho dos blocos é definido pelo utilizador e diz-se ter uma malha mais apertada ou menos apertada consoante o mesmo for menor ou maior, respectivamente. Ao conjunto dos blocos é habitual chamar-se de grid. Evidentemente que o estudo não só pode ser feito no plano como no espaço 3D passando a existir a coordenada da profundidade. Neste caso a grid passaria a ter mais de uma camada de blocos (Figura 9).

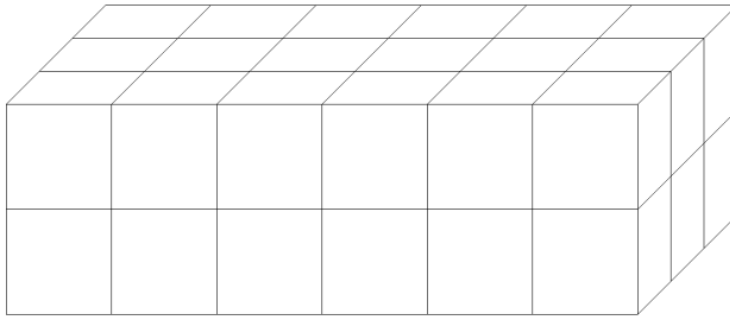


Figura 9 - Grid de duas camadas de 18 blocos cada.

Geralmente são utilizados os caracteres **I**, **J** e **K** para determinar o número de blocos que existe em cada direcção. Se o ângulo que a grid fizer em relação ao eixo **y** for zero então **I** corresponde a **x**, **J** a **y** e **K** a **z** (isto se cada unidade em **x**, **y**, **z** corresponder a um bloco em **I**, **J**, **K**, mas falaremos disso depois).

## VARIOGRAFIA

O variograma é uma ferramenta estatística de grande importância na geoestatística dado que nos dá uma medida da continuidade espacial mas comecemos pela análise inicial aos estatísticos básicos. Imaginemos que fazemos um histograma da amostragem da nossa horta e que o resultado é o da Figura 10.

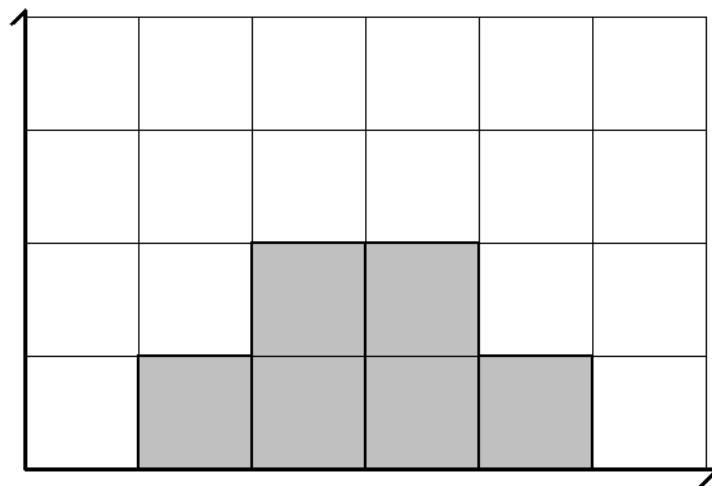


Figura 10 - Histograma dos dados amostrados da horta.

Imaginemos agora que o vizinho do lado da horta faz a mesma análise à sua horta e o seu histograma resulta num igual ao nosso. Quando foi feita a análise espacial da horta do vizinho resultou no mapa da Figura 11.



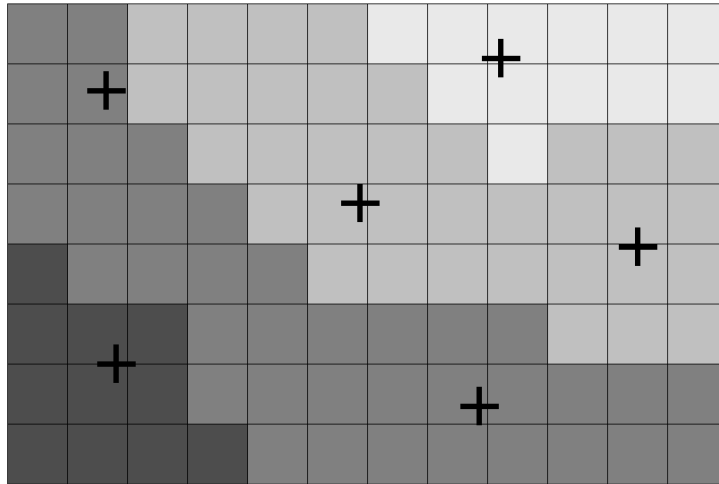


Figura 11 - Mapa da dispersão do fósforo na horta do vizinho.

Este caso nada tem a ver com o da nossa horta, no entanto os seus histogramas são iguais. De maneira a conseguir completar a informação dada pelos estatísticos básicos temos o variograma. O variograma é calculado pela seguinte expressão:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [Z(x_{\alpha}) - Z(x_{\alpha} + h)]^2$$

Onde  $x$  corresponde às coordenadas da posição  $\alpha$ ,  $Z(x_{\alpha})$  o valor atribuído à localização (no nosso caso seria valor de fósforo na localização  $x_{\alpha}$ ),  $h$  a distância a que se pretende efectuar uma estimação do mesmo valor,  $N(h)$  o número de pares de pontos para cada valor de  $h$ . Para efectuar este cálculo é necessário, mais uma vez, discretizar o espaço considerado e por esse motivo  $h$  vai corresponder a intervalos com diferentes distâncias. Aplicando isto no caso da horta e começando pela amostra superior direita ficamos com intervalos de variografia como os da Figura 12.

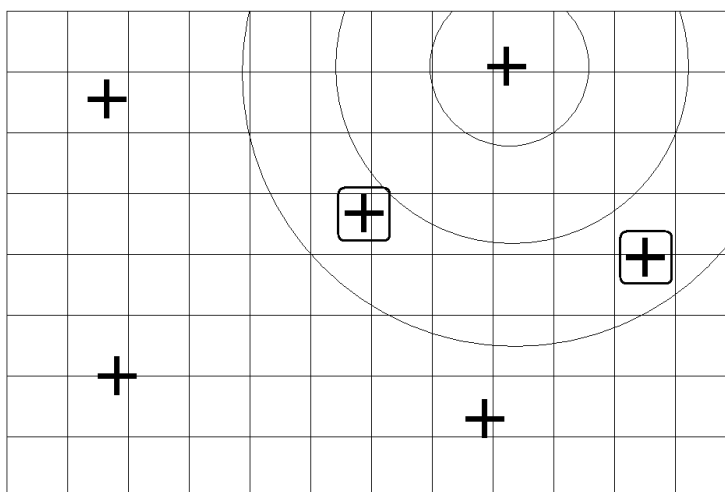


Figura 12 - Intervalos de variografia que poderiam ser aplicados à horta.

No primeiro e segundo intervalos considerados não existem amostra com que fazer o cálculo pelo ficarão intervalos vazios. No terceiro intervalo considerado já existem duas amostras

disponíveis e por isso o valor do variograma diferente de zero. Repetindo o mesmo processo para todas as amostras e considerando todos os intervalos de distância  $h$  no final teríamos que perfazer a sua média por intervalo (porque o intervalo três da segunda amostra pode dar um valor diferente do intervalo três da terceira amostra, por exemplo). O resultado do estudo de variografia seria o da Figura 13.

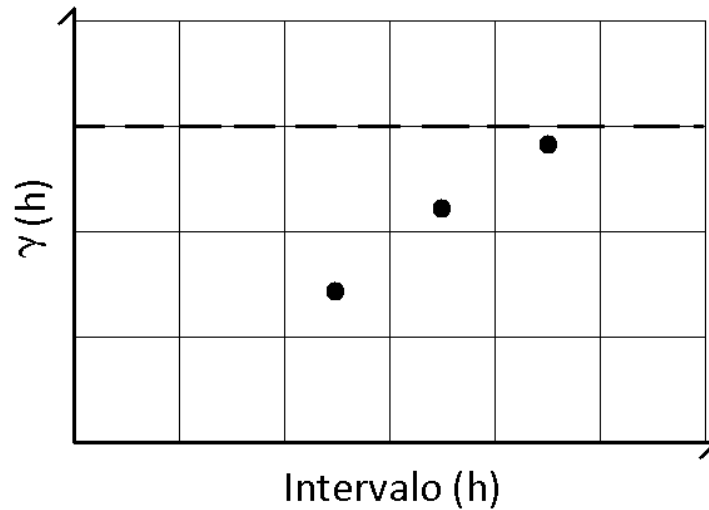


Figura 13 - Variograma omni-direccional do estudo feito à horta.

Neste caso o  $\gamma(h)$  resultou em médias diferentes para cada intervalo de distância  $h$  (mais adiante iremos explicar a análise da variografia). Este tipo de variograma é chamado de variograma omni-direccional porque não existe limite angular ou intervalo angular que delimite uma direcção preferencial de estudo. A dispersão espacial de uma variável, como é o fósforo na nossa horta, pode ser isotrópica como é o exemplo da Figura 14 onde não existe direcção preferencial.

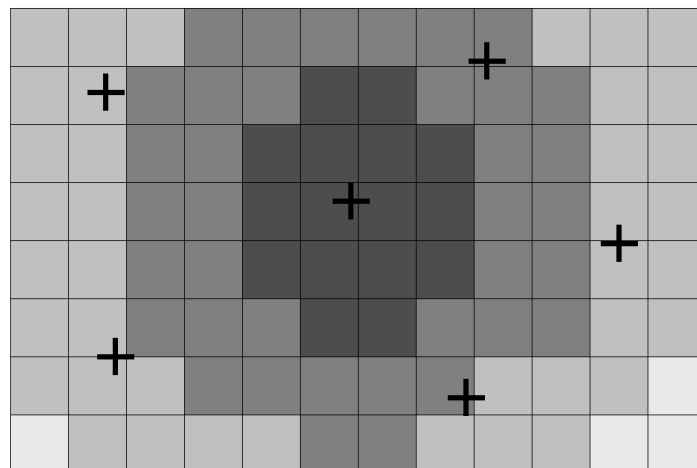


Figura 14 - Isotropia na dispersão espacial de uma variável.

Ou anisotrópica como é o exemplo da Figura 15 onde a direcção preferencial é claramente da esquerda para a direita pois apresenta muito maior continuidade que de cima para baixo.

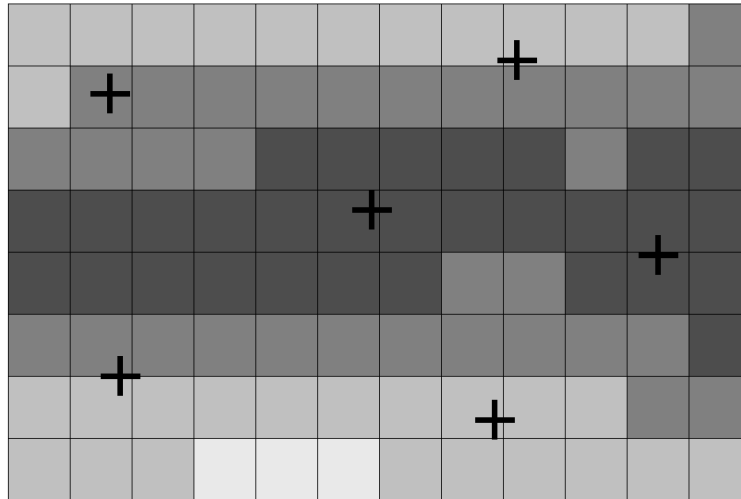


Figura 15 - Anisotropia na dispersão espacial de uma variável onde se nota muito maior continuidade da esq. para a dir. que de cima para baixo.

Sendo o comportamento da variável isotrópico então um variograma omni-direccional faz sentido mas se não for é necessário especificar em que direcção se faz o estudo. Para isso precisamos de alguns parâmetros essenciais como é o caso do tamanho de cada intervalo e da tolerância angular que existe em relação à direcção que se pretende estudar como se pode ver na Figura 16.

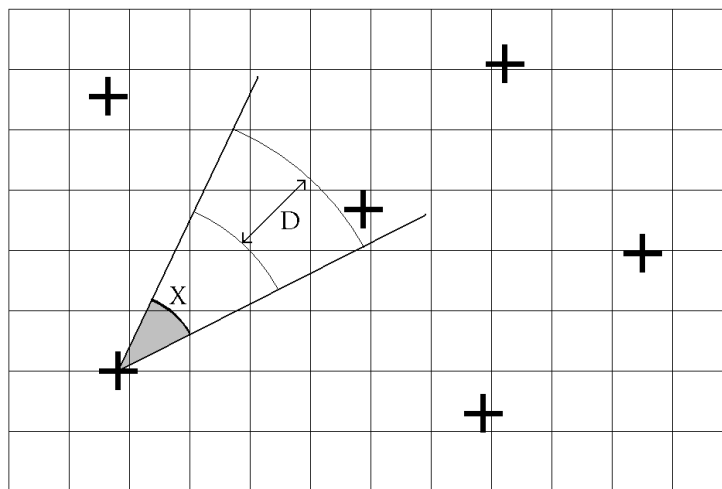


Figura 16 - Intervalos de variografia com tamanho D e tolerância angular X.

Assim temos um variograma para a direcção escolhida. Se fizermos vários variogramas para várias direcções ficamos com uma noção da continuidade espacial que existe em cada direcção (isto se existirem dados amostrados que o permitam, podem existir direcções onde não existem quaisquer amostras para serem estudadas). Dado que estes cálculos serão efectuados pelo software mais à frente abordaremos mais parâmetros de variografia.

## MODELOS DE VARIOGRAMAS

O estudo de variografia dá-nos informação sobre a continuidade mas não nos é possível estimar blocos em distâncias para a qual não existem pontos que apareçam no variograma. Para resolver este problema é necessário impor uma função ou modelo que corresponda o mais possível aos pontos do variograma e consiga reproduzir novos pontos em distâncias onde não existe nenhum. Na Figura 17 está um modelo ajustado (embora não muito bem ajustado) aos pontos do variograma omni-direccional da nossa horta. Com a inclusão desta curva no variograma passamos a ter qualquer  $\gamma(h)$  (que no nosso caso podemos interpretar como sendo o valor indirecto de fósforo) a qualquer distância  $h$  porque o modelo é contínuo e por isso sempre possível de se usar.

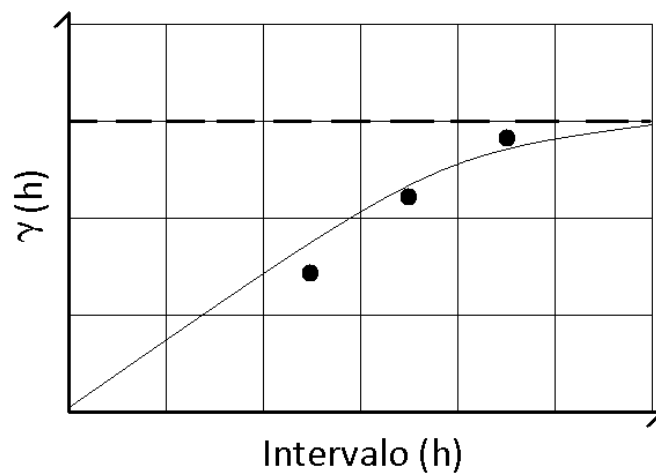


Figura 17 - Modelo ajustado ao variograma omni-direccional da horta.

Os modelos de adequação a variogramas mais usados são o exponencial, gaussiano e potencial. As diferenças que têm de uns para os outros estão no tipo de curva que os caracteriza. O modelo exponencial e potencial têm apenas uma concavidade, o modelo gaussiano por outro lado tem duas perfazendo uma forma de “s” no variograma. Na Figura 18 estão ilustradas as diferenças mais óbvias entre estes três modelos.

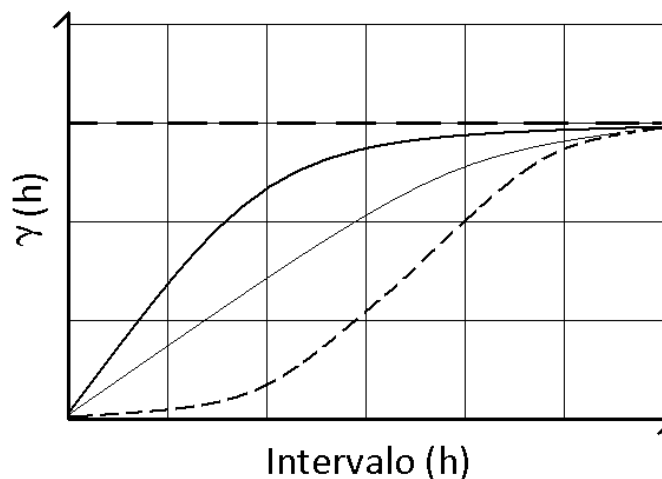


Figura 18 - Modelo exponencial (esq.), potencial (centro) e gaussiano (dir.) ajustados num variograma.

Qualquer software de geoestatística já vem preparado para conseguir calcular os modelos a partir dos parâmetros amplitude, patamar e efeito pepita (Figura 19).

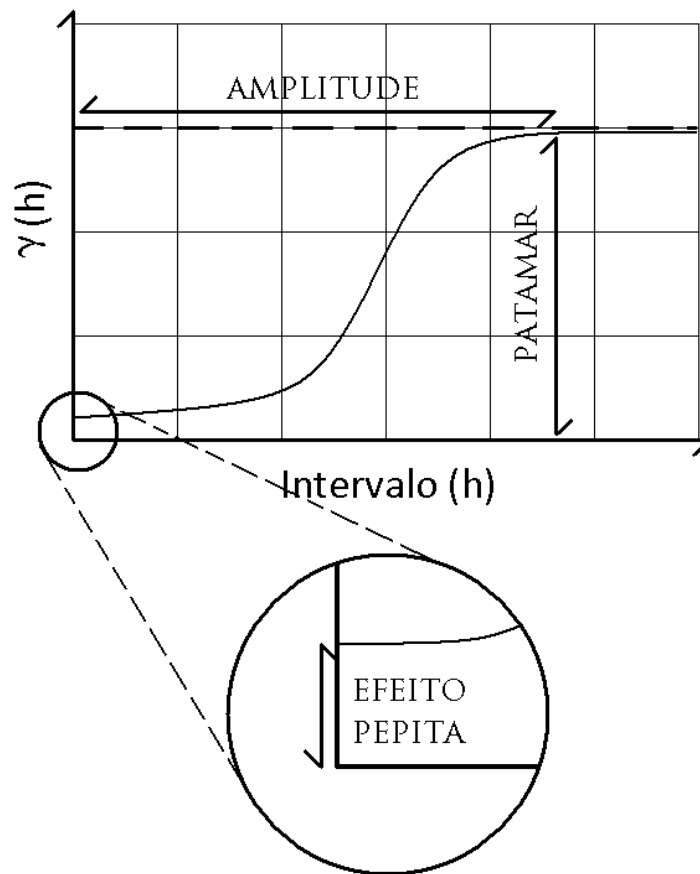


Figura 19 - Parâmetros patamar, amplitude e efeito pepita na adequação de um modelo ao variograma.

A amplitude é a distância da origem ao ponto onde o modelo toca o patamar. O patamar corresponde ao  $\gamma(h)$  para o qual acima do mesmo já não existe qualquer correlação (geralmente considera-se a variância da amostragem como valor do patamar, embora este procedimento depende muito de outros factores que neste documento não iremos abordar). O efeito pepita é por vezes inserido e principalmente quando são usados modelos gaussianos, para evitar más estimações devido a descontinuidades nas zonas mais próximas da origem. Na maioria dos casos, usando o modelo exponencial ou potencial este efeito é negligenciável e por isso sem necessidade de ser aplicado. Para que estes modelos possam ser utilizados para estimação é usado o conceito de elipse (em 2D) ou elipsóide (em 3D). Em termos práticos o elipsóide (como doravante o tratarei para qualquer caso) faz a função para a direcção que o modelo faz para a distância. Para descrever o elipsóide apenas precisamos de 2 (no caso 2D) ou 3 (3D) direcções definidas com modelos mas ele consegue interpolar qualquer outra direcção a partir dessas. Na Figura 20 temos uma ilustração da aparência do elipsóide em 2D para o caso isotrópico e anisotrópico.

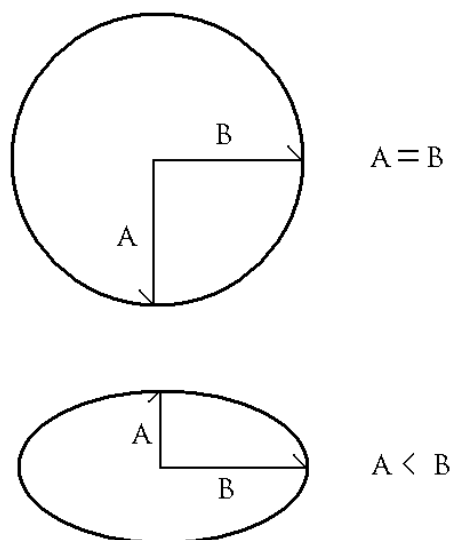


Figura 20 - Elipsóide (elipse ou esfera por ser um caso 2D) de um caso isotrópico (cima) e anisotrópico (baixo).

Os comprimentos dos eixos A e B são correspondentes às amplitudes dos modelos respectivos dos variogramas das direcções do “A” e do “B”. No caso isotrópico as distâncias são iguais, no caso anisotrópico as distâncias são diferentes. Para fazer um elipsóide é necessário que as direcções que o definem sejam perpendiculares entre si. Se optarmos por utilizar uma direcção de  $45^\circ$  em relação ao norte então a outra direcção terá que ser necessariamente  $-45^\circ$  pois  $45 - (-45) = 90$ , condição necessária da perpendicularidade. Na Figura 21 está exemplificada a interpolação de todas as direcções com base nas direcções ortogonais de definição. Sem o elipsóide o software de geoestatística não teria como saber a amplitude para direcções diferentes daquelas que são as de definição do elipsóide.

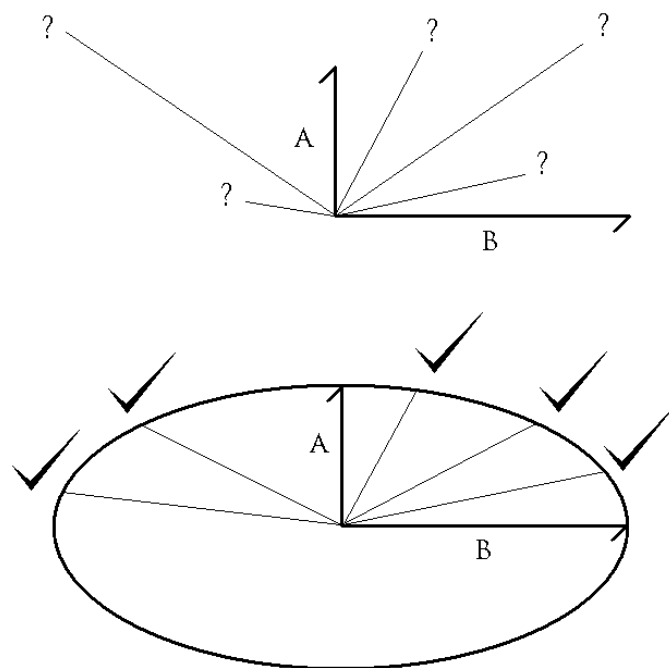


Figura 21 - Necessidade do elipsóide para interpolar direcções não conhecidas para o modelo.

Até agora os casos apresentados foram sempre bidimensionais mas é importante compreender que a direcção é, regra geral, dada por coordenadas esféricas o que implica sempre dois parâmetros: rotação ( $\varphi$ ) e inclinação ( $\theta$ ). Cada direcção é representada, por este motivo, por  $(\varphi, \theta)$  no qual se o nosso caso de estudo for no plano horizontal então  $\theta$  é sempre zero. Se as nossas direcções de estudo na horta (2D) forem a  $45^\circ$  e a  $-45^\circ$  então os nossos variogramas direccionais serão  $(45,0)$  e  $(-45,0)$ . Entenda-se que ambos os ângulos têm, normalmente um intervalo de  $]90, -90]^\circ$  pois a direcção é lida nos dois sentidos como se pode ver na Figura 22.

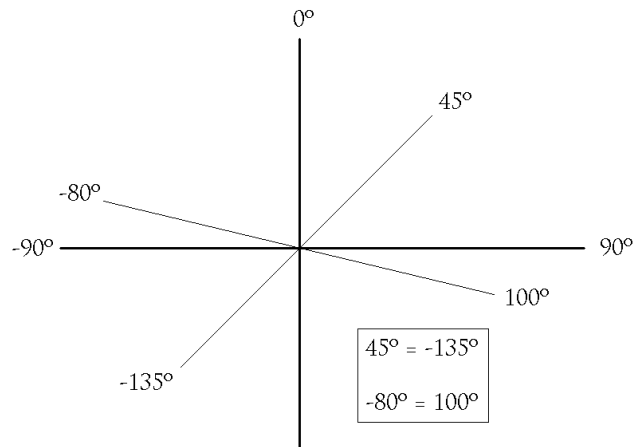


Figura 22 - Ângulo é lido nos dois sentidos da mesma direcção e por esse motivo  $45^\circ$  é igual a  $-135^\circ$ , por exemplo.

O ângulo de  $45^\circ$  é igual ao de  $-135^\circ$  e o de  $-80^\circ$  ao de  $100^\circ$ . Assim quando pretendemos inserir uma direcção o fazemos somente num intervalo de  $180^\circ$  como é o caso do  $]90, -90]^\circ$ .

## ESTIMAÇÃO POR KRIGAGEM

Krigagem é um método de estimação à semelhança do inverso do quadrado da distância, polígonos de influência, etc. Vamos tentar usar um exemplo prático e simples para explicar o que acontece na krigagem. Na Figura 23 temos um caso no qual temos três amostras para estimar a restante área por método de krigagem.

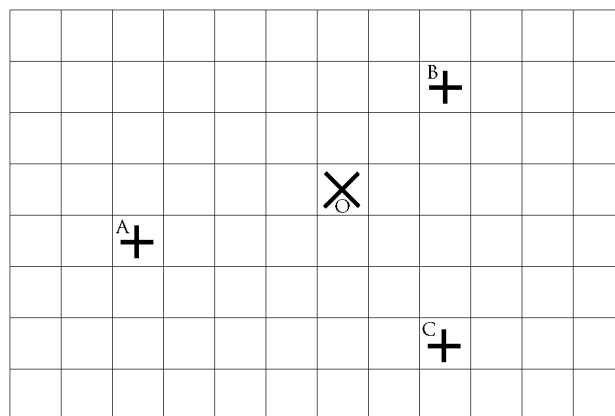


Figura 23 - Amostragem de três pontos que irão servir para calcular um quarto (ponto O) a partir do método krigagem.

Para calcular o ponto O (e depois os restantes pelo mesmo processo) vamos fazer um estudo de variografia à amostragem ao qual adaptamos um modelo potencial a dois pontos de variografia num variograma omni-direccional por o estudo das várias direcções revelar que se trata de um caso isotrópico (Figura 24).

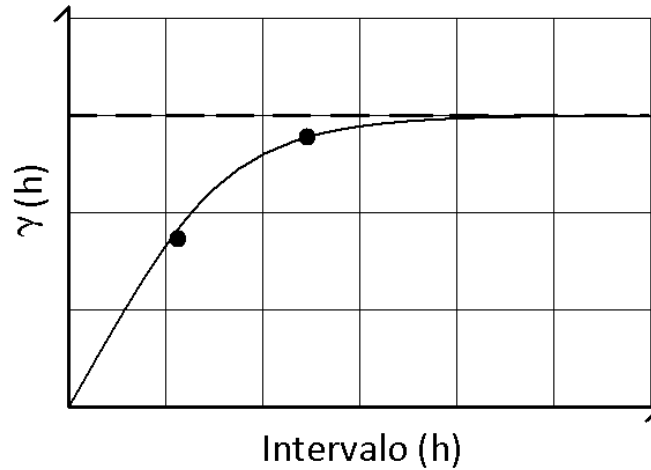


Figura 24 - Modelo potencial ajustado a dois pontos do variograma.

O nosso patamar corresponde à variância e a amplitude resultante é de  $\omega$ . Agora já nos é possível saber o valor de  $\gamma(h)$  para qualquer distância das amostras ao ponto a estimar mas ainda falta saber o valor da amplitude para qualquer direcção que pretendamos. Assim fazemos o elipsóide a partir de duas quaisquer direcções ortogonais usando a mesma amplitude para as duas (caso contrário não seria isotrópico) como se pode ver na Figura 25.

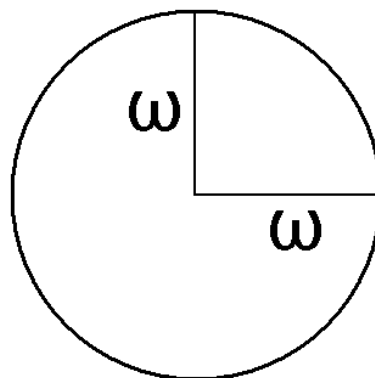


Figura 25 - Construção de elipsóide com duas direcções de definição ortogonais com a mesma amplitude  $\omega$ .

Agora já temos o valor de  $\gamma(h)$  para qualquer distância em qualquer direcção. Sabendo que as três amostras vão ser usadas para estimar o ponto O é necessário atribuir um peso a cada uma delas de maneira a que a direcção, e conseqüentemente o modelo, bem como a distância estejam consideradas na estimação do ponto O. Para fazer isto vamos comparar as amostras entre si para determinar o peso que cada uma deverá ter no ponto a ser estimado. Assim aplicando o seguinte sistema (existe ainda o multiplicador de Lagrange “ $\lambda$ ” para otimizar os resultados do sistema, mas dado que se trata de um exemplo para fins de explicação vamos evitar complicações):



$$\begin{array}{lll} W_A\gamma(h_{AA}) & W_B\gamma(h_{AB}) & W_C\gamma(h_{AC}) = \gamma(h_{AO}) \\ W_A\gamma(h_{BA}) & W_B\gamma(h_{BB}) & W_C\gamma(h_{BC}) = \gamma(h_{BO}) \\ W_A\gamma(h_{CA}) & W_B\gamma(h_{CB}) & W_C\gamma(h_{CC}) = \gamma(h_{CO}) \end{array}$$

Com o acrescento da linha para evitar erros (o somatório dos pesos não pode dar maior ou menor que 1):

$$W_A + W_B + W_C = 1$$

Onde  $W_X$  é o ponderador correspondente à amostra  $X$ . O  $\gamma(h_{EV})$  é o valor de variograma para a distância  $h$  entre  $E$  e  $V$  com o modelo nessa mesma direcção. Após descobrir-se os ponderadores para cada amostra considerada na estimação calculamos o ponto  $O$  com a seguinte equação:

$$P_O = W_A P_A + W_B P_B + W_C P_C$$

Onde  $P_X$  é o valor de fósforo na localização  $X$ . Agora já sabemos o valor estimado no ponto  $O$  (ou bloco). Repetindo o mesmo processo para todos os blocos ficamos com a nossa estimação por krigagem. Existe mais de um tipo de krigagem mas aqui apenas se pretende ensinar o princípio básico da mesma. Seguidamente iremos fazê-lo em recurso a software especializado.

## INPUT DE DADOS NUM SOFTWARE DE GEOESTATÍSTICA

Antes de começarmos a trabalhar no software é importante perceber com que tipo de input o mesmo funciona. Existem dois grandes tipos de informações que podemos ceder a um software de geoestatística e modelação. Input de “point-data” e input de “grid”. Começemos pelo primeiro. Point-data são pontos referenciados espacialmente com um ou mais valores de variáveis. No caso da nossa horta os point-data seriam as amostras, apresentando as coordenadas  $x$ ,  $y$  e o respectivo valor (leia-se percentagem ou rácio) de fósforo na localização onde cada uma se encontra. O excerto abaixo tem três colunas onde a primeira é a coordenada em  $x$ , a segunda a coordenada em  $y$  e a terceira é a percentagem de fósforo em cada uma das localizações.

0.9	1.0	0.15
1.0	3.2	0.20
3.0	2.3	0.07
3.8	0.7	0.21
4.0	3.4	0.30
5.0	2.0	0.17

O ficheiro de input é formato ASCII e portanto editável em software como o bloco de notas, por exemplo. O ficheiro teria de ter, então, apenas a informação do excerto acima com possível adição de cabeçalho que alguns softwares exigem para funcionarem (mais à frente falaremos disto). Se utilizarmos a informação do excerto e a computarmos num gráfico de

dispersão com folha de cálculo, por exemplo, vamos ter o resultado da Figura 26 onde é especificado apenas a localização das amostras como já anteriormente tínhamos visto (neste exemplo não está especificado teor, se estivesse cada bola correspondente à amostra teria uma cor correspondente a um teor de fósforo).

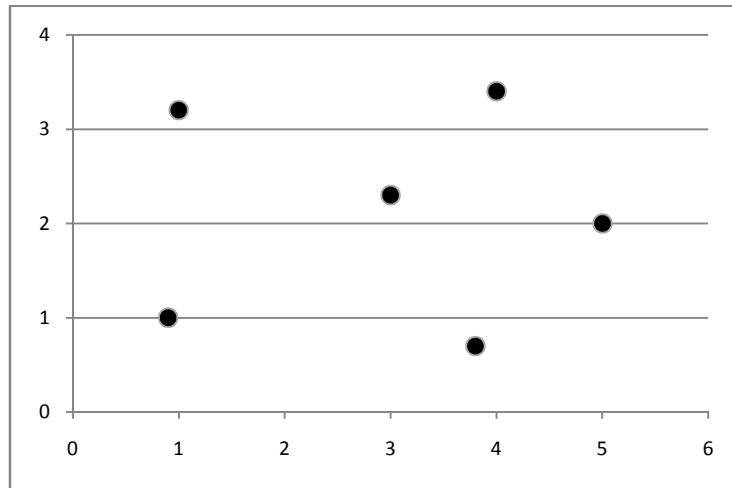


Figura 26 - Computação do ficheiro dos point-data das amostras da horta (sem evidenciação do teor de fósforo).

Se quiséssemos meter mais variáveis nas mesmas localizações do point-data da horta (porque poderiam ter resultado outros elementos contaminantes na análise das amostras) teríamos que ter mais uma coluna. Mais à frente explicaremos detalhadamente como fazer a krigagem em recurso a software mas após termos dado como input o point-data das amostras recolhidas da horta (correspondente ao excerto), criado uma grid que considerasse a área coberta pelas amostras, e feito o respectivo estudo de variografia resultou no mapeamento da Figura 27 onde o azul representa os baixos teores e os vermelhos altos teores.

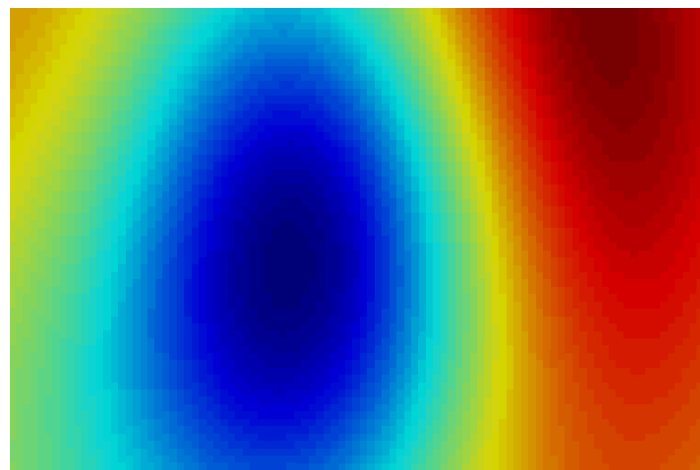


Figura 27 - Mapa estimado por krigagem dos teores de fósforo na horta onde o azul corresponde a baixo teor e o vermelho a alto teor.

Após feita a krigagem no software utilizado precisamos dar entrada deste resultado noutra software e por esse motivo exportamos a grid da horta (em que cada bloco tem um valor de fósforo proveniente da krigagem). O ficheiro de output poderá vir no formato que se mostra no seguinte excerto:

0.16  
0.17  
0.19  
0.22  
0.23  
0.24  
0.23  
0.22  
...

Vem apenas uma coluna com correspondência dos valores de fósforo em cada bloco sem qualquer referência da localização dos mesmos e por isso sem possibilidade organizar os dados no espaço. Regra geral os softwares têm um algoritmo que especifica a que posição de uma grid corresponde cada um dos valores de fósforo que passo agora a explicar. No caso da horta criamos uma malha ou grid com 98 blocos no eixo do  $x$  e 64 no eixo do  $y$ . Sendo que  $64 * 98 = 6272$  então temos 6272 blocos na nossa grid e por isso 6272 valores de fósforo no nosso ficheiro de output. Se fizermos deste ficheiro um de input num software de geoestatística e especificarmos que o input é para ser dado numa grid com as mesmas características acima dadas então o software irá recolocar todos os valores nos blocos respectivos sem necessidade de ter colunas para as coordenadas. O algoritmo usado é, geralmente, o seguinte. Primeiro é percorrida linha do eixo  $x$  ou  $I$ , depois a do  $y$  ou  $J$ , e finalmente a do  $z$  ou  $K$  como se vê na Figura 28.

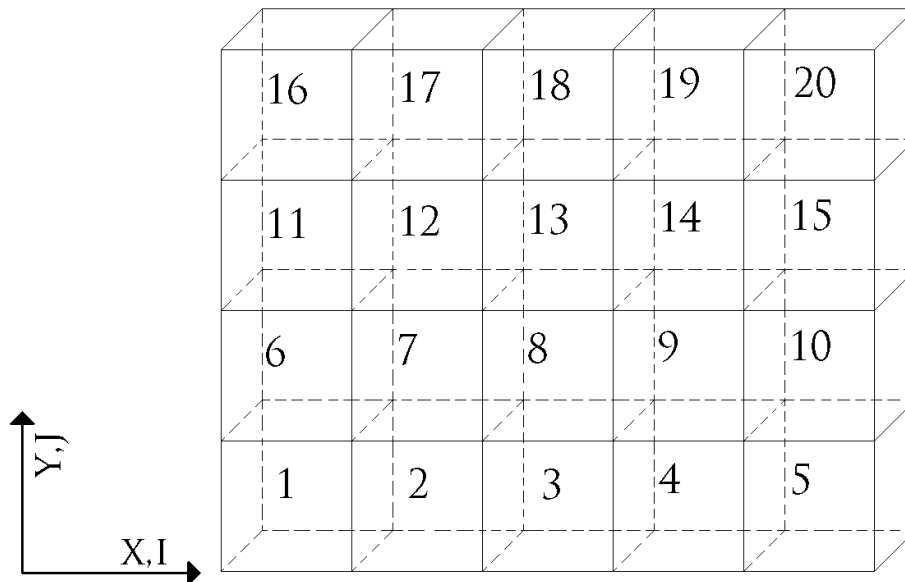


Figura 28 - Algoritmo comum utilizado na reorganização dos ficheiros de grid.

O primeiro valor que aparecer no ficheiro do output da horta corresponderá à posição  $(x,y)$ ,  $(1,0)$ , o segundo valor à posição  $(2,0)$ , o terceiro à  $(3,0)$  e finalmente até completar um linha do eixo  $x$  que fica na posição  $(98,0)$ . O valor a seguir ficará na posição  $(1,1)$  e depois  $(2,1)$ , e por ai

adiante até ter completado a primeira (ou única no caso da horta por não existir profundidade e por isso o número de blocos em z é 1) camada. Se o caso fosse tridimensional então após completar a primeira camada o software iria completar a segunda como se vê na Figura 29 (e se z tivesse mais de 2 blocos iria continuar fazendo camadas até todas estarem completas).

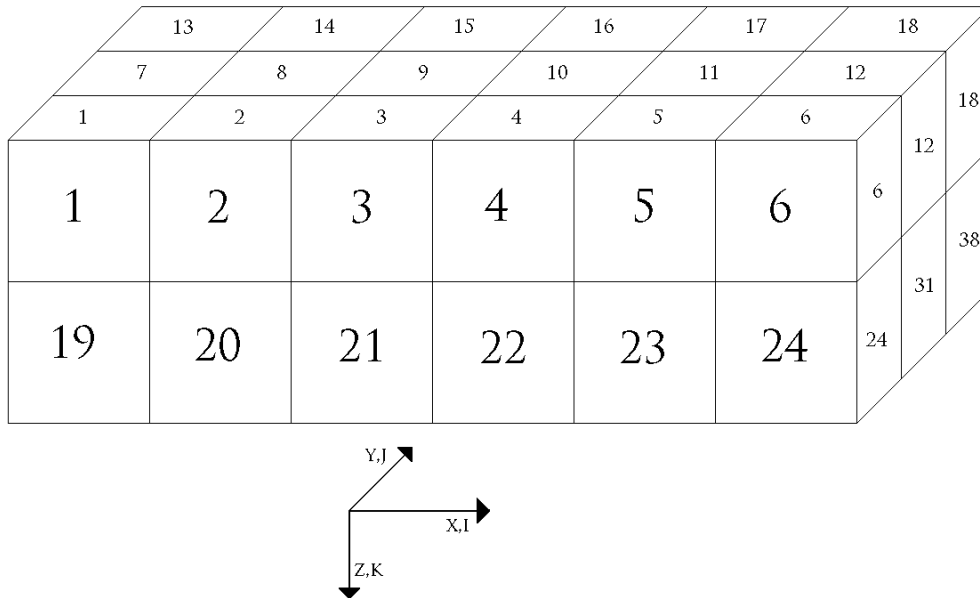


Figura 29 - Reorganização dos ficheiros de grid num caso tridimensional.

Não sendo os únicos estes tipos de ficheiros são os mais comuns.

## EXERCÍCIO DE KRIGAGEM

Ao longo deste exercício vamos explicar ao pormenor todo o processo da krigagem, desde o tratamento inicial dos dados, à construção dos variogramas e finalmente a estimação. Todo o exercício é feito à mão (quanto muito utilizando a calculadora) por isso podes tu também ir buscar um lápis e papel para acompanhar. Por agora não vai haver computadores.

O variograma (semi-variograma) é o gráfico (e/ou fórmula) que descreve a diferença esperada de valor entre pares de amostras dada uma certa orientação. Vamos agora fazer os variogramas experimentais ou calculados. Imaginemos que temos um jazigo de ferro do qual foi feita uma amostragem (Figura 30).

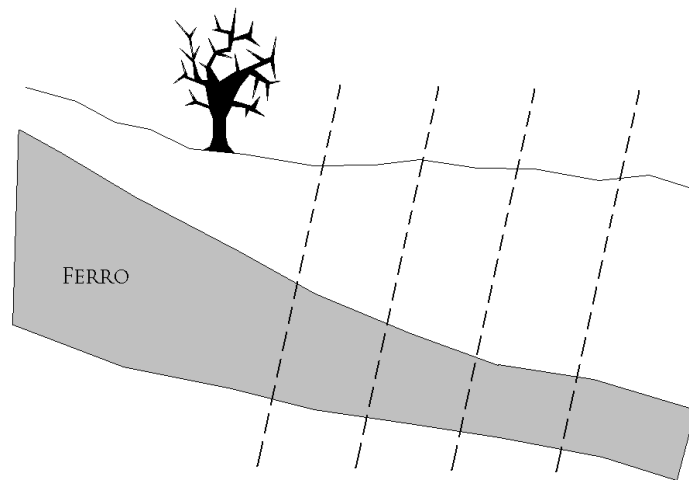


Figura 30 - Jazigo de ferro a ser estudado com a amostragem ilustrada a tracejado.

A malha das sondagens é regular, com distâncias entre amostras iguais (Figura 31).

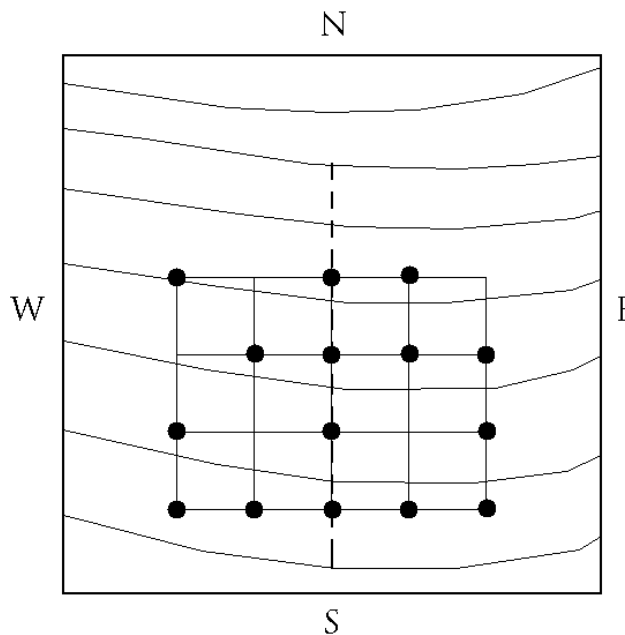


Figura 31 - Localização das sondagens na planta do terreno.

Foram tirados os teores, não considerando a profundidade, para fazer uma estimação bidimensional sobre a planta (Figura 32). As amostragens foram feitas a 100 metros de distância umas das outras na direcção “Este-Oeste” e “Norte-Sul”.

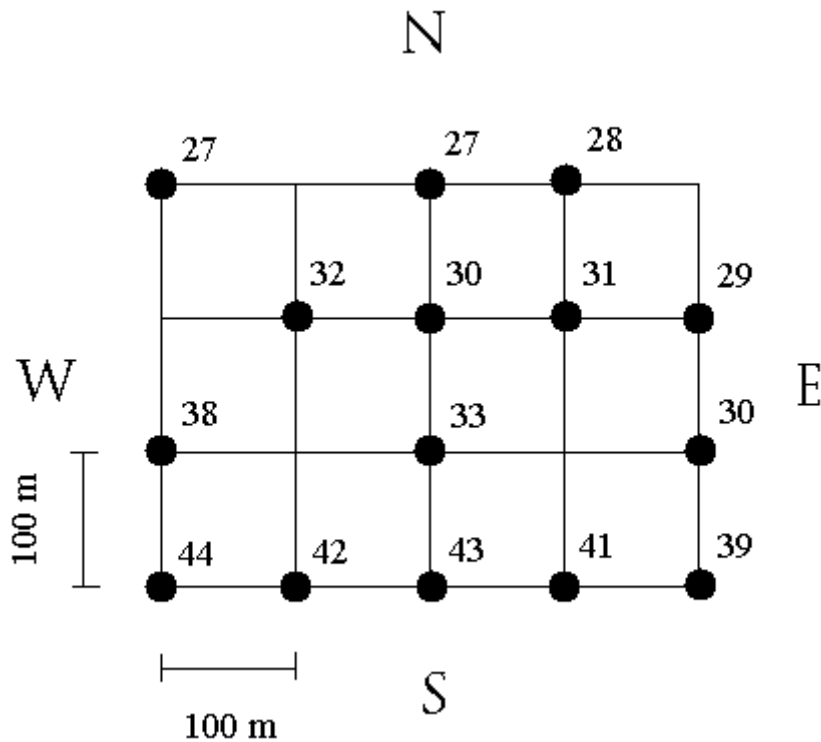


Figura 32 - Distância entre sondagens e teor de ferro em cada uma delas.

Vamos considerar o cálculo do variograma na direcção “Este-Oeste” para a distância de 100 metros. Os pares de amostras que existem à distância de 100 metros são os da Figura 33.

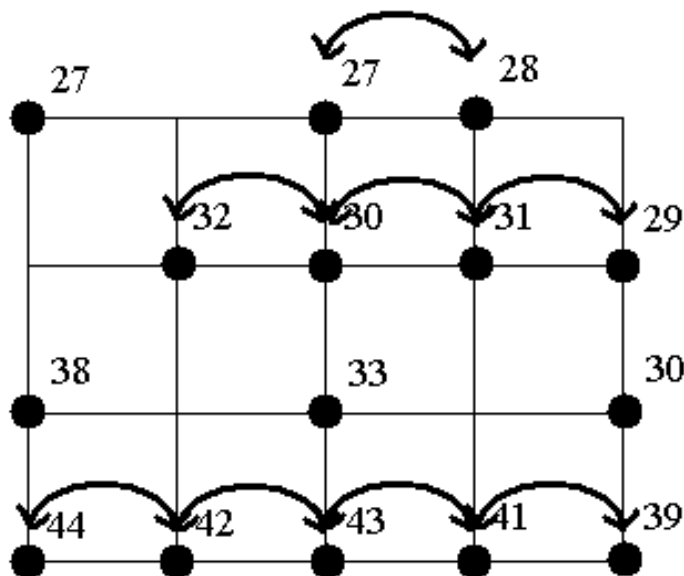


Figura 33 - Pares de pontos na direcção "Este-Oeste" à distância de 100 metros.

Considerando que a fórmula do variograma é:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [Z(x_{\alpha}) - Z(x_{\alpha} + h)]^2$$

Então o nosso cálculo irá ser, para oito pares de pontos:

$$\gamma(100) = [(27 - 28)^2 + (32 - 30)^2 + (30 - 31)^2 + (31 - 29)^2 + (44 - 42)^2 + (42 - 43)^2 + (43 - 41)^2 + (41 - 39)^2] / (2 * 8)$$

$$\gamma(100) = 1,1875$$

Consideremos agora a mesma direção para a distância de 200 metros (Figura 34).

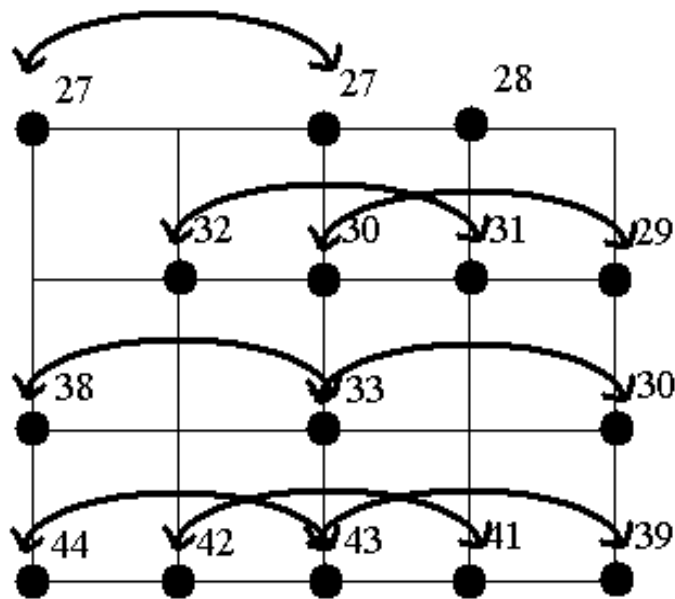


Figura 34 - Pares de pontos na direção "Este-Oeste" à distância de 200 metros.

O cálculo do variograma para oito pares de pontos irá ser:

$$\gamma(200) = [(27 - 27)^2 + (32 - 31)^2 + (30 - 29)^2 + (38 - 33)^2 + (33 - 30)^2 + (44 - 43)^2 + (42 - 41)^2 + (43 - 39)^2] / (2 * 8)$$

$$\gamma(200) = 3.375$$

Considerando a distância de 300 metros (Figura 35):

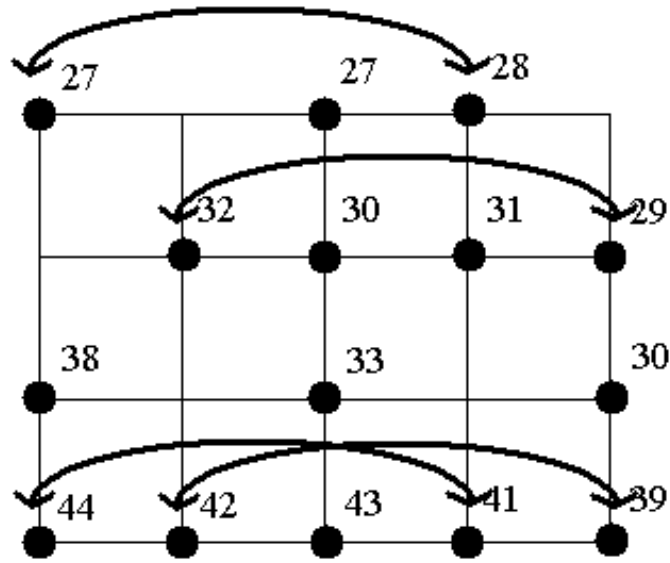


Figura 35 - Pares de pontos na direção "Este-Oeste" à distância de 300 metros.

O cálculo do variograma irá ser para quatro pares de pontos:

$$\gamma(300) = [(27 - 28)^2 + (32 - 29)^2 + (44 - 41)^2 + (42 - 39)^2] / (2 * 4)$$

$$\gamma(300) = 3,5$$

Fazendo apenas para três classes à distância de 100, 200 e 300 metros ficamos com o variograma experimental da Figura 36 (repare-se que tem o número de pares de pontos para cada distância).

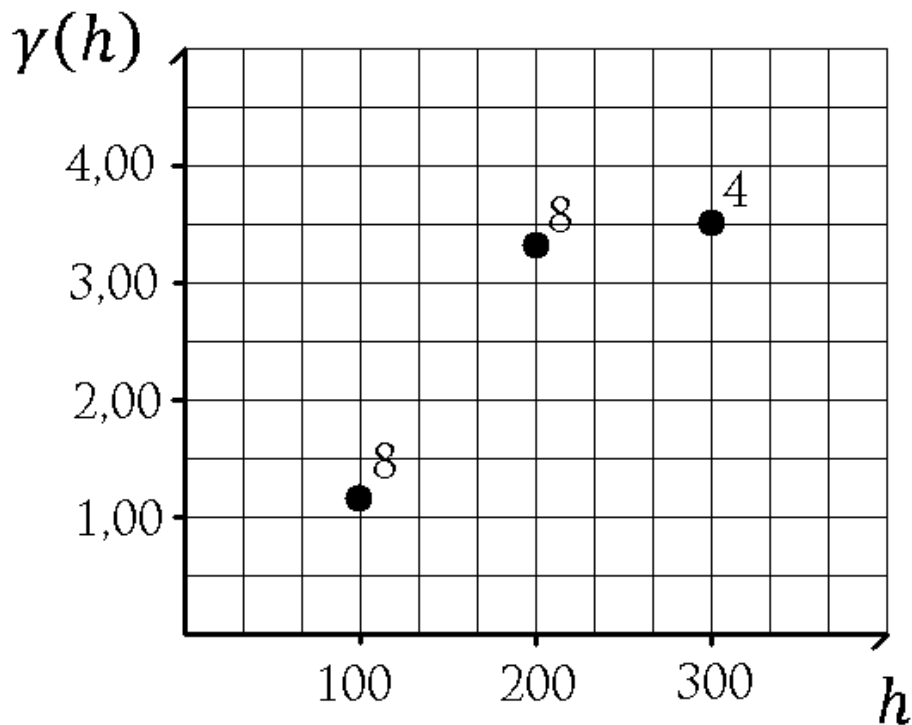


Figura 36 - Variograma experimental da direção "Este-Oeste".



Consideremos agora a direcção "Norte-Sul" para a distância de 100 metros (Figura 37).

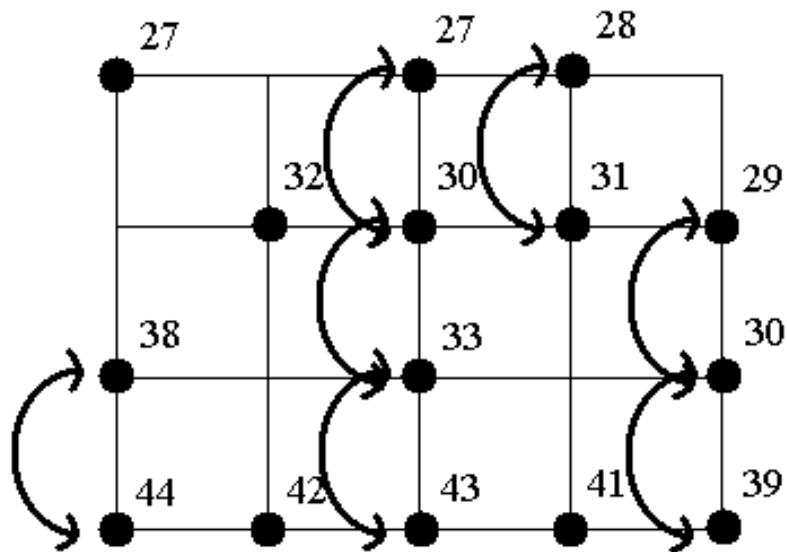


Figura 37 - Pares de pontos na direcção "Norte-sul" à distância de 100 metros.

O cálculo do variograma para sete pares de pontos é:

$$\gamma(100) = [(38 - 44)^2 + (27 - 33)^2 + (30 - 33)^2 + (33 - 43)^2 + (28 - 31)^2 + (29 - 30)^2 + (30 - 39)^2] / (2 * 7)$$

$$\gamma(100) = 19.4286$$

Considerando a distância de 200 metros na mesma direcção (Figura 38).

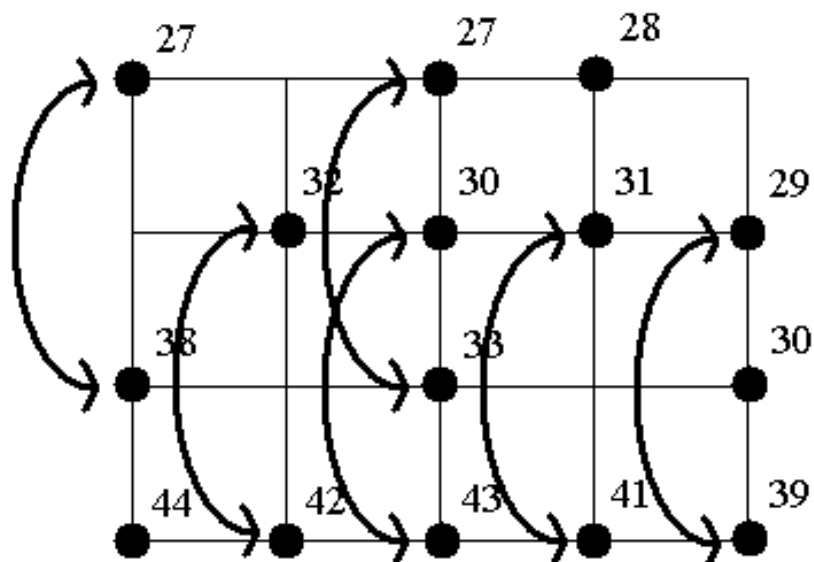


Figura 38 - Pares de pontos na direcção "Norte-sul" à distância de 200 metros.

O cálculo do variograma para seis pares de pontos é:

$$\gamma(200) = [(27 - 38)^2 + (32 - 42)^2 + (27 - 33)^2 + (30 - 43)^2 + (31 - 41)^2 + (29 - 39)^2] / (2 * 6)$$

$$\gamma(200) = 48,83(3)$$

Finalmente temos a distância de 300 metros na direcção "Norte-Sul" (Figura 39).

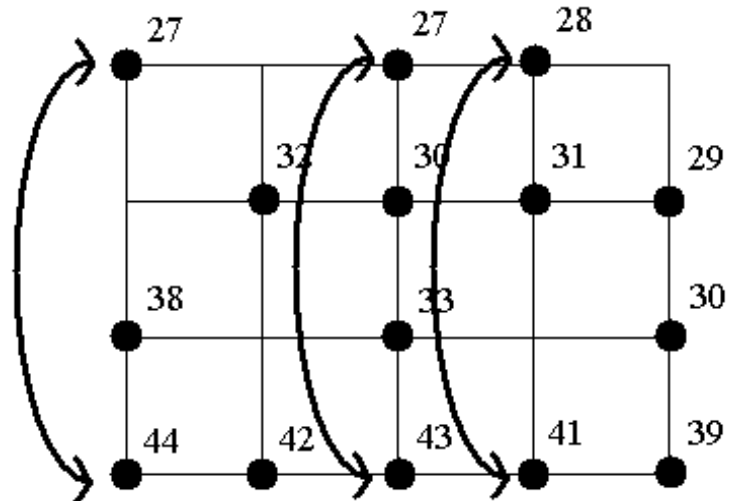


Figura 39 - Pares de pontos na direcção "Norte-sul" à distância de 300 metros.

O cálculo do variograma para três pares de pontos é:

$$\gamma(300) = [(27 - 44)^2 + (27 - 43)^2 + (28 - 41)^2] / (2 * 3)$$

$$\gamma(300) = 119$$

O variograma experimental vai ser então o da Figura 40.

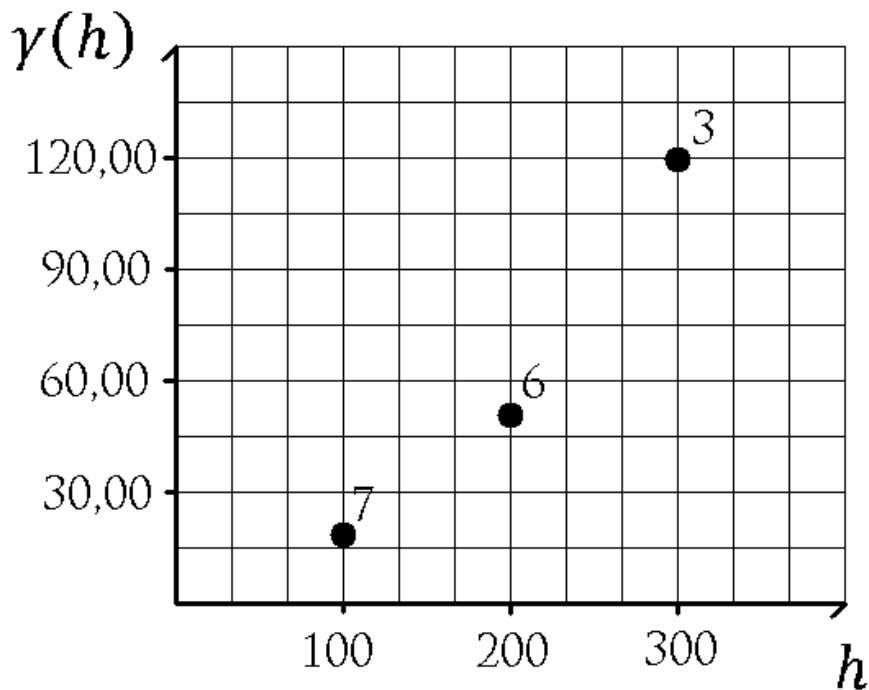


Figura 40 - Variograma experimental da direcção "Norte-sul".

Podíamos continuar a estudar outras direcções com outras distâncias (para um estudo mais completo era o que devia ser feito) mas em favor da brevidade vamos ficar por estas direcções. Agora é necessário completar os nossos variogramas com os estatísticos básicos. Para fazer os estatísticos básicos começamos por usar os nossos dados iniciais (Tabela 1).

<b>X</b>	<b>Y</b>	<b>Fe</b>
1	4	27
3	4	27
4	4	28
5	3	29
3	3	30
5	2	30
4	3	31
2	3	32
3	2	33
1	2	38
5	1	39
4	1	41
2	1	42
3	1	43
1	1	44

Tabela 1 - Dados iniciais do nosso exercício.

A partir destes fizemos classes de teor de ferro e a frequência de existia em cada uma delas (Tabela 2).

<b>Classes</b>	<b>Frequência</b>
<b>[27, 31]</b>	7
<b>]31, 35]</b>	2
<b>]35, 39]</b>	2
<b>]39, 44]</b>	4

Tabela 2 - Classes de teores e respectivas frequências.

A partir daqui já conseguimos construir um histograma (Figura 41).

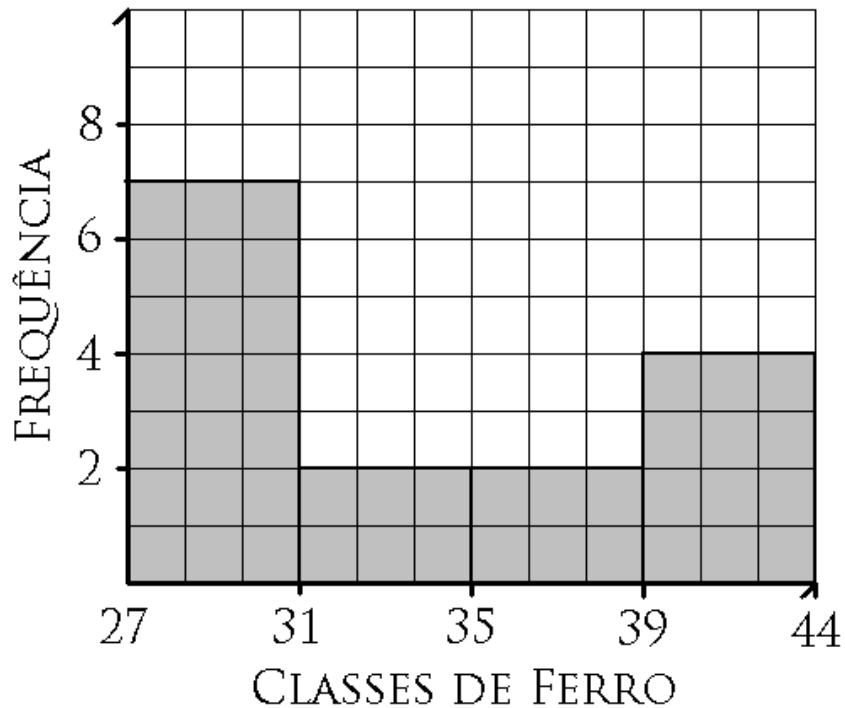


Figura 41 - Histograma dos nossos dados iniciais.

O histograma revela uma distribuição curiosa mas dado a escassez de dados não vamos assumir nada. A nossa amostragem é de malha regular e tem boa representatividade na região que pretendemos estudar, por esse motivo vamos usar a variância como patamar para os nossos modelos de variografia. A variância é dada por (sendo  $X$  a variável de amostragem):

$$Var(X) = E((X - \mu)^2)$$

No qual  $\mu = E(X)$  e  $E(X)$  o valor esperado. Em casos como o nosso, no qual tratamos todos os valores como tendo igual probabilidade, o valor esperado é a média aritmética.

$$E(X) = \frac{1}{n} \sum_{i=1}^n X_i$$

E  $\mu$ , portanto, a média aritmética da nossa amostragem. Assim a variância vai ser a média aritmética do quadrado das diferenças de todas as amostras para a média.

$$Var(X) = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$$

No nosso caso a média aritmética é igual a:

$$E(X) = \frac{(27 + 27 + 28 + 29 + 30 + 30 + 31 + 32 + 33 + 38 + 39 + 41 + 42 + 43 + 44)}{15}$$

$$E(X) = 34,2667$$

A partir da média calculamos a variância:

$$Var(X) = [(27 - 34,2667)^2 + (27 - 34,2667)^2 + (28 - 34,2667)^2 + (29 - 34,2667)^2 + \dots + (43 - 34,2667)^2 + (44 - 34,2667)^2]/15$$

$$Var(X) = 38,49524$$

Temos agora o histograma, média e variância como estatísticos básicos. Ao metermos a variância nos variogramas experimentais ficamos com o patamar (Figura 42).

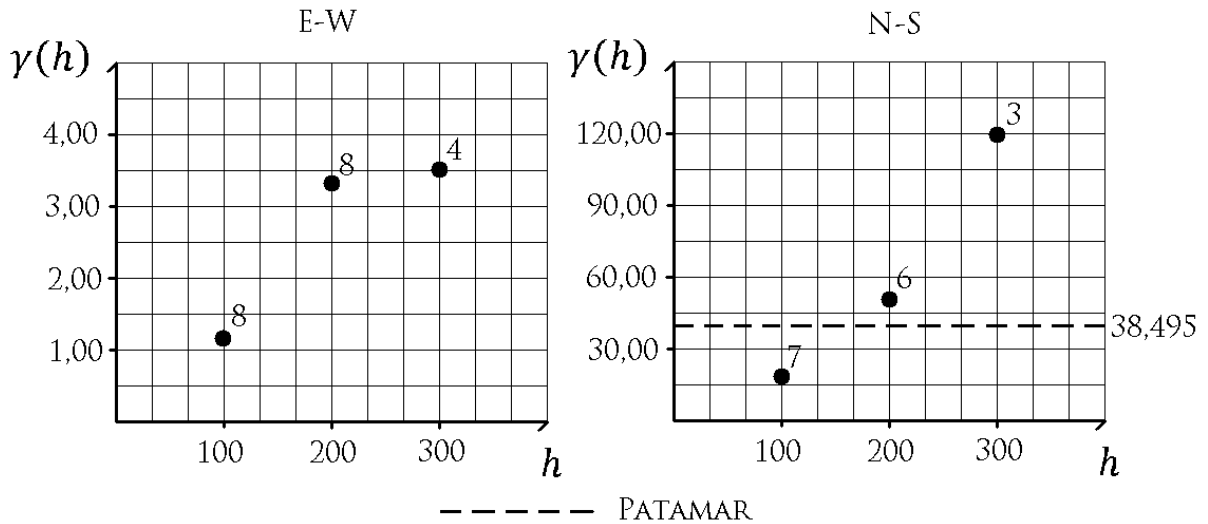


Figura 42 - Variogramas experimentais com patamar inserido.

Repare-se que na direcção “Norte-Sul” o patamar está presente a tracejado mas na direcção “Este-Oeste” não. Ele existe nesta direcção, no entanto os valores dos variogramas são demasiado baixos para que o patamar seja visto a esta escala. Por comparação com as duas direcções podemos dizer que a continuidade na direcção “Este-Oeste” é muito superior à direcção “Norte-Sul”. Ao fazermos uma nova escala para o variograma da “Este-Oeste” ficamos com o resultado da Figura 43.

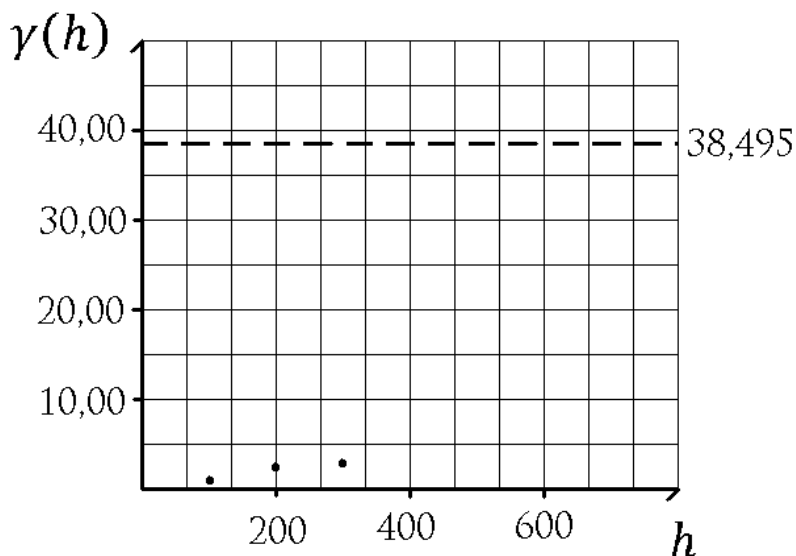


Figura 43 - Variograma da direcção "Este-Oeste" com uma nova escala que permite ver o patamar.

Agora é necessário ajustar um modelo aos variogramas experimentais. Existem muitas funções possíveis de se usar em variografia mas vamos comentar apenas algumas das mais utilizadas. Se considerarmos que  $C$  é o patamar e  $a$  a amplitude então os modelos terão que ser definidos com estas constantes (muito embora no ajuste do modelo num software o que se varia é a amplitude, mas para todos os efeitos as funções do modelo são de variável  $h$ , melhor dizendo distância). A título de exemplo vamos explorar as hipóteses modelo *Gaussiano* e modelo *Esférico*. A fórmula do modelo gaussiano é a seguinte:

$$\gamma(h) = C \left( 1 - e^{\left(-3\frac{h^2}{a^3}\right)} \right)$$

Uma maneira de ajustar o modelo manualmente ao variograma é invertendo a função para descobrir qual a amplitude  $a$  sabendo  $C, h, \gamma(h)$ . Assim ficamos com:

$$\gamma(h) = C \left( 1 - e^{\left(-3\frac{h^2}{a^3}\right)} \right) \equiv \ln \left( 1 - \frac{\gamma(h)}{C} \right) = -3 \frac{h^2}{a^2} \equiv a = - \frac{h}{\sqrt{\ln \left( 1 - \frac{\gamma(h)}{C} \right) / 3}}$$

Evidentemente numa altura em que podemos fazer isto em recurso a um computador não é razoável (excepto para fins pedagógicos) ajustar o modelo o manualmente. Na Figura 44 temos o aspecto visual do modelo gaussiano.

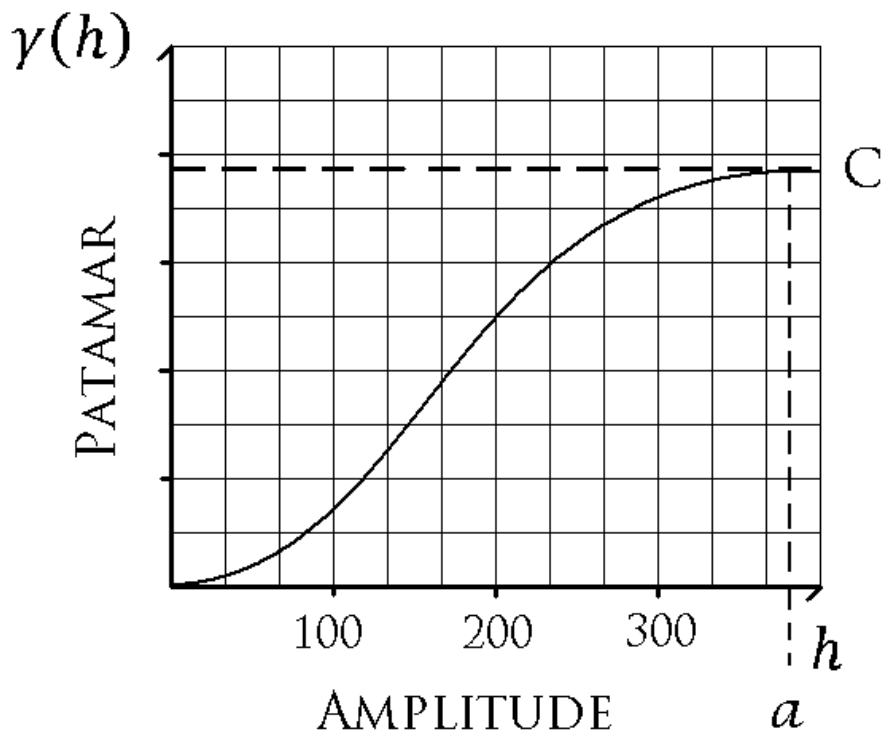


Figura 44 - Aspecto do modelo gaussiano no gráfico de variograma.

Num software o ajuste do modelo é feito fazendo variar a amplitude (dado que o patamar é fixo e que o ajuste é, regra geral, a olho para os pares de ponto que dependem da distância e

variograma). A variação da amplitude em relação ao exemplo da figura anterior vai resultar em funções idênticas às da Figura 45 (a tracejado as variadas, a preenchido a original).

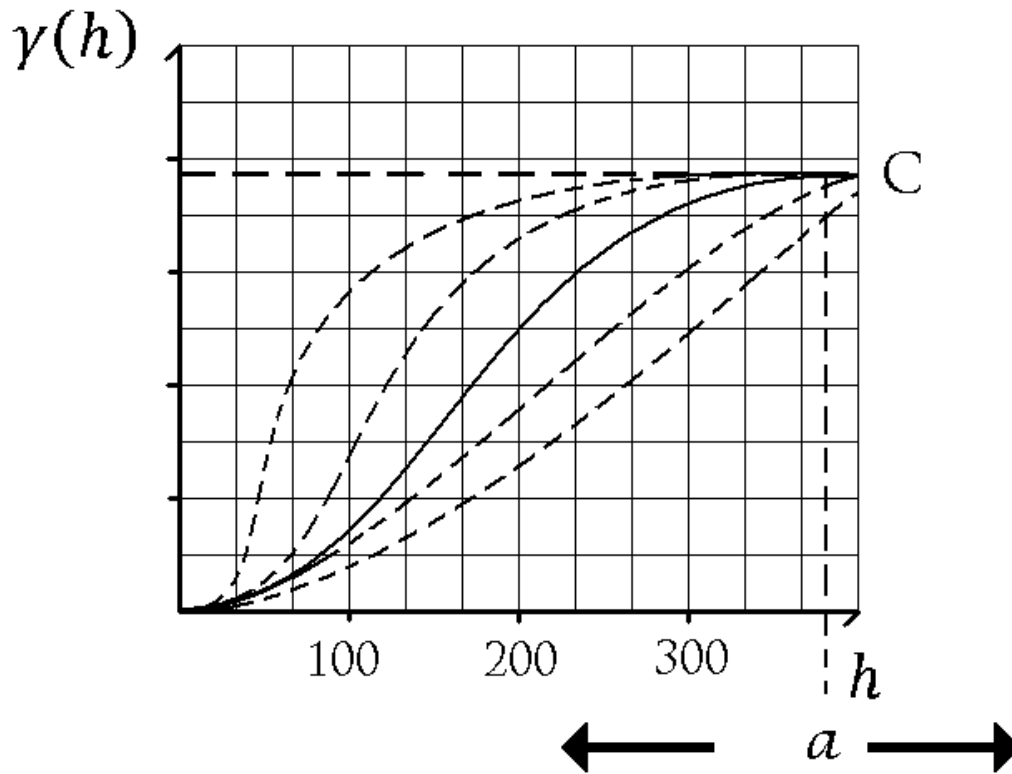


Figura 45 - Ajuste do modelo gaussiano variando a amplitude.

Se por outro lado tentarmos a aplicação da exponencial a sua fórmula é a seguinte:

$$\gamma(h) = C \left( 1 - e^{\left(-\frac{3h}{a}\right)} \right)$$

A inversão desta função resulta nas seguintes operações:

$$\gamma(h) = C \left( 1 - e^{\left(-\frac{3h}{a}\right)} \right) \equiv \ln \left( 1 - \frac{\gamma(h)}{C} \right) = -3 \frac{h}{a} \equiv a = -3 \frac{h}{\ln \left( 1 - \frac{\gamma(h)}{C} \right)}$$

O aspecto visual da exponencial é o da Figura 46.

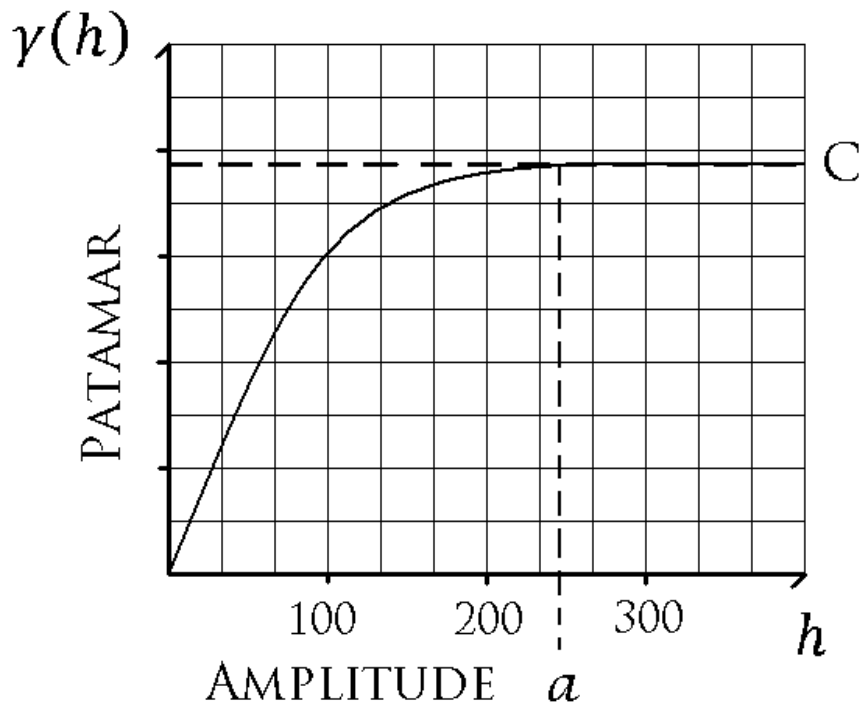


Figura 46 - Aspecto do modelo exponencial no gráfico de variograma.

Se no software fizemos o ajuste variando a amplitude então vamos ter funções idênticas às da Figura 47 (a tracejado as variadas, a preenchido a original).

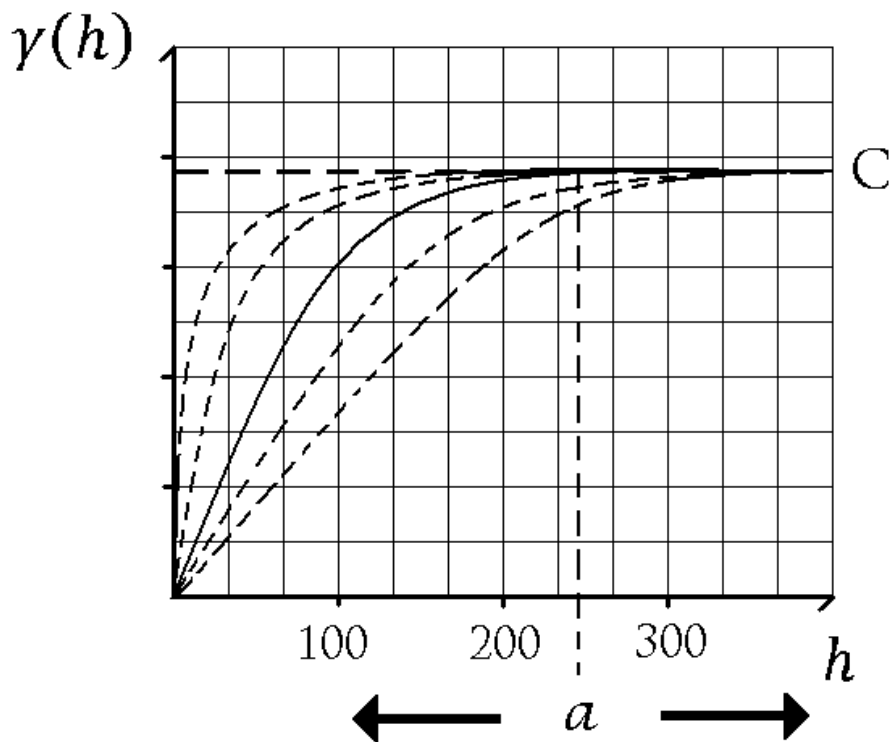


Figura 47 - Ajuste do modelo exponencial variando a amplitude.



Continuando o nosso exercício e dado que os nossos pares de pontos abaixo do patamar são poucos vamos ajustar este modelo manualmente experimentando o gaussiano e exponencial. Os nossos dados são os da Tabela 3 (distância, variogramas e patamar).

h	N-S			E-W		
	100	200	300	100	200	300
<b>19,4286</b>	48,8333	119,0000	<b>1,1875</b>	<b>3,3750</b>	<b>3,5000</b>	
<b>C</b>	<b><u>38,4952</u></b>					

Tabela 3 - Resumo dos dados obtidos com o estudo de variografia.

Apenas os valores a negrito estão abaixo do patamar (negrito e sublinhado) e por isso na direcção "N-S" apenas podemos considerar um ponto, na direcção "E-W" teríamos que considerar os três não fosse o facto de estarem todos tão próximos entre si e distantes do patamar. Assim na direcção "E-W" também só vamos considerar os pares de ponto à distância 200 metros. Começando pela direcção norte-sul e recuperando as fórmulas da inversão das funções modelos ficamos com:

$$a = -3 \frac{100}{\ln\left(1 - \frac{19,4286}{38,495}\right)} = 983,169 \text{ para o modelo esférico}$$

$$a = - \frac{100}{\sqrt{\ln\left(1 - \frac{19,4286}{38,495}\right) / 3}} = 313,555 \text{ para o modelo gaussiano}$$

Na direcção Este-Oeste temos (usando a distância de 200 metros):

$$a = -3 \frac{200}{\ln\left(1 - \frac{3,375}{38,495}\right)} = 15056,531 \text{ para o modelo esférico}$$

$$a = - \frac{200}{\sqrt{\ln\left(1 - \frac{3,375}{38,495}\right) / 3}} = 1735,312 \text{ para o modelo gaussiano}$$

O elipsóide de variografia ficaria como na Figura 48, nos dois casos considerados.

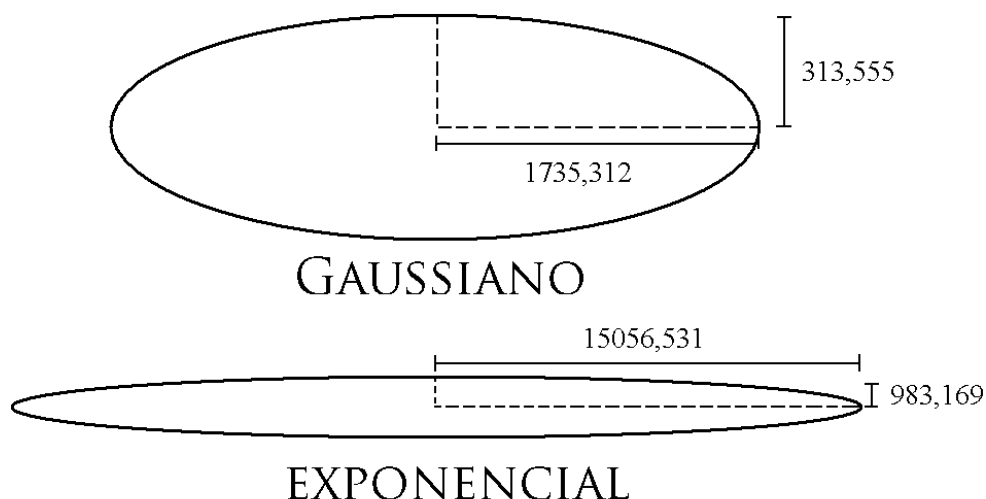


Figura 48- Elipsóide de variografia resultantes da inserção do modelo gaussiano (cima) e exponencial (baixo).

As relações de anisotropia são:

$$\text{Relação de anisotropia} = R_a = \frac{\text{direcção maior}}{\text{direcção menor}}$$

$$R_a = \frac{1735,312}{313,555} = 5,534 \text{ no caso gaussiano}$$

$$R_a = \frac{15056,531}{983,169} = 15,314 \text{ no caso exponencial}$$

O que implica que o eixo maior no caso gaussiano é 5,534 vezes maior que o eixo menor, e no caso exponencial, 15,314 vezes maior que o seu correspondente eixo menor. Dada a extensão dos nossos dados e porque o nosso variograma experimental tem poucos pontos a grande distância do patamar vamos utilizar o modelo gaussiano para prosseguir na estimação. Para fazer o nosso elipsóide (elipse neste caso) de variografia vamos usar a fórmula da elipse.

$$\frac{x^2}{b^2} + \frac{y^2}{c^2} = 1 \quad \text{ou} \quad \frac{x^2}{b^2} + \frac{y^2}{c^2} + \frac{z^2}{d^2} = 1 \quad \text{se fosse um elipsóide}$$

Onde  $b$ ,  $c$  e  $d$  são as amplitudes das direcções “Estes-Oeste ( $x$ )”, “Norte-Sul ( $y$ )” e “base-topo ( $z$ , que no nosso caso não existe)”. Assim, para que consigamos retirar a amplitude em qualquer direcção tendo já as direcções ortogonais calculadas nos variogramas podemos usar coordenadas elipsoidais:

$$x = b \cdot \cos(\beta) \cdot \cos(\lambda)$$

$$y = c \cdot \cos(\beta) \cdot \sin(\lambda)$$

$$z = d \cdot \sin(\beta)$$

Onde  $\beta$  e  $\lambda$  são os ângulos que se verifica na Figura 49.

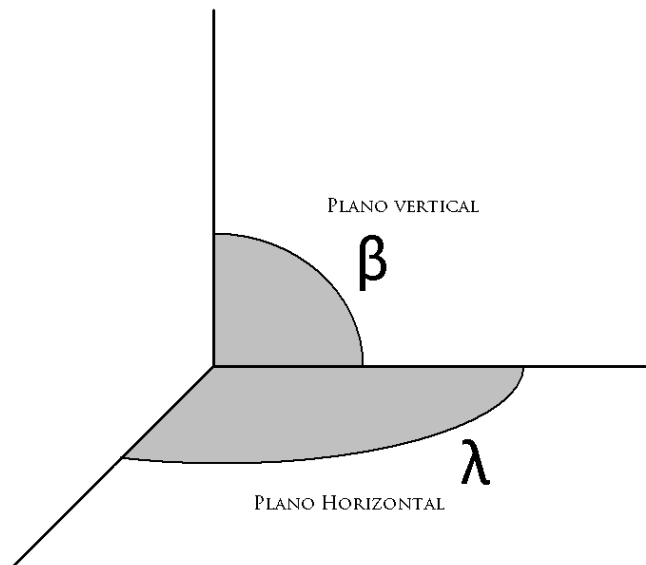


Figura 49 - Ilustração dos significados das letras em relação aos ângulos.

Para saber a amplitude numa dada direcção basta que se aplique a seguinte fórmula:

$$a = \sqrt{(b \cdot \cos(\beta) \cdot \cos(\lambda))^2 + (c \cdot \cos(\beta) \cdot \sin(\lambda))^2 + (d \cdot \sin(\beta))^2}$$

No nosso caso dado que se trata de uma elipse (2D) e não um elipsóide (3D) esta fórmula é simplificada para (dado que  $\beta$  e  $d = 0$  e por isso  $\cos(\beta) = 1$ ):

$$a = \sqrt{(b \cdot \cos(\lambda))^2 + (c \cdot \sin(\lambda))^2}$$

Então se pretendermos saber qual a amplitude do modelo a uma determinada direcção, como por exemplo  $45^\circ$ , aplicamos a fórmula:

$$a = \sqrt{(313,555 \cdot \cos(45))^2 + (1735,312 \cdot \sin(45))^2} = 1485.742$$

Após tirar a amplitude da direcção  $45^\circ$  podemos fazer a sua aplicação na fórmula do modelo (neste caso gaussiano) e temos a fórmula de cálculo de variograma para todas as distâncias nesta direcção:

$$\gamma(h) = C \left( 1 - e^{\left(-3 \frac{h^2}{a^2}\right)} \right) = 38,495 \left( 1 - e^{\left(-3 \frac{h^2}{1485.742^2}\right)} \right)$$

E está feito o nosso elipsóide de variografia. Neste momento podemos calcular o variograma para qualquer direcção utilizando as seguintes fórmulas:

$$a = \sqrt{(313,555 \cdot \cos(\lambda))^2 + (1735,312 \cdot \sin(\lambda))^2}$$

$$\gamma(h) = 38,495 \left( 1 - e^{\left(-3 \frac{h^2}{a^2}\right)} \right)$$

Já temos tudo o que é necessário para fazer a estimação, falta apenas definir como a mesma vai ser feita. No nosso caso vamos fazer uma grid com malha larga para que os cálculos não sejam excessivos (Figura 50).

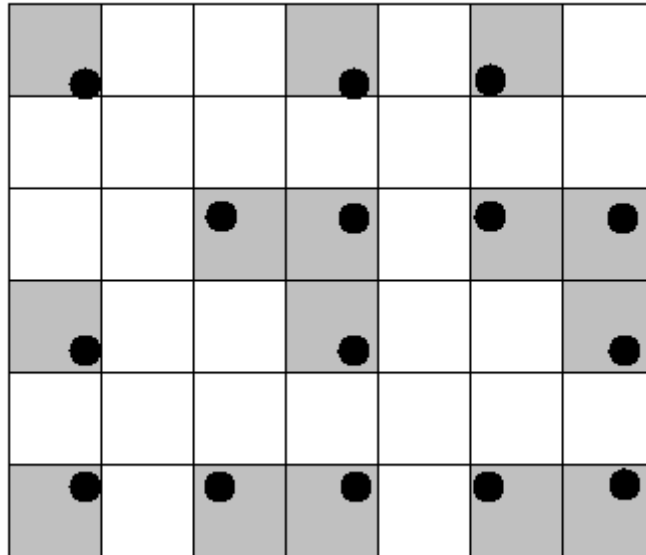


Figura 50 - Inserção da grid com os dados amostrados (blocos a cinzento).

Apenas os blocos sombreados contêm amostras, os restantes necessitam ser estimados. Um factor que tem de ser tido em conta na krigagem é o elipsóide de procura. Se definirmos um elipsóide de 200 metros na direcção “Este-Oeste” e 100 metros na direcção “Norte-Sul” vamos ter uma procura semelhante à da Figura 51.

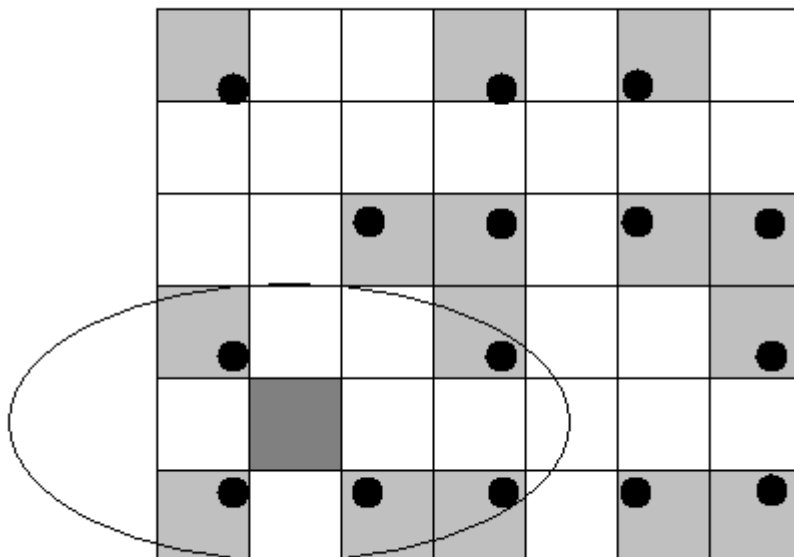


Figura 51 - Grid do jazigo com acção do elipsóide de procura aquando a estimação de um bloco.

Neste caso pretendemos estimar o bloco cinzento-escuro e o elipsóide de procura capta 5 amostras que vão ser usadas na estimação. Para exemplificar o cálculo de um bloco não vamos utilizar este dado que o cálculo seria extenso e moroso por serem demasiadas amostras consideradas. Vamos portanto calcular o bloco evidenciado na Figura 52 onde o elipsóide de procura capta três amostras.

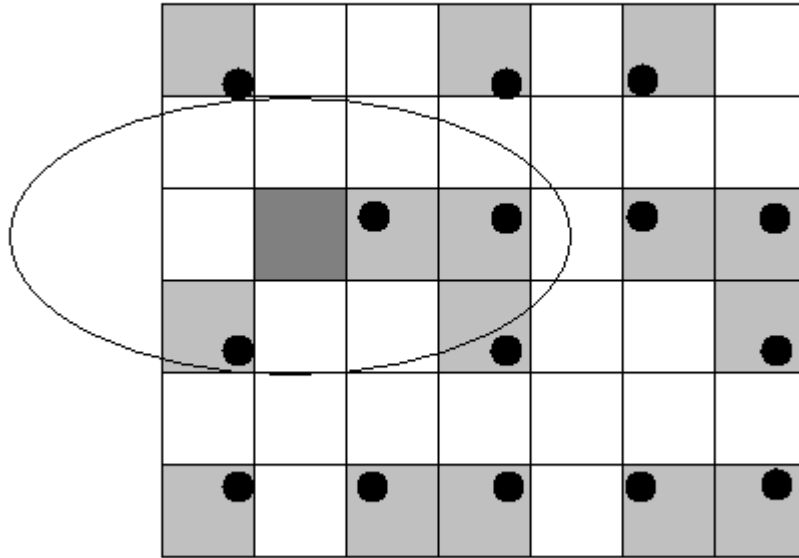


Figura 52 - Bloco que pretendemos calcular neste exercicio e respectivo raio de procura que capta 3 amostras para serem usadas na estimação.

Para simplificar vamos considerar que das três amostras consideradas duas estão na direcção “Este-Oeste” em relação ao bloco a estimar e uma a  $45^\circ$  (por já termos as amplitudes calculadas nestas direcção, se pretendêssemos maior pormenor seria razoável verificar as direcções correctas e correspondentes amplitudes (Figura 53)).

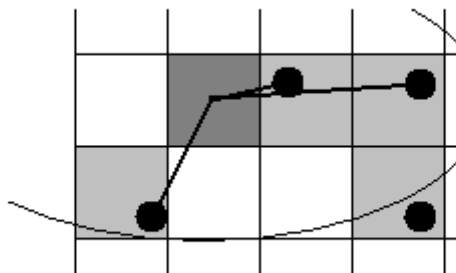


Figura 53 - Direcções das amostras consideradas na estimação em relação ao bloco a estimar.

Sabemos que estas três amostras vão ser consideradas para o cálculo do bloco evidenciado e até sabemos os variogramas nas direcções que o bloco faz com as amostras. O que precisamos

de saber agora é qual o peso que cada amostra vai ter no cálculo deste bloco, dado que estão a distâncias diferentes (se fosse pesos iguais o bloco mais longínquo teria a mesma influência que o mais perto o que poderia comprometer a nossa estimação. De facto assume-se que as zonas mais próximas do nosso bloco são também mais idênticas.) Assim tendo três amostras consideradas temos que ter três pesos diferentes que na sua totalidade terão que ser 1. Assim ficamos com a linha seguinte (em que o número 1,2 e 3 corresponde às amostras reais das esquerda para a direita (das que estão dentro do elipsóide de procura) e 0 à amostra a ser estimada):

$$w_1 + w_2 + w_3 = 1$$

O resto do sistema ficará:

$$w_1\gamma(h_{11}) + w_2\gamma(h_{12}) + w_3\gamma(h_{13}) = \gamma(h_{10})$$

$$w_1\gamma(h_{21}) + w_2\gamma(h_{22}) + w_3\gamma(h_{23}) = \gamma(h_{20})$$

$$w_1\gamma(h_{31}) + w_2\gamma(h_{32}) + w_3\gamma(h_{33}) = \gamma(h_{30})$$

O que na totalidade resulta de um sistema de 4 linhas para 3 variáveis ( $w_1, w_2, w_3$  visto que os parâmetros restantes podem ser todos calculados). Para que se compreenda o sistema acima é necessário perceber que  $\gamma(h_{pZ})$  é o variograma na direcção  $PZ$  e com distância,  $h$ , igual à distância entre estes dois pontos (Figura 54).

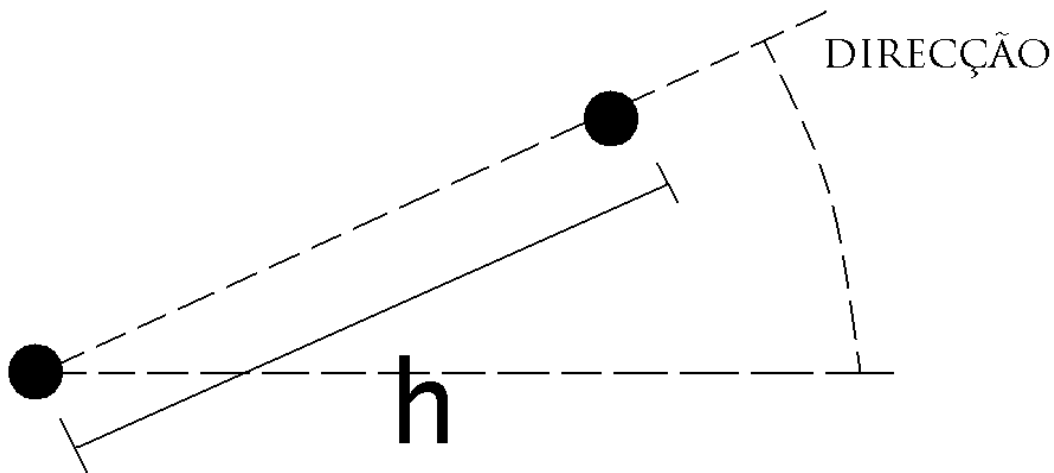


Figura 54 - Explicação da noção de distância e direcção no símbolo do variograma utilizado no sistema de estimação de pesos para as amostras.

Para simplificação de cálculos consideremos (apenas consideramos duas direcções em todos os cálculos, claro que para ser o mais bem feito possível teríamos que ir ver as amplitudes para as direcções correctas e as distâncias certas pois também estas foram muito arredondadas):

$$\gamma(h_{10}) = 38,495 \left( 1 - e^{\left( -3 \frac{100^2}{1485.742^2} \right)} \right) = 0,519$$

$$\gamma(h_{20}) = 38,495 \left( 1 - e^{\left( -3 \frac{100^2}{1735,312^2} \right)} \right) = 0,382$$

$$\gamma(h_{30}) = 38,495 \left( 1 - e^{\left( -3 \frac{200^2}{1735,312^2} \right)} \right) = 1,503$$

$$\gamma(h_{11}) = 0$$

$$\gamma(h_{22}) = 0$$

$$\gamma(h_{33}) = 0$$

$$\gamma(h_{12}) = \gamma(h_{21}) = 38,495 \left( 1 - e^{\left( -3 \frac{150^2}{1485,742^2} \right)} \right) = 1,159$$

$$\gamma(h_{13}) = \gamma(h_{31}) = 38,495 \left( 1 - e^{\left( -3 \frac{250^2}{1485,742^2} \right)} \right) = 3,134$$

$$\gamma(h_{23}) = \gamma(h_{32}) = 38,495 \left( 1 - e^{\left( -3 \frac{100^2}{1735,312^2} \right)} \right) = 0,382$$

O que nos deixa com o sistema:

$$w_2 1,159 + w_3 3,134 = 0,519$$

$$w_1 1,159 + w_3 0,382 = 0,382$$

$$w_1 3,134 + w_2 0,382 = 1,503$$

$$w_1 + w_2 + w_3 = 1$$

A resolução do sistema pode ser feita por método de gauss (Figura 55):

$$\left[ \begin{array}{ccc|c} 0 & 1,159 & 3,134 & 0,519 \\ 1,159 & 0 & 0,382 & 0,382 \\ 3,134 & 0,382 & 0 & 1,503 \\ 1 & 1 & 1 & 1 \end{array} \right] \begin{array}{l} \times (1 / 3,034) \\ \times (2,704) \end{array}$$

$$\left[ \begin{array}{ccc|c} 0 & 1,159 & 3,134 & 0,519 \\ 1,159 & 0 & 0,382 & 0,382 \\ 0 & 0 & -2,062 & 0,29 \\ 1 & 1 & 1 & 1 \end{array} \right]$$

Figura 55 - Cálculo por método de Gauss dos pesos das amostras na estimação do bloco.

Assim a partir da terceira linha podemos dizer que:

$$w_3 = \frac{0,29}{-2,062} = -0,14$$

Sabendo  $w_3$  e usando a primeira linha dizemos que:

$$w_2 = \frac{(0,519 + 3,136 * 0,14)}{1,159} = 0,826$$

E finalmente como precisamos de uma solução que respeite a última linha fazemos a aplicação desta:

$$w_1 = 1 - 0,826 + 0,14 = 0,314$$

Agora já sabemos os pesos que cada uma das amostras tem na estimação do bloco que estamos a considerar. Assim o valor final do bloco tem:

$$\begin{aligned} teor_0 &= w_1 teor_1 + w_2 teor_2 + w_3 teor_3 = \\ teor_0 &= 0,314 teor_1 + 0,826 teor_2 - 0,14 teor_3 \end{aligned}$$

Relembrando os teores das amostras reais (Figura 56):

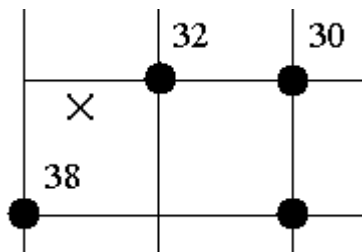


Figura 56 - Teores das amostras consideradas na estimação.

Temos que:

$$teor_0 = 0,314 * 38 + 0,82 * 32 - 0,14 * 30 = 34,164$$

Já está. Conseguimos fazer uma estimação do teor utilizando o método de krigagem. Note-se que foram dados pesos mais elevados às amostras mais perto da zona a estimar (1 e 2) e ainda maior aquela que se encontra na direcção de maior continuidade ou maior amplitude (2). Para completar esta estimação teríamos que fazer estes cálculos para todos os blocos da nossa grid. E está terminado o nosso exercício de krigagem.



# INTRODUÇÃO AO S-GEMS

O S-GeMS foi desenvolvido com dois objectivos. O primeiro é providenciar um software de interface amigável com múltiplas ferramentas geoestatísticas. O seu fácil manuseamento provem do seu aspecto gráfico sóbrio e simples e a possibilidade de visualizar directamente o resultado das operações num ambiente tridimensional interactivo. O segundo objectivo é a criação de um software cujas funcionalidades pudessem ser melhoradas. O código do S-GeMS é aberto.

O site oficial do software é <http://sgems.sourceforge.net/> e existe um manual disponível para o mesmo (disponível no conjunto onde vem este documento), bem como tutoriais como é o caso do de Geoffrey Bohling do Kansas Geological Survey, Universidade do Kansas (<http://people.ku.edu/~gbohling/BoiseGeostat/>).

Este manual apenas pretende dar o passo inicial no mundo da geoestatística usando o software S-GeMS. Para operações mais completas deverá consultar bibliografia adicional.

## INTERFACE DO S-GEMS

O interface do S-GeMS é composto por três grandes secções: o painel de algoritmos o painel de visualização e o painel de comandos. A partir do painel de comandos é possível introduzir por código os comandos que se podem fazer com teclado e rato a partir dos painéis algoritmos e visualização. Por esse motivo, e em favor da brevidade, não vamos ensinar a linguagem do painel de comandos neste documento. Tudo será feito através do interface. Ao entrarmos no software S-GeMS (poderá ser necessário maximizar para que o interface ocupe o ecrã todo) deparamos com algo idêntico à Figura 57 onde o painel de algoritmos e visualização estão assinalados a vermelho.

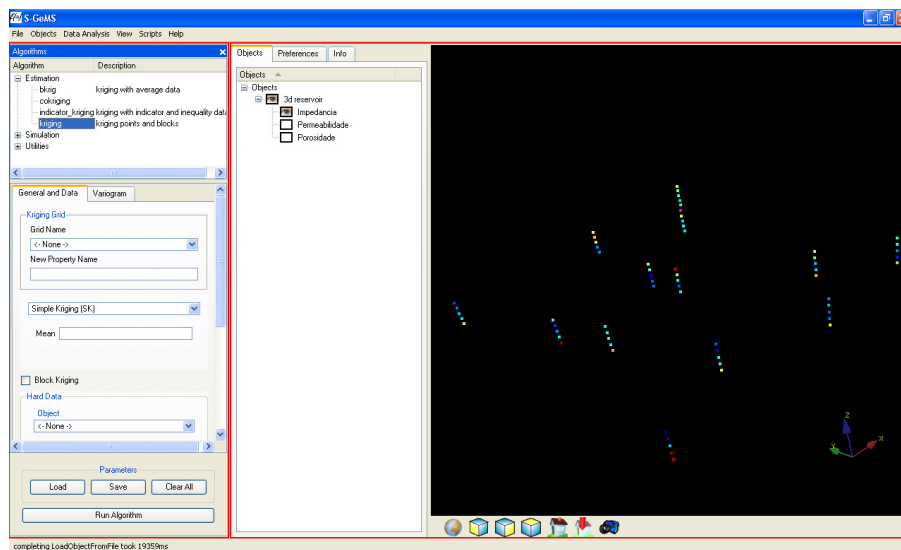


Figura 57 - Interface do S-GeMS.

## PAINEL DE ALGORITMOS

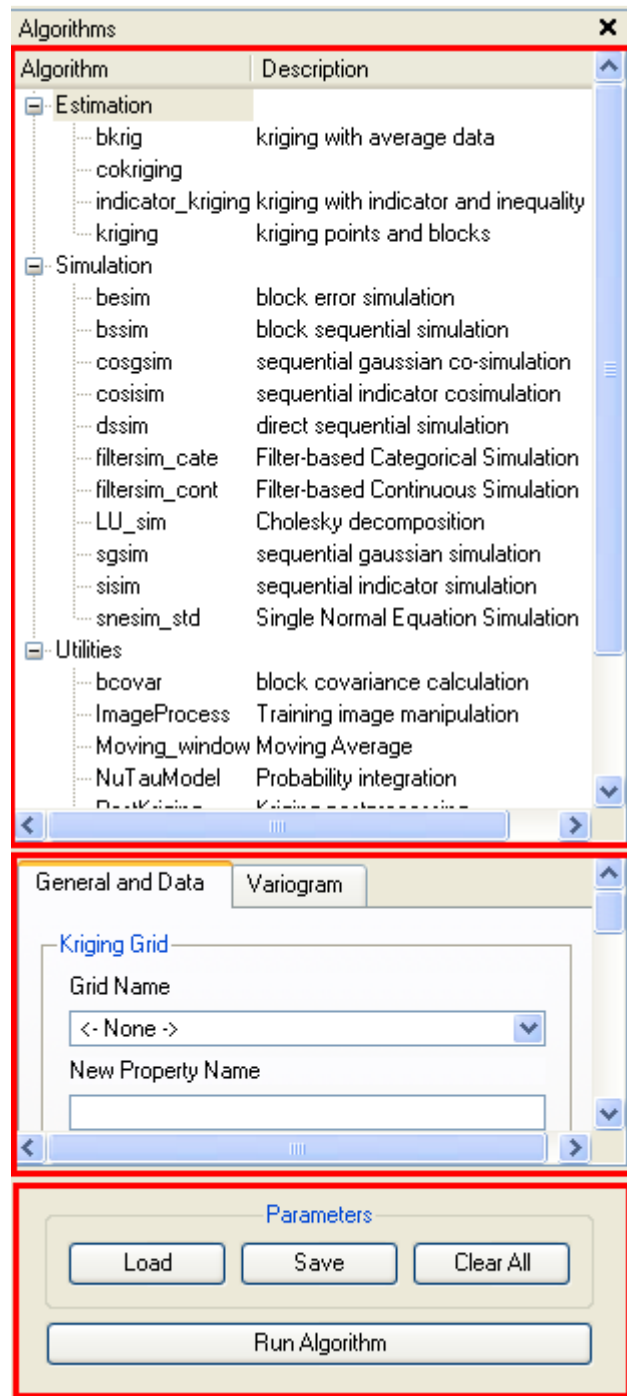



Figura 58 - Painel de algoritmos: listagem de funcionalidade (cima), especificação dos parâmetros da funcionalidade (centro) e comandos básicos (baixo).

Começamos pelo painel de algoritmos que tem 3 subdivisões (Figura 58): a listagem das funcionalidades, a especificação dos parâmetros da funcionalidade (esta subdivisão só aparece após ter sido seleccionada uma funcionalidade) e os comandos básicos (salvar, carregar, apagar e correr os parâmetros introduzidos na funcionalidade escolhida). Na listagem de funcionalidades aparece as funções de estimação, simulação e restantes aplicações. Ao

carregarmos no símbolo  de cada uma delas aparece cada uma das funcionalidades individualmente (Figura 59).

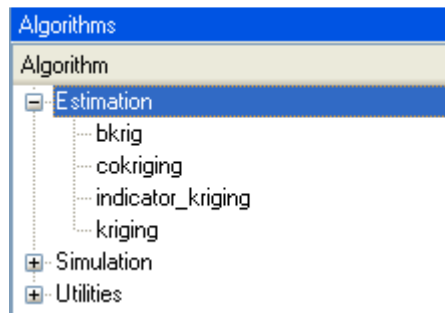


Figura 59 - Selecção das funcionalidades no painel de algoritmos.

A subdivisão de comandos básicos (Figura 60) tem quatro comandos ao seu dispor. A função de “Load” serve para carregar um ficheiro que contenha a informação dos parâmetros que pretendemos introduzir na nossa funcionalidade de escolha. A função de “Save” serve para salvar os parâmetros que introduzimos manualmente. A função de “Clear All” serve para apagar todos os parâmetros que introduzimos. O comando de “Run Algorithm” serve para correr a funcionalidade com os parâmetros que lhe foram impostos. O resultado aparecerá no painel de visualização que mais à frente falaremos.

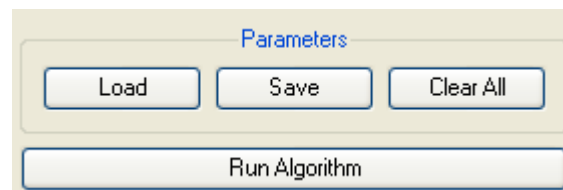


Figura 60 - Subdivisão de comandos básicos.

A subdivisão da especificação dos parâmetros da funcionalidade escolhida apresenta os pontos de preenchimento dos parâmetros dependendo da função que se escolheu fazer. A título de exemplo vamos ver o caso da krigagem. Ao escolher-se a funcionalidade krigagem aparece a sua subdivisão (Figura 61) a qual está dividida no sector “General and Data” (o primeiro que aparece) e sector “Variogram” (que podemos escolher no topo da subdivisão). Para que a krigagem funcione teremos que preencher todos os campos. Algumas das informações vamos buscar a outras ferramentas do S-GeMS que falaremos mais à frente. Por agora repare-se que aqui metemos todos os parâmetros necessários ao cálculo da estimação começando pela grid onde a vamos fazer (“Grid Name”), dando um nome à variável que se vai estimar (não precisa de ser igual à dos pontos amostrados, se bem que, por questões de organização, é conveniente), a escolha do tipo de krigagem (na figura está o “Simple Kriging (SK)” ou krigagem simples, outros podem ser seleccionados), a escolha dos dados iniciais a partir do qual se faz a estimação (“Hard Data”) onde se escolhe o nome do objecto (“Object”) e a respectiva propriedade que se pretende estudar (“Property”) e finalmente os dados do elipsóide de procura (Search Ellipsoid) que mais à frente falaremos. Depois de completos os dados deste sector é necessário completar os dados do sector “Variogram” entre os quais o efeito pepita (“Nugget Effect”), o número de estruturas dos nossos variogramas (“Nb of Structures”) e a

especificação de cada uma delas (geralmente só se usa uma) com o patamar (“Contribution”), o tipo de modelo utilizado (Type) e o elipsóide de variografia que mais à frente veremos.

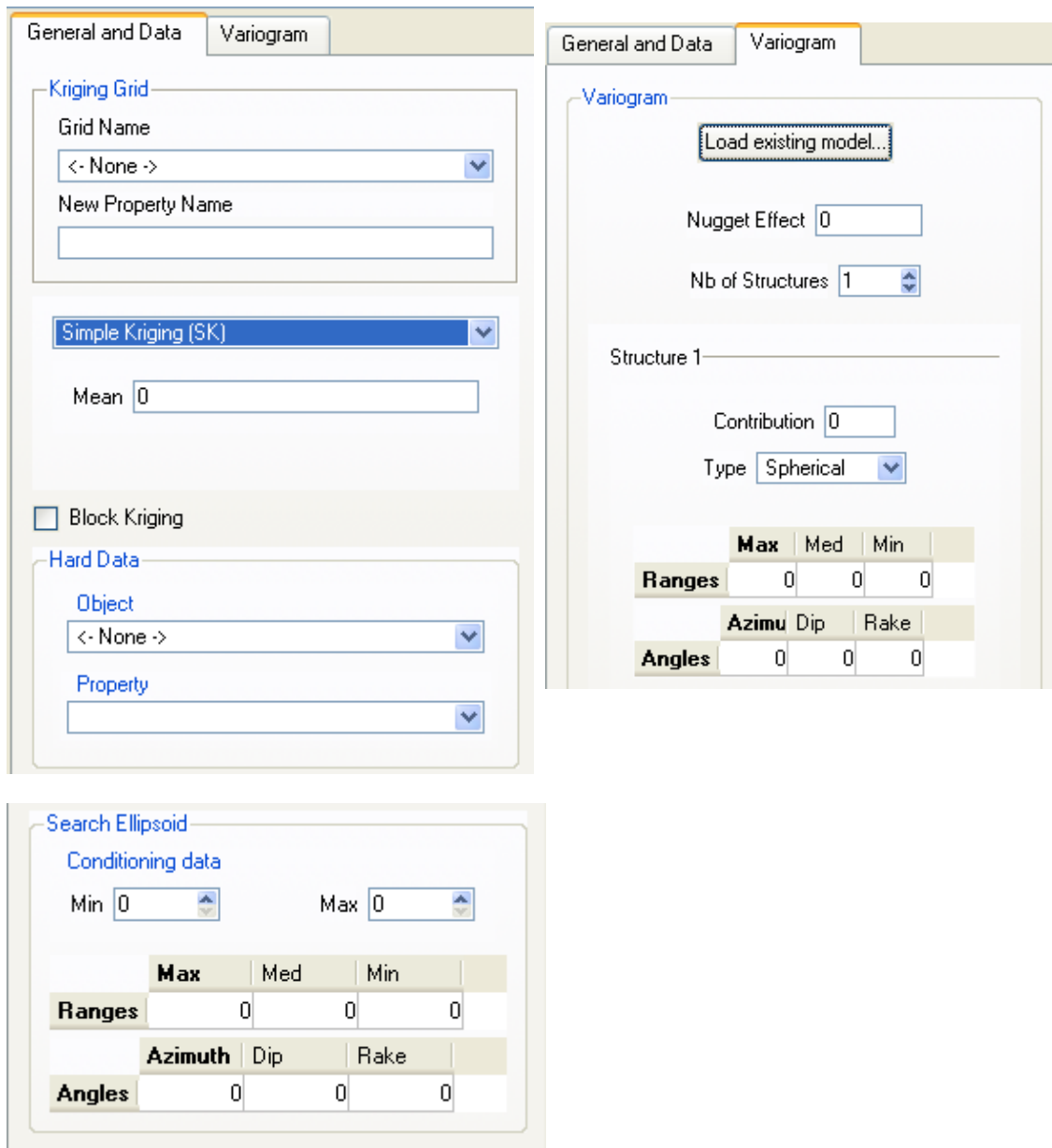


Figura 61 - Especificação dos parâmetros da funcionalidade, neste caso krigagem.

## PAINEL DE VISUALIZAÇÃO

O painel de visualização tem duas subdivisões: as especificações e escolhas dos objectos (já vamos explicar o que é um objecto) e a visualização propriamente dita (Figura 62).

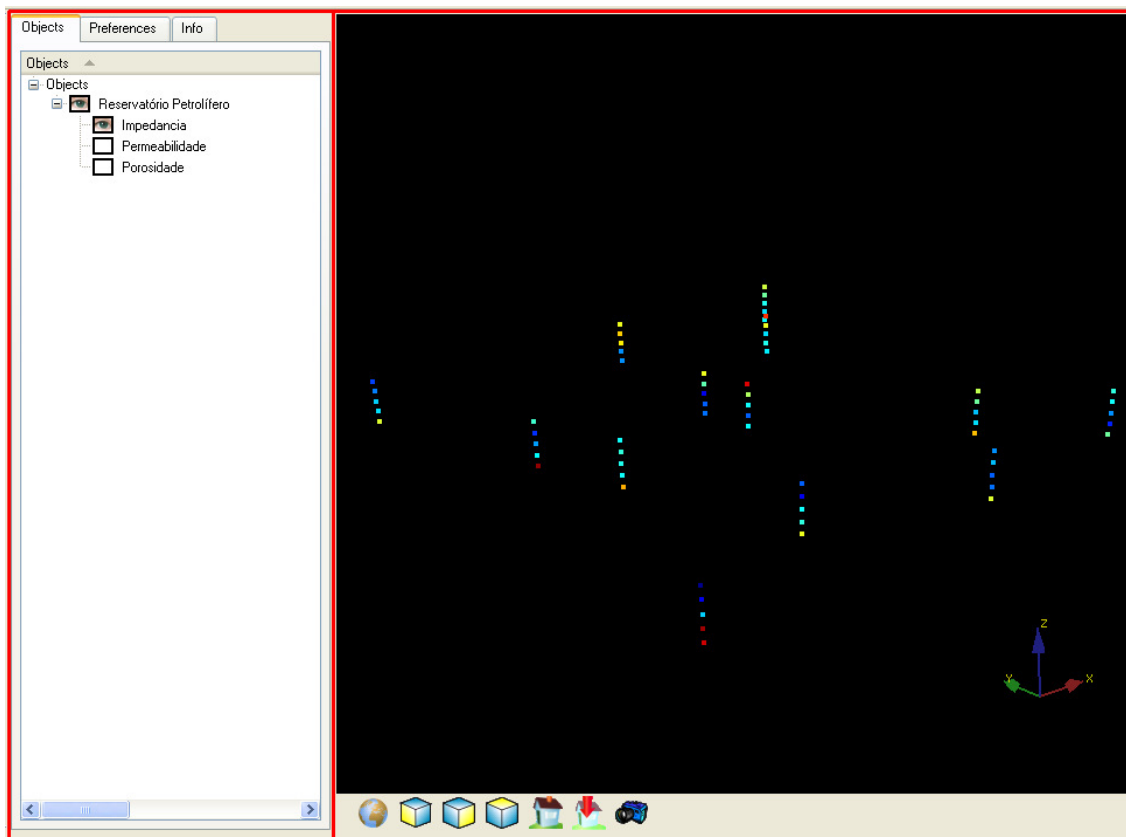


Figura 62 - Painel de visualização.

Começemos por ver a subdivisão das especificações e escolha dos objectos. Na Figura 63 temos um caso de um conjunto de point data que foi inserido no S-GeMS que vem com três variáveis: impedância, permeabilidade e porosidade. A este conjunto foi dado o nome de “Reservatório Petrolífero”. Quando seleccionamos um esse fica escolhido com um símbolo de um olho. Isto significa que este é o objecto que estamos a visualizar no visualizador tridimensional. O objecto é um conjunto de poin-data, grids, etc.



Figura 63 - Selecção do objecto a visualizar.

Dentro dos objectos existem propriedades (ou variáveis que no nosso caso é a impedância, permeabilidade e porosidade). Se fossemos a fazer a krigagem que vimos anteriormente este seria o objecto (“Object”) escolhido como dados de estimação (“Hard Data”) e a propriedade (“Property”) qualquer uma das três variáveis que se pretenda estudar. Tal como podemos

verificar esta subdivisão tem três sectores: objectos (“Objects”), preferências (“Preferences”) e informações (“Info”). O primeiro sector é o que acabamos de explicar. O terceiro (“Info”) dá-nos informações como podemos ver na Figura 64 (número de pontos, valores mínimos e máximos de **x**, **y** e **z**).

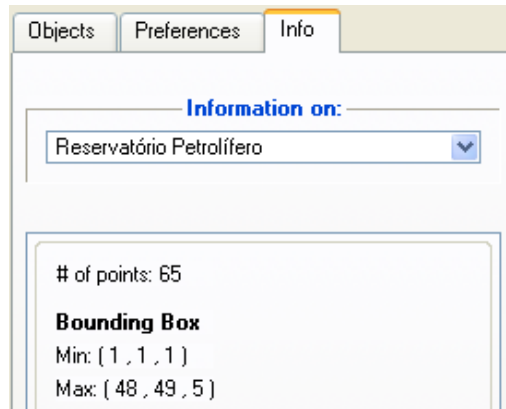


Figura 64 - Informações do objecto.

O segundo sector é o das preferências (“Preferences”) onde seleccionando o objecto sobre o qual queremos ver as preferências (“Preferences for:”) bem como a propriedade (“Property”) podemos escolher o tamanho dos point-data no visualizador (“Point-Size”), o tipo de escala que pretendemos (“Colormap”) e o valor mínimo e máximo que pretendemos visualizar (“Min” e “Max”, por defeito vem o mínimo e máximo do conjunto).

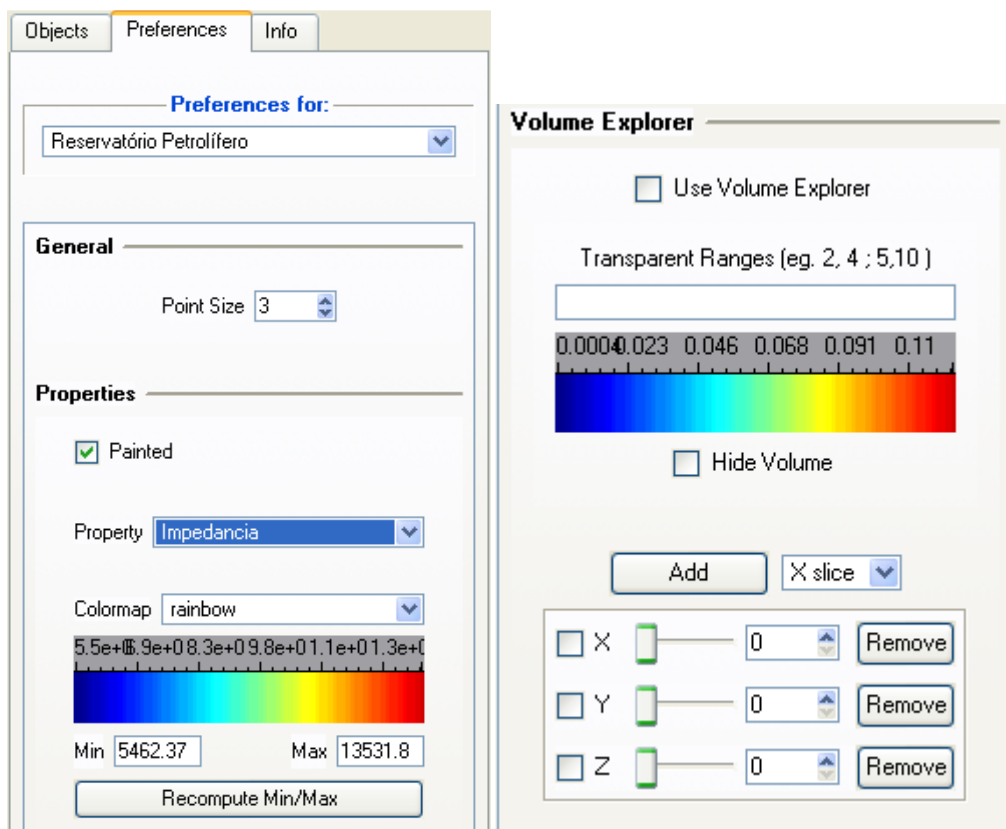



Figura 65 - Barra de preferência no painel de visualização.

Se o objecto estudado for uma grid (e por isso tem volume) vai existir uma terceira opção que é o explorador (“Volume Explorer”). Usando o volume Explorer podemos fazer cortes nas direcções **x**, **y** e **z** para estudar a ou as propriedades dessa grid que de outra forma não haveria maneira de as visualizar (para ver os cortes é necessário escolher a opção “Use Volume Explorer” e depois a “Hide Volume” para esconder a grid e apenas ficarem visíveis os cortes). O visualizador é bastante intuitivo e apenas com o rato é possível rodar o objecto e fazer zoom in ou zoom out (se o rato tiver scroll). Para percorrer o espaço sem rodar basta carregar no Ctrl+Right Click e andar com o rato na direcção que se pretende. Para fazer zoom in ou zoom out é carregar no Ctrl+Shift+Right Click e mover o rato. Existe mais uma série de botões com perspectivas de visualização que não vale a pena explicar aqui (são facilmente reconhecidas as suas funções com a experimentação) por isso vou evidenciar apenas um, o “Camera Snapshot”(  ) que permite retirar o que está no visualizador para um ficheiro de imagem(ao carregarmos vai aparecer a Figura 66).

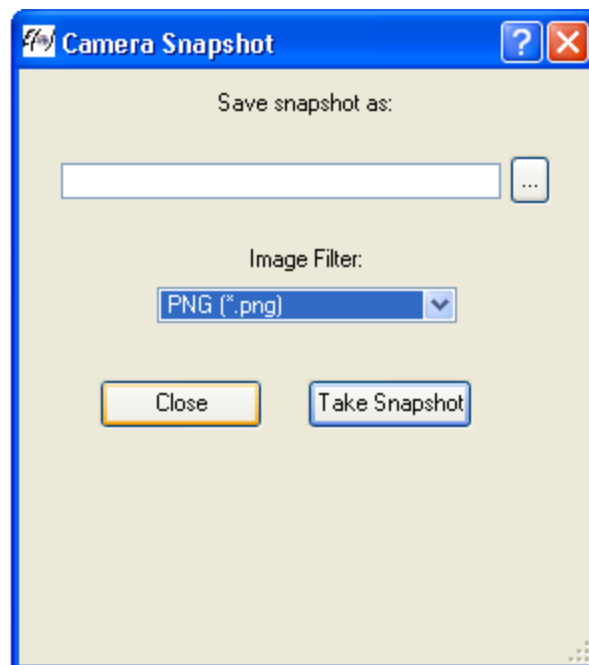


Figura 66 - Criação de ficheiro de imagem com o resultado do painel de visualização.

## COMANDOS BÁSICOS NO S-GEMS

### INSERIR DADOS NO S-GEMS

Para inserir dados para o S-GeMS basta fazer um Load Object (Ctrl+L) ou simplesmente leva-lo com o rato para dentro do visualizador do S-GeMS. Ao ser feito isto vai aparecer o seguinte menu:

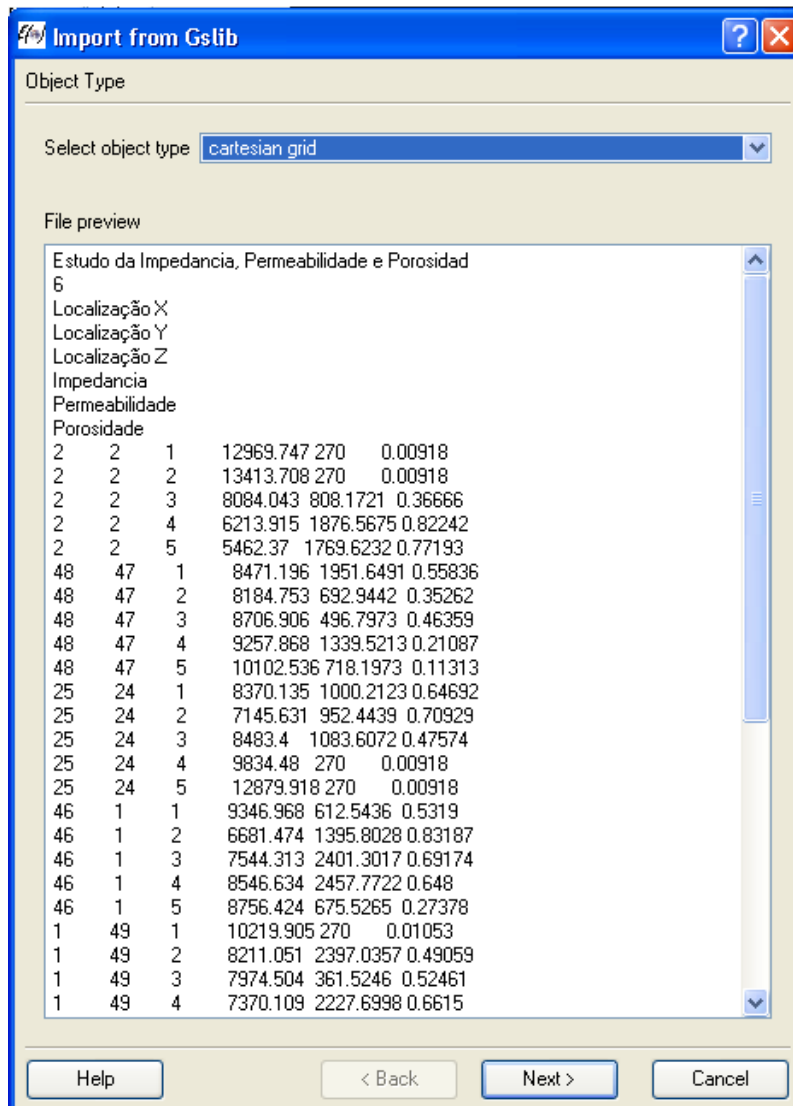


Figura 67 - Menu de input de ficheiros.

Neste menu podemos ver uma visualização do ficheiro que vamos utilizar e é necessário escolher que tipo de objecto estamos a meter ("Select Object Type"). Neste caso é um conjunto point-data e por isso a opção correcta é "point set".



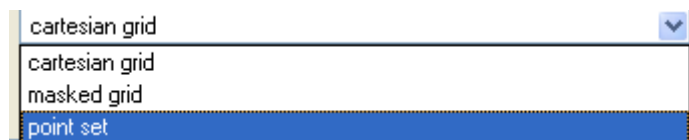


Figura 68 - Escolha do tipo de input.

Depois de escolher a correcta designação carregamos no “Next” e seguimos para o menu a seguir.

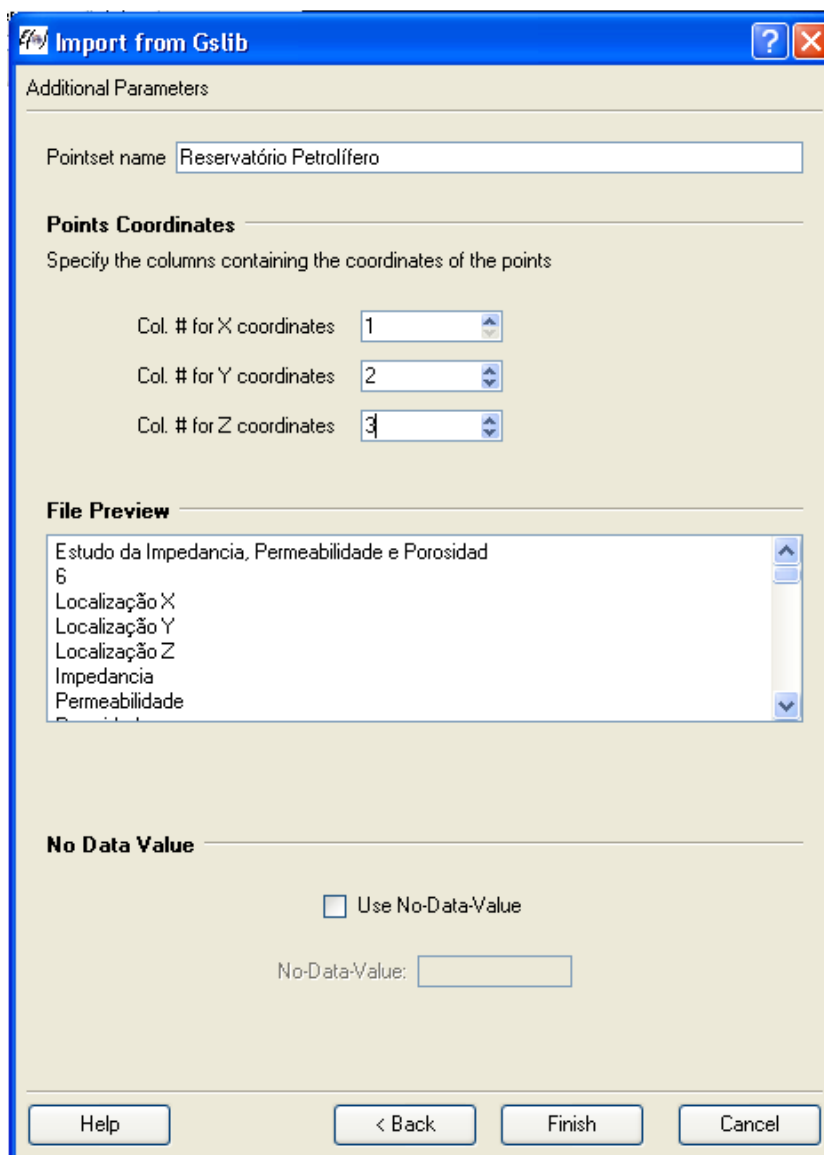


Figura 69 - Menu de input de point-data.

Neste menu escolhemos o nome que queremos dar ao nosso objecto (“Pointset name”) e os locais apropriados para onde está a coluna das coordenadas. Neste caso o nome dado é “Reservatório Petrolífero” e a localização das coordenadas **x**, **y** e **z** é coluna 1, 2 e 3, respectivamente. A partir daqui já podemos trabalhar sobre eles, basta que seleccionemos a sua visualização carregando no objecto e respectiva propriedade que pretendemos observar.

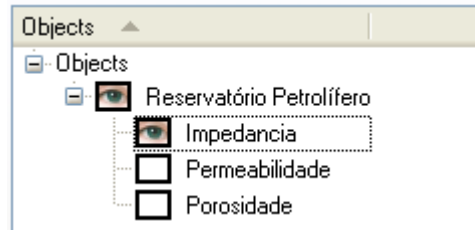


Figura 70 - Escolha do objecto a visualizar no visualizador.

O resultado no visualizador será imediato e terá o seguinte aspecto:

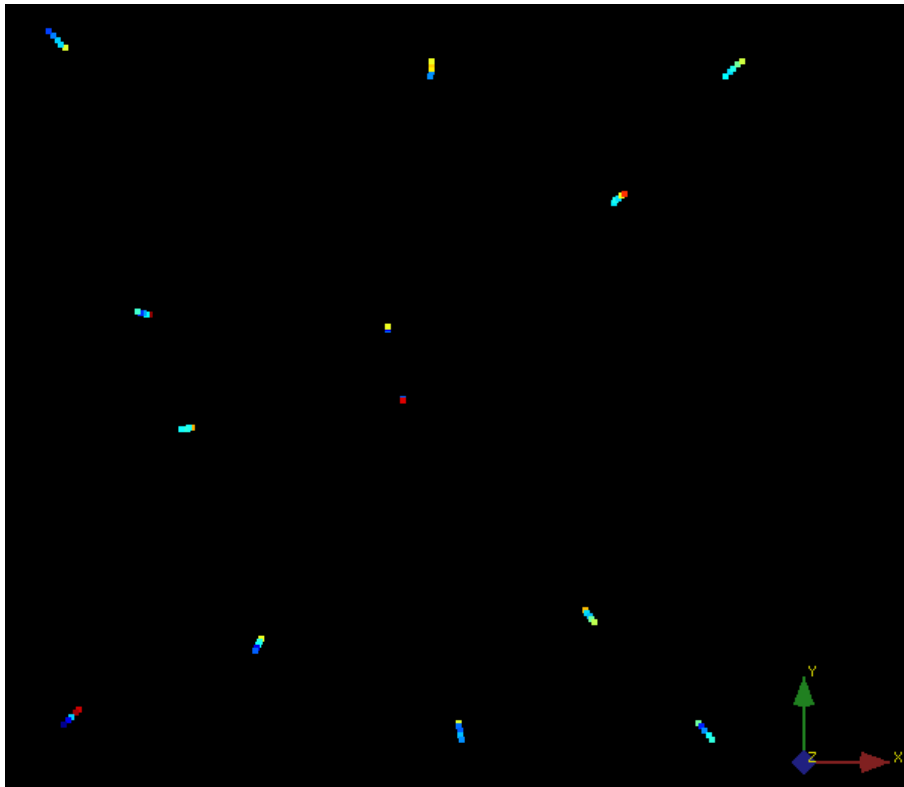


Figura 71 - Resultado da visualização do input.

Para salvar qualquer objecto que tenhamos inserido ou criado no S-GeMS basta ir à barra superior e seleccionar “Objects” e depois “Save Object”.

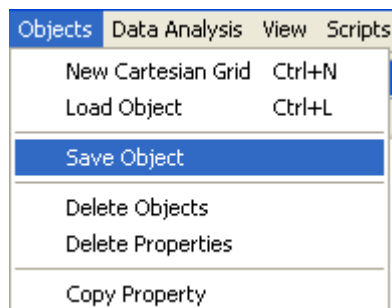


Figura 72 - Guardar objecto.

## FERRAMENTAS DO S-GEMS

A maioria das ferramentas no S-GeMS pode ser encontrada na barra superior no “Data Analysis”.

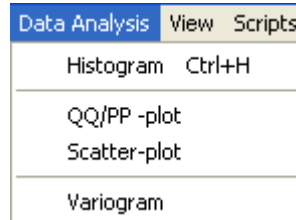


Figura 73 - Barra do "Data Analysis".

Aqui podemos ter acesso ao menu de histogramas (“Histogram”, Ctrl+H), menu do QQ/PP-plot, Scatterplot (gráfico de dispersão) e menu de variogramas (“Variogram”). No menu de histogramas podemos ver o histograma de qualquer propriedade de cada objecto. No caso da Figura 74 foi seleccionado o objecto “Reservatório Petrolífero” e a propriedade “Impedância”. É possível meter os valores mínimos e máximos bem como mudar o tipo de visualização de função de probabilidade (“pdf” – o histograma propriamente dito) ou de função cumulada de probabilidade (“cdf”) no “Plot type”. Se formos para a divisão “Display Options” podemos editar os eixos e mete-los em escala logaritmica. Repare-se que no menu de histogramas é possível ver (na parte inferior) os dados estatísticos referentes à propriedade estudada, inclusive a variância que posteriormente poderá ser utilizada como patamar na variografia.

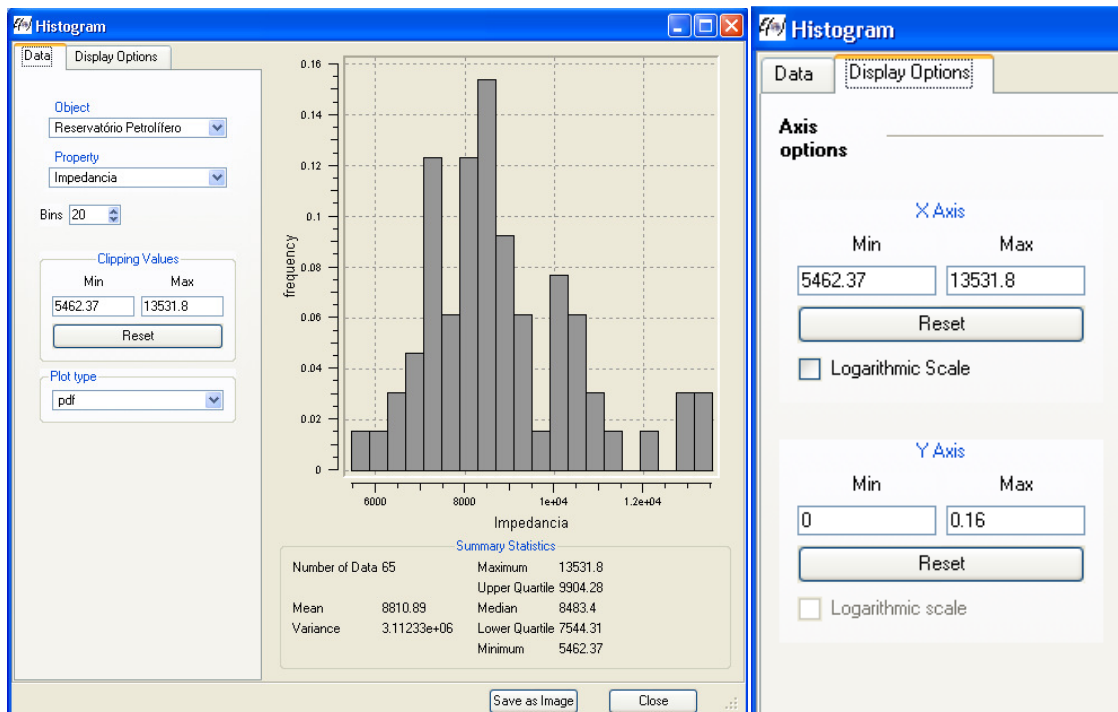


Figura 74 - Menu de histograma.

Se utilizarmos o menu de gráfico de correlação (“Scatterplot”) vamos deparar com um menu semelhante. Mais uma vez é necessário escolher o objecto que pretendemos estudar e comparar as propriedades dentro do mesmo (não é possível estudar variáveis de objectos diferentes, para o fazermos é necessário meter todas as que queremos dentro do mesmo ficheiro). No caso da Figura 75 escolhemos comparar as propriedades “impedância” e “Porosidade” do objecto “Reservatório Petrolífero”, para além disto é possível inserir a linha de regressão linear com a opção “Show Linear Regression Line” (na figura não dá para ver a opção, para se ver é necessário fazer scroll down no sector “Data”, mas o resultado é a linha vermelha por cima do gráfico de dispersão). O sector “Display Options” é em tudo semelhante ao menu de histogramas e novamente volta a ser possível dizer quais os máximos e mínimos para cada eixo e mete-los em escala logarítmica.

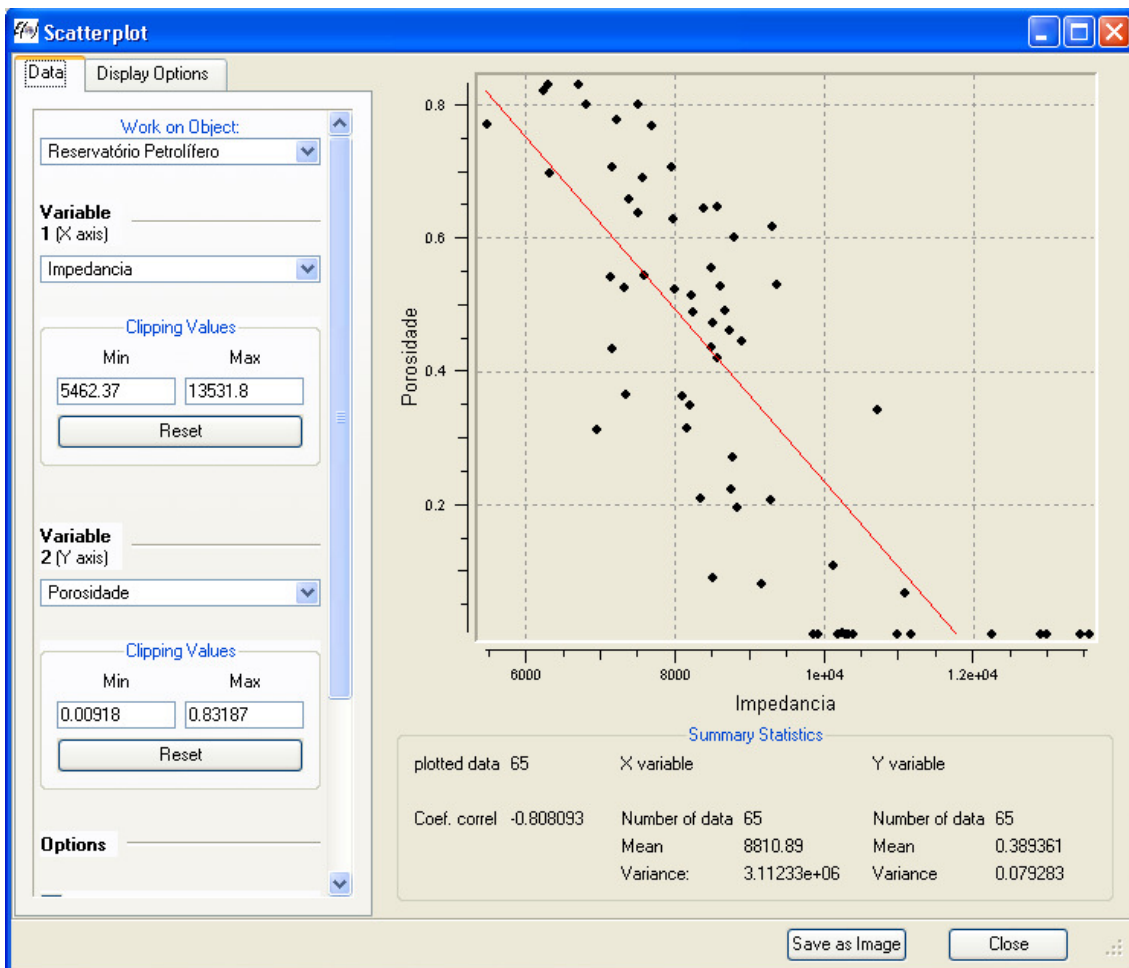


Figura 75 - Menu de gráfico de dispersão.

O QQ/PP-plot é uma maneira de conseguir comparar as funções de distribuição de diferentes propriedades em diferentes objectos. Por esse motivo temos de escolher o objecto (mesmo que seja igual) para cada propriedade a ser comparada. Para além disso escolhemos a visualização como sendo QQ-plot ou PP-plot e podemos limitar o intervalo considerado no “Clipping Values”. No caso da Figura 76 temos a visualização QQ-plot seleccionada e as propriedades “Impedância” e “Permeabilidade” do objecto “Reservatório Petrolífero”. O “Display Options” é idêntico aos dois casos abordados anteriormente.

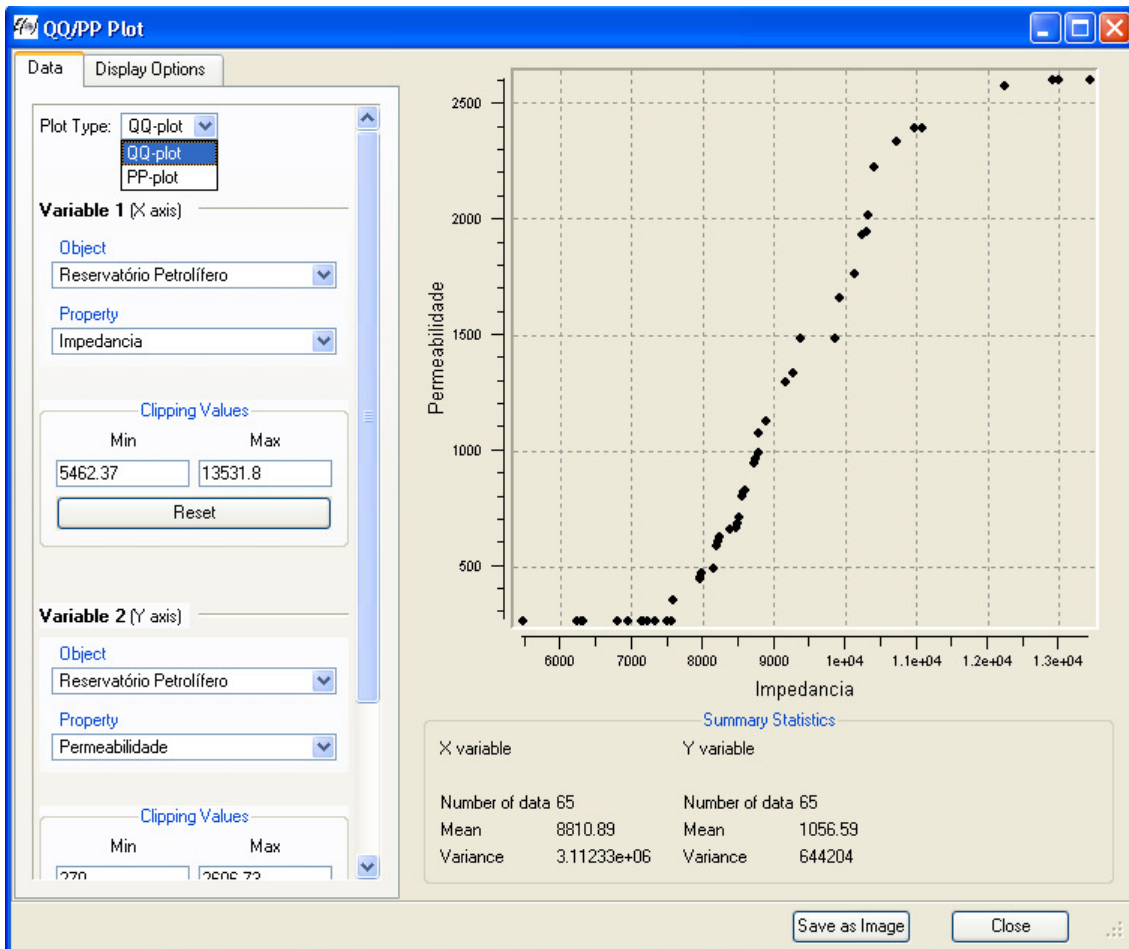


Figura 76 - Menu de QQ/PP-plot.

A variografia, dado o seu teor um pouco mais complexo, será explicada no tutorial prático. O tutorial prático será composto por dois exercícios. Um primeiro explicado de forma intensiva (o caso da horta) e um segundo para acompanhamento (caso do reservatório petrolífero).

# UTILIZAÇÃO DO SOFTWARE S-GeMS NA ESTIMAÇÃO DO TEOR DE CONTAMINANTES NUMA HORTA

O software S-GeMS é um programa open-source disponível gratuitamente online e tem a finalidade de resolver problemas associados a dispersão espacial de variáveis utilizando métodos geoestatísticos. O site do software é <http://sgems.sourceforge.net/> onde está disponível para download o executável para instalação (opção Windows e sistemas Linux).

## DADOS DA HORTA



Figura 77- Local da horta comunitária biológica.

Um grupo de alunos de engenharia numa universidade decidiu usar um dos terrenos no campus da mesma para começar um projecto de horta comunitária biológica (Figura 77). Dada a necessidade de os vegetais cultivados estarem o mais incólume possível de contaminações decidiu-se fazer uma análise química ao terreno da horta. Devido ao uso inadequado de produtos químicos nas vizinhanças foram encontrados vestígios significativos de fósforo, arsénio e cobre. Os recursos são limitados e é necessário estimar a quantidade de contaminantes e zonas do terreno a tratar para posterior cultivo. Para isso foram feitas 16 amostragens em pontos diferentes com malha na qual se tentou obter uma que representasse a área total da horta (Figura 78). Dada a possibilidade de infiltração no terreno dos mesmos contaminantes devido a águas de chuva optou-se por amostrar às diferentes profundidades de 87, 87.5 e 88 metros (no qual 88 representa a cota da superfície).

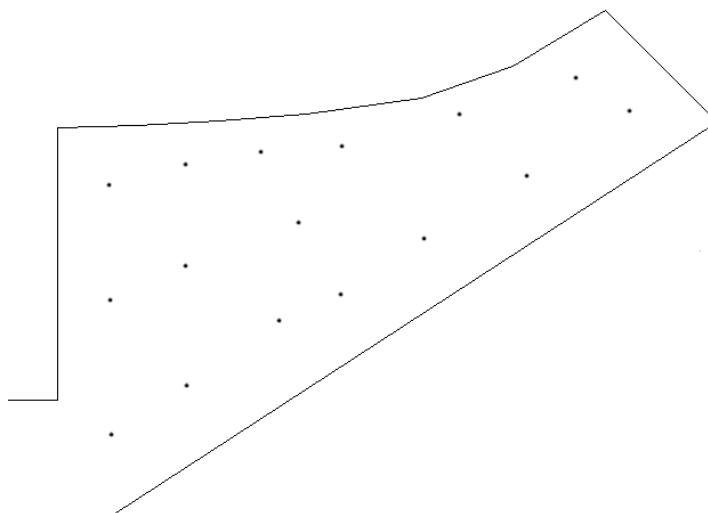


Figura 78 - Localização das amostragens no terreno da horta.

Os resultados das análises químicas foram compilados em ficheiros ASCII com 6 colunas na qual a primeira é localização em **x**, a segunda em **y**, terceira em **z**, quarta teor em fósforo, quinta em arsénio e sexta em cobre. Neste caso o software utilizado para se fazer a estimação é o S-GeMS pelo que os ficheiros de input necessitam de um cabeçalho. O cabeçalho é formado primeiro pelo nome do projecto, depois pelo número de variáveis (incluindo coordenadas), e finalmente pelo nome, em diferentes linhas, da cada variável como se pode ver no excerto do ficheiro de input abaixo.

```

Estudo na horta comunitária biológica
6
X
Y
Z
Fósforo
Arsénio
Cobre
487987.93 4287445.15 88 0.1200 0.0048 0
487987.93 4287445.15 87.5 0.0635 0.0002 0
487987.93 4287445.15 87. 0.0278 0.0000 0
... ..

```

Estes são os nossos dados iniciais para o estudo das contaminações no terreno da horta.

## PREPARAÇÃO DOS DADOS

Após serem digitalizados em formato ASCII com a composição apresentada no capítulo anterior abrimos o software S-GeMS. Podemos inserir os dados de duas maneira: agarrando o ficheiro com o rato e metendo-o dentro do interface do S-GeMS ou indo ao **Objects** na barra superior e fazer **Load Object** **Ctrl+L** (como se pode ver também existe o atalho Ctrl+L). Seleccionamos o ficheiro das amostragens e abrimos. Apareceu uma nova janela com um **File preview** onde podemos ver o ficheiro que acabamos de inserir, e um **Select object type**, neste ultimo temos três hipóteses. Apenas nos interessa o **point set** pois o ficheiro que

inserirmos é, de facto, um ficheiro de point-data (caso quiséssemos inserir uma grid teríamos que escolher `cartesian grid` ). Carregamos no  e seguimos para a especificação dos nossos dados. O `Pointset name` é o campo onde se mete o nome dos dados que estamos a inserir e neste caso ficou como “Amostras da Horta”. Os campos seguintes são a especificação das colunas das coordenadas **x**, **y** e **z** (repare-se que neste caso temos de meter coluna 3 para o **z**, se o caso fosse bidimensional deixaríamos com a expressão “NA (2d)” no mesmo espaço). Existe apenas mais uma escolha a fazer, o  `Use No-Data-Value` . Os No-Data são pontos em que existe a localização (o point) mas não existe um valor atribuído (o data). No caso da horta isto não sucede mas imaginemos que em todas as localizações teríamos feito a análise química ao fósforo, arsénio, mas ao cobre em apenas algumas. Existirão pontos onde existem valores de fósforo e arsénio mas não cobre, sendo por isso, o cobre, nesta localização, um No-Data. Geralmente opta-se por escolher um número que não tem hipótese de ser confundido com outro no mesmo ficheiro como é o caso de, por exemplo, o “-99.99”. Assim em algumas localizações existiria o valor de fósforo, o valor de arsénio e o -99.99 para o cobre. Seleccionando  `Use No-Data-Value` aparece uma caixa para especificar o valor do No-Data (no caso acima seria -99.99). Daqui para a frente o software não iria considerar os pontos No-Data na estimação do cobre. No caso da nossa horta não existem valores No-Data por isso podemos seguir em frente carregando no  . Na secção de objectos no interface do S-GeMS (Figura 79) seleccionamos o “Amostras da Horta”, que acabamos de criar, e qualquer uma das variáveis se quisermos que os point-data no visualizador sejam coloridos à escala.

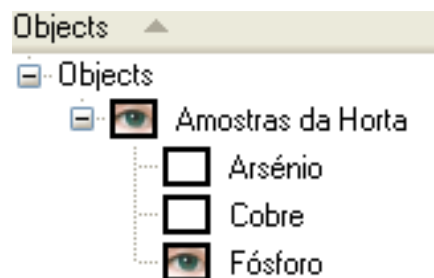


Figura 79- Secção de objectos no interface do S-GeMS (Note-se que o S-GeMS tirou os nomes das variáveis directamente do cabeçalho do ficheiro).

Repare-se que no visualizador a escolha é actualizada de imediato. Podemos rodar a imagem, aproximar, etc., consoante a preferência e necessidade do utilizador. Na Figura 80 está a visualização do fósforo, em perspectiva, a partir dos point-data.



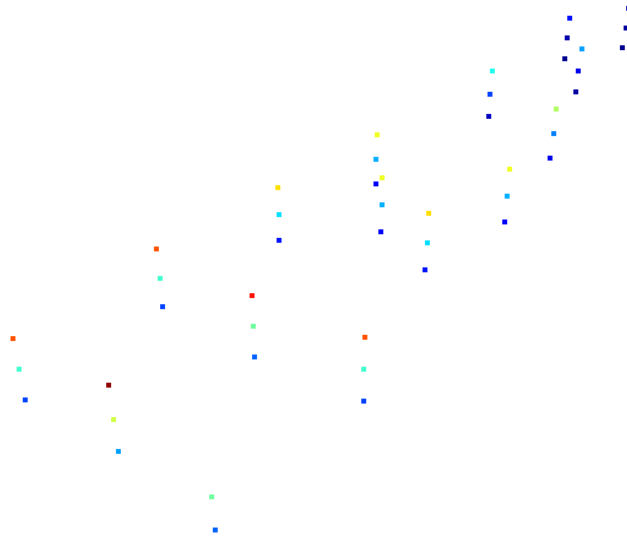


Figura 80 - Amostras da Horta com cores representativas da quantidade de fósforo em cada localização.

A inserção dos dados está feita mas temos um problema que podemos evitar. As grids criadas pelo S-GeMS são sempre em relação aos eixos **x**, **y** e **z** e a sua forma é sempre paralelepípedica. Se construirmos uma grid nestas condições vamos ter muito espaço que vai ser estimado e onde não existe amostras, nem sequer terreno para a horta (Figura 81). Para minimizar este problema vamos aplicar uma rotação de coordenadas às localizações das amostras.

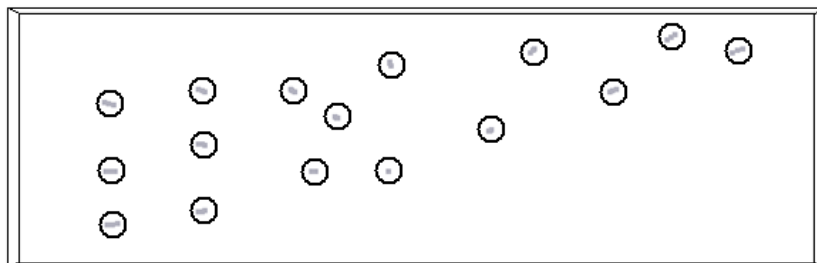


Figura 81 - Grid aplicada às amostras na disposição em que se encontram naturalmente.

Para fazermos a rotação (neste caso no plano **xOy**) vamos aplicar as fórmulas:

$$X_f = D * \text{Cos}(tg^{-1} \left( \frac{Y_i}{X_i} \right) + \eta)$$

$$Y_f = D * \text{Sen}(tg^{-1} \left( \frac{Y_i}{X_i} \right) + \eta)$$

$$D = \sqrt{X_i^2 + Y_i^2}$$

Onde  $X_f$  e  $Y_f$  são as coordenadas x e y depois da rotação.  $X_i$  e  $Y_i$  as coordenadas x e y antes da rotação.  $\eta$  o ângulo de rotação que pretendemos impor sobre as coordenadas. Em recurso a folha de cálculo foram aplicadas estas equações a todas as localizações com um  $\eta = -\pi/16$ . O resultado foi o da Figura 82.



Figura 82 - Inserção da grid nas amostras com rotação.

Gravámos um novo ficheiro com as três colunas das novas localizações e as três respectivas de valores das variáveis fósforo, arsénio e cobre. Voltamos novamente a inserir um ficheiro de dados, desta com as coordenadas transformadas, e damos o nome de “Amostras da Horta Transformadas”.

## ESTATÍSTICOS DESCRITIVOS

É importante tentar reconhecer padrões nos nossos dados e perceber se um ou mais fenómenos tiveram envolvidos na dispersão dos contaminantes. O primeiro passo foi analisar a distribuição dos dados a partir de histogramas. Para o fazer fomos ao [Data Analysis](#) e seleccionamos a opção [Histogram Ctrl+H](#) (atalho Ctrl+H). Apareceu o menu de histogramas e seleccionamos o objecto “Amostras da Horta”, analisando as propriedades fósforo, arsénio e cobre, uma de cada vez resultando na Figura 83, Figura 84 e Figura 85.

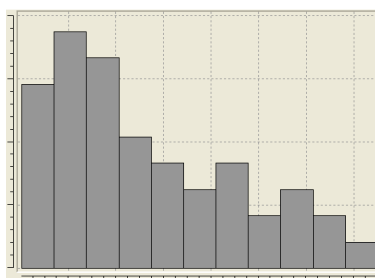


Figura 83 - Histograma do fósforo.

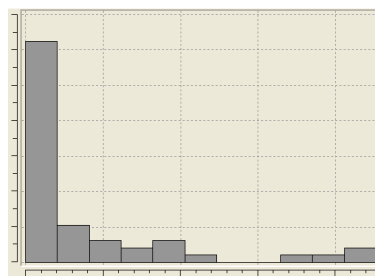


Figura 84 - Histograma do arsénio.

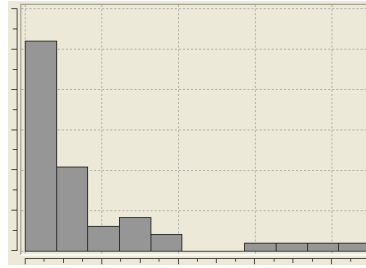


Figura 85 - Histograma do cobre.

Todas as distribuições são assimétricas com maior frequência nos teores mais baixos. A oscilação no caso do cobre e arsénio é muito mais acentuada que no caso do fósforo. Talvez por estarem presentes no terreno à mais tempo tenham tido tempo de ter uma maior diluição. Para além disso parecem existir duas populações no caso do arsénio e do cobre, os valores mais baixos e mais frequentes, e os valores mais altos e menos frequentes. Possivelmente correspondem a dois despejos diferentes. Se tentarmos fazer uma cronologia a partir dos histogramas vemos que inicialmente foram despejados os contaminantes arsénio e cobre agora já muito diluídos e com teores baixos, posteriormente foi o fósforo com uma diluição menor, e finalmente um novo despejo de arsénio e cobre quase sem diluição. Uma maneira de suportar esta teoria é ver os gráficos de dispersão seleccionando novamente [Data Analysis](#) e depois `Scatter-plot`. Mais uma vez escolhemos o objecto “Amostras da Horta” e comparamos O fósforo com o arsénio e o cobre, e o arsénio com o cobre resultando na Figura 86, Figura 87, Figura 88.

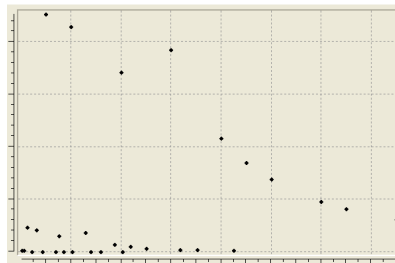


Figura 86 - Scatterplot de fósforo vs arsénio.

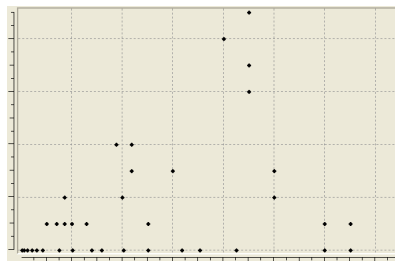


Figura 87 - Scatterplot de fósforo vs cobre.

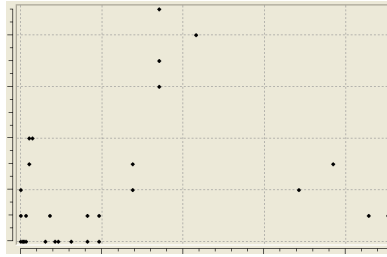


Figura 88 - Scatterplot de arsénio vs cobre.

Não existe relação aparente entre fósforo e cobre e tal como previsto na relação entre arsénio e cobre se denotam duas populações. Curiosamente o mesmo acontece entre o arsénio e o fósforo o que leva a pensar que a hipótese cronológica comentada anteriormente pode estar incompleta. Apesar de tudo temos informações suficientes para construir um conhecimento significativo sobre a história do terreno da horta. A seguir faremos a estimação.

## ESTUDO DE VARIOGRAFIA

Depois de termos aplicada uma rotação nas coordenadas das amostras da horta, o terreno ficou alinhado com os eixos do software pelo que as direcções de variografia mais óbvias são de azimutes  $0^{\circ}$  e  $90^{\circ}$  e de inclinação  $90^{\circ}$  (porque as amostras foram retiradas na direcção estritamente vertical em profundidade). Mais uma vez vamos ao [Data Analysis](#) mas agora seleccionamos a opção [Variogram](#). Aparece o menu de variografia e nós pretendemos fazer um estudo de variografia para cada contaminante separadamente. Assim seleccionamos a opção “Compute variograms from scratch” (fazer os variogramas desde o início, a outra opção serve para ir buscar um ficheiro de variogramas já feitos) e “Amostras da Horta Transformadas” no Grid name (é o conjunto de pontos que se pretende estudar, eles chamam grid a tudo). O head property e tail property são o fósforo. Carrega-se no “NEXT” e passamos para a definição dos nossos intervalos de variografia. O S-GeMS pede os parâmetros número de intervalos (number of lags), tamanho do intervalo (lag separation), tolerância no tamanho (lag tolerance), direcção (azimuth, dip), tolerância angular (tolerance), limite de direcção (bandwidth), e limite superior e inferior (head cut off e tail cut off). Na Figura 89 estão os parâmetros pedidos pelo software.

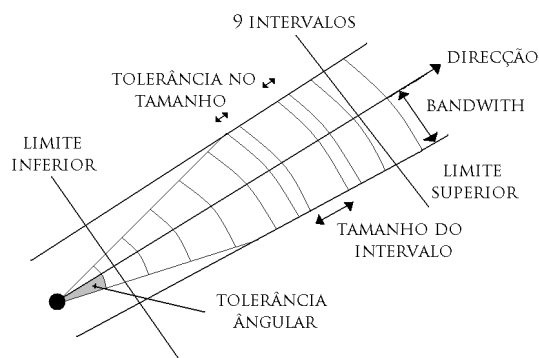


Figura 89 - Parâmetros pedidos no software S-GeMS.

A tolerância do tamanho é um acrescento ao tamanho normal para que se existirem amostras próximas do limite, também essas sejam consideradas. O bandwidth faz o limite em distância da tolerância angular pois a partir de certo ponto um ângulo pequeno dá origem a uma extensão muito grande. Os limites superiores e inferiores são semelhantes a garantem que a o estudo de variografia só começa a partir de certo ponto e termina quando chegar a outro ponto. Para fazer o estudo de variografia ao fósforo no menu seguinte pedimos 44 intervalos com tamanho de 0.5 e tolerância de tamanho 0.1 (os intervalos têm de ser pequenos caso contrário não conseguimos fazer um estudo adequado em profundidade dado que os furos só têm 1 metro de extensão). As direcções são (90,0); (0,0); (0,90) que são representativas dos eixos **x**, **y** e **z** do S-GeMS. As tolerâncias angulares para as direcções **x** e **y** são de 15<sup>o</sup> (note-se que é para cada lado, no total perfaz 30<sup>o</sup>), para **z** ficou apenas 1<sup>o</sup> para que sejam considerados apenas os valores imediatamente abaixo. Não queremos distâncias limite (para fazer isto basta, para o bandwidth, meter um número grande de maneira a não limitar nada na extensão da horta). Fazemos “NEXT” e vemos os nossos variogramas. É possível arrumar as janelas como se tratasse de um interface Windows. Precisamos de ajustar os modelos às várias direcções. O S-GeMS constrói directamente o elipsóide a partir da caixa da Figura 90.

The image shows a software interface for adjusting a variogram model. At the top, there is a text box for 'Sill (Contribution)' containing the value '001632'. Below it is a dropdown menu for 'Type' set to 'Exponential'. A section titled 'Ranges' contains three rows: 'Max' with a slider and a text box containing '25', 'Med' with a slider and a text box containing '18', and 'Min' with a slider and a text box containing '2'. Each row has a 'Reset' button to its right. At the bottom, there is a section for 'Angles' with three text boxes, each containing the value '0'.

Figura 90 - Ajustamento do modelo ao elipsóide de variografia.

O “Sill” corresponde ao patamar e dado que as nossas amostragens estão bem distribuídas optámos por meter a variância do fósforo (este dado está no histograma do fósforo no S-GeMS). O “Type” corresponde ao tipo de modelo que neste caso ficou como o exponencial. Os ranges equivale às amplitudes na direcção “Max”, “Med” e “Min” que correspondem à maior, média e menor direcção de amplitude. Na secção abaixo “Angles” especificamos qual é a direcção maior e neste caso é (0,0,0). O resultado da adequação dos modelos está na Figura 91.

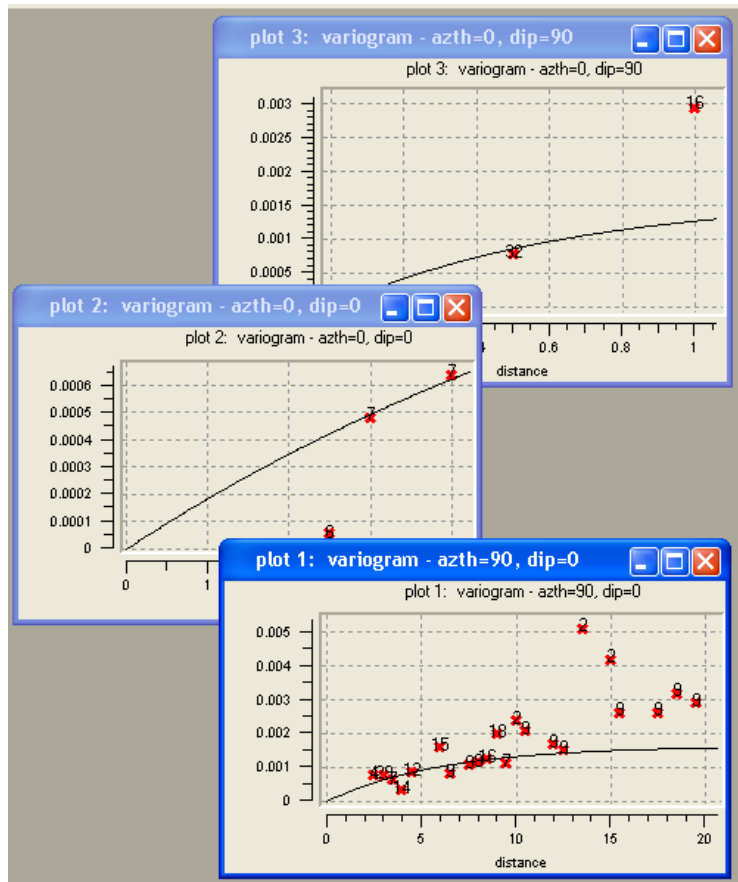


Figura 91 - Resultado do estudo de variografia feito ao fósforo.

## KRIGAGEM

Temos o elipsóide de variografia feito e as amostras por onde estimar. A única coisa que nos falta é a discretização do espaço, melhor dizendo a grid. Com a transformação das coordenadas a extensão em  $x$ , já considerando a margem é de cerca de 22 metros. Em  $y$  cerca de 7 metros. Uma maneira de fazer as contas da grid é perceber quantos blocos vão haver por metros. Se houver quatro blocos por metro vamos ter:

$$4 * 22 = 88 \text{ em } x$$

$$4 * 7 = 28 \text{ em } y$$

Os números não são elevados e podemos apertar a malha e mesmo assim ter um custo computacional razoável num computador normal. Então optamos por fazer oito blocos por metro ficando com:

$$8 * 22 = 176 \text{ em } x$$

$$8 * 7 = 56 \text{ em } y$$

$$8 * 1 = 8 \text{ em } z$$

Com um total de  $176 * 56 * 8 = 78848$  blocos no nosso modelo. E cada bloco irá ter um tamanho (em **x**, **y** e **z**)  $1/8 = 0,125$  metros. Agora que está definida a grid vamos construí-la no software indo ao **Objects** e escolhendo a opção **New Cartesian Grid Ctrl+N** (com atalho Ctrl+N). Aparece o menu de construção da grid. O “grid name” é o nome do novo objecto grid e ficou como “Grid da Horta”. O número de células em **x**, **y** e **z** ficaram com 176, 44 e 8 blocos mantendo um rácio satisfatório entre os lados da grid (é suficiente para cobrir as amostras eficientemente). O tamanho do bloco ficou igual em todos os lados com o valor de 0,125. As coordenadas do canto inferior esquerdo da grid foram tiradas após análise de um gráfico criado em folha de cálculo (basta fazer um gráfico de dispersão **x vs y**) e os valores são 1315048 para **x**, 4109856 para **y** e 87 para **z** (Figura 92).

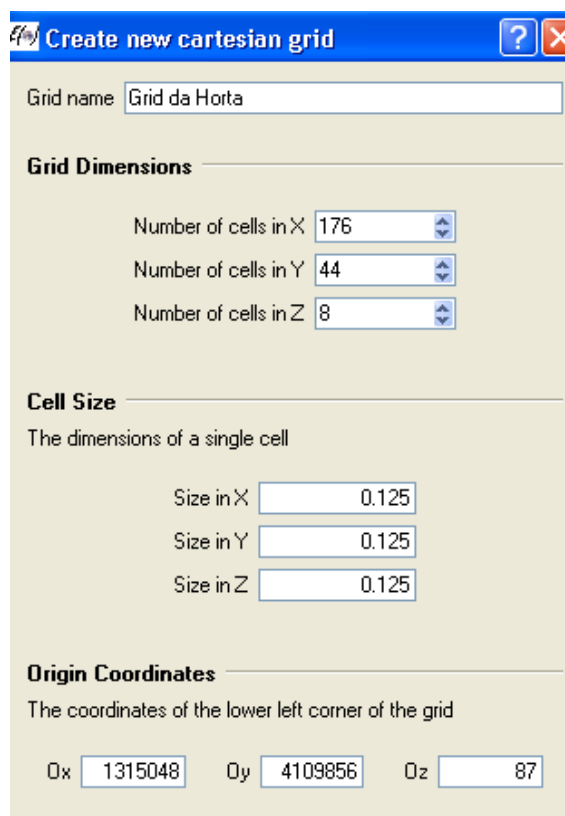


Figura 92 - Menu da grid completo com os dados da nossa horta.

O resultado foi o da Figura 93.

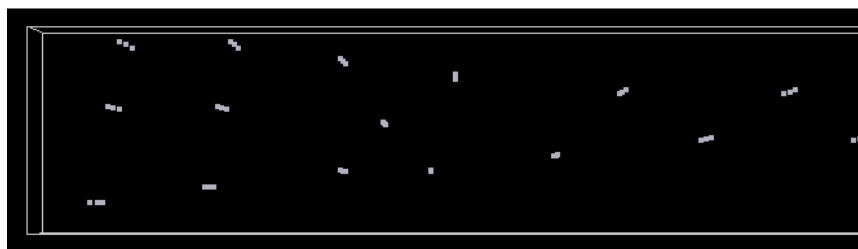


Figura 93 - Grid e point data no visualizador do S-GeMS.

Agora já podemos fazer a estimação por krigagem. Na secção de “Algorithms” no interface do S-GeMS carregamos na opção “Estimation” e seguidamente na opção “Kriging” (Figura 94).

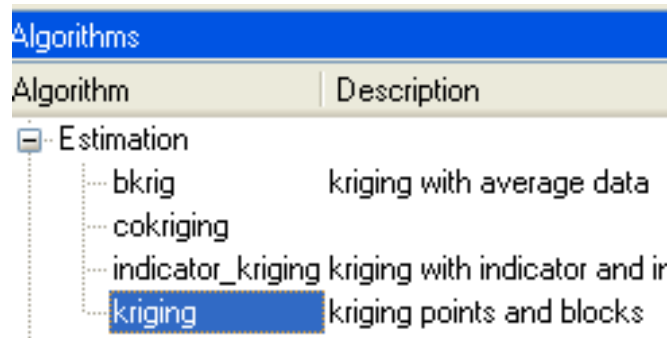


Figura 94 - Caixa de algoritmos do S-GeMS.

Em baixo desta secção aparece a caixa da krigagem. O “Grid name” é a grid que acabamos de criar com o nome “Grid da Horta” e a “New Property name” é a variável que a mesma vai estimar e por isso metemos fósforo. Seguidamente está a opção entre krigagem simples e krigagem normal. Não iremos explicar neste momento a diferença mas se for escolhida a simples é necessário inserir a média pretendida. No nosso caso optámos por krigagem normal (ordinary kriging). De seguida temos de dizer a partir de que pontos vão ser feitas as estimações com os campos “object” e “property”, o primeiro é “Amostra da Horta Transformadas” e o segundo o “fósforo”. Nesta página ainda falta definir o elipsóide de procura que nada tem a ver com o elipsóide de variografia. Este elipsóide existe para limitarmos a estimação dos pontos às amostras mais próximas ou não. Neste caso o terreno da horta é pequena e não necessitamos de estar a limitar. Metemos uma extensão muito superior aos limites do terreno com 50 em cada eixo do elipsóide e ângulo (0,0,0) correspondente. Consideramos as 6 amostras como o mínimo para calcular cada um dos blocos e 16 como o máximo (Figura 95).

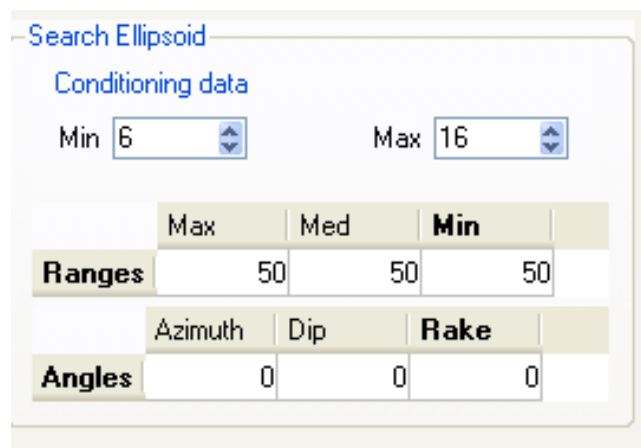


Figura 95 - Elipsóide de procura para a estimação.

De seguida passamos para a página seguinte ( General and Data Variogram ) e inserimos os dados a que chegamos no estudo de variografia com efeito pepita de 0 e apenas uma estrutura (por agora não explicaremos como fazer mais do que uma mas é possível meter mais de um modelo em cada variograma). O “Contribution” é equivalente ao “Sill” usado nos variogramas (patamar) que no nosso caso utilizamos a variância. Mais uma vez o valor é de 0,001632 com exponencial no “Type” e elipsóide de variografia como está na Figura 96.



Contribution 001632			
Type Exponential			
	Max	Med	Min
<b>Ranges</b>	25	18	2
	Azimuth	Dip	Rake
<b>Angles</b>	0	0	0

Figura 96 - Elipsóide de variografia inserido na pasta de variografia da krigagem.

De seguida carregamos em Run Algorithm e deixamos o cálculo acabar.

## ANÁLISE DOS RESULTADOS

O resultado da estimação, em planta, é o da Figura 97.



Figura 97 - Teor de fósforo no terreno da horta.

A análise da planta é insuficiente mas com o visualizador do S-GeMS é possível meter o modelo em qualquer perspectiva (Figura 98).

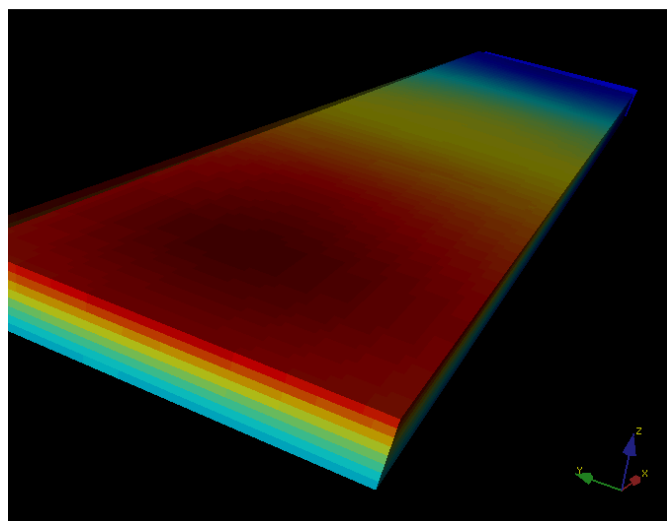


Figura 98 - Modelo da horta em perspectiva no visualizador do S-GeMS.

Ainda assim é necessário estudar o interior do modelo e deste modo isto não é possível. Para o fazer vamos ao **Objects** **Preferences** **Info** numa das secções do interface do S-GeMS e nesse menu vamos à parte respeitante ao "Volume Explorer". E escolhemos a opção "use volume Explorer" e também a opção "hide volume" (vai apagar o modelo e somente os cortes vão ficar visíveis). De seguida escolhemos em que eixo pretendemos fazer cortes e andamos com a barra respectiva de maneira a percorrer o volume em **x**, **y** e **z**. Se necessário é possível inserir mais de um corte no mesmo eixo fazendo um "Add" ao eixo que se pretende acrescentar. Na Figura 99 estão estes tópicos preenchidos e foi feito um "Add" ao **y** ficando com duas opções de corte.

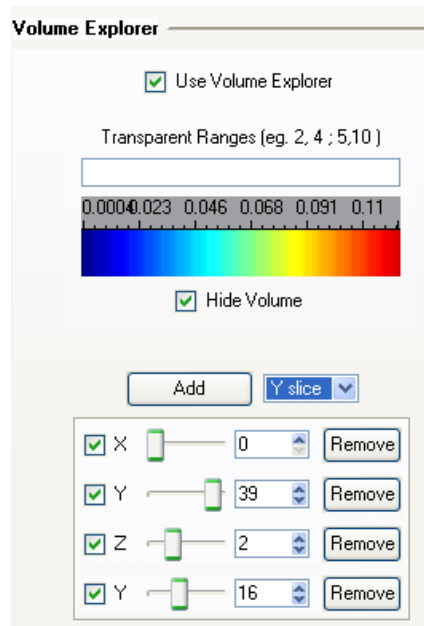


Figura 99 - Secção do "Volume Explorer" do S-GeMS.

O resultado destas operações vão ter o resultado da Figura 100 onde se pode ver os cortes especificados na Figura 99 no visualizador do S-GeMS.

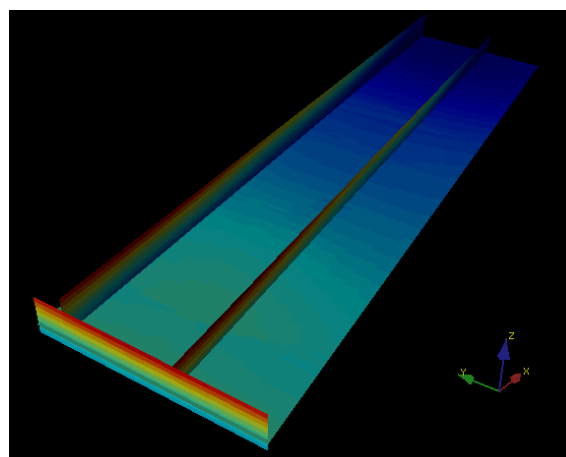


Figura 100 - Visualização dos cortes no visualizador do S-GeMS.

O teor de fósforo é maior na zona da esquerda do modelo da horta e parece ir perdendo importância quanto maior a profundidade.

# ESTIMAÇÃO DA IMPEDÂNCIA NUM RESERVATÓRIO PETROLÍFERO

Estes dados foram cedidos para a cadeira de Modelação de Reservatório Petrolíferos no Mestrado em Engenharia Geológica e Mineira no IST. Os dados são sintéticos mas pretendem emular um contexto de estudo possível.

## DADOS DO RESERVATÓRIO

Os dados do reservatório foram inseridos no software S-GeMS com o cabeçalho no excerto abaixo.

Estudo da Impedancia, Permeabilidade e Porosidade

6

Localização X

Localização Y

Localização Z

Impedancia

Permeabilidade

Porosidade

2 2 1 12969.747 270 0.00918

2 2 2 13413.708 270 0.00918

.... ..

É um conjunto de point data com coordenadas em **x**, **y** e **z** (conjunto tridimensional). Representam o estudo feito em vários poços ao longo de várias profundidades.

## ESTATÍSTICOS DESCRITIVOS

Os resultados no estudo de histogramas foram os da Figura 101.

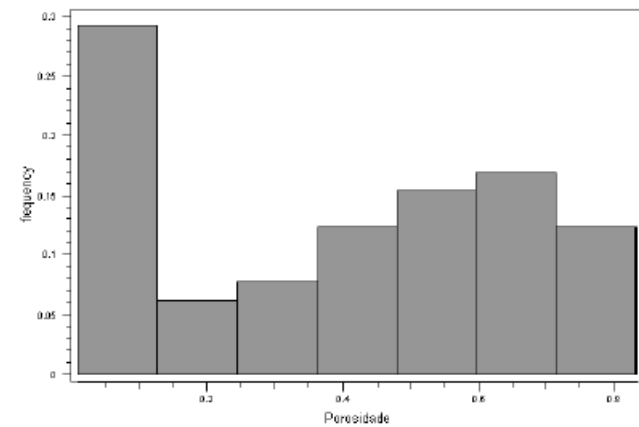
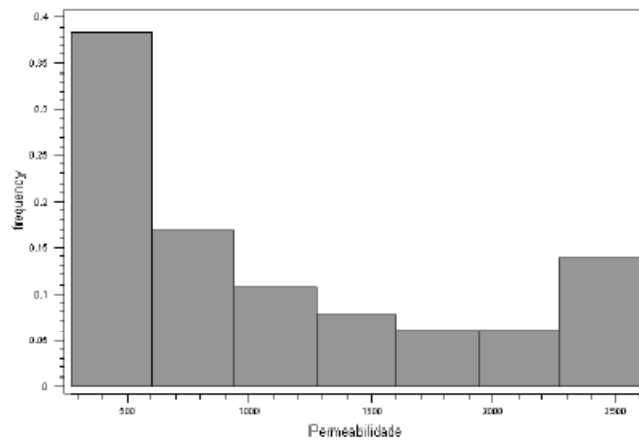
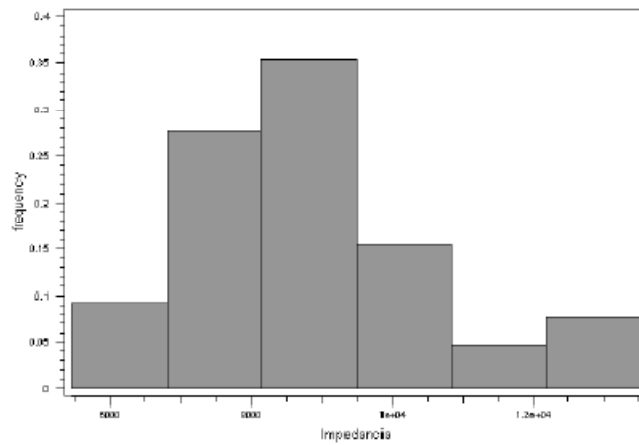


Figura 101 - Histogramas da Impedância (cima), permeabilidade (centro) e porosidade (baixo).

As funções de distribuição de probabilidade são consideravelmente diferentes e pelo menos no caso da permeabilidade e porosidade parece existir duas populações diferentes. A análise aos gráficos de dispersão revelou que não parece existir diferenciação significativa para considerar duas populações (Figura 102).

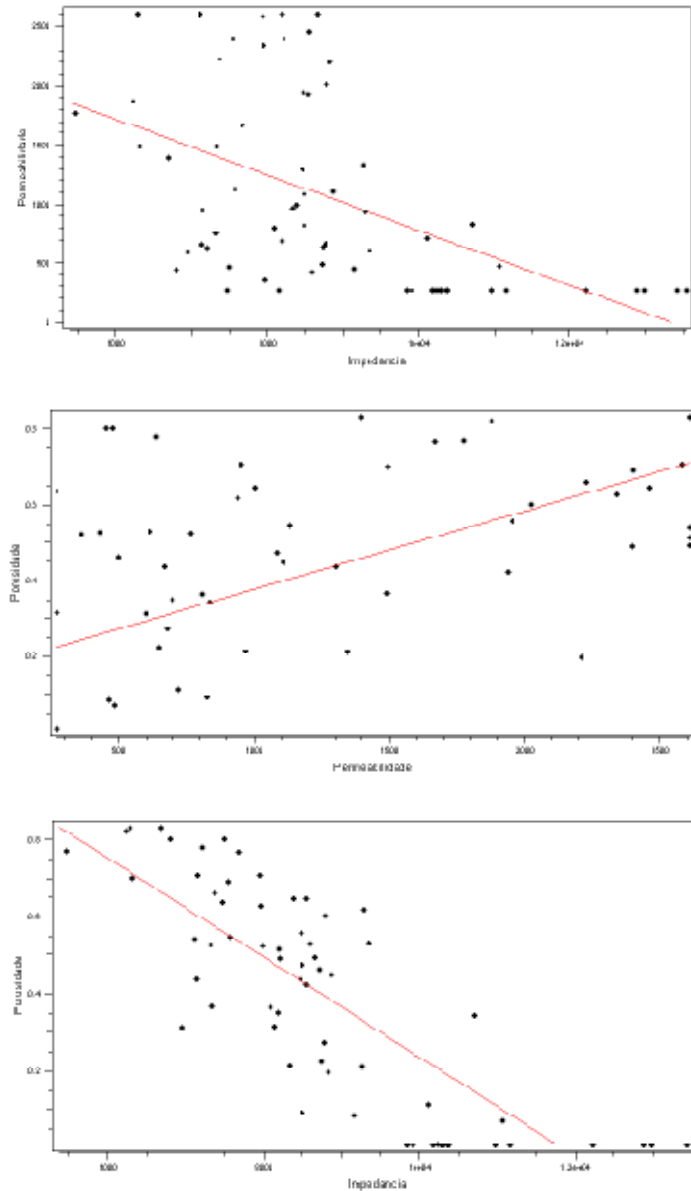


Figura 102 - Gráficos de dispersão "Permeabilidade vs Impedância" (cima), "Porosidade vs Permeabilidade" (centro), "Porosidade vs Impedância" (baixo).

O resumo dos estatísticos resultou na tabela da Figura 103.

	<i>Impedância</i>	<i>Permeabilidade</i>	<i>Porosidade</i>
<b>Média</b>	8810.89	1056.59	0.389361
<b>Variância</b>	3.11733e+06	644704	0.079783
<b>Máximo</b>	13531.80	2606.73	0.83187
<b>Mínimo</b>	5462.37	270	0.00918
<b>Número de dados</b>	65	65	65

Figura 103 - Tabela dos estatísticos descritivos.

Os coeficientes de correlação para todas as propriedades estão na Figura 104.

	Impedância	Permeabilidade	Porosidade
Impedância	1	-0.513664	-0.808093
Permeabilidade	-0.513664	1	0.592357
Porosidade	-0.808093	0.592357	1

Figura 104 - Tabela dos coeficientes de correlação.

## ESTUDO DE VARIOGRAFIA

Para fazermos o nosso estudo de variografia fomos ver quais os valores mínimos para cada uma das coordenadas (Figura 105).

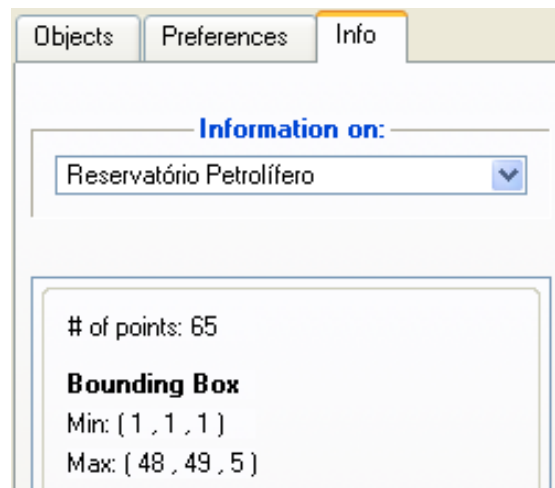


Figura 105 - Informações sobre o objecto "Reservatório Petrolífero".

De seguida fomos à secção "Variogram" e preenchemos os campos do menu inicial como na Figura 106.



Figura 106 - Campos iniciais da variografia.

Como os dados estão a cerca de 50 unidades de comprimento no plano  $xOy$  e 5 em  $z$  é necessário que o tamanho das classes de variografia seja pequeno o suficiente para fazer bons histogramas no plano horizontal e vertical. Por esse motivo os campos foram preenchidos como Figura 107.

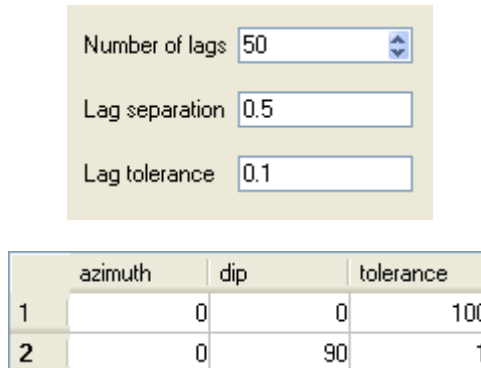


Figura 107 - Preenchimento dos campos de variografia.

Desta maneira (com tamanho de 0.5) mesmo no eixo  $z$  vão existir 10 classes de variografia. Repare-se que apenas duas direcções são consideradas (plano horizontal e eixo vertical). Na direcção horizontal (ângulo de tolerância superior a 90) é um variograma omni-direccional. O resultado é o da Figura 108.

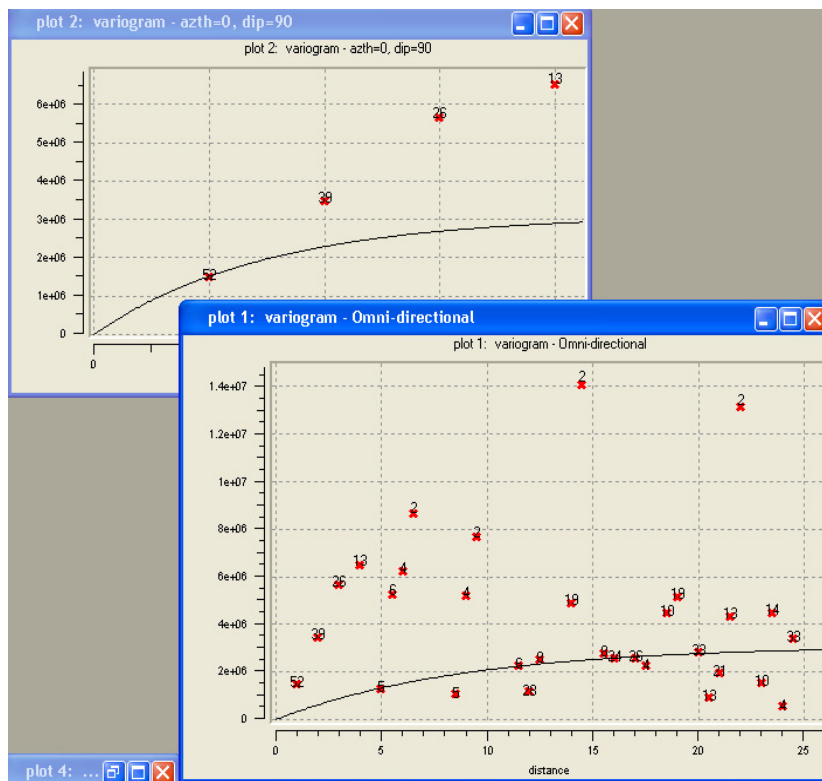


Figura 108 - Variogramas vertical (cima) e omni-direccional (baixo).

Note-se que o variograma omni-direccional tem muito ruído entre o qual os pontos que captou verticalmente. Se fossem muitos o ficheiro inicial teria que ser modificado para conseguirmos um bom variograma mas neste caso é fácil identificar os pontos pertencentes ao plano vertical por comparação. Eliminando esses pontos e inserindo o patamar com valor da

variância retirado do menu de histograma da impedância ficamos com o modelo da Figura 109 no plano horizontal.

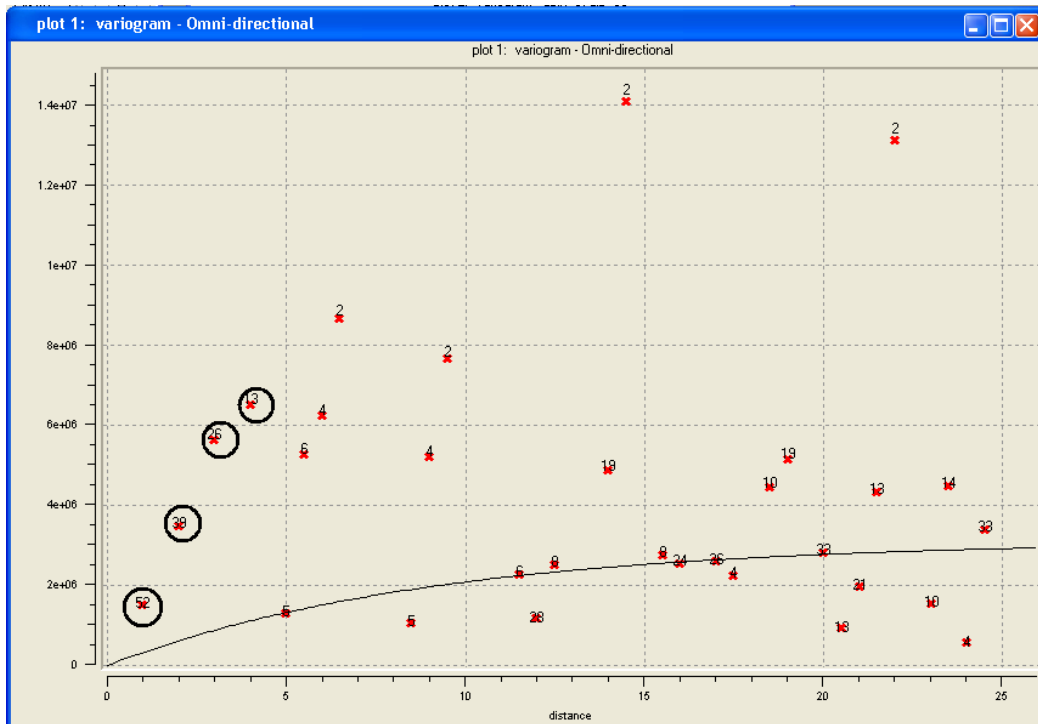


Figura 109 - Modelo do variograma omni-direcional onde foram desconsiderados os pontos do eixo vertical. O ruído assume-se ser devido também aos pontos do eixo vertical noutros poços.

O nosso elipsóide de variografia ficou como na Figura 110.

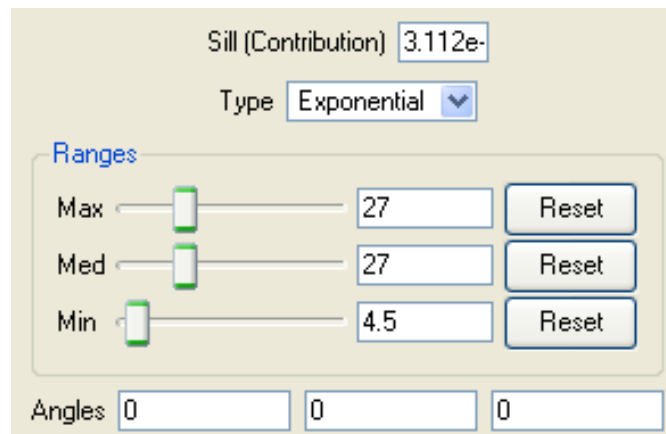


Figura 110 - Elipsóide de variografia.

O modelo utilizado foi o exponencial sem efeito pepita. Para fazermos a krigagem da impedância precisamos de uma grid. A grid foi criada tendo em conta os valores retirados das informações do objecto "Reservatório Petrolífero". Para isto seleccionamos a opção "New Cartesian Grid" e preenchemos os campos como se encontram na Figura 111.



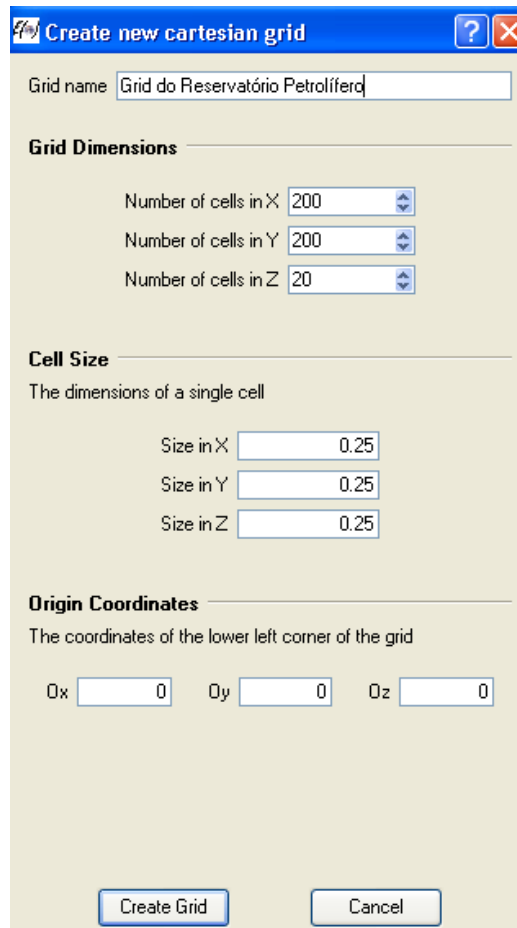


Figura 111 - Preenchimento dos dados no menu de criação da grid.

Foi dado o nome de “Grid do Reservatório Petrolífero” com número de células de 200, 200 e 20 para **x**, **y** e **z**, respectivamente. As dimensões dos blocos ficaram iguais para todas as direcções com 0.25 unidades de comprimento. Assim:

$$0,25 * 200 = 50$$

$$0,25 * 20 = 5$$

Que são aproximadamente as dimensões do campo considerado pelas amostras (50\*50\*5). No visualizador podemos confirmar se a grid ficou bem inserida ou não. Neste caso podemos constatar que sim a partir da Figura 112.

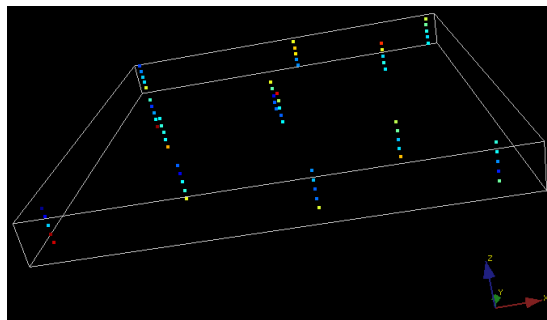


Figura 112 - Inserção da grid com as amostras.

## KRIGAGEM DA IMPEDÂNCIA

Para fazer a krigagem da impedância fomos buscar esta funcionalidade ao painel de algoritmos e preenchemos como se apresenta na Figura 113 e Figura 114.

General and Data Variogram

**Kriging Grid**

Grid Name  
Grid do Reservatório Petrolífero

New Property Name  
Impedância

Ordinary Kriging (OK)

Block Kriging

**Hard Data**

Object  
Reservatório Petrolífero

Property  
Impedancia

**Search Ellipsoid**

Conditioning data

Min 6 Max 30

	Max	Med	Min
<b>Ranges</b>	70	70	20

	Azimuth	Dip	Rake
<b>Angles</b>	0	0	0

Figura 113 - Parâmetros de krigagem (1).

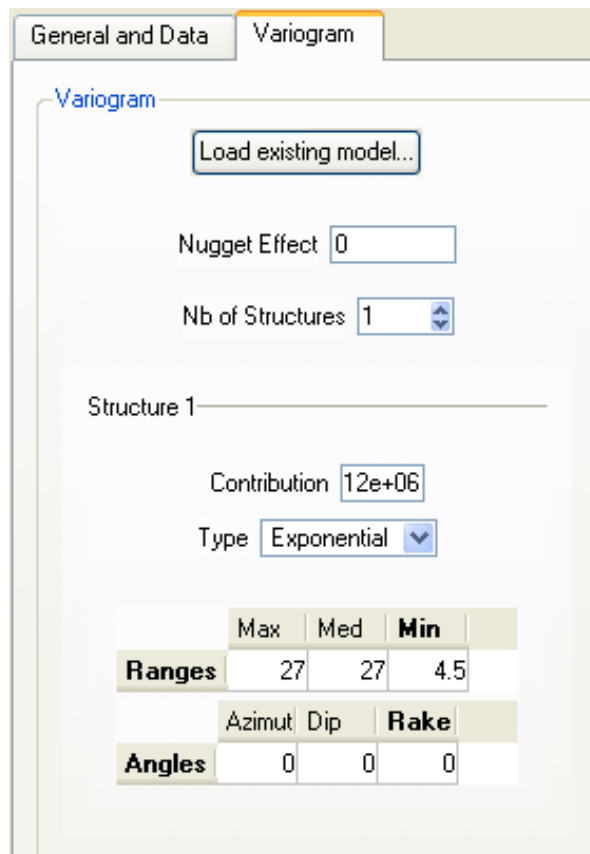


Figura 114 - Parâmetros de krigagem (2).

Fazemos o “Run Algorithm” (deve demorar algum tempo visto a grid ser grande) e o resultado pode ser visto após selecção como se apresenta na Figura 115.

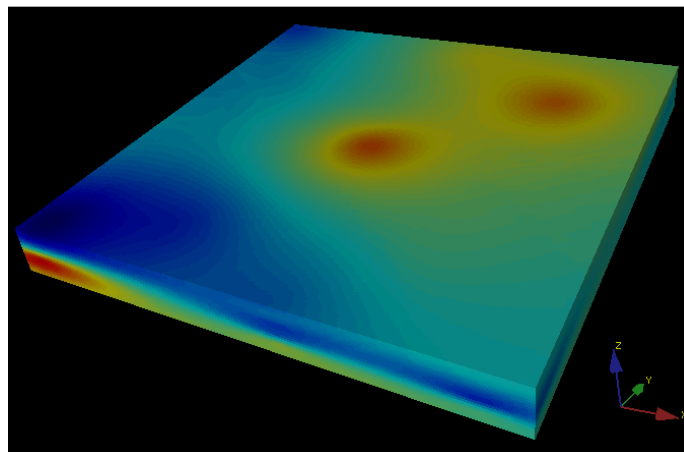


Figura 115 - Resultado da krigagem da impedância.

Foi uma krigagem bem sucedida. Para melhor análise da estimação foram feitos alguns cortes no “Volume Explorer” na zona de “Preferences”. O preenchimento do “Volume Explorer” é como o que está na Figura 116.

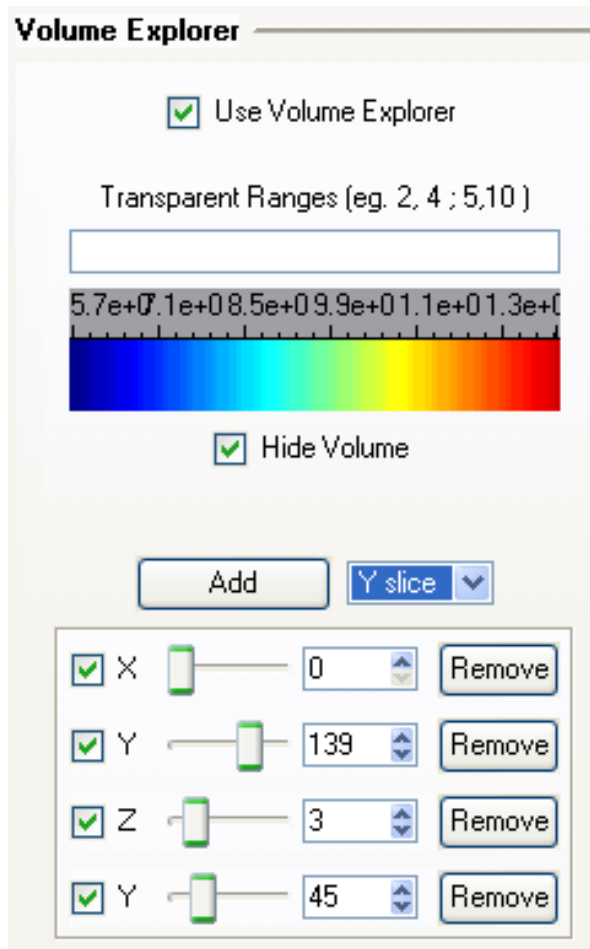


Figura 116 - Preenchimento do "Volume Explorer" para a estimação da impedância.

O resultado dos cortes é o da Figura 117.

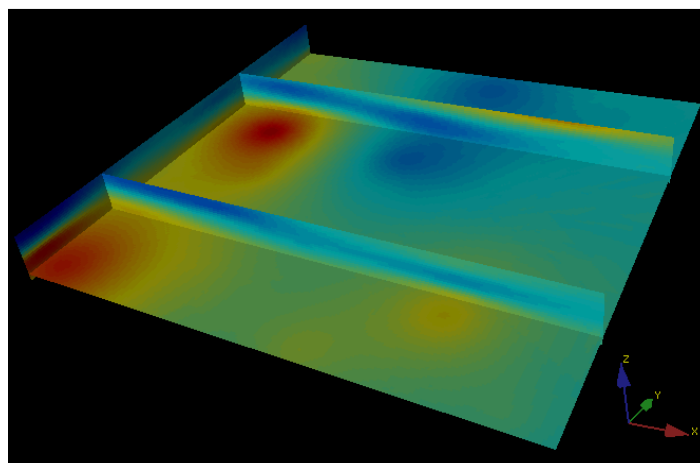


Figura 117 - Resultado dos cortes no "Volume Explorer".

O valor da impedância varia sendo mais forte no topo no canto superior direito e na base canto inferior esquerdo.

## EXERCÍCIO 1

Tenta fazer a estimação do teor de fósforo na primeira horta.

- Bidimensional.
- Seis amostras.
- Variografia pouco representativa.
- Esquema de coordenadas simples.
- Uma variável.

O estudo de variografia é pouco importante dado a escassez de dados que existe e o tamanho da horta. Faz apenas uma estimação por krigagem sem grandes atenções aos pormenores (ficheiro “Horta do meu quintal” na pasta de “exemplos sintéticos”).

## EXERCÍCIO 2

É o exemplo que utilizamos para fazer o processo de krigagem à mão mas também se pode fazer a partir do software. Faz a estimação do teor de ferro num jazigo.

- Bidimensional.
- Quinze amostras.
- Forte anisotropia.
- Malha quase regular.
- Uma variável.

O estudo de variografia é importante pois a anisotropia é considerável. Poderá haver alguma dificuldade em ajustar o modelo aos variogramas devido à escassez de pontos. A malha é regular e por isso existem direcções com muita facilidade em fazer os variogramas experimentais (“Jazigo de ferro” na pasta de “exemplos sintéticos”).

## EXERCÍCIO 3

Faz a análise estatística e estimação da porosidade, impedância e permeabilidade em recurso a vários poços num jazigo petrolífero.

- Tridimensional.
- Vários poços.
- Variografia precisa.
- Relações entre variáveis importantes.
- Esquema de coordenadas simples.
- 3 variáveis.

É um exemplo feito para bater certo. Deverá existir grande facilidade em fazer variogramas no plano  $xOy$  e alguma dificuldade em  $Oz$ . Existem relações entre as variáveis (“Reservatório Petrolífero” na pasta de “exemplos sintéticos”).

## EXERCÍCIO 4

Caso da nossa horta!!! Ainda falta fazer a estimação do arsénio e do cobre e também é melhor voltar a verificar a do fósforo.

- Tridimensional.
- Muitas amostras.
- Variografia razoável.
- 3 variáveis.
- Poderá ser melhorado com transformação de coordenadas.

Tem alguma complexidade mas nada que já não tenhas aprendido a fazer. No exemplo foi feita a transformação de coordenadas. Podes tentar com as coordenadas originais ou com as coordenadas transformadas. Sabes que mais?!, tenta com as duas (“Horta do Dept de Minas” e “Horta do Dept de Minas trans” na pasta de “exemplos sintéticos”).

## EXERCÍCIO 5

Este caso é real, não vou dizer de onde, nem do quê, mas trata-se de um jazigo e precisa de alguém que o estude. Por agora já deverás estar à vontade para mexer num caso que te poderia aparecer à frente em qualquer centro de geoestatística do mundo. Estás por tua conta (“Jazigo” na pasta de “exemplos reais”).

Boa sorte...

## BIBLIOGRAFIA

BARNES, R. *VARIOGRAM TUTORIAL*. GOLDEN SOFTWARE, INC.

BOHLING, G. (2007). *INTRODUCTION TO GEOSTATISTICS - HYDROGEOPHYSICS: THEORY, METHODS, AND MODELING*. EUA: KANSAS GEOLOGICAL SURVEY.

BOHLING, G. (2007). *S-GEMS TUTORIAL NOTES - HYDROGEOPHYSICS: THEORY, METHODS, AND MODELING*. EUA: KANSAS GEOLOGICAL SURVEY.

CLARK, I. (1979). *PRATICAL GEOSTATISTICS*. ESCÓCIA: GEOSTOKOS LIMITED.

MERKS, J. W. *SAMPLING AND STATISTICS: A SYNOPSIS OF TOOLS AND TECHNIQUES*. CANADA.

REMY, N., BOUCHER, A., & WU, J. (2006). *SGEMS USER'S GUIDE*.

SOARES, A. (2006). *GEOESTATÍSTICA PARA AS CIÊNCIAS DA TERRA E DO AMBIENTE*. PORTUGAL: IST PRESS.