



**UNIVERSIDADE TÉCNICA DE LISBOA
INSTITUTO SUPERIOR TÉCNICO**

**Desenvolvimento de um Sistema Geoestatístico para o
Tratamento de Imagem**

**Maria Aline da Silva Gomes Camacho Baião
(Mestre)**

Dissertação para obtenção do grau de Doutor em Ciências de Engenharia

Orientador: Doutor Henrique José Figueiredo Garcia Pereira

Co-Orientador: Doutor José António Almeida

Júri

Presidente: Reitor da Universidade Técnica de Lisboa

Vogais: Doutor Henrique José Figueiredo Garcia Pereira

Doutor António Jorge Gonçalves de Sousa

Doutor José António Almeida

Doutor Joaquim Eduardo de Sousa Góis

Doutor Pedro Miguel Berardo Duarte Pina

Setembro 2005

Resumo

Nesta dissertação é apresentado um *package*, construído em Visual C++ para ambiente Windows, que articula os procedimentos geoestatísticos de cálculo com entradas e saídas sob a forma de imagem com o objectivo de contribuir para o ajustamento dos métodos geoestatísticos às especificidades dos dados retirados de imagens.

Assim, dada uma imagem, o sistema efectua o tratamento geoestatístico solicitado e apresenta os resultados sob a forma de imagem. No sistema desenvolvido, foi ainda privilegiada a modelação automática do variograma (2D) como procedimento principal para a produção do modelo global e o recurso à opção manual, ao contrário do habitual, apenas para casos de existência de informação pericial minorando, dessa forma, a intervenção do utilizador que apenas dialoga com o sistema através de janelas auto-explicativas.

O algoritmo de modelação automática recorre aos algoritmos genéticos como metodologia para optimização e aos resultados do teste da validação cruzada para quantificação do grau de ajuste do modelo possibilitando ainda a detecção automática de anisotropias.

O *software* foi testado com três casos de estudo típicos apresentando-se os resultados que ilustram as suas potencialidades no domínio das aplicações envolvendo tratamento de imagem.

Palavras chave: imagem geoestatística modelação automática algoritmo genético

Abstract

This thesis describes a package build for Windows environments with Visual C++ that merges geostatistic procedures with input and output image in order to contribute to adjust geostatistical methods to the specificities of digital image data.

The system processes the required geostatistical image treatment showing the results by an output image. An automated model fitting is also included in the software as a procedure to produce the global model so the user only have to dialog with the system through user-friendly windows. The manual fit is available but is treated as a technique for fitting only when expert knowledge is known.

The automated fitting is based on a genetic algorithm approach and uses the results of cross validation to validate the model. Within this procedure the user can also calculate anisotropy without any special skills.

The software has been tested with three case studies and the results are presented to illustrate the system performance in what concerns the image treatment applications.

KEY-WORDS: image geostatistic automated fit genetic algorithm

Agradecimentos

Ao meu orientador e co-orientador.

A todos os que, de diversas formas, contribuíram para a realização desta tarefa.

Esta dissertação contém na contra-capá um CD de instalação do software implementado

Índice de Matérias

| | |
|--|----------|
| Índice de Figuras | xi |
| Índice de Tabelas | xvii |
| 1. Introdução | 1 |
| 1.1. Colocação do problema e objectivos | 1 |
| 1.2. Especificidade dos casos em estudo | 5 |
| 1.3. Organização da dissertação | 6 |
| 2. Metodologia | 8 |
| 2.1. Estratégia da metodologia implementada | 8 |
| 2.2. Revisão de metodologias geoestatísticas | 9 |
| 2.2.1. Medidas de continuidade espacial e variabilidade | 13 |
| 2.2.1.1. Variograma experimental | 13 |
| 2.2.1.2. Covariância espacial | 14 |
| 2.2.2. Modelos teóricos para o ajustamento do modelo experimental | 16 |
| 2.2.2.1. Efeito de pepita | 16 |
| 2.2.2.2. Modelo Esférico | 17 |
| 2.2.2.3. Modelo exponencial | 18 |
| 2.2.2.4. Modelo gaussiano | 18 |
| 2.2.2.5. Modelo de potência | 19 |
| 2.2.3. Estimação por <i>krigagem</i> normal e <i>cokrigagem</i> colocalizada | 20 |
| 2.2.4. Simulação geoestatística | 24 |
| 2.3. Conceitos relativos a imagens digitais | 27 |
| 2.3.1. Resolução de imagens digitais | 30 |

| | |
|---|-----------|
| 2.3.2. Georeferenciação de imagens | 31 |
| 2.3.3. Modelos de codificação de cor | 32 |
| 2.3.4. Alguns formatos de armazenamento de imagens matriciais | 34 |
| 2.3.4.1. Formato BMP | 35 |
| 2.3.4.2. Formato GIF | 35 |
| 2.3.4.3. Formato JPG | 35 |
| 2.3.4.4. Formato TIFF | 35 |
| 2.3.4.5. Formato PNG | 36 |
| 2.3.5. As operações sobre imagens e o conceito de vizinhança | 37 |
| 2.3.5.1. Vizinhança de um <i>pixel</i> | 37 |
| 2.3.5.2. Operações sobre imagens | 37 |
| 2.3.6. Reescalamento de imagens | 38 |
| 2.3.6.1. Vizinhança mais próxima | 38 |
| 2.3.6.2. Interpolação bilinear | 39 |
| 2.3.6.3. Interpolação bicúbica | 40 |
| | |
| 3. Ajustamento automático de variogramas experimentais por um algoritmo genético | 41 |
| | |
| 3.1. O problema do ajustamento dos variogramas experimentais e testes de validação | 41 |
| | |
| 3.2. Conceitos relativos a procura de soluções para problemas de otimização | 43 |
| | |
| 3.3. Computação evolucionária | 45 |
| | |
| 3.3.1. Algoritmos genéticos | 46 |
| | |
| 3.3.2. Programação evolucionária | 47 |
| | |
| 3.3.3. Estratégias evolutivas | 48 |
| | |
| 3.4. Algoritmos evolutivos | 49 |
| | |
| 3.4.1. Operadores usados | 49 |

| | |
|--|----|
| 3.4.1.1. Selecção | 49 |
| 3.4.1.2. Recombinação | 50 |
| 3.4.1.3. Mutação | 50 |
| 3.4.1.4. Substituição | 51 |
| 3.4.2. Inicialização dos parâmetros de um algoritmo evolutivo | 51 |
| 3.4.3. Vantagens e desvantagens dos algoritmos evolutivos | 52 |
| 3.5. Componentes do algoritmo genético implementado | 53 |
| 3.5.1. População | 54 |
| 3.5.2. Critério de paragem | 56 |
| 3.5.3. Avaliação | 56 |
| 3.5.4. Inicialização | 57 |
| 3.5.5. Selecção | 57 |
| 3.5.7. Mutação | 58 |
| 3.5.8. Exemplo de aplicação e contraste com o sistema tradicional manual | 58 |
| 3.5.9. Estudo comparativo com outros <i>softwares</i> disponíveis | 67 |
| 4. Descrição do Sistema e Manual do Utilizador | 70 |
| 4.1. Síntese das principais funcionalidades | 70 |
| 4.2. Manual do utilizador | 75 |
| 4.2.1. Entradas e <i>saídas</i> de dados | 77 |
| 4.2.1.1. Ficheiros de dados admissíveis | 77 |
| 4.2.1.2. Descrição dos ficheiros de entrada | 78 |
| 4.2.1.3. Ficheiros de saída | 80 |
| 4.2.1.4. Validações e escolha de variáveis/unidades de medida no ficheiro de dados | 81 |
| 4.2.1.5. Visualização dos dados de entrada | 83 |
| 4.2.2. Variografia e ajustamento do modelo teórico | 85 |

| | |
|---|-----|
| 4.2.2.1. Análise estrutural | 85 |
| 4.2.2.2. Gráficos de superfície | 85 |
| 4.2.2.3. Gráficos direccionais | 89 |
| 4.2.2.4. Diagramas de dispersão | 96 |
| 4.2.2.5. Ajustamento da função teórica ao variograma experimental | 99 |
| 4.2.3. Estimação | 103 |
| 4.2.3.1. Ordinary <i>Krigagem</i> | 103 |
| 4.2.3.2. <i>Cokrigagem</i> colocalizada | 106 |
| 4.2.4. Simulação Sequencial Gaussiana | 109 |
| 4.3. Alguns aspectos particulares da implementação | 112 |
| 4.3.1. Ambiente de desenvolvimento | 112 |
| 4.3.2. Diagrama de classes | 113 |
| 4.3.3. Descrição dos algoritmos mais relevantes | 115 |
| 4.3.3.1. Entrada de dados | 115 |
| 4.3.3.2. Cálculos de Estatísticos | 118 |
| 4.3.3.3. Representações gráficas | 120 |
| 4.3.3.4. Modelação | 121 |
| 4.3.3.5. Validação cruzada, <i>Krigagem</i> e <i>cokrigagem</i> colocalizada, simulação | 124 |
| 4.3.3.6. <i>Krigagem</i> de imagens | 125 |
| 4.3.3.7. Simulação de imagens | 126 |
| 4.3.4. Principais estruturas intermédias | 127 |
| 4.3.5. Bibliotecas e módulos importados | 128 |

| | | |
|---------------|---|------------|
| 5. | Estudo de casos | 130 |
| 5.1. | Reescalonamento de imagens | 130 |
| 5.1.1. | Fotografia em tons de cinzento | 133 |
| 5.1.2. | Fotografia a cores | 136 |
| 5.1.3. | Imagem de satélite | 142 |
| 5.1.4. | Análise dos resultados | 148 |
| 5.2. | Inferência espacial | 150 |
| 5.2.1. | Análise Estatística | 153 |
| 5.2.2. | Análise variográfica | 154 |
| 5.2.3. | Estimação por <i>krigagem</i> | 155 |
| 5.2.4. | Simulação de imagens | 156 |
| 5.2.5. | Análise dos resultados | 158 |
| 5.3. | Reprodução de padrões de cor | 160 |
| 5.3.1. | Amostra 1 | 161 |
| 5.3.2. | Amostra 2 | 175 |
| 6. | Conclusões | 193 |
| | Referências | 197 |
| | Apêndice I | |
| | Tabela comparativa de métodos de estimação | 202 |
| | Apêndice II | |
| | Resultados obtidos com uma metodologia bottom-up | 208 |

Índice de Figuras

Capítulo 2. Metodologia

| | |
|---|----|
| Fig. 1 Efeito de pepita puro | 17 |
| Fig. 2 Modelo esférico | 17 |
| Fig. 3 Modelo exponencial | 18 |
| Fig. 4 Modelo gaussiano | 19 |
| Fig. 5 Modelo de potência | 20 |
| Fig. 6 – Simulação de um valor ZI | 25 |
| Fig. 7 Informação de cor em hexadecimal com 6 caracteres | 29 |
| Fig. 8 Imagem matricial (imagem de satélite) e sua ampliação (2x) | 30 |
| Fig. 9 Imagens com 72 e 300 <i>ppi</i> com a mesma área de ocupação | 31 |
| Fig. 10 Síntese aditiva | 32 |
| Fig. 11 Informação RGB | 33 |
| Fig. 12 Informação RGB em imagem <i>grayscale</i> | 34 |
| Fig. 13 Tipos de vizinhanças | 37 |
| Fig. 14 <i>Pixels</i> a preencher na imagem da esquerda por ampliação de 30% da imagem da direita | 38 |
| Fig. 15 Ampliação de 300% pelo método da vizinhança mais próxima | 39 |
| Fig. 16 Vizinhança de um <i>pixel</i> | 39 |
| Fig. 17 Ampliação de 300% pelo método da interpolação bilinear | 40 |
| Fig. 18 Ampliação de 300% pelo método da interpolação bicúbica | 40 |

Capítulo 3. Ajustamento automático de variogramas experimentais por um algoritmo genético

| | |
|---|----|
| Fig. 19 Curva ajustada a um conjunto de pontos | 43 |
| Fig. 20 Algoritmo evolucionário | 45 |
| Fig. 21 Fluxograma do algoritmo genético implementado | 54 |
| Fig. 22 Cromossoma | 55 |
| Fig. 23 Histograma da população de 100 elementos com lei uniforme | 58 |
| Fig. 24 Histograma da população de 100 elementos com lei gaussiana | 59 |
| Fig. 25 Histograma da população de 100 elementos com lei lognormal | 59 |
| Fig. 26 Variograma experimental e modelo teórico ajustado à população com lei uniforme | 60 |
| Fig. 27 Variograma experimental e modelo teórico ajustado à população com lei gaussiana | 60 |
| Fig. 28 Variograma experimental e modelo teórico ajustado à população com lei lognormal | 60 |
| Fig. 29 População com lei uniforme: representação dos valores reais vs valores estimados para os modelos de variograma com $a = 18$ (modelo do ajustamento manual), $a = 8,2$ (modelo a que corresponde o menor EM) e $a = 15$ (modelo a que corresponde o menor EQM) e o modelo ajustado pelo algoritmo genético | 64 |
| Fig. 30 População com lei gaussiana: representação dos valores reais vs valores estimados para os modelos de variograma com $a = 20$ (modelo do ajustamento manual), $a = 8,5$ (modelo a que corresponde o menor EM) e $a = 17$ (modelo a que corresponde o menor EQM) e o modelo ajustado pelo algoritmo genético | 65 |
| Fig. 31 População com lei lognormal: representação dos valores reais vs valores estimados para os modelos de variograma com $a = 22$ (modelo do ajustamento manual), $a = 14,9$ (modelo a que corresponde o menor EM) e $a = 20$ (modelo a que corresponde o menor EQM) e o modelo ajustado pelo algoritmo genético | 65 |
| Fig. 32 Resultados do ajustamento com o algoritmo genético para a população com lei uniforme (esquerda: 10 lags; direita 33 lags) | 66 |
| Fig. 33 Resultados do ajustamento com o algoritmo genético para a população com lei gaussiana (esquerda: 10 lags; direita 33 lags) | 66 |
| Fig. 34 Resultados do ajustamento com o algoritmo genético para a população com lei lognormal (esquerda: 10 lags; direita 33 lags) | 66 |

Capítulo 4. Descrição do Sistema e Manual do Utilizador

| | |
|--|----|
| Fig. 35 Diagrama de funções | 71 |
| Fig. 36 Opções disponíveis no menu principal | 76 |
| Fig. 37 Caixas de diálogo para a escolha do ficheiro | 77 |
| Fig. 38A Malha irregular 2D | 78 |

| | |
|--|----|
| Fig. 38B Malha regular 2D | 79 |
| Fig. 39 Exemplo de uma <i>imagem</i> matricial em formato bitmap com 190x142 <i>pixels</i> | 79 |
| Fig. 40 Extracto do ficheiro de resultados da <i>krigagem</i> (*.out) | 80 |
| Fig. 41 Caixa de diálogo 2D/3D (malha com <i>coordenadas</i> explícitas) | 81 |
| Fig. 42 Estatísticas iniciais (malha com <i>coordenadas</i> explícitas) | 82 |
| Fig. 43 Caixa de diálogo 2D/3D (caso de <i>ficheiro</i> de imagem ou regular 2D) | 83 |
| Fig. 44 Representação gráfica de uma estrutura de <i>dados</i> 2D com <i>coordenadas</i> explícitas | 83 |
| Fig. 45 Representação gráfica de uma estrutura de <i>dados</i> 3D no plano YOZ | 84 |
| Fig. 46 Representação gráfica de uma estrutura de <i>dados</i> 2D com <i>coordenadas</i> implícitas | 84 |
| Fig. 47 Caixa de diálogo <i>Settings for Variography</i> | 85 |
| Fig. 48 Janela para estruturas de dados com <i>coordenadas</i> explícitas | 85 |
| Fig. 49 Janela para estruturas de dados com <i>coordenadas</i> explícitas | 86 |
| Fig. 50 Parâmetros: Superfície | 86 |
| Fig. 51 Janela Surface com os resultados do variograma para dados com <i>coordenadas</i> explícitas | 87 |
| Fig. 52 Janela Surface com os resultados do variograma para dados sem <i>coordenadas</i> explícitas | 87 |
| Fig. 53 Exemplo dos resultados sob a forma de tabela das superfícies do variograma para amostras com <i>coordenadas</i> explícitas | 88 |
| Fig. 54 Exemplo dos resultados sob a forma de tabela das superfícies do variograma para amostras sem <i>coordenadas</i> explícitas | 88 |
| Fig. 55 Janela Directional (caso malha 2D com <i>coordenadas</i> explícitas) | 89 |
| Fig. 56 Janela Directional (caso malha 3D sem <i>coordenadas</i> explícitas) | 90 |
| Fig. 57 Parâmetros – Direccional - estruturas de dados com <i>coordenadas</i> explícitas | 91 |
| Fig. 58 Exemplo dos resultados sob a forma de gráfico direccional do variograma para amostras com <i>coordenadas</i> explícitas | 92 |
| Fig. 59 Exemplo dos resultados sob a forma de gráfico direccional do variograma para amostras sem <i>coordenadas</i> explícitas | 93 |
| Fig. 60 Parâmetros – Direccional malha (6x5) sem <i>coordenadas</i> explícitas | 93 |
| Fig. 61 Exemplo dos resultados sob a forma de tabela do variograma para amostras com <i>coordenadas</i> explícitas | 94 |
| Fig. 62 Exemplo dos resultados sob a forma de tabela do variograma para amostras sem <i>coordenadas</i> explícitas | 95 |

| | |
|---|-----|
| Fig. 63 Diagrama de dispersão para o Lag 2 (estrutura com coordenadas explícitas) | 96 |
| Fig. 64 Exemplo dos resultados sob a forma de gráfico de dispersão do variograma para amostras com coordenadas explícitas | 97 |
| Fig. 65 Exemplo dos resultados sob a forma de gráfico de dispersão do variograma para amostras sem coordenadas explícitas | 97 |
| Fig. 66 Janela Fitting Model | 99 |
| Fig. 67 Janela Fitting após o Fit All (imagem) | 100 |
| Fig. 68 Janela Fitting Model após o Auto Fit | 102 |
| Fig. 69 Janela de interface da opção Ordinary Kriging | 103 |
| Fig. 70 Janela de resultados de estimação (casos não imagem) | 105 |
| Fig. 71 Janela de resultados de estimação (caso de imagem) | 106 |
| Fig. 72 Janela interface <i>Cokrigagem</i> colocalizada | 107 |
| Fig. 73 Janela Resultados de <i>Cokrigagem</i> colocalizada | 109 |
| Fig. 74 Janela interface simulação sequencial gaussiana | 110 |

Capítulo 5. Estudo de casos

| | |
|--|-----|
| Fig. 75 Diagrama da metodologia proposta para a ampliação de imagens com uma banda | 131 |
| Fig. 76 Diagrama da metodologia proposta para a ampliação de imagens com três bandas | 132 |
| Fig. 77 Fotografia em tons de cinzento | 133 |
| Fig. 78 Variogramas experimentais e modelo teórico omnidireccional | 133 |
| Fig. 79 Variogramas experimentais e modelo teórico, direcção 0°, 45°, 90° e 135° | 134 |
| Fig. 80 Foto inicial, ampliação a 30% com <i>pixels</i> “sem cor” e resultado da <i>krigagem</i> | 135 |
| Fig. 81 (da esquerda para a direita) Ampliações a 30% obtidos pelo método da vizinhança mais próxima, interpolação bilinear e bicúbica | 135 |
| Fig. 82 Fotografia a cores | 136 |
| Fig. 83 Variogramas experimentais e modelo teórico, canal <i>Red</i> , modelo omnidireccional | 136 |
| Fig. 84 Variogramas experimentais e modelo teórico, canal <i>Red</i> , direcção 0°, 45°, 90° e 135° | 137 |
| Fig. 85 Variogramas experimentais e modelo teórico, canal <i>Green</i> , modelo omnidireccional | 138 |
| Fig. 86 Variogramas experimentais e modelo teórico, canal <i>Green</i> , direcção 0°, 45°, 90° e 135° | 138 |
| Fig. 87 Variogramas experimentais e modelo teórico, canal <i>Blue</i> , modelo omnidireccional | 139 |

| | |
|--|-----|
| Fig. 88 Variogramas experimentais e modelo teórico, canal Blue, direcção 0°, 45°, 90° e 135° | 140 |
| Fig. 89 Foto inicial, ampliação a 30% com <i>pixels</i> “sem cor” e resultado da <i>krigagem</i> | 141 |
| Fig. 90 (da esquerda para a direita) Ampliações a 30% com o método da vizinhança mais próxima, interpolação bilinear e interpolação bicúbica | 141 |
| Fig. 91 Composição RGB de 3 <i>bandas</i> de uma imagem de satélite Landsat TM | 142 |
| Fig. 92 Variogramas experimentais e modelo teórico, canal Red, modelo omnidireccional | 142 |
| Fig. 93 Variogramas experimentais e modelo teórico, canal Red, direcção 0°, 45°, 90° e 135° | 143 |
| Fig. 94 Variogramas experimentais e modelo teórico, canal Green, modelo omnidireccional | 144 |
| Fig. 95 Variogramas experimentais e <i>modelo</i> teórico, canal Green, direcção 0°, 45°, 90° e 135° | 144 |
| Fig. 96 Variogramas experimentais e modelo teórico, canal <i>Blue</i> , modelo omnidireccional | 145 |
| Fig. 97 Variogramas experimentais e modelo teórico, canal Blue, direcção 0°, 45°, 90° e 135° | 146 |
| Fig. 98 Imagem inicial, ampliação a 30% com <i>pixels</i> “sem cor” e resultado da <i>krigagem</i> | 147 |
| Fig. 99 (da esquerda para a direita) ampliações a 30% com método da vizinhança mais próxima, interpolação bilinear e interpolação bicúbica | 147 |
| Fig. 101 Diagrama síntese da metodologia proposta para a estimação de uma imagem média da precipitação e parametrização da incerteza por simulação | 151 |
| Fig. 101 Diagrama síntese da metodologia <i>proposta</i> para a estimação de uma imagem média da precipitação e parametrização da incerteza por simulação | 151 |
| Fig. 102 Modelo digital de terreno do Algarve | 152 |
| Fig. 103 Estatísticas univariadas da precipitação | 153 |
| Fig. 104 Estatísticas univariadas das cotas para as 36 localizações das estações meteorológicas | 153 |
| Fig. 105 Estatística bivariada entre <i>a</i> precipitação e as cotas | 154 |
| Fig. 106 Modelo teórico de variograma da precipitação total | 155 |
| Fig. 107 Estimação da pluviosidade média na região do Algarve. Esq) <i>Krigagem</i> normal; dir) <i>cokrigagem</i> colocalizada | 156 |
| Fig. 108 Variância de <i>krigagem</i> da pluviosidade estimada na região do Algarve. Esq) <i>Krigagem</i> normal; dir) <i>cokrigagem</i> colocalizada | 156 |
| Fig. 109 Simulação da pluviosidade na região do Algarve. a) b) e c) 3 imagens simuladas; d) variância local das 100 realizações | 157 |
| Fig. 110 Simulação da pluviosidade na região do Algarve. a) b) e c) 3 imagens co-simuladas com o modelo digital do terreno; d) variância local das 100 realizações | 158 |

| | |
|---|-----|
| Fig. 111 Diagrama da metodologia proposta para a simulação de imagens de padrões de rochas ornamentais | 160 |
| Fig. 112 Amostra de mármore | 161 |
| Fig. 113 Variogramas experimentais e modelo teórico, canal Red, modelo omnidireccional | 161 |
| Fig. 114 Variogramas experimentais e modelo teórico, canal <i>Red</i> , direcção 0°, 45°, 90° e 135° | 162 |
| Fig. 115 Variogramas experimentais e <i>modelo</i> teórico, canal Green, modelo omnidireccional | 162 |
| Fig. 116 Variogramas experimentais e modelo teórico, canal Green, direcção 0°, 45°, 90° e , 135° | 163 |
| Fig. 117 Variogramas experimentais e modelo teórico, canal Blue, modelo omnidireccional | 163 |
| Fig. 118 Variogramas experimentais e modelo teórico, canal <i>Blue</i> , direcção 0°, 45°, 90° e , 135° | 164 |
| Fig. 119 3 imagens simuladas, canal Red | 165 |
| Fig. 120 3 imagens simuladas, canal Green | 166 |
| Fig. 121 – 3 imagens simuladas, canal Blue | 166 |
| Fig. 122 Amostra inicial e 3 amostras simuladas | 167 |
| Fig. 123 Variogramas experimentais do canal <i>Red</i> sobrepostos ao modelo teórico dos dados | 168 |
| Fig. 124 Variogramas experimentais do canal Green sobrepostos ao modelo teórico dos dados | 169 |
| Fig. 125 Variogramas experimentais do canal <i>Blue</i> sobrepostos ao modelo teórico dos dados | 170 |
| Fig. 126 Amostra de mármore | 175 |
| Fig. 127 Variogramas experimentais e modelo teórico, canal Red, modelo omnidireccional | 175 |
| Fig. 128 Variogramas experimentais e modelo teórico, canal Red, direcção 0°, 45°, 90° e , 135° | 176 |
| Fig. 129 Variogramas experimentais e modelo teórico, canal Green, modelo omnidireccional | 176 |
| Fig. 130 Variogramas experimentais e modelo teórico, canal Green, direcção 0°, 45°, 90° e 135° | 177 |
| Fig. 131 Variogramas experimentais e modelo teórico, canal Blue, modelo omnidireccional | 177 |
| Fig. 132 Variogramas experimentais e modelo <i>teórico</i> , canal <i>Blue</i> , direcção 0°, 45°, 90° e 135° | 178 |
| Fig. 133 3 imagens simuladas, canal Red | 179 |
| Fig. 134 3 imagens simuladas, canal Green | 180 |
| Fig. 135 3 imagens simuladas, canal Blue | 180 |
| Fig. 136 Amostra inicial e 3 amostras simuladas | 181 |
| Fig. 137 Variogramas experimentais do canal Red sobrepostos <i>ao</i> modelo teórico dos dados | 182 |
| Fig. 138 Variogramas experimentais do canal <i>Green</i> sobrepostos ao modelo teórico dos dados | 183 |
| Fig. 139 Variogramas experimentais do canal <i>Blue</i> sobrepostos ao modelo teórico dos dados | 183 |

Índice de Tabelas

Capítulo 3. Ajustamento automático de variogramas experimentais por um algoritmo genético

| | |
|--|----|
| Tabela 1. Resultados da validação cruzada para vários parâmetros do variograma e população com lei uniforme | 61 |
| Tabela 2. Resultados da validação cruzada para vários parâmetros do variograma e população com lei gaussiana | 62 |
| Tabela 3. Resultados da validação cruzada para vários parâmetros do variograma e população com lei lognormal | 63 |

Capítulo 5. Estudo de casos

| | |
|---|-----|
| Tabela 4. Tabela de correlações entre os canais de cor Red, Green e Blue | 164 |
| Tabela 5. Comparação entre frequências absolutas e relativas para intervalos de cor relativamente à imagem original e imagem simulada (#1) | 171 |
| Tabela 6. Comparação entre frequências absolutas e relativas para intervalos de cor relativamente à imagem original e imagem simulada (#2) | 172 |
| Tabela 7. Comparação entre frequências absolutas e relativas para intervalos de cor relativamente à imagem original e imagem simulada (#3) | 173 |
| Tabela 8. Frequência de <i>pixels</i> (em %) convertidos em novos intervalos de cor e intervalos de cor que desaparecem nas imagens simuladas | 174 |
| Tabela 9. Comparação dos modelos teóricos da imagem original e das imagens simuladas | 174 |
| Tabela 10. Tabela de correlações entre os canais de cor <i>Red</i> , <i>Green</i> e <i>Blue</i> | 178 |
| Tabela 11. Comparação entre frequências absolutas e relativas para intervalos de cor relativamente à imagem original e imagem simulada (#1) | 185 |
| Tabela 12. Comparação entre frequências absolutas e relativas para intervalos de cor relativamente à imagem original e imagem simulada (#1) (continuação) | 186 |

| | |
|--|------------|
| Tabela 13. Comparação entre frequências absolutas e relativas para intervalos de cor relativamente à imagem original e imagem simulada (#2) | 187 |
| Tabela 14. Comparação entre frequências absolutas e relativas para intervalos de cor relativamente à imagem original e imagem simulada (#2) (continuação) | 188 |
| Tabela 15. Comparação entre frequências absolutas e relativas para intervalos de cor relativamente à imagem original e imagem simulada (#3) | 189 |
| Tabela 16. Comparação entre frequências absolutas e relativas para intervalos de cor relativamente à imagem original e imagem simulada (#3) (continuação) | 190 |
| Tabela 17. Frequência de <i>pixels</i> (em %) convertidos em novos intervalos de cor e intervalos de cor que desaparecem nas imagens simuladas | 191 |
| Tabela 18. Comparação dos modelos teóricos da imagem original e das imagens simuladas | 191 |

1. Introdução

1.1. Colocação do problema e objectivos

As imagens constituem uma importante fonte de informação em quase todos os domínios. De facto, com o avanço das tecnologias informáticas, a Aquisição e o Tratamento de Imagens em formato digital tornaram-se duas etapas, subsequentes e indispensáveis, à interpretação de numerosos fenómenos em quase todas as áreas científicas e tecnológicas, nomeadamente, na Física, Geografia, Ciências Médicas, Engenharia e, mais especificamente, nas Ciências da Terra e do Ambiente.

Para o Tratamento de Imagens digitais podem, entre outras, utilizar-se metodologias geoestatísticas bastando, para tal, que a imagem seja formalizada como a realização de uma função aleatória num domínio espacial, dotada de uma estrutura que possa ser expressa sob um sistema de autocorrelações dependente do vector distância euclideana.

A Geoestatística (ciência que aborda os fenómenos naturais sob o ponto de vista estatístico) foi utilizada pela primeira vez em 1962, por G. Matheron, do Centro de Geoestatística de Fontainebleau, em França [MA 65]. Nascia assim um novo ramo da ciência que viria a sofrer uma rápida evolução e teria importantes repercussões em diversas outras áreas. Tal como na maioria das ciências aplicadas, o nascimento e o crescimento da geoestatística beneficiaram da rápida evolução das ciências da computação, em que a criação de ferramentas e de algoritmos (*software*) tão necessários para a geoestatística se desenvolveu em paralelo com a evolução da tecnologia (*hardware*).

Inicialmente, a geoestatística restringiu-se ao domínio estrito da avaliação de reservas geológicas e mineiras, ou seja, regionalizações que descrevem fenómenos geológicos a partir de planos de amostragem pontuais e muito limitados ([DA 77], [JH 78]).

A extensão da geoestatística às aplicações onde é necessária a integração de dados sob a forma de imagens é mais recente. As primeiras aplicações surgem na década de 80 do século XX, principalmente na área da caracterização de reservatórios petrolíferos, constituindo uma forma

de integrar informação de elevada densidade espacial, por exemplo proveniente de prospecção geofísica, em algoritmos de simulação. Posteriormente surgem novas aplicações na área do ambiente, principalmente com a utilização de imagens de satélite e fotografia aérea em problemas de caracterização e monitorização.

A integração de dados pontuais e imagens em algoritmos geoestatísticos consagrou as designações *informação primária* (ou principal) e *informação secundária* (ou suplementar) que reflectem o grau de incerteza associada a cada um destes dois tipos de dados. Os dados resultantes de análises de laboratório ou de outras medições designam-se por informação primária por representarem informação mais fidedigna, que em contrapartida é tipicamente escassa, de aquisição morosa e de elevado custo. A informação secundária – que, em contrapartida, é abundante, de aquisição rápida e de baixo custo - surge geralmente sob a forma de imagens (imagens de detecção remota, prospecção geofísica, prospecção sísmica, modelo digital do terreno) e a sua intervenção é normalmente indirecta, ou seja, por correlação com a informação primária.

A articulação de imagens de um dado fenómeno com a informação pontual obtida a partir de planos de amostragem convencionais, desde que exista correlação local entre os dois tipos de informação, constitui um dos avanços mais importantes da geoestatística nos algoritmos de estimação por *krigagem* e simulação, nomeadamente, ao permitir a inferência indirecta, mas de forma muito mais contínua, de uma variável medida em escassos locais. Neste âmbito são muito utilizados os algoritmos de estimação baseados em métodos factoriais [SOU 00], que possibilitam a visualização dos dados em espaços de dimensão mais reduzida e os métodos de estimação por *cokrigagem* colocalizada, ou seja, aqueles em que são conhecidos os valores de um atributo indirecto em todas as localizações a estimar.

As imagens, não só pelo seu maior volume de informação, comparativamente a uma amostragem pontual, mas também pelo seu arranjo numa malha regular de linhas e colunas, exigem a reformulação dos programas informáticos subjacente às ferramentas geoestatísticas, onde intervêm, por vezes, importantes simplificações. Por exemplo, o cálculo de uma função experimental que quantifique a estrutura espacial de uma imagem é uma tarefa muito facilitada, quando comparada com a correspondente a um conjunto de dados em malha irregular, pois o arranjo prévio em malha regular define à priori as direcções e os passos

possíveis, eliminando-se a questão, sempre trabalhosa, da escolha das tolerâncias lineares e angulares mais adequadas.

Em termos geoestatísticos, a estrutura espacial é traduzida por uma única função global da distância vectorial entre os pares de pontos amostrais, que é designada por *variograma*. Quando o seu cálculo diz respeito aos valores de uma imagem, esta função é definida como a variância dos acréscimos da variável representada (por exemplo, uma escala de cores) para diferentes passos \vec{h} , onde \vec{h} é o vector que liga entre si todos os pares de pontos. O variograma quantifica o modo como a variabilidade da variável depende da direcção (o que se designa por anisotropia) e do módulo do vector \vec{h} (sendo de esperar que a variabilidade dos acréscimos cresça com a distância entre eles, pelo menos até um certo limite, designado por amplitude). Dado que o variograma é uma função global, estimado como a esperança matemática dos quadrados das diferenças entre os valores da variável quando o vector \vec{h} percorre a totalidade do campo onde a imagem se desenvolve, é possível, através deste estatístico experimental, sintetizar a variabilidade do arranjo espacial dos *pixels* da imagem, e indirectamente, a dimensão média dos objectos nela representados.

A partir do modelo que descreve este estatístico experimental, é possível estimar o valor da variável em localizações desconhecidas, seleccionando e pesando os valores dos pontos que se encontram na vizinhança dessa localização, em função da estrutura global da imagem e da sua correlação espacial com a localização do domínio a estimar. Este procedimento é designado por *krigagem* (em linguagem geoestatística) e corresponde ao conceito de *Best Linear Unbiased Estimator* (BLUE).

Um ponto crucial deste procedimento de estimação é o ajustamento do modelo teórico ao variograma experimental. Este ajustamento, na tradição da geoestatística, é feito manualmente com intervenção pericial, aproveitando o conhecimento que existe sobre o fenómeno subjacente à variável a estimar. Principalmente nos casos em que essa variável é obtida em imagens, em que há um acesso “completo” à sua realização até à escala representada da imagem, o ajustamento manual, sempre moroso porque envolve um longo processo de tentativa e erro (*trial and error*), pode ser substituído, com grande vantagem, por um processo automático.

Num processo puramente automático (por exemplo baseado em otimização de uma qualquer característica), o modelo escolhido deve passar por um procedimento de *validação cruzada*, antes de ser utilizado posteriormente nos processos de estimação ou simulação. O teste mais utilizado para a validação cruzada consiste em retirar um a um os pontos experimentais, estimá-los por *krigagem* através dos vários modelos candidatos que representam matematicamente o variograma e decidir posteriormente qual o “melhor” modelo segundo critérios de minimização calculados a partir dos desvios dos erros (simples e/ou quadráticos) entre os valores reais e os correspondentes valores estimados. Através deste procedimento, é possível quantificar globalmente a qualidade do modelo, abrindo assim uma via para a validação de um processo de automatização.

Todas as especificidades referidas como particularidades e problemas na integração de imagens nos casos de estudo em Ciências da Terra, com solução por aplicação de metodologias Geoestatísticas, fomentaram a motivação que esteve na génese do trabalho agora concretizado.

Em resumo, pretende-se, com esta dissertação, contribuir para a selecção e adaptação dos métodos geoestatísticos mais apropriados para o tratamento de **imagens digitais**.

A investigação subjacente ao trabalho aqui apresentado foi complementada com a concepção de um sistema para ambiente Windows, a sua programação em Visual C++, e o desenho de uma articulação coerente entre a entrada dos dados e a saída dos resultados sob a forma de imagens e os procedimentos geoestatísticos de cálculo. O resultado é um sistema informático, com funcionamento semi-automático, que foi designado por *MODIMAV – sistema de modelação geoestatística para imagens com ajustamento semi-automático dos variogramas*, contemplando os algoritmos geoestatísticos de modelação de variogramas, estimação e simulação com maior potencial no Tratamento de Imagem.

Uma das soluções estudadas e implementadas neste *software* merecedora de particular destaque é o processo de modelação automática do variograma, substituindo o tradicional “ajustamento manual”, que mesmo assim é mantido no sistema aqui apresentado como procedimento alternativo a recorrer quando exista informação pericial relevante (em especial, quando esta conduz a resultados contraditórios com o ajustamento automático). A solução proposta é baseada em *algoritmos genéticos* a qual - para além do seu rigoroso conteúdo teórico e elevada performance informática – se revelou de grande utilidade prática, por evitar

todos os procedimentos em que o utilizador era chamado a escolher o tipo de modelo e a fornecer os respectivos parâmetros.

Com o sistema aqui desenvolvido, pode trabalhar-se uma imagem através de ferramentas geoestatísticas com intervenção mínima do utilizador, que apenas dialoga com o sistema através de janelas auto-explicativas. Refira-se também que, dada a atrás referida especificidade geométrica dos dados de partida, o sistema está direccionado para a modelação automática ou semi-automática do variograma que traduz, numa função única (sem exigir intervenção do utilizador para efectuar parametrizações), os aspectos mais importantes da estrutura espacial global representada pela imagem. Desta forma, a detecção de relações entre os elementos da imagem que intervêm na criação dos *outputs* finais, pode ser feita sem qualquer intervenção do utilizador.

1.2. Especificidade dos casos em estudo

Para exemplificar a escolha e a implementação dos algoritmos geoestatísticos incorporados no programa desenvolvido (*MODIMAV*), foram seleccionados três exemplos funcionais de procedimentos típicos baseados em imagens, que se passam a apresentar, salientando as especificidades relativas ao suporte dos dados.

O primeiro exemplo refere-se à ampliação por mudança de escala de uma imagem, com o objectivo de realçar determinados aspectos que não eram aparentes na situação de partida. Relativamente aos procedimentos tradicionais de ampliação, neste sistema informático, a actualização dos valores após ampliação é feita utilizando um estimador geoestatístico baseado em *Krigagem Normal (Ordinary Kriging)*. Neste exemplo funcional pretende-se apenas estimar valores de uma variável numa malha regular mais ou menos densa do que a da imagem de partida, correspondente à mudança das dimensões de apresentação da imagem, sem que haja qualquer intervenção de informação secundária.

O segundo exemplo diz respeito à caracterização global e local, por estimação e por simulação, dos valores da precipitação para toda a região do Algarve, a partir de medições pontuais de estações de monitorização, utilizando o Modelo Digital do Terreno (MDT) como uma imagem secundária correlacionada localmente com a pluviometria. Neste segundo

exemplo funcional, a imagem tem o estatuto de informação *secundária* e é utilizada para complementar o estudo de um atributo *principal*, que é a precipitação medida pontualmente.

O terceiro exemplo consiste na simulação de padrões de tonalidade e heterogeneidade da cor em chapas de rochas ornamentais, condicionada a parâmetros estatísticos (histograma e variograma dos canais de cor vermelho, verde e azul) extraídos a partir de imagens de referência (*training images*) obtidas ‘por catálogo’. Esta funcionalidade e correspondente rotina informática têm potencialidades para serem aplicadas em programas de arquitectura em que o efeito da escolha de determinado padrão da rocha ornamental pode ser antecipada, pela referida simulação, à escala do edificio a projectar. Neste exemplo funcional são capturados parâmetros estatísticos descritivos das características da imagem e são reproduzidas novas imagens, com extensão diferente relativamente à imagem de partida, isto é, toda a informação tem o estatuto de principal.

Finalmente, é importante realçar que estes três exemplos, seleccionados para ilustrar as metodologias consideradas nesta dissertação, têm como denominador comum a intervenção de imagens sob diferentes perspectivas, com intervenção *principal* ou *secundária*. Foi a partir destes pressupostos que o sistema informático *MODIMAV* foi desenvolvido, pretendendo-se, por um lado, avaliar as necessidades de ajustamento dos métodos geoestatísticos a esta situação e, por outro, demonstrar as potencialidades de tais métodos na Análise de Imagem e a sua integração nas Ciências da Terra.

1.3 Organização da dissertação

O corpo principal da dissertação é composto por seis capítulos.

Após o enquadramento do problema e a caracterização dos casos em estudo já efectuados neste capítulo, introduzem-se, no 2º capítulo, os fundamentos teóricos relativos aos conceitos geoestatísticos adaptados à imagem digital e necessários para a compreensão do sistema construído.

No 3º capítulo, após a introdução de conceitos genéricos sobre optimização e computação evolutiva, descreve-se detalhadamente o algoritmo genético utilizado bem como alguns pormenores da sua implementação e testes de desempenho.

No 4º capítulo, faz-se a descrição detalhada do sistema desenvolvido (*MODIMAV*) e apresenta-se o manual do utilizador.

No 5º capítulo aplica-se o sistema a três casos de estudo típicos, para ilustrar as suas potencialidades.

Finalmente no 6º e último capítulo retiram-se algumas conclusões e propõem-se linhas de investigação futura para o aperfeiçoamento do sistema proposto.

2. Metodologia

Neste capítulo apresentam-se os conceitos teóricos relativos às metodologias geoestatísticas que foram consideradas neste trabalho (estimação e simulação) e, também, alguns conceitos relativos à caracterização e tratamento de imagens digitais.

No que concerne às metodologias geoestatísticas apresenta-se o conceito de Variável Regionalizada como a realização de uma Função Aleatória e apresentam-se as hipóteses restritivas na inferência estatística. Introduce-se a função covariância espacial e o variograma como ferramentas reveladoras da continuidade espacial, referem-se os conceitos de variogramas experimentais univariados indicando as fórmulas utilizadas para o respectivo cálculo e referem-se as funções teóricas mais utilizadas nas Ciências da Terra para o ajustamento dos variogramas experimentais. Finalmente apresenta-se o formalismo teórico das metodologias de estimação por *krigagem* e simulação geoestatística por Simulação Sequencial Gaussiana.

Relativamente a conceitos sobre as imagens, caracterizam-se as imagens digitais, descrevem-se os modelos de codificação de cor, enumeram-se alguns dos formatos mais utilizados no seu armazenamento, introduz-se o conceito de georeferenciação e, finalmente, introduz-se a questão do reescalonamento, descrevendo as técnicas mais comuns para a sua execução.

2.1. Estratégia da metodologia implementada

A metodologia adoptada e implementada no sistema informático *MODIMAV*, desenvolvido no âmbito deste trabalho, assenta em duas grandes etapas: a primeira, corresponde à análise estrutural semi-automática, que tem como resultado o modelo teórico que melhor se ajusta ao variograma experimental segundo as várias direcções; a segunda, diz respeito à estimação por *krigagem* e à simulação geoestatística, donde resultam mapas de valores da variável numa malha não amostrada, obtidos a partir da estrutura espacial caracterizada pelo modelo teórico.

O sistema *MODIMAV*, como novo programa informático aqui desenvolvido, pode ser visto como uma extensão do que é descrito em Pannatier [PA 96], ajustado à especificidade das imagens, em particular, às adaptações necessárias para as entradas correspondentes a ficheiros de *pixels*.

Neste sistema, a modelação teórica dos variogramas experimentais é efectuada semi-automaticamente, com a possibilidade de proceder a ajustes manuais para responder a pressupostos periciais, o que se revela como um importante avanço relativamente ao sistema descrito em Pannatier [PA 96]. A modelação automática recorre a um *algoritmo genético*; o desempenho do ajustamento é avaliado pela função inversa da diferença de quadrados entre os valores observados e os valores estimados por *krigagem* normal num teste de validação cruzada.

Na etapa subsequente, e consoante o objectivo do estudo, procede-se à estimação por *krigagem* (nas variantes *krigagem normal* e/ou *cokrigagem colocalizada*) ou à simulação de imagens. Estes algoritmos de estimação e de simulação foram implementados no sistema, por adaptação das respectivas fontes FORTRAN do sistema GSLIB [DJ 98].

2.2. Revisão de metodologias geoestatísticas

O conceito de Geoestatística (estudo estatístico de um fenómeno natural) foi utilizado pela primeira vez em 1962 por G. Matheron na École des Mines de Paris para resolver problemas de avaliação de reservas mineiras. Nascia assim um novo ramo da ciência que viria a sofrer uma rápida evolução e disseminação em diversas outras áreas que exigem o tratamento de dados cujo posicionamento espacial é determinante para a sua interpretação, como por exemplo, Cartografia Temática, Caracterização de Georecursos tipicamente não mineiros (petrolíferos, hídricos, florestais, cinegéticos e piscícolas) e caracterização da Qualidade do Ambiente (contaminação do solo, água e ar).

Para descrever e interpretar os fenómenos espaciais que se desenvolvem nesses domínios é necessário considerar variáveis $z(x)$, dependentes de uma localização x com uma, duas ou três dimensões.

Segundo Journal e Huijbregts [JH 78], Pereira [PE 79] e Sousa e Muge [SM 90], estas variáveis $z(x)$, são designadas por Variáveis Regionalizadas e apresentam simultaneamente:

- Um carácter *aleatório*, manifestado pela irregularidade e variação imprevisível de ponto para ponto;
- Um carácter *estruturado*, manifestado pelas relações existentes entre os seus valores nos diferentes pontos do espaço onde o fenómeno ocorreu (é de esperar que, em pontos vizinhos, os valores das variáveis se relacionem mais do que em pontos mais afastados).

Segundo Pereira [PE 79], destacam-se quatro características qualitativas destas variáveis:

- I) Localização – a variável está definida numa determinada zona do espaço que se designa por campo geométrico. Cada variável só faz sentido se for previamente delimitada à região onde ela é definida e amostrada.
- II) Continuidade – variação espacial da variável que, apesar da complexidade das flutuações verificadas, manifesta em média continuidade, ou seja:

$$\lim_{x \rightarrow x_0} E\{z(x) - z(x_0)\} = 0$$

Na prática, existem situações em que a continuidade não é empiricamente aparente para grandes passos da amostragem. O limite para o qual tende a diferença de quadrados da variável, quando os pontos se aproximam teoricamente até à distância zero, $\lim_{x \rightarrow x_0} E\{(z(x) - z(x_0))^2\} = C_0$, é designado tradicionalmente por efeito de pepita (C_0).

- III) Anisotropia – ocorre quando a estrutura espacial numa determinada direcção é diferente da que se verifica noutra direcção qualquer. É quantificada pela relação entre as características da variável para diferentes direcções, em geral, perpendiculares.
- IV) Fenómenos de transição – no interior do campo geométrico da variável podem surgir estruturas particulares onde exista continuidade no interior a cada unidade e variações bruscas entre unidades.

Em resumo, todas as variáveis dependentes do espaço onde exista um sistema de relações entre os valores tomados em diferentes pontos (para além de um certo carácter aleatório), podem ser tratadas como Variáveis Regionalizadas e analisadas segundo o formalismo desenvolvido pela Geoestatística.

Nesse formalismo, para interpretar a Variável Regionalizada $z(x)$, utiliza-se a noção de Função Aleatória (variável aleatória com uma infinidade de componentes), denotada por $Z(x)$.

Sendo x_1, \dots, x_n as coordenadas dos pontos do espaço onde existem medidas da Variável Regionalizada $z(x)$, o conjunto dos valores da variável nesses pontos $z(x_1), \dots, z(x_n)$ é tomado como a realização de uma Função Aleatória a n dimensões.

A lei de densidade de probabilidade a n dimensões $f(z(x_1), \dots, z(x_n))$ só poderia ser obtida empiricamente se houvesse um número significativo de realizações da função aleatória, ou seja, se fosse possível repetir as experiências que conduzem, cada uma delas, a uma variável regionalizada. Ora, é absolutamente impossível obter tal densidade de probabilidade a partir de uma única experiência (a amostragem que dá origem ao conjunto de dados experimentais disponíveis).

A impossibilidade de induzir a lei de distribuição da Função Aleatória obriga à formulação de algumas hipóteses restritivas para que a inferência estatística seja possível [MA 78].

Segundo Journel e Huijbregts [JH 78], Pereira [PE 79], Soares [SO 83] e Sousa e Muge [SM 90] consideram-se as seguintes hipóteses restritivas:

- Estacionaridade de 1ª ordem – a função aleatória é estacionária se a sua lei de distribuição de probabilidade for invariante por translação, ou seja, $[Z(x_1), \dots, Z(x_n)]$ tem a mesma lei de distribuição de probabilidade que $[Z(x_1 + \vec{h}), \dots, Z(x_n + \vec{h})]$, onde \vec{h} é um vector qualquer que desloca a variável no interior do campo.

Se a função densidade de probabilidade é invariante por translação, também o será o primeiro momento da função aleatória.

Assim verifica-se que $E\{Z(x)\} = m(x) = E\{Z(x + \vec{h})\} = m = \text{constante}$, o que implica que a média é invariante por translação.

- Estacionaridade de 2ª ordem – existe estacionaridade de 2ª ordem se a covariância $C(\vec{h}) = E\left\{Z(x + \vec{h}) \cdot Z(x) - m(x)^2\right\}$ e a média $m(x)$ não dependerem de x , ou seja do ponto de apoio, mas apenas do vector que liga os pontos utilizados no seu cálculo.
- Hipótese intrínseca – considerando os incrementos $Z(x + \vec{h}) - Z(x)$ e admitindo que não dependem de x , obtêm-se:

$$E\left\{Z(x + \vec{h}) - Z(x)\right\} = m(\vec{h}), \text{ deriva linear que se supõe nula, e}$$

$$E\left\{\left[Z(x + \vec{h}) - Z(x)\right]^2\right\} = 2\gamma(\vec{h}), \text{ variograma experimental.}$$

Verificando-se a hipótese intrínseca, o variograma $\gamma(\vec{h})$, sendo independente do ponto de apoio, é suficiente para definir a estrutura da variável regionalizada no espaço.

A função variograma é o instrumento básico da geoestatística onde assenta toda a teoria que permite a estimação e simulação de variáveis regionalizadas a partir da sua estrutura espacial. Com a modelação dos variogramas experimentais, obtidos para as direcções consideradas preferenciais, obter-se-á uma função representativa das características espaciais da variável regionalizada em toda a área considerada.

Os modelos geoestatísticos utilizados na estimação ou na simulação de fenómenos relacionados dependentes de variáveis quantitativas estruturadas espacialmente resultam essencialmente da teoria das funções aleatórias auto-correlacionadas.

A escolha de um determinado modelo geoestatístico deve ter por base o estudo e o conhecimento do fenómeno em causa, incorporando-se a sua componente espacial na respectiva caracterização, que pode servir para aplicações distintas nas áreas do planeamento, do ordenamento, da valorização e monitorização do recurso natural. A validação destes modelos geoestatísticos é efectuada *à posteriori*, mediante o cruzamento entre os resultados obtidos pela sua modelização e a realidade, expressa quantitativa ou qualitativamente (pela plausibilidade científica dos *outputs* obtidos).

2.2.1. Medidas de continuidade espacial e variabilidade

Nesta fase abordam-se resumidamente os instrumentos geoestatísticos de medida da continuidade espacial das características da variável estudada, quer seja amostrada por uma imagem ou por um outro plano de amostragem, em geral, irregular. Estes instrumentos cumprem fundamentalmente dois objectivos: por um lado, caracterizam e quantificam o modo como uma dada variável se dispersa no espaço, fornecendo o seu grau de continuidade espacial e o seu modo de dispersão (ou seja, a sua análise estrutural), e por outro, servem de base aos processos de inferência por estimação e por simulação geoestatística.

Foram considerados neste trabalho apenas medidas de continuidade simples, ou seja, aquelas que se referem a uma única variável, excluindo assim as medidas cruzadas entre pares de variáveis ([IS 89], [PA 96]).

De acordo com a formalização apresentada em Pannatier [PA 96], foram consideradas as seguintes medidas de continuidade espacial e variabilidade: variograma experimental simples, variograma normalizado simples, covariância simples, correlação simples e madograma simples.

2.2.1.1. Variograma experimental

O variograma experimental ou semivariograma para o vector \vec{h} , $\gamma(\vec{h})$, onde $|\vec{h}|$ é a distância euclidiana entre os pontos x_i e $x_i + \vec{h}$ é calculado pela seguinte expressão, onde $N(\vec{h})$ é o número de pares de pontos segundo a direcção de \vec{h} e distanciados de $|\vec{h}|$:

$$\gamma(\vec{h}) = \frac{1}{2N(\vec{h})} \sum_{i=1}^{N(\vec{h})} \left[z(x_i + \vec{h}) - z(x_i) \right]^2 \quad [1]$$

Pode ainda definir-se o variograma normalizado para o vector \vec{h} , $\gamma_n(\vec{h})$, que é calculado pela seguinte fórmula:

$\gamma_n(\vec{h}) = \frac{\gamma(\vec{h})}{\sigma_{-\vec{h}} \cdot \sigma_{+\vec{h}}}$ onde $\gamma(\vec{h})$ é o variograma calculado segundo [1], e

$$\sigma_{-\vec{h}}^2 = \frac{1}{N(\vec{h})} \sum_{i=1}^{N(\vec{h})} z^2(x_i) - m_{-\vec{h}}^2 \quad \text{e} \quad \sigma_{+\vec{h}}^2 = \frac{1}{N(\vec{h})} \sum_{i=1}^{N(\vec{h})} z^2(x_i + \vec{h}) - m_{+\vec{h}}^2 \quad \text{são as}$$

variâncias esquerda e direita para o vector \vec{h} , respectivamente, com:

$$m_{-\vec{h}} = \frac{1}{N(\vec{h})} \sum_{i=1}^{N(\vec{h})} z(x_i) \quad \text{e} \quad m_{+\vec{h}} = \frac{1}{N(\vec{h})} \sum_{i=1}^{N(\vec{h})} z(x_i + \vec{h}) \quad \text{as médias esquerda e}$$

direita, para o vector \vec{h} , respectivamente.

O variograma normalizado, quando representado graficamente, apresenta a mesma forma do variograma não normalizado a menos de um factor de escala – o patamar é a unidade.

Uma outra medida de continuidade é o chamado madograma, que para o vector \vec{h} , é calculado pela seguinte fórmula:

$$M(\vec{h}) = \frac{1}{2N(\vec{h})} \sum_{i=1}^{N(\vec{h})} |z(x_i) - z(x_i + \vec{h})|$$

Trata-se de uma função equivalente ao semi-variograma onde o quadrado das diferenças é substituído pelo respectivo módulo. É pouco utilizado pela inconsequência teórica em termos de utilização nos estimadores, podendo ser utilizado apenas para comparação de resultados.

2.2.1.2. Covariância espacial

A covariância espacial para o vector \vec{h} , onde $|\vec{h}|$ é a distância euclidiana entre os pontos x_i e $x_i + \vec{h}$, é calculada experimentalmente pela seguinte fórmula:

$$C(\vec{h}) = \frac{1}{N(\vec{h})} \sum_{i=1}^{N(\vec{h})} z(x_i)z(x_i + \vec{h}) - m_{-\vec{h}} m_{+\vec{h}} \quad [2]$$

sendo $m_{-\vec{h}} = \frac{1}{N(\vec{h})} \sum_{i=1}^{N(\vec{h})} z(x_i)$ e $m_{+\vec{h}} = \frac{1}{N(\vec{h})} \sum_{i=1}^{N(\vec{h})} z(x_i + \vec{h})$ as médias esquerda e direita, para o vector \vec{h} , respectivamente.

A covariância espacial $C(\vec{h})$ relaciona-se com $\gamma(\vec{h})$ da seguinte forma [IS 89]:

$$\gamma(\vec{h}) = C(0) - C(\vec{h}) \text{ onde } C(0) = C(|\vec{h}| = 0) \text{ é a variância da população.}$$

O correlograma, que é uma medida normalizada da correlação espacial, para o mesmo vector \vec{h} é calculado pela seguinte fórmula:

$$\rho(\vec{h}) = \frac{C(\vec{h})}{\sigma_{-\vec{h}} \sigma_{+\vec{h}}} \text{ onde } C(\vec{h}) \text{ é a covariância calculada segundo [2] e as variâncias,}$$

esquerda e direita, para \vec{h} calculadas de acordo com as fórmulas enunciadas na secção 2.2.1.1.

O correlograma relaciona-se com o variograma experimental da seguinte forma [IS 89]:

$$\gamma(\vec{h}) = \rho(0) - \rho(\vec{h}) \text{ onde } \rho(0) \text{ é o coeficiente de correlação.}$$

Também o correlograma relaciona-se [PA 96] com o variograma normalizado da seguinte forma:

$$\rho(\vec{h}) = 1 - \gamma_n(\vec{h})$$

2.2.2. Modelos teóricos para o ajustamento do modelo experimental

Os variogramas experimentais calculados para diferentes direcções do espaço, a partir da informação disponível (por exemplo, uma imagem ou um conjunto irregular de amostras), constituem uma representação da estrutura de continuidade da variável em estudo.

Após o cálculo do variograma experimental, há que modelá-lo através de uma função que deverá resumir toda a informação estrutural importante. O variograma experimental pode ser interpretado como um estimador do variograma teórico, calculado apenas para um número discreto de pontos em função da densidade e orientação dos pontos amostrados.

Tal como em Pannatier [PA 96], foram usadas funções definidas positivas para constituir os seguintes modelos do variograma: efeito de pepita, esférico, exponencial, gaussiano e potência. O modelo final pode ser constituído por uma função ou por uma combinação linear de funções, visto que a combinação linear de funções definidas positivas continua a ser uma função definida positiva. Estes modelos podem ser isotrópicos ou anisótropos, dependendo de como variam direcionalmente.

2.2.2.1. Efeito de pepita

O efeito de pepita (c_0) é um indicador da irregularidade do fenómeno em estudo. Pode ser interpretado como a variabilidade à pequena escala não abrangida pela malha de amostragem e também como a variabilidade à escala da amostra induzida por erros não sistemáticos de amostragem, que adicionam à estrutura do fenómeno um ruído aleatório. O efeito de pepita, sendo o valor do variograma para $|\vec{h}| = 0$, infere-se pela intersecção da recta que passa pelos primeiros pontos do variograma com o eixo das ordenadas. Na fig. 1 representa-se graficamente uma situação de efeito de pepita puro cuja expressão geral é a seguinte (graficamente c_0 é uma medida da descontinuidade na origem):

$$\gamma^*(|\vec{h}|) = \begin{cases} 0 & \text{se } |\vec{h}| = 0 \\ c_0 & \text{se } |\vec{h}| > 0 \end{cases}$$

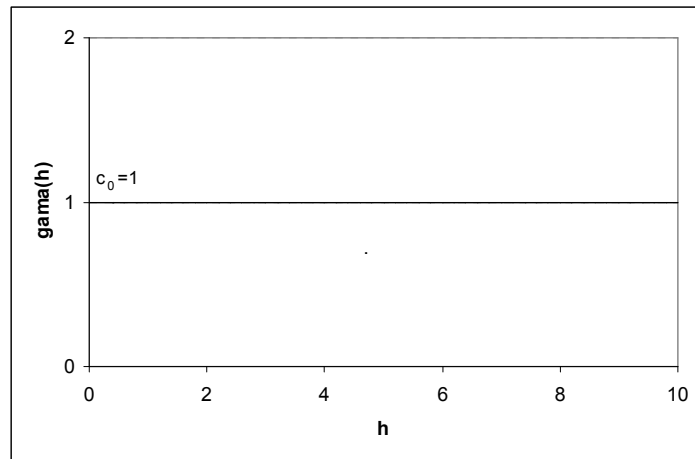


Fig. 1 Efeito de pepita puro

2.2.2.2. Modelo Esférico

O modelo esférico é um dos modelos mais utilizados na geoestatística. Apresenta patamar e o crescimento é linear junto à origem. Na fig. 2 representa-se graficamente a forma de um modelo esférico, cuja expressão geral é a seguinte (c é o patamar e a é a amplitude):

$$\gamma^*(|\vec{h}|) = \begin{cases} c \left[1.5 \frac{|\vec{h}|}{a} - 0.5 \left(\frac{|\vec{h}|}{a} \right)^3 \right] & |\vec{h}| \leq a \\ c & |\vec{h}| > a \end{cases}$$

A amplitude a designa a distância a partir da qual a auto-correlação se dissipa, quando o variograma atinge o patamar.

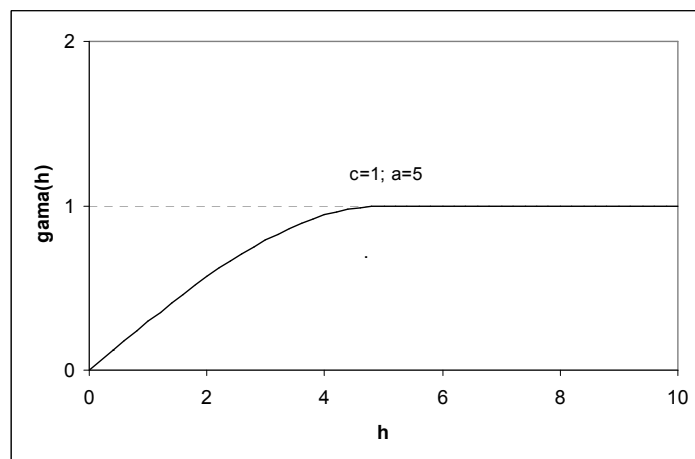


Fig. 2 Modelo esférico

2.2.2.3. Modelo exponencial

O modelo exponencial cresce mais rapidamente junto à origem e tende assintoticamente para o patamar, sendo a amplitude em função da sua expressão definida pela distância onde se atinge aproximadamente 95% do patamar. Na fig. 3 representa-se graficamente um modelo exponencial, cuja expressão geral é a seguinte (sendo c o patamar e a a amplitude):

$$\gamma^* (|\vec{h}|) = c \left[1 - e^{-\frac{3|\vec{h}|}{a}} \right]$$

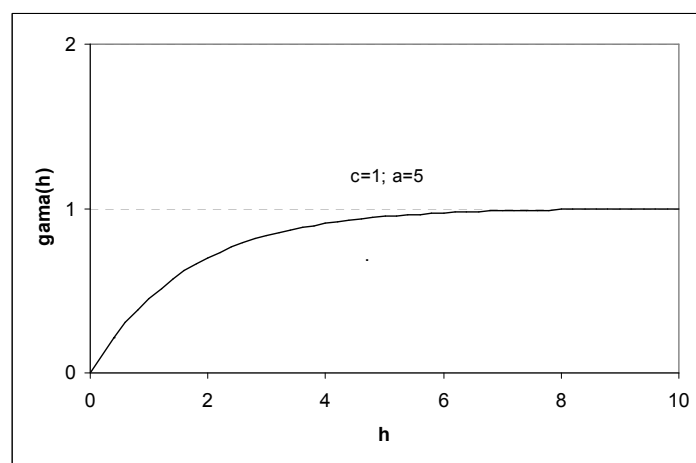


Fig. 3 Modelo exponencial

2.2.2.4. Modelo gaussiano

O modelo gaussiano é geralmente usado para o ajustamento do comportamento de variáveis extremamente contínuas. A forma da função reflecte este facto, pois exhibe comportamento parabólico junto da origem e apresenta um ponto de inflexão para uma distância ligeiramente superior. Tal como o modelo exponencial, tende assintoticamente para o patamar, sendo a amplitude definida pela distância onde se atinge também aproximadamente 95% do patamar. Na fig. 4 representa-se graficamente um modelo gaussiano cuja expressão geral é a seguinte (sendo c o patamar e a a amplitude):

$$\gamma^*(|\vec{h}|) = c \left[1 - e^{-3 \left(\frac{|\vec{h}|}{a} \right)^2} \right]$$

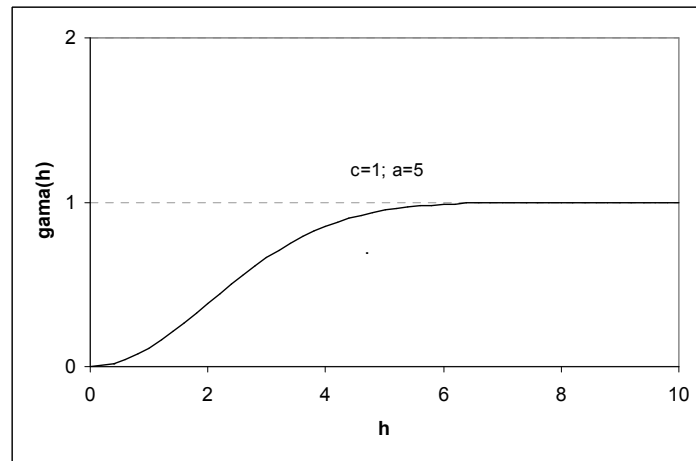


Fig. 4 Modelo gaussiano

2.2.2.5. Modelo de potência

O modelo de potência é utilizado em fenómenos não estacionários e descreve variáveis cuja continuidade é evidenciada a qualquer distância (não existe amplitude nem patamar). A forma da função é variável e depende do expoente considerado. Na fig. 5 representa-se graficamente um modelo exponencial cuja expressão geral é a seguinte,

$$\gamma^*(|\vec{h}|) = \begin{cases} c |\vec{h}|^a & 0 < a < 2 \\ \text{não definido} & a > 2 \end{cases}$$

Neste tipo de modelo c é um factor de escala e a um expoente que pode variar entre 0 e 2. No caso de $a=1$ (expoente) o modelo é uma recta com declive igual a c .

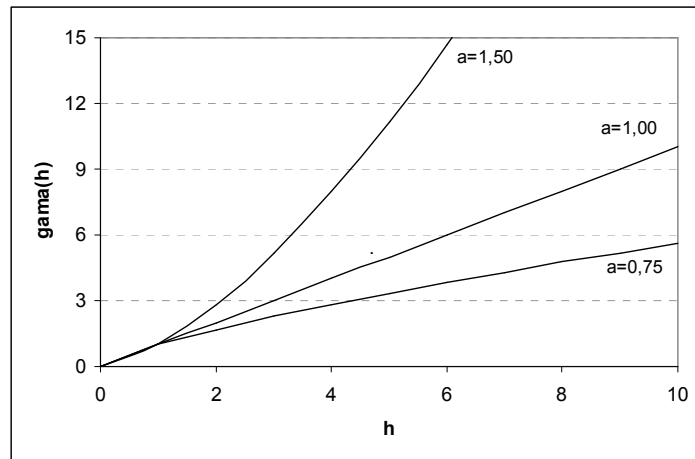


Fig. 5 Modelo de potência

A escolha de um destes modelos baseia-se no comportamento do variograma experimental nos valores próximos da origem, dado que é nessa região que o número de pares de pontos é mais significativo. Estes modelos caracterizam o padrão da continuidade da nossa variável.

Por vezes são utilizados uma combinação de modelos para explicar um único fenómeno, partindo do pressuposto que qualquer combinação linear de variogramas definidos positivos, com coeficientes positivos é ainda um modelo definido positivo.

2.2.3. Estimação por *krigagem* normal e *cokrigagem* colocalizada

A estimação, em geoestatística, é a avaliação de uma grandeza a partir de informação existente [SM 87], num local não amostrado. O resultado do processo de estimação depende de um processo que compreende as relações entre o variograma calculado, a localização geográfica das amostras e o ponto a estimar.

O estimador geoestatístico é um estimador dito linear, pois a estimação de uma grandeza num determinado local e suporte é feita com base na combinação linear de valores conhecidos dessa mesma grandeza devendo os ponderadores obedecer aos dois seguintes critérios [SM 87]:

- Condição de não enviesamento universal, ou seja, a esperança matemática dos desvios entre os valores reais e os estimados deve ser nula;

- Condição de variância mínima, ou seja, os desvios entre os valores reais e os estimados devem apresentar variância mínima.

Segundo [SO 83] o estimador geoestatístico (*krigagem*) aparece, pela primeira vez, empregue por George Matheron em 1965, em homenagem a Donald Krige, que tinha tido a intuição de que era necessário ponderar diferencialmente o valor das amostras para estimar as reservas de ouro das minas da África do Sul.

O estimador linear de *Krigagem* Normal (ou *ordinary kriging*), $[z(x_0)]_{KO}^*$, para um suporte x_0 obedecendo aos critérios anteriormente referidos, traduz-se na seguinte expressão [LA 03]:

$$[z(x_0)]_{KO}^* = \sum_{\alpha=1}^n \lambda_{\alpha}^{KO} Z(x_{\alpha}), \text{ sendo,}$$

$[z(x_0)]_{KO}^*$ – valor estimado pela *krigagem* normal para a variável regionalizada $Z(x)$ na localização e suporte x_0 ;

$z(x_{\alpha})$ – valores da variável nos pontos conhecidos x_{α} que intervêm na estimação de x_0 ;

λ_{α}^{KO} – ponderadores que afectam cada valor $z(x_{\alpha})$;

n – número de pontos x_{α} que intervêm na estimação.

Os dois critérios anteriormente enunciados conduzem às seguintes condições, respectivamente:

$$1) \sum_{\alpha=1}^n \lambda_{\alpha}^{KO} = 1$$

2) Se o erro de estimação for expresso por $\varepsilon = [z(x_0)]_{KO}^* - z(x_0)$, então por aplicação do formalismo de Lagrange obtém-se:

$$Var\{\varepsilon\} = Var\{z(x_0)\} + 2\mu \left(\sum_{\alpha=1}^n \lambda_{\alpha}^{KO} - 1 \right)$$

Após desenvolvimento, chega-se à seguinte condição em função da covariância espacial:

$$Var\{\varepsilon\} = C(0) + \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha}^{KO} \lambda_{\beta}^{KO} C(x_{\alpha}, x_{\beta}) - 2 \sum_{\alpha=1}^n \lambda_{\alpha}^{KO} C(x_{\alpha}, x_0) + 2\mu \left(\sum_{\alpha=1}^n \lambda_{\alpha}^{KO} - 1 \right)$$

A minimização da variância do erro efectua-se igualando a zero as $(n+1)$ derivadas parciais em ordem aos ponderadores λ_{α} e ao parâmetro de Lagrange μ . Assim sendo, o cálculo dos estimadores resulta da resolução do seguinte sistema de $(n+1)$ equações, conhecido por Sistema de *Krigagem*:

$$\begin{cases} \sum_{\alpha=1}^n \lambda_{\alpha}^{KO} C(x_{\alpha}, x_{\beta}) + \mu = C(x_{\beta}, x_0), & \beta = 1, \dots, n \\ \dots \\ \sum_{\alpha=1}^n \lambda_{\alpha}^{KO} = 1 \end{cases}$$

Neste sistema $C(x_{\alpha}, x_{\beta})$ representam as covariâncias espaciais entre as amostras utilizadas na estimação e $C(x_{\beta}, x_0)$ a covariância espacial entre as amostras e a localização a estimar.

De acordo com [SM 87] a aplicação generalizada deste tipo de estimadores deve-se, principalmente, a utilizarem apenas o modelo da função variograma para o seu cálculo e à não dependência da variância de estimação do valor da variável regionalizada, podendo por esse facto, ser calculado se forem conhecidos os valores dos estimadores.

Se, para além de uma variável principal, existir uma variável secundária com uma densidade de amostragem superior, esta pode ser incorporada no modelo de estimação desde que exista correlação entre as duas variáveis, designando-se este procedimento por *cokrigagem* (extensão da *krigagem* a duas variáveis). Este estimador segue os mesmos critérios que a *krigagem* e portanto o seu cálculo resumir-se-á, novamente, a um problema de minimização de variância de estimação.

No caso particular de sobre-abundância da variável secundária em relação ao número de amostras da variável principal (por exemplo, nos casos em que a variável secundária é conhecida em todas as localizações) uma solução possível consiste em reter somente o valor da variável secundária que está localizada no ponto a estimar, designando-se este tipo de estimação por *cokrigagem colocalizada*.

Assim, e ainda segundo [SO 83], se considerarmos $Z_1(x)$ a variável principal cujos valores são conhecidos em n_1 amostras e $Z_2(x)$ a variável secundária disponível em todos os nós da malha a estimar, o estimador de $Z_1(x)$ e $x_0 [Z_1(x_0)]^*$, através da *cokrigagem colocalizada* é definido por:

$$[Z_1(x_0)]_{CK}^* = \sum_{i=1}^{n_1} a_i Z_1(x_i) + b_0 Z_2(x_0)$$

Seguindo a mesma dedução apresentada anteriormente para a *krigagem* normal, o sistema de *cokrigagem* colocalizada é expresso por um sistema com $(n_1 + 2)$ equações:

$$\begin{cases} \sum_{i=1}^{n_1} a_i C_{Z_1}(x_i, x_j) + b_0 C_{Z_1 Z_2}(x_i, x_j) + \mu = C_{Z_1}(x_j, x_0) & j = 1, \dots, n_1 \\ \dots \\ \sum_{i=1}^{n_1} a_i C_{Z_1 Z_2}(x_j, x_i) + b_0 + \mu = C_{Z_1 Z_2}(0) \\ \sum_{i=1}^{n_1} a_i + b_0 = 1 \end{cases}$$

Neste sistema $C_{Z_1 Z_2}(x_i, x_j)$ são covariâncias espaciais cruzadas entre $Z_1(x_i)$ e $Z_2(x_j)$ e $C_{Z_1 Z_2}(0)$ é a covariância espacial cruzada entre $Z_1(x)$ e $Z_2(x)$ para distância zero, que é o coeficiente de correlação entre as duas variáveis.

Assumindo a aproximação de Markov, a função covariância espacial cruzada $C_{Z_1 Z_2}(x_i, x_j)$ pode ser inferida através da seguinte relação em termos das covariâncias simples ou correlogramas [AJ 94]:

$$\begin{aligned} C_{Z_1 Z_2}(\vec{h}) &\approx \frac{C_{Z_1 Z_2}(0)}{C_{Z_1 Z_1}(0)} C_{Z_1 Z_1}(\vec{h}) \\ \rho_{Z_1 Z_2}(\vec{h}) &\approx \rho_{Z_1 Z_2}(0) \rho_{Z_1 Z_1}(\vec{h}) \end{aligned}$$

Esta aproximação permite que a inferência da variável primária seja feita à custa da covariância espacial da variável primária e do coeficiente de correlação entre a variável secundária e a variável primária (ρ_{z_1} e ρ_{z_2}):

$$C_{z_1 z_2}(\vec{h}) = \rho_{z_1 z_2} \cdot C_{z_1 z_1}(\vec{h})$$

2.2.4. Simulação geoestatística

A simulação geoestatística é um processo estocástico espacial destinado a gerar imagens que reproduzem a distribuição espacial e a incerteza associada às diferentes variáveis tratadas nas Ciências da Terra. Esta classe de modelos tem como objectivo reproduzir imagens da realidade que reflectem as seguintes características do conjunto amostral: variância, lei de distribuição de probabilidade e auto-correlação espacial. O resultado é um conjunto de imagens equiprováveis com a mesma distribuição espacial dos dados experimentais que garante a reprodução dos estatísticos e a variabilidade espacial das amostras (histograma e variograma ou covariância espacial). Normalmente, com a simulação não se pretende obter a imagem média ou a mais provável das características de um dado fenómeno regionalizado (que é o objectivo da estimação), mas sim um conjunto de imagens equiprováveis, que permitem visualizar diferentes comportamentos extremos das características internas ou morfológicas desse fenómeno, e simultaneamente quantificar a incerteza local dessas características.

Um modelo de simulação pretende reproduzir, na imagem simulada, a variabilidade do fenómeno em estudo, tendo por base os seguintes estatísticos: função distribuição de $Z(x) - F_Z(z) = \text{prob}\{Z(x) < z\}$ e variograma $\gamma(\vec{h})$ que reproduz a continuidade espacial [SO 00].

Deste modo, se designarmos por $z_s(x)$ o conjunto de valores simulados e por $z(x_\alpha)$, $x_\alpha = 1, \dots, n$, os n valores experimentais, a imagem simulada deve cumprir as seguintes condições teóricas:

1. Para qualquer valor de z : $\text{prob}\{Z(x_\alpha) < z\} = \text{prob}\{Z_c(x) < z\}$;

2. $\vec{\gamma}(h) = \vec{\gamma}_s(h)$, sendo $\vec{\gamma}(h)$ os variogramas dos valores experimentais e $\vec{\gamma}_s(h)$ os variogramas dos valores simulados;
3. Se a simulação for condicional, coincidência com os valores experimentais, $Z(x_\alpha) = Z_s(x_\alpha)$, ou seja, em qualquer ponto experimental x_α , o valor $Z(x_\alpha)$ e o valor simulado $Z_s(x_\alpha)$ são iguais. Isto significa que a imagem simulada passa pelos pontos experimentais e apresenta a mesma variabilidade.

Existem vários modelos geoestatísticos de simulação, sendo apenas abordados neste trabalho o da *Simulação Sequencial Gaussiana* e *Cosimulação Sequencial Gaussiana*.

O método de Simulação Sequencial Gaussiana permite a simulação de valores de uma variável uma malha regular de N pontos e sintetiza-se nos seguintes passos:

1. Escolha de uma sequência aleatória de pontos x_u numa malha regular para simulação;
2. Para a primeira localização x_u , simulação de um valor $Z(x_u)$ a partir da função de distribuição cumulativa (conforme se ilustra na fig. 6) $F(Z(x_u)|(n))$ que é condicionada a n valores observados. Após simulado, esse valor é utilizado como um novo valor a juntar aos valores conhecidos, passando a dispor-se de $n+1$ valores.

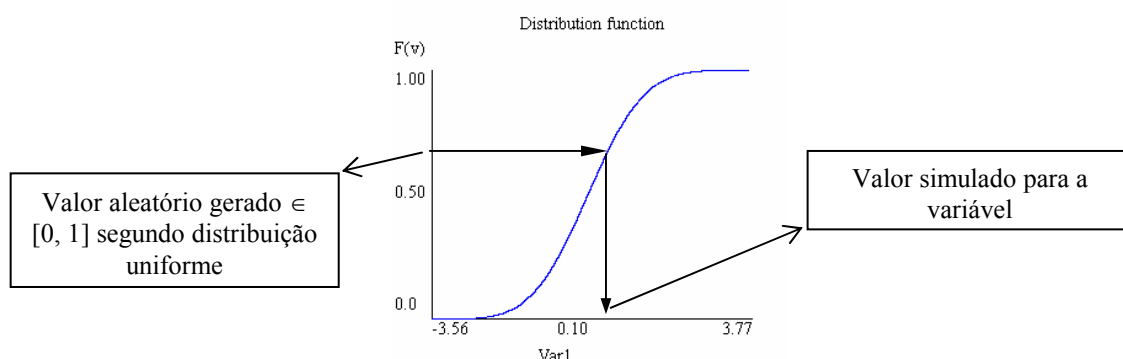


Fig. 6 – Simulação de um valor Z_1

3. Para outra localização x_f , simulação de um novo valor $Z(x_f)$ a partir da função de distribuição cumulativa $F(Z(x_f) | (n+1))$.

4. Repetir o passo 3 para todas as N localizações.

O passo crucial deste método consiste em estimar localmente as funções de distribuição cumulativas, $F(Z(x_0) | (n + i))$.

As funções de distribuição cumulativas são estimadas localmente do seguinte modo:

1. Transformação dos dados de partida para um histograma com lei Gaussiana;
2. Definir aleatoriamente um conjunto de valores amostrados de uma malha de pontos $x_u, u=1, N$, para simular.
3. Estimar a média e variância local de $z(x_u)$, identicamente, respectivamente, com o estimador de simples *krigagem* $z(x_u)^*$ e estimar a variância $\sigma_{sk}^2(x_u)$ condicionada aos dados experimentais $z(x_i)$ e aos valores simulados previamente $z^s(x_i)$.
4. Definir o intervalo de $Fz(z)$ para serem amostrados, usando a função cumulativa de distribuição Gaussiana e simular o valor $z^s(x_u)$ da função cumulativa de distribuição $Fz(z)$, pelo procedimento de Monte Carlo, e considerando um valor p da distribuição uniforme $U(0,1)$.
5. Quando todos os nós da malha definida estiveram simulados, transformar os valores simulados para a lei de distribuição original.

Uma das possibilidades de extensão deste algoritmo é a simulação conjunta das variáveis originais. Assim, pode-se simular simultaneamente a K variáveis, em que a simulação é condicionada à simulação de uma variável pré-simulada ([GHJ 93], [GO 97]).

Supondo que temos duas variáveis, $Z_1(x)$ e $Z_2(x)$, se escolhermos a primeira variável $Z_1(x)$, supondo que é a mais importante ou a que apresenta uma maior evidência de continuidade espacial, o algoritmo da Cosimulação Sequencial Gaussiana considera a estimação local da média e da variância com um procedimento de *cokrigagem* colocalizada.

2.3. Conceitos relativos a imagens digitais

Dado que as aplicações do sistema proposto incidem sobre imagens digitais, torna-se necessário discutir alguns conceitos básicos referentes ao Tratamento de Imagem e delimitar as condições em que tais imagens foram utilizadas neste trabalho.

É difícil ousar uma definição de imagem pois, sendo inerente a este conceito a ideia de representação, teremos que tomar em conta a subjectividade da interpretação humana quando pretendemos fazer uma aproximação coerente a essa representação.

No entanto, poderemos sempre associar a uma imagem:

- a) À extracção de informações relevantes relativas aos objectos ou às relações entre os objectos representados (é comum dizer-se que “uma imagem vale por mil palavras”);
- b) À existência de um suporte que possibilitará a troca de informações, ou seja, o seu armazenamento, processamento e difusão.

A diversidade de aplicações na área da imagem é uma consequência do conjunto de informações associado a uma imagem e do avanço das novas tecnologias que permitem, com equipamentos cada vez mais sofisticadas e a menor custo, a representação de objectos do “mundo real” não captáveis pela visão humana.

Essa “visão computacional” é materializada nas imagens digitais, ou seja, imagens que podem ser armazenadas em memórias digitais recorrendo, para a sua representação, à codificação em *bits* (*binary digit*) que, por sua vez, são agrupados em *bytes* (*binary term*), possibilitando, dessa forma, a sua utilização para o armazenamento e processamento digital.

As imagens digitais que são expressas por uma variável numa malha regular designam-se por imagens matriciais (ou *raster*) e as que são geradas através da aplicação de quaisquer algoritmos (imagens geradas, por exemplo, no CAD¹ ou em SIG²) são denominadas imagens vectoriais.

¹ *Computer Aiding Design*

² *Sistemas de Informação Geográfica*

Nesta dissertação só vão ser trabalhadas as imagens em estrutura matricial, isto é, aquelas que são expressas por uma variável numa malha regular, não sendo consideradas as imagens vectoriais.

Uma imagem matricial é constituída por uma matriz de pontos correspondendo aos *pixels* da imagem. Cada ponto (*pixel*) corresponde a um valor particular de luminosidade/cor formando, no seu conjunto, um mapa de pontos (*bitmap*).

Por convenção a dimensão de uma imagem matricial é dada pelo produto do número de linhas pelo número colunas da matriz que a constitui entendendo o número de linhas como o número de *pixels* na horizontal e o número colunas como o número de *pixels* na vertical.

São exemplos de imagens matriciais quase todas as imagens resultantes da digitalização através de um *scanner*, da gravação de imagens de satélite e ou das câmaras digitais (fotográficas ou de vídeo). É sempre possível gravar uma imagem como um ficheiro *.bmp* ainda que as imagens com alta resolução gerem ficheiros *bmp* de tamanho considerável. A maior vantagem da utilização de imagens matriciais está na possibilidade do manuseamento de cada *pixel*.

A digitalização é o processo de conversão de uma imagem contínua numa imagem digital (imagem matricial) correspondente a uma matriz $A[m,n]$ onde m e n são inteiros positivos representando, respectivamente, o número de colunas e o número de linhas da matriz A . Os elementos $a[i,j]$ da matriz $A[m,n]$ contêm informação sobre cor (ver fig. 7³), profundidade e/ou tempo.

³ Imagem obtida em Color Picker, <http://www.asahi-net.or.jp/%7EAFX6M-FJMY/>, <http://www.asahi-net.or.jp/%7EAFX6M-FJMY/java09e.html>

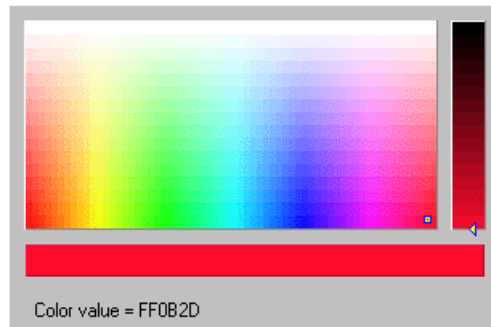


Fig. 7 Informação de cor em hexadecimal com 6 caracteres

As imagens matriciais podem conter qualquer número de cores, que são descritas através da sua codificação em canais. Nesta introdução apenas se pretende abordar os tipos de imagens matriciais que irão ser considerados neste trabalho e que são os seguintes:

- Imagens *grayscale* – são imagens constituídas apenas por gradações de cinzentos (incluindo o preto e o branco) que usam apenas um canal de cor. O nível de cinzentos existente numa imagem é normalmente uma potência de grau 2 onde o expoente corresponde ao número de *bits* usados para a representação dos níveis de brilho. Se o expoente for 1, a imagem designa-se por binária, sendo os níveis de cinzento 0 e 1, ou seja, preto e branco.
- Imagens *coloridas* – usam três ou mais canais de codificação da cor. O nível de cada canal de cor também é codificado por uma potência de grau 2.

Este tipo de imagens, quando ampliadas (ver fig. 8) ou reduzidas, podem perder qualidade uma vez que, para efectuar este tipo de operações, ter-se-á que criar *pixels* inexistentes ou redistribuir os existentes.



Imagem original



Fig. 8 Imagem matricial (imagem de satélite) e sua ampliação (2x)

2.3.1. Resolução de imagens digitais

Todas as imagens digitais - por exemplo, as imagens de satélite e imagens resultantes de digitalizações - são armazenadas como conjuntos de binários ou inteiros que representam gamas de valores existentes no espectro electromagnético (por exemplo, intensidade de luz reflectida, comprimentos de onda de energia, etc.).

Em [YGV 98] define-se imagem digital como uma função bidimensional $i(x,y)$ da intensidade da imagem i no ponto de coordenadas (x,y) , podendo a intensidade (geralmente representada por um número real e só excepcionalmente - por exemplo, no caso da ressonância magnética - por um número complexo) depender de variáveis como a profundidade, a cor e o tempo de exposição.

Cada imagem é constituída por conjuntos de objectos que podem ser agrupados, por necessidade de operar em zonas mais localizadas, no que vulgarmente se designa por regiões. Por exemplo, pode destacar-se numa região uma “mancha visual” para permitir, por ampliação, a percepção do objecto ou objectos aí representados (este tipo de abordagem é muito comum para o processamento de imagens de satélite).

A resolução de uma imagem (número de *pixels* ou *picture cells* por unidade de dimensão) é a medida do detalhe que um equipamento pode capturar ou reproduzir. Por exemplo, a resolução de uma máquina fotográfica digital ou de um *scanner* (expressa em número de *pixels* horizontais por número de *pixels* verticais) é o número de *pixels* que podem ser capturados. A resolução de um ecrã (*pixels per inch* (PPI)) ou de uma impressora (*dots per inch* (DPI)) é o número de *pixels* ou pontos que podem ser visualizados ou impressos.

Note-se que é possível alterar a resolução de uma imagem sem alterar a dimensão da área que é representada. Na fig. 9 visualizam-se duas imagens⁴ representadas para a mesma dimensão (938x1250) mas que têm, respectivamente, resoluções originais de 72 e 300 *ppi* determinando assim que, a imagem da direita apresente maior clareza e detalhe.

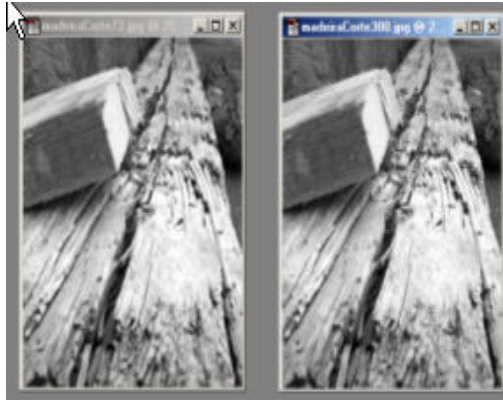


Fig. 9 Imagens com 72 e 300 *ppi* com a mesma área de ocupação

As imagens em alta resolução (no mínimo com 300 *ppi*) possibilitam, obviamente, a visualização com maior detalhe e maior gradação de cores que em baixa resolução. Todavia, o aumento de resolução na representação visual de uma imagem obtida em baixa resolução (cerca de 72 *ppi*) apenas determina a utilização de maior número de *pixels* sem qualquer melhoria da qualidade da imagem, pois os *pixels* inseridos não existiam na imagem original e, como tal, correspondem a repetições ou estimações de cor.

A dimensão ocupada por um ficheiro de imagem, sem compressão, é proporcional à sua dimensão em *pixels* donde, em geral, imagens com maior numero de *pixels* conduzem a um maior detalhe na impressão mas, conseqüentemente, exigem mais espaço para armazenamento e a processamentos mais morosos. Assim, é necessário estabelecer um compromisso entre a qualidade suficiente para capturar todos os dados necessários para uma certa aplicação e o tamanho do ficheiro que armazena esses dados.

Note-se ainda que o tamanho de um ficheiro depende, também, do tipo de formato de armazenamento e do número de *bits* utilizados para armazenar a cor (profundidade de cor).

2.3.2. Georeferenciação de imagens

As imagens produzidas por sensores remotos, sejam fotografias aéreas ou imagens de satélite, apresentam uma série de distorções espaciais por causas diversas, tais como, variações de altitude, posição e velocidade de uma plataforma e desvio posicional do terreno, devido à

⁴ Imagem obtida em *real-eyes a.k.a Franck JULIEN*, <http://sxc.hu/browse.phtml?f=view&id=36555>

topografia do mesmo. O registo bruto desses sensores não apresentam, portanto, precisão cartográfica quanto ao posicionamento dos elementos lá representados.

Para o processamento desse tipo de imagens é muitas vezes necessária a obtenção precisa das coordenadas referentes aos elementos aí representados que, pelas distorções já referidas, não poderão ser obtidas tendo apenas em conta os *pixels* e a escala que representam.

A correcção geométrica ou georeferenciação é o ajuste efectuado nas imagens obtidas por sensores remotos. Usualmente, este ajuste é realizado através da indicação do sistema de referência (por exemplo, latitude/longitude), dos sistemas de projecção (por exemplo, Gauss e UTM) e das coordenadas da imagem (limites inferior, superior, esquerdo e direito) podendo, no entanto, existir situações em que a georeferenciação é realizada através de outra imagem georeferenciada ou mapa cartográfico.

Existem diversas ferramentas interactivas que realizam de forma semi-automática a georeferenciação, incorporando, em muitas situações, Sistemas de Informação Geográfica (SIG) que possibilitam a utilização da imagem de um terreno a uma determinada escala, para se fazerem medições de comprimentos, áreas e direcções, desde que sejam previamente eliminadas as distorções.

2.3.3. Modelos de codificação de cor

Em periféricos de saída do tipo ecrã e de entrada do tipo *scanner*/câmara digital, a cor é obtida pela síntese aditiva da modelação da intensidade de três feixes de luz monocromáticos (cores fundamentais): vermelho, verde e azul.

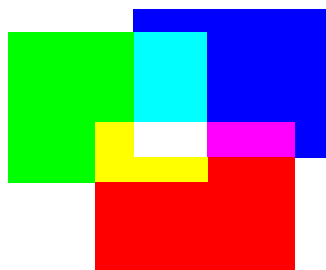


Fig. 10 Síntese aditiva

Na fig. 10 pode observar-se que a luz branca aparece na confluência da intensidade dos três feixes e as três cores primárias, ciano, magenta e amarelo aparecem, respectivamente, no cruzamento do verde com o azul, no cruzamento do vermelho com o azul e no cruzamento do verde com o vermelho.

No armazenamento digital das imagens do tipo matricial recorre-se a ficheiros onde uma cor é descrita por valores em três canais (*Red Green Blue* - RGB). Esses valores, compreendidos entre 0 e 255 por canal (*8-bits* por canal), representam a intensidade de cada uma das cores.

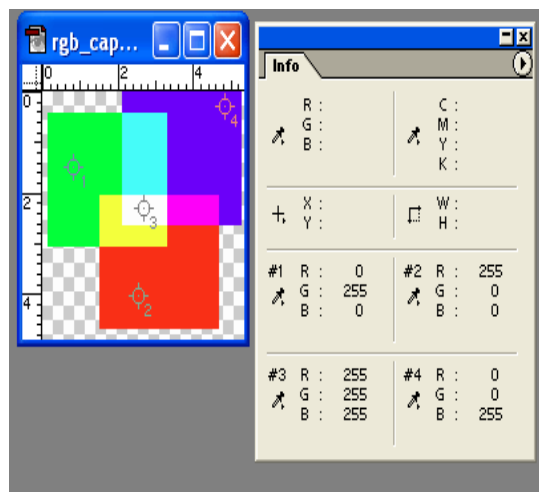


Fig. 11 Informação RGB

Assim, na imagem da fig. 11, obtida através do *Adobe Photoshop®*, pode ver-se que o vermelho corresponde ao RGB (255,0,0), o verde ao RGB (0,255,0) e o azul ao RGB (0,0,255). Finalmente, o branco resultante da síntese aditiva corresponde ao RGB (255,255,255).

Nas imagens *grayscale* onde se dispõe apenas de um único canal de cor existirá uma graduação de 0 a 255 tons de cinza onde o preto corresponde ao valor zero e o branco ao 255. Estas imagens, quando representadas em modelos de três canais do tipo RGB, apresentam os mesmos valores para os três canais (ver fig. 12).

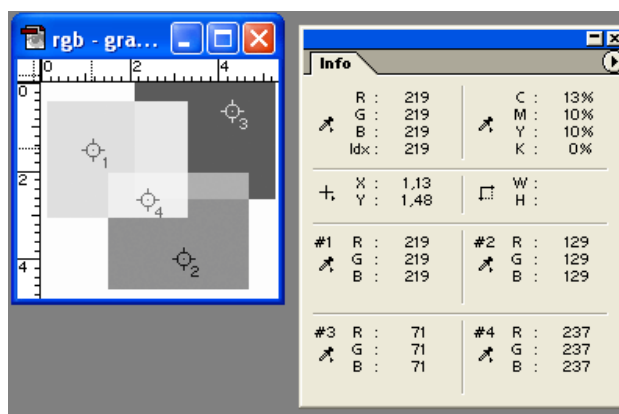


Fig. 12 Informação RGB em imagem *grayscale*

A síntese subtractiva das cores, criação de qualquer cor pela mistura de cores primárias (ciano, magenta e amarelo) é muito comum na quadricomia, técnica actual muito utilizada para impressão onde, para obtenção da imagem final, se efectuam quatro impressões de cores transparentes, três correspondendo às cores complementares das primárias e uma quarta, ao preto, “para reforçar as sombras e ajudar a avivar o recorte da imagem” [NR 01].

Assim, uma imagem digital que se pretenda impressa em quadricomia, terá que ser convertida para um formato que utilize quatro canais de cor (*Cian Magenta Yellow Black - CMYK*), onde o valor em cada canal representa a saturação (0-100%) da cor respectiva. Da combinação das quatro cores com valores de saturação a 100% obter-se-á o preto e, no caso de todos os seus valores de saturação serem 0%, obter-se-á o branco puro.

Geralmente, os *softwares* de preparação de imagem para impressão, por exemplo o *Adobe Photoshop®*, incorporam ferramentas para conversão automática de modo RGB para CMYK pois se, por um lado, o “olho humano vê biliões de cores, o ecrã do computador atinge os 16 milhões, um filme fotográfico 10 a 15 milhões e a impressão apenas 6 mil cores” [HP PAG]. Esta circunstância impede a reprodução perfeita das imagens coloridas.

2.3.4. Alguns formatos de armazenamento de imagens matriciais

Apresenta-se seguidamente uma pequena resenha de formatos de armazenamento de imagens matriciais [JM 00], referidos aqui por serem de uso corrente em sistemas Windows ou por serem muito utilizados na troca de imagens entre diferentes sistemas operativos (Windows®, Machintosh® e UNIX®).

2.3.4.1. Formato BMP

O formato BMP é o formato padrão para o Microsoft Windows[®], e é suportado por todos os *softwares* gráficos. Destacam-se aqui os formatos *bmp* com 8 bits (correspondendo a um só canal) e de 24 bits (correspondendo a 3 canais), pois serão estes os mais utilizados nesta dissertação para armazenamento de imagem.

2.3.4.2. Formato GIF

O formato GIF (*Graphic Interchange Format*) foi desenvolvido, originalmente, pela CompuServe[®], e é muito utilizado no ambiente da Internet pois utiliza um método de compressão que permite preservar a imagem original.

2.3.4.3. Formato JPG

O formato JPG (*Joint Photographic Experts Group format*) é um formato a que se recorre apenas quando se pretende usar um número elevado de cores em imagens de baixa resolução, dado que a compressão efectuada não preserva as imagens em alta resolução. A maioria dos utilizadores da *Web* aceita o formato JPG por este permitir que a compressão corresponda à qualidade pretendida.

2.3.4.4. Formato TIFF

O formato TIFF (*Tagged Image File Format*) é um dos formatos de imagem desenvolvido há mais tempo que possibilita alguma compressão sem perda de qualidade. Sendo um formato reconhecido pela Machintosh[®], Windows[®] e Unix[®] é, por conseguinte, um formato conveniente para a transferência de ficheiros.

2.3.4.5. Formato PNG

O formato PNG (*Portable Network Graphics*) é bastante recente (1995) [HP PNG] e apareceu como uma alternativa ao formato GIF (CompuServe®) e como uma extensão do formato TIFF [HP PNG]. Este formato possibilita a existência de imagens com profundidade de cor de 48 *bits* (16 por canal) e suporta o armazenamento, entre outros, do factor *gamma* (curva entre a intensidade luminosa e o contraste), o que permite a correcção de potenciais distorções na imagem quando visualizada em diferentes ecrãs. Tal como o GIF, o PNG também permite uma compressão que preserva a imagem original.

2.3.5. As operações sobre imagens e o conceito de vizinhança

Nesta secção, após a introdução do conceito de vizinhança de um *pixel* e da definição do tipo de operações passíveis de realizar sobre uma imagem, introduz-se a problemática do reescalonamento.

2.3.5.1. Vizinhança de um *pixel*

Para a caracterização das operações sobre imagens digitais é necessário, em primeiro lugar, atender ao conceito de vizinhança.

A vizinhança é um conceito fundamental para o processamento digital de imagem, destacando-se, pela sua utilização [YGV 98], alguns tipos de vizinhança rectangulares (fig. 13):

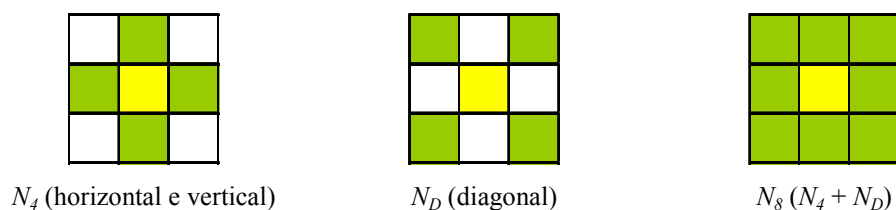


Fig. 13 Tipos de vizinhanças

2.3.5.2. Operações sobre imagens

As operações sobre imagens digitais, que transformam uma imagem de entrada numa imagem de saída, podem ser classificadas em três categorias [YGV 98]:

- Pontuais – se o valor de saída numa dada coordenada depende apenas do valor de entrada nessa coordenada;
- Locais – se o valor de saída numa dada coordenada depende apenas dos valores existentes na vizinhança do valor de entrada nessa coordenada;
- Globais – se o valor de saída numa dada coordenada depende de todos os valores de entrada existentes na imagem.

2.3.6. Reescalonamento de imagens

O reescalonamento ou redimensionamento de uma imagem implica o aumento ou diminuição do seu número de *pixels*. Para diminuir o número de *pixels*, é necessário retirar informação da imagem; pelo contrário, para aumentar o número de *pixels*, ter-se-á que criar novos *pixels* baseados nos valores dos *pixels* existentes

Quando se faz uma ampliação, por exemplo a 30%, duma imagem (fig. 14) intercalar-se-ão novos *pixels* (30 em cada 100) cujos valores terão que ser obtidos à custa dos existentes, usando técnicas ou métodos de interpolação.

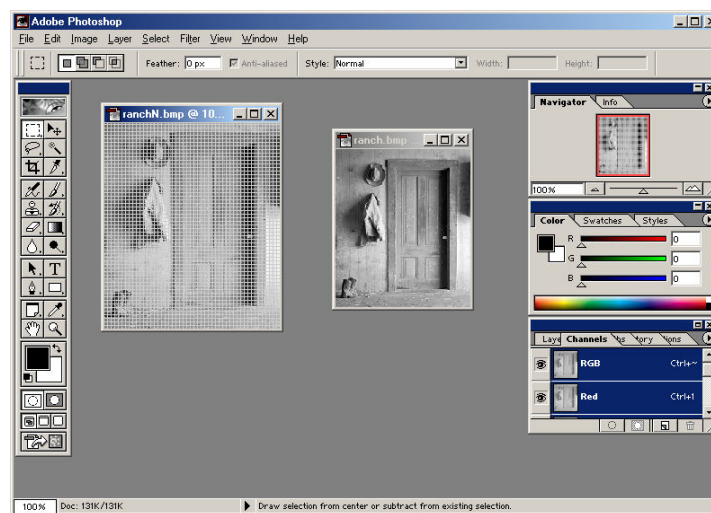


Fig. 14 *Pixels* a preencher na imagem da esquerda por ampliação de 30% da imagem da direita

De seguida, descrevem-se os três métodos que, segundo a literatura consultada ([SUN 99], [HG 03] *et al.*), são os mais utilizados pelos *softwares* de *design* gráfico em redimensionamento de imagem (métodos disponíveis para reescalonamento, por exemplo, no Adobe Photoshop®). Refira-se que uma das funcionalidades práticas do sistema *MODIMAV* implementado neste trabalho é proceder à utilização do interpolador *krigagem* para alterar a dimensão de uma imagem, tal como é ilustrado num dos exemplos de estudo.

2.3.6.1. Vizinhança mais próxima

Cada *pixel* (i', j') da imagem redimensionada estima-se à custa do valor do *pixel* mais próximo (i, j) da imagem inicial. Assim, se a imagem original tiver como dimensões l e n e a

imagem transformada as dimensões l' e n' , a localização do *pixel* estimado por (i, j) será dada por $(i', j') = (i' = i.l'/l, j' = j.n'/n)$ sendo l'/l e n'/n inteiros calculados por defeito.

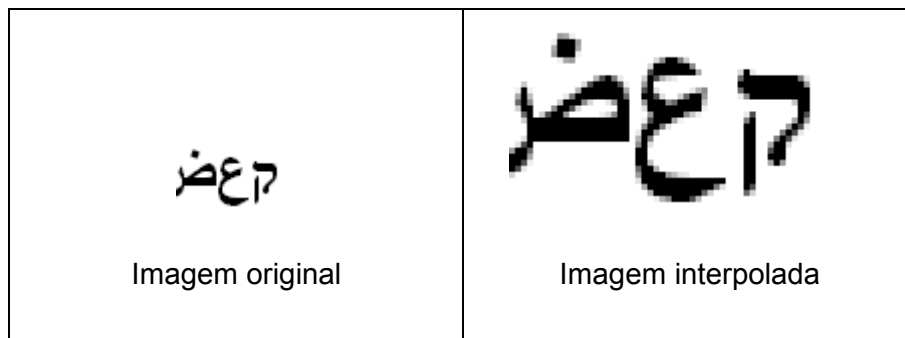


Fig. 15 Ampliação de 300% pelo método da vizinhança mais próxima

Destas três técnicas descritas esta é a mais rápida em tempo de CPU mas, como não tem em conta a influência de outros *pixels* na proximidade, gera, usualmente, efeitos de *aliasing*⁵ (ampliação de 300% na fig. 15).

Note-se que esta imagem (fig. 15) foi usada como exemplo para ilustrar os aspectos aqui referidos pela sua configuração recortada que permite salientar os aspectos das operações consideradas.

2.3.6.2. Interpolação bilinear

A interpolação bilinear, interpolação de 1ª ordem, efectua duas interpolações lineares consecutivas (uma segundo a linha e outra segundo a coluna dos *pixels* da imagem) estimando cada *pixel* de acordo com a vizinhança rectangular (2x2) (ver fig. 16) do *pixel* mais próximo e correspondente da imagem original.

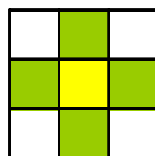


Fig. 16 Vizinhança de um *pixel*

⁵ Como as imagens são criadas usando *pixels* quadrados as figuras não quadradas parecem possuir recortes.

A interpolação bilinear é uma técnica que requer mais tempo de CPU que a técnica da vizinhança mais próxima (obriga a mais cálculos matemáticos) mas produz imagens de melhor qualidade ainda que, por vezes, com alguns efeitos de esbatimento (ampliação de 300% na fig. 17).

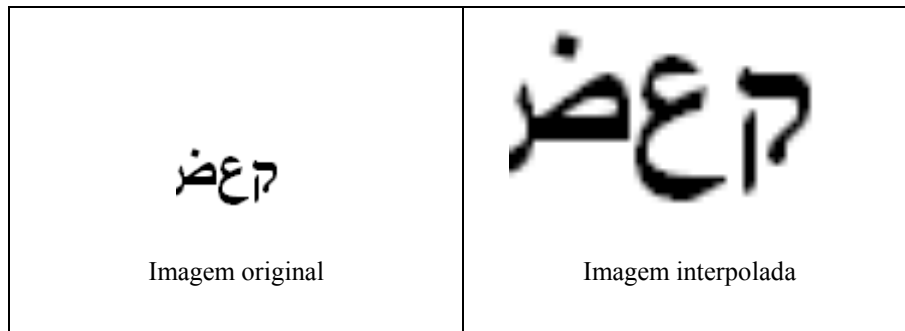


Fig. 17 Ampliação de 300% pelo método da interpolação bilinear

2.3.6.3. Interpolação bicúbica

Esta interpolação, como a designação indica, já não é linear mas sim cúbica e a vizinhança considerada será 4×4 *pixels*, ou seja, efectuar-se-á uma interpolação cúbica para os 16 *pixels* com menor distância das coordenadas do *pixel* original.

Note-se que interpolação bicúbica usa uma maior vizinhança e um modelo matemático mais sofisticado, com o objectivo de preservar os detalhes da imagem (ampliação de 300% na fig. 18) embora aumente o tempo de CPU para a construção da imagem redimensionada.

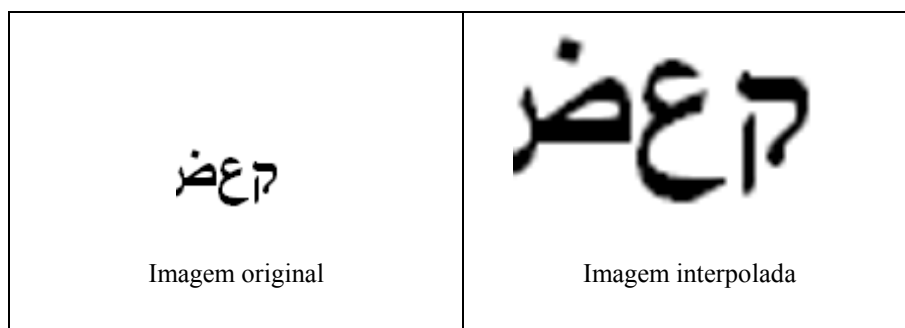


Fig. 18 Ampliação de 300% pelo método da interpolação bicúbica

3. Ajustamento automático de variogramas experimentais por um algoritmo genético

Neste capítulo apresentam-se os conceitos subjacentes e o algoritmo genético desenvolvido para o ajustamento automático de variogramas experimentais a uma função teórica.

Neste âmbito, descrevem-se sucintamente métodos que, na óptica de problemas de optimização, permitem o ajustamento automático com funções pré-definidas, introduz-se a computação evolucionária, caracterizam-se os seus principais métodos de implementação e as vantagens da sua utilização na resolução de problemas desse tipo, apresenta-se o algoritmo desenvolvido acompanhado de três exemplos sintéticos que ilustram o seu desempenho em comparação com a modelação manual tradicional de variogramas experimentais e, finalmente, compara-se o *software* aqui desenvolvido com outros *packages* disponíveis para atingir os mesmos objectivos.

3.1. O problema do ajustamento dos variogramas experimentais e testes de validação

No ajustamento dos variogramas experimentais, existe sempre um certo grau de incerteza sobre as hipóteses de estacionaridade que foram assumidas, modelos teóricos utilizados no ajustamento e seus parâmetros o que, em conjunto, se reflecte na qualidade da estimação por *krigagem*.

Esta incerteza global, que representa o que podemos designar por erro de estimação ($E=Z(x_a)-Z^*(x_a)$), com $Z(x_a)$ o valor amostrado e o respectivo valor estimado $Z^*(x_a)$, pode ser avaliada por um teste de validação cruzada ("cross validation") [IS 89].

No teste de validação cruzada, cada ponto real (x_α) é sucessivamente excluído e o seu valor é estimado a partir das restantes amostras, sendo o erro de estimação calculado pela média entre o valor real e o estimado em cada ponto.

Este teste permite avaliar a qualidade do método de estimação face a vários parâmetros de escolha subjectiva, nomeadamente, o modelo teórico do variograma e respectivos parâmetros considerados no ajustamento (por exemplo, direcções principais e relações de anisotropia, efeito de pepita), e o número de amostras a considerar na estimação por *krigagem* de cada ponto.

Normalmente, na avaliação dos resultados do teste de validação cruzada, consideram-se dois parâmetros de decisão: o erro médio (EM) e o erro quadrático médio (EQM):

$$EM = \frac{1}{n} \sum_{\alpha=1}^n (Z(x_\alpha) - Z^*(x_\alpha))$$

$$EQM = \frac{1}{n} \sum_{\alpha=1}^n (Z(x_\alpha) - Z^*(x_\alpha))^2$$

O erro médio (EM) é a diferença entre a média dos valores reais e a média dos valores estimados. Mede o enviesamento da estimação, ou seja, o desequilíbrio dos valores sub-estimados ou sobre-estimados. O erro médio varia, teoricamente, entre $-\infty$ e $+\infty$ e pretende-se que seja tão próximo de zero quanto possível.

O erro quadrático médio (EQM) quantifica a dispersão da nuvem de pontos, num diagrama bivariado que representa os valores reais *versus* os valores estimados. Pelo facto de considerar a diferença ao quadrado, ignora as situações de sub-estimação e sobre-estimação. Também este parâmetro deve ser o mais próximo de zero.

O maior problema que se coloca na avaliação dos resultados de um teste de validação cruzada é o número de parâmetros que é necessário ajustar (pelo menos três ou quatro) e o facto de poderem considerar-se pelo menos dois critérios de qualidade, nem sempre concordantes.

3.2. Conceitos relativos a procura de soluções para problemas de otimização

Num típico problema de otimização, pretende-se minimizar (ou eventualmente maximizar) uma função, designada genericamente por *função objectivo*, que depende de um conjunto de parâmetros.

Por exemplo, na fig. 19 ilustra-se um problema simples de otimização, onde é representada a função $F(X)$ do tipo polinomial de grau 2 que melhor se ajusta ao conjunto de n pontos (x_i, y_i) , isto é, $F(X)$ é tal que a soma das distâncias entre todos os pontos (x_i, y_i) e a função $F(X)$ é mínima.

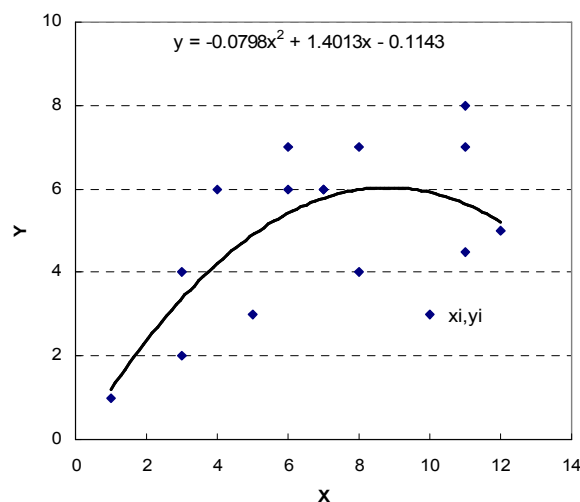


Fig. 19 Curva ajustada a um conjunto de pontos

Consoante as especificidades de cada caso de Figura considerado, os problemas de otimização podem, para além da simples pesquisa de um máximo ou mínimo local ou global de uma dada função, ter necessidade de considerar restrições, *a priori*, nos valores permitidos para as variáveis independentes [PTVF 02].

Para identificar o valor óptimo de uma qualquer função objectivo destacam-se três classes de métodos de otimização ([HP KUSIAK]):

- *Métodos enumerativos* – a pesquisa de um óptimo num espaço finito consiste em calcular o valor da função para cada ponto desse espaço e reter a situação óptima. Estes métodos são extraordinariamente simples mas têm a desvantagem de testarem as soluções de forma sistemática e exaustiva. Por isso tornam-se pouco eficientes à medida que aumenta a dimensão do espaço de pesquisa, podendo

mesmo tornar-se computacionalmente inviáveis no caso de espaços de pesquisa de grandes dimensões.

- *Métodos deterministas* – estes métodos baseiam-se na programação matemática e geram uma solução obtida através do refinamento de pesquisas anteriores. Nesta categoria, destacam-se os métodos de *hill climbing* ou o método de gradiente, em que o cálculo do valor óptimo é efectuado através de um processo iterativo em que se parte de um ponto inicial e se procura, na sua vizinhança, o ponto que conduz a um gradiente máximo. O processo é repetido a partir do novo ponto até que o gradiente mude de sinal.

Estes métodos permitem a rápida convergência para a solução, se for considerado um “bom” valor de pesquisa inicial. Todavia, apenas possibilitam a pesquisa de um óptimo local na vizinhança desse ponto, o que pode conduzir à sua ineficiência em espaços onde existam vários óptimos locais, não se encontrando pois a solução que conduz a um óptimo global. Esta desvantagem, pode ser minorada, quando se utiliza a computação paralela para pesquisa simultânea em múltiplos pontos do espaço.

- *Métodos estocásticos (ou aleatórios)* – são métodos que percorrem aleatoriamente o espaço do problema, guardando as melhores soluções encontradas. Em métodos puramente aleatórios, a questão da eficiência, para espaços de grande dimensão é semelhante à que foi referida anteriormente para os métodos enumerativos. Contudo, se forem usadas técnicas de pesquisa que possibilitem o uso do conhecimento adquirido nos resultados anteriores para a escolha dos valores seguintes, estes métodos podem ser mais eficientes do que os métodos deterministas pois, ao contrário destes, são capazes de ultrapassar as regiões de mínimos/máximos locais pela sua natureza aleatória e probabilística.

Nesta categoria destacam-se os algoritmos de *Simulated Annealing* e de *Computação Evolucionária*, sendo estes últimos os que foram considerados neste trabalho na resolução do problema proposto.

3.3. Computação evolucionária

A computação evolucionária é uma área recente de investigação no domínio da optimização inspirada na teoria da evolução de Darwin como algoritmo de selecção e reprodução [AS 04]. Neste tipo de métodos, os "modelos computacionais de processos evolutivos são os elementos chave no design e implementação de sistemas *computer-based-problem-solving*" [SJBFG 93].

Tal como na teoria da evolução de Darwin, todos os modelos existentes na área da computação evolucionária têm em comum a ideia de população (conjunto de indivíduos/cromossomas) e os processos de selecção e reprodução com base no melhor desempenho (maior capacidade de sobrevivência).

Tipicamente, um algoritmo evolucionário é codificado da seguinte forma [SJBFG 93] (fig. 20):

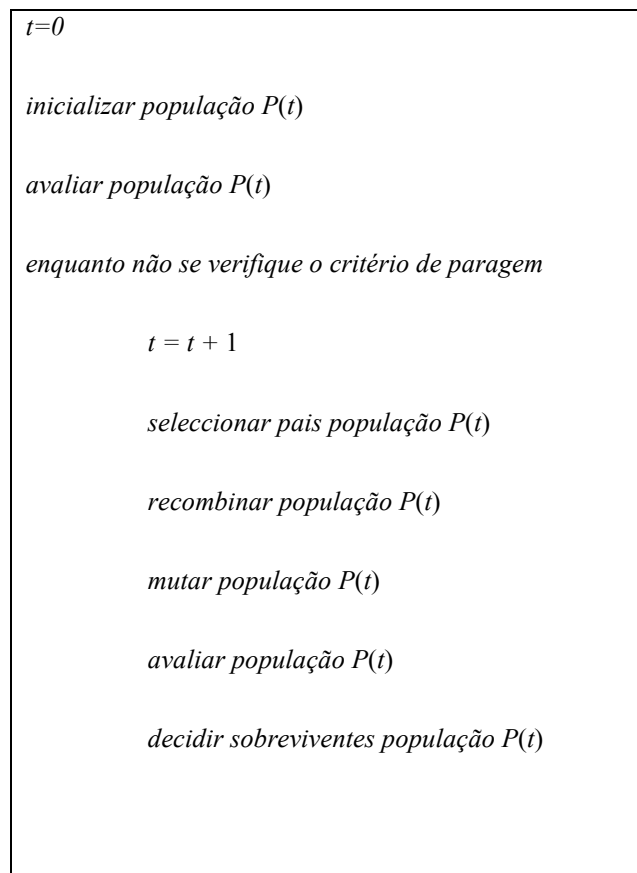


Fig. 20 Algoritmo evolucionário

Num algoritmo evolucionário considera-se que:

- A população $P(t)$ é constante ao longo das gerações (t), sendo inicializada aleatoriamente ou a partir de técnicas que permitam restringir o campo de busca da solução;
- A selecção é efectuada em dois passos: decisão sobre quem serão os progenitores e quantos filhos terão;
- A reprodução é realizada pelo processo de recombinação (troca de material genético) e/ou mutação (introdução de informação genética) com o objectivo de criar indivíduos diferentes dos pais;
- A sobrevivência, tal como na teoria da evolução de Darwin, apenas é possível aos mais aptos/capazes, ou seja, aqueles que tiverem melhor desempenho (melhores classificações de acordo com a função de avaliação).

Descrevem-se, seguidamente, de forma sintética, três modelos (algoritmos genéticos, programação evolucionária e estratégias evolutivas) que resultam de diferentes abordagens na implementação dos algoritmos evolutivos, detalhando-se especialmente os algoritmos genéticos que foram implementados no sistema *MODIMAV* desenvolvido no âmbito deste trabalho.

É de notar que esta classificação em três modelos deriva de factores históricos relacionados com o aparecimento de diferentes linhas de investigação [TA 00] e que as grandes diferenças entre eles advêm fundamentalmente dos processos de recombinação e selecção utilizados.

3.3.1. Algoritmos genéticos

Os algoritmos genéticos, inspirados no processo de evolução natural, foram desenvolvidos por Holland ([HJ 75]) com base no seguinte paradigma: “sendo uma população constituída por n indivíduos, serão seleccionados n progenitores que, por recombinação, gerarão n filhos. Os filhos sofrem mutações e sobrevivem substituindo todos os pais ou apenas os pais menos aptos”.

Os algoritmos genéticos usaram, inicialmente, uma população que representa um conjunto de soluções para um problema, codificada através de conjuntos de parâmetros (genes) que são agrupados em cadeias de valores designados por indivíduos ou cromossomas (designando-se por alelo, cada valor possível para um gene).

Tal como na ciência genética, o conjunto de parâmetros representados no cromossoma (material genético de um indivíduo) designa-se por genótipo que, por sua vez, contém a informação necessária para construir um organismo (fenótipo), ou seja, “os valores dos seus genes e o ambiente que o rodeia”[TA 00]. A aptidão do indivíduo pode ser calculada a partir do cromossoma usando uma função de avaliação.

Para a codificação dos genes é utilizada com frequência a representação binária, embora existam autores que, em problemas onde se pretende um elevado grau de precisão, consideram inadequado este tipo de codificação (entre outros [CH 04] e [MSB 91]), pois a utilização dessa representação obrigará a um esforço computacional muito elevado.

Tal como é defendido em [CH 04] e [MSB 91], foram utilizados nesta dissertação os números reais na codificação do conjunto de parâmetros que representam as possíveis soluções.

Os algoritmos genéticos são referenciados na literatura, quase sempre, como os mais eficientes em problemas de optimização, no que diz respeito à detecção de máximos e mínimos globais (entre outros, [SJBFG 93] e [CH 04]).

3.3.2. Programação evolucionária

Um algoritmo de programação evolucionária [FOW 66] é muito semelhante a um algoritmo genético, sendo a principal diferença relativa à forma como os descendentes são gerados. Neste caso, não existe qualquer estratégia recombinatória ou de cruzamento.

Esta opção é justificada por se considerar que “o processo de mutação poder permitir, se necessário, perturbações semelhantes às da recombinação”[SJBFG 93].

Neste tipo de algoritmos, em consequência da não utilização da recombinação, é possível usar qualquer tipo de representação para os elementos da sua população, pois esta apenas será limitada pelo universo do problema e pela escolha adequada do operador de mutação.

Assim, se a população for constituída por n indivíduos serão seleccionados n pais e serão gerados, através de mutação, n filhos (um por cada progenitor). De seguida, os $n+n$ indivíduos serão avaliados e n sobreviverão, sendo a escolha dos sobreviventes efectuada com base numa função probabilística que permite apenas a sobrevivência dos que revelam melhor desempenho.

3.3.3. Estratégias evolutivas

Segundo [TA 00], as estratégias evolutivas são introduzidas por Rechenberg em 1973 e, ao contrário da programação evolucionária, utilizam um operador de recombinação/cruzamento.

Neste tipo de algoritmo, os filhos são gerados por pares de progenitores através de recombinação e mutação. Se considerarmos uma população com n indivíduos que serão todos progenitores, o número de filhos gerados é, em geral, superior a n , sendo a população seguinte (nova geração) constituída, em geral, da seguinte forma:

- Ou sobrevivem apenas n filhos;
- Ou sobrevivem, entre ascendentes e descendentes, os n indivíduos mais aptos, ou seja, com melhor desempenho.

Note-se que este tipo de modelo também se diferencia do anterior na forma como é efectuada a selecção embora, tal como na programação evolucionária e ao contrário dos algoritmos genéticos, a mutação seja adaptativa [SJBF 93], ou seja, cada variável em cada indivíduo pode mudar de acordo com uma determinada função de distribuição de probabilidade. Esta informação suplementar possibilita a sua evolução, ou seja, a adaptação da mutação ao longo da pesquisa.

O processo de recombinação/selecção também diverge do que é utilizado nos algoritmos genéticos, ou seja, enquanto nos algoritmos genéticos a mutação é aplicada a indivíduos seleccionados do espaço de progenitores, no caso das estratégias evolutivas toda a população é submetida à mutação e a selecção da nova população é efectuada posteriormente. É assim possível, com os algoritmos genéticos, ter indivíduos na nova população que já existiam na população anterior.

3.4. Algoritmos evolutivos

Nesta secção especificam-se os operadores usados, descreve-se a inicialização dos parâmetros de um algoritmo evolutivo e conclui-se referindo as vantagens e desvantagens da utilização deste tipo de algoritmos.

3.4.1. Operadores usados

Seguidamente destacam-se algumas técnicas mais usadas na implementação destes operadores.

3.4.1.1. Selecção

A selecção permite dirigir a pesquisa para áreas de maior desempenho e, tal como foi referido anteriormente, é realizada em duas etapas do algoritmo evolutivo: na fase de escolha dos progenitores e na escolha dos sobreviventes.

Descrevem-se, de seguida, algumas metodologias para a selecção de indivíduos:

- *Selecção aleatória* – escolhem-se aleatoriamente os indivíduos da população. Note-se que, se o único processo de selecção for este, a população não converge pois não se introduz melhoria na qualidade da população.
- *Torneio* – escolhem-se aleatoriamente k indivíduos, e o que tiver melhor desempenho será o eleito. Esta técnica é muito utilizada com $k = 2$;
- *Selecção pela média* – escolhem-se os indivíduos cujos desempenhos sejam superiores aos da média da população;
- *Roleta* – a nova população é obtida por tiragens, com ou sem reposição, considerando a probabilidade de cada indivíduo sobreviver como proporcional ao seu desempenho;
- *Elitista* – escolhe-se o que apresenta melhor desempenho.

3.4.1.2. Recombinação

A recombinação, no caso de codificação de variáveis reais, usa basicamente dois métodos:

- A troca de valores entre os parâmetros;
- A média dos valores dos parâmetros.

Quando se referem algoritmos genéticos com representações por cadeias de *bits*, a recombinação poderá ser:

- *Com um ponto* – cada descendente recebe informação genética de cada um dos progenitores baseada num ponto de corte. Por exemplo, ao recombinar AX com BY obter-se-iam descendentes da forma AY e BX.
- *Multi-ponto* – corresponde à generalização do caso anterior;
- *Uniforme* – o descendente recebe a mesma informação genética de cada progenitor.

3.4.1.3. Mutação

Com a mutação, podem naturalmente obter-se indivíduos (genes) com melhor ou pior desempenho do que a população existente. No caso de indivíduos com pior desempenho, estes serão eliminados em gerações futuras; caso contrário, serão preservados, garantindo assim uma certa diversidade na população.

As técnicas de mutação mais utilizadas são:

- Reinicialização de um indivíduo;
- Troca aleatória de dois indivíduos;
- Inversão de um *bit* em caso de codificação por *bits*;

Um parâmetro importante é a probabilidade da mutação que corresponde a uma taxa indicando a frequência com que esta se efectuará.

3.4.1.4. Substituição

O objectivo da substituição (sobrevivência) é decidir se a nova população é apenas constituída pelos novos indivíduos ou se existem indivíduos da geração anterior que aí devam permanecer. Note-se que, os "filhos" não são obrigatoriamente melhores que os progenitores.

Em primeiro lugar, deve referir-se que a taxa de substituição indica a percentagem de filhos que devem substituir a geração anterior.

Existem várias metodologias para processar a substituição, nomeadamente:

- *Uniforme* – escolhem os melhores filhos na percentagem indicada pela taxa de substituição;
- *Torneio* – um pai e um filho vão a torneio até que a percentagem de filhos indicada pela taxa de substituição se encontra na nova população.

3.4.2. Inicialização dos parâmetros de um algoritmo evolutivo

Num algoritmo evolutivo do tipo genético, é necessário ponderar os valores a atribuir aos seguintes parâmetros: tamanho da população, taxa de mutação, taxa de recombinação e número de gerações.

De Jong (1975) [LF 00], numa tentativa de padronização dos valores para os parâmetros acima referidos, sugeriu os seguintes valores para esses parâmetros:

- Tamanho da população constante: 50 - 100;
- Taxa recombinatória: 0,6 - 0,9;
- Taxa de mutação: 0,01 – 0,001.

Shaffer *et al.* (1989) [LF 00] prosseguem o estudo de De Jong fazendo experiências com seis tamanhos de população (10; 20; 30; 50; 100; 200), 10 taxas de recombinação (0,05; 0,15; 0,25; . . . ; 0,95), sete taxas de mutação (0,001; 0,002; 0,005; 0,01; 0,02; 0,05; 0,1) e dois operadores de recombinação. Com os testes realizados sugerem que ocorre:

- Uma relação inversa entre o tamanho da população e a taxa de mutação;
- Um bom desempenho com os seguintes parâmetros:
 - a) tamanho da população (constante): 20 - 30;
 - b) taxa recombinação: 0,75 – 0,95;
 - c) taxa de mutação: 0,005 – 0,01.

Não devendo tomar-se estes valores como padrão (uma vez que o domínio de aplicação condiciona o valor a utilizar para os parâmetros), os valores apontados podem, contudo, servir de indicadores para uma primeira aproximação experimental no domínio a aplicar.

3.4.3. Vantagens e desvantagens dos algoritmos evolutivos

Os algoritmos evolutivos, ao contrário dos métodos tradicionais que, iterativamente, vão refinando uma solução, operam horizontalmente sobre todo o espaço de busca onde sejam menos afectados pelos máximos ou mínimos locais [CH 04]. Por outro lado, devido à natureza paralela da sua pesquisa estocástica, também são menos sensíveis às condições iniciais. Por outro lado, a existência de uma população (em vez de um único ponto do espaço do problema) possibilita que a solução não esteja dependente de um único valor e o uso de uma função de avaliação permite que a solução não dependa dos valores obtidos, mas sim da adaptação, determinada pela função de avaliação, desses valores ao tipo de problema de optimização.

Como desvantagens aponta-se que, apesar destes métodos serem metodologias robustas para a pesquisa de óptimos globais, a convergência será mais rápida e mais precisa se for utilizado um algoritmo tradicional com um bom valor inicial. Isto significa que os algoritmos evolutivos são mais lentos devido à sua necessidade de avaliar a correcção das sub-soluções. Note-se ainda que, devido à natureza estocástica destes, a solução será sempre uma estimativa e não um valor exacto.

3.5. Componentes do algoritmo genético implementado

No sistema informático *MODIMAV* desenvolvido no âmbito deste trabalho implementou-se um algoritmo genético para a modelação automática dos variogramas.

Neste algoritmo, cujo fluxograma é apresentado na figura 21, usaram-se as seguintes definições para as suas componentes (cujos valores são justificados, posteriormente, nos sub pontos 3.5.1 a 3.5.9):

- *População* – cada cromossoma (solução) é constituído por seis genes (parâmetros), usando inteiros e reais na sua codificação (existindo 60 indivíduos para cada solução possível);
- *Critério de paragem* – 370 gerações;
- *Inicialização* – aleatória, mas condicionada a limites numéricos (por exemplo, o parâmetro “efeito de pepita” não pode ser superior à variância);
- *Avaliação* – foi considerada uma função baseada no quadrado da diferença entre os valores reais (valores calculados no variograma experimental) e os valores estimados, ou seja, os valores calculados através do modelo teórico em avaliação;
- *Seleção* – o número de descendentes é igual ao número de progenitores, escolhidos aleatoriamente entre os mais aptos e englobando sempre o mais apto da geração anterior;
- *Recombinação* – troca de parâmetros do mesmo tipo entre progenitores e descendentes. Taxa de cruzamento: 50%;
- *Mutação* - por reinicialização aleatória de parâmetros com uma taxa de mutação de 50%.

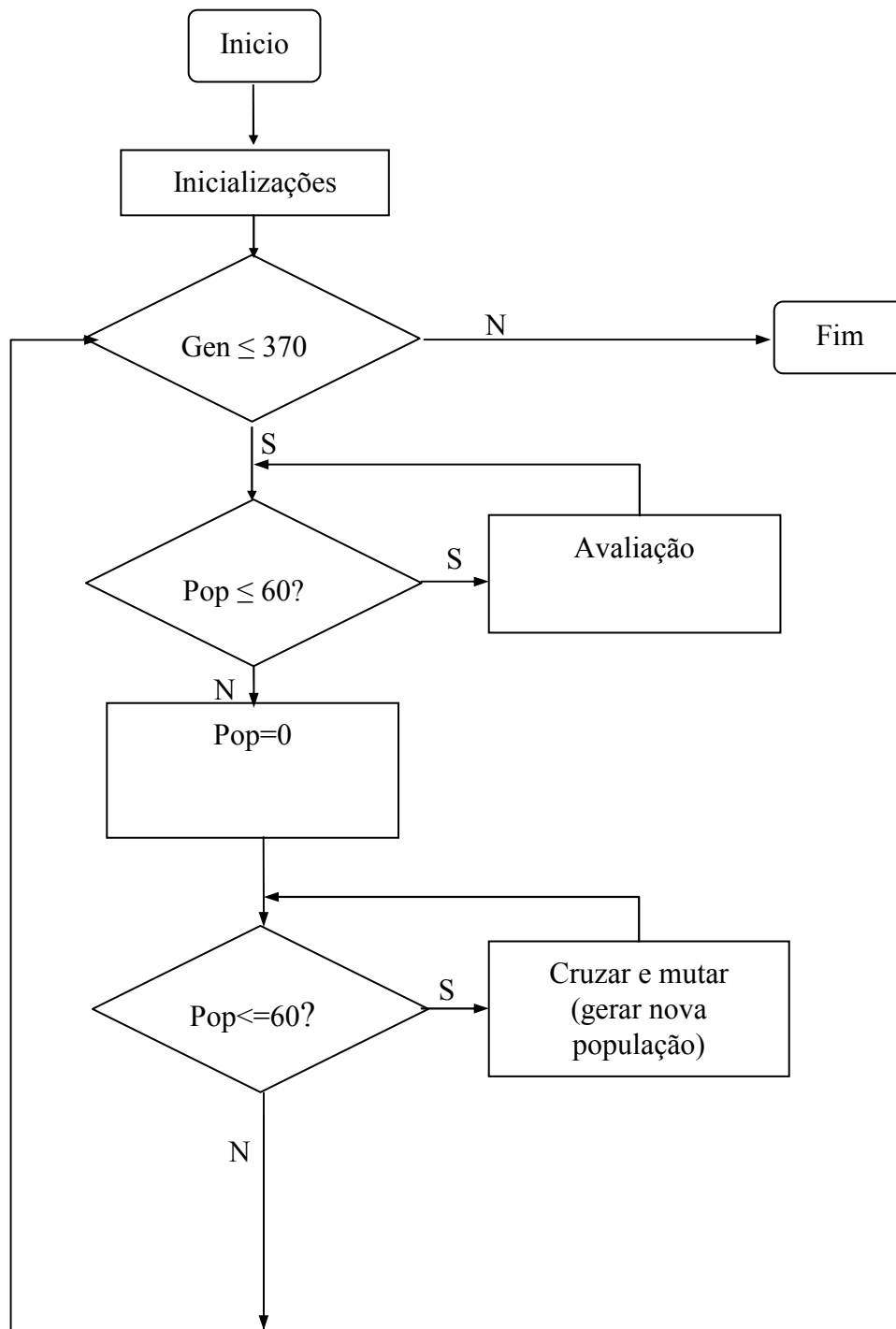


Fig. 21 Fluxograma do algoritmo genético implementado

3.5.1. População

Considerando que cada indivíduo da população deve representar uma solução do problema, teremos, neste caso, o conjunto de soluções constituído pelos modelos teóricos referidos no cap. 2, designadamente, pelas cinco funções (efeito de pepita, esférico, exponencial,

gaussiana e potência, com os parâmetros denotados por a e c e de acordo com as fórmulas constantes na secção 2.2.2 do capítulo 2). Estas funções resultaram apenas em quatro tipos de modelos possíveis (o efeito de pepita c_0 foi adicionado como constante a cada função). Assim, cada modelo é associado a um efeito de pepita (c_0) e a um par de parâmetros, constituído por um patamar ou contribuição da variância (c) e uma amplitude ou expoente (a) consoante a estrutura.

Para além do efeito de pepita (c_0), foram consideradas apenas duas estruturas e cada solução (cromossoma) foi codificada em seis parâmetros (ou genes) contendo a seguinte informação (fig. 22):

- Tipo da função ($tipoF$) que determina os quatro tipos de soluções possíveis (uma para cada tipo de modelo teórico);
- Efeito de pepita (c_0);
- Patamar ou contribuição da variância (c_1) e amplitude ou expoente (a_1) relativos à primeira estrutura;
- Patamar ou contribuição da variância (c_2) e amplitude ou expoente (a_2) relativos à segunda estrutura.

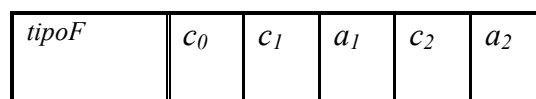


Fig. 22 Cromossoma

Os parâmetros foram codificados de acordo com as seguintes condições:

- O tipo de função é codificado por um número inteiro com as seguintes modalidades: 0-esférico, 1-exponencial, 2- gaussiano, 3-potência);
- Todos os outros parâmetros são codificados por números reais positivos.

Considerou-se a população constituída por 60 indivíduos (60 de cada tipo de função candidata a solução) depois de um estudo de sensibilidade que revelou que não há melhoria no algoritmo através de um incremento de 50 a 100% na dimensão da população.

Deve notar-se que o algoritmo anteriormente descrito será chamado duas vezes para a obtenção do modelo teórico. Da primeira vez utiliza-se o conceito de paralelismo, ou seja, serão processados paralelamente os mesmos procedimentos em 3 populações diferentes (correspondentes, respectivamente, às funções esférica, exponencial e gaussiana) que conduzem à obtenção de um dado modelo teórico. A segunda intervenção do algoritmo corresponde apenas à pesquisa numa população referente a um modelo de potência e, finalmente, decidir-se-á, automaticamente, pelo modelo que obtiver melhor ajuste.

Note-se que se optou por duas invocações do algoritmo genético por se considerar que a pesquisa do modelo teórico obedece a critérios diferentes. Assim, no primeiro caso a variância condicionará os valores do efeito de pepita e do patamar e a função apenas será ajustada aos pontos situados abaixo ou sobre a variância.

3.5.2. Critério de paragem

Um bom critério de paragem seria a detecção de uma solução de “qualidade” o que, neste caso concreto, exigiria uma bateria computacionalmente ineficiente de testes sistemáticos de validação cruzada sobre o modelo ajustado. Assim, optou-se por considerar como critério de paragem um número de gerações (370), a partir da qual não se detectou melhoria significativa do erro calculado.

3.5.3. Avaliação

A função de avaliação ($fAval$) que quantifica a aptidão de cada indivíduo $F(x)$ é a seguinte:

$$fAval(F(x)) = \frac{1}{\sum_i (f(x_i) - F(x_i))^2 + 1} \quad \text{onde } P_i(x_i, y_i = f(x_i)) \text{ são os pontos do(s)}$$

variograma(s) experimental(ais) calculado(s). Quando o algoritmo genético é aplicado pela

primeira vez apenas são considerados os pontos do(s) variograma(s) experimental(ais) que se situem abaixo da variância.

Quando $\sum_i (f(x_i) - F(x_i))^2$ é muito grande, esta função tenderá para zero e, no caso ideal, ou seja, quando os valores estimados coincidirem com os experimentais, tenderá para um.

3.5.4. Inicialização

A população de partida é escolhida aleatoriamente, ou seja, os genes em cada cromossoma são aleatoriamente inicializados, sendo os seus valores condicionados por algumas restrições óbvias (por exemplo, c e c_0 não podem ser superiores à variância).

3.5.5. Selecção

A selecção retirará de uma geração o conjunto dos progenitores que será constituído por tantos quantos os descendentes que vão ser gerados.

Ao conjunto de progenitores pertencerá:

- O melhor indivíduo dessa geração (apenas será progenitor na população a que pertence!) (conceito de elitismo);
- Todos os indivíduos escolhidos aleatoriamente em função de valores de selecção.

3.5.6. Recombinação ou cruzamento

A recombinação é efectuada aos pares e consiste na troca de valores entre parâmetros de progenitores e de descendentes usando uma taxa de cruzamento de 50%.

3.5.7. Mutação

Após a recombinação, a população é sujeita ao processo de mutação, ao qual, no caso concreto deste sistema, se atribuiu a uma taxa de 50%, tendo-se ainda optado por realizar a mutação reiniciando aleatoriamente os parâmetros.

3.5.8. Exemplo de aplicação e contraste com o sistema tradicional manual

Foram efectuados vários testes ao algoritmo apresentando-se, de seguida, um exemplo que ilustra o seu comportamento no ajuste de um modelo teórico a um variograma experimental.

Para avaliar a qualidade do modelo de ajustamento automático de variogramas consideraram-se por simulação geostatística três conjuntos sintéticos de 100 amostras cada, a que correspondem leis de distribuição uniforme, normal e lognormal. As amostras estão georeferenciadas, aleatoriamente, numa área de 100x100 unidades de distância, exibindo continuidade espacial bem definida e quantificada a cerca de 20% da dimensão da área estudada. Os histogramas destas leis de distribuição apresentam-se nas figuras 23, 24 e 25.

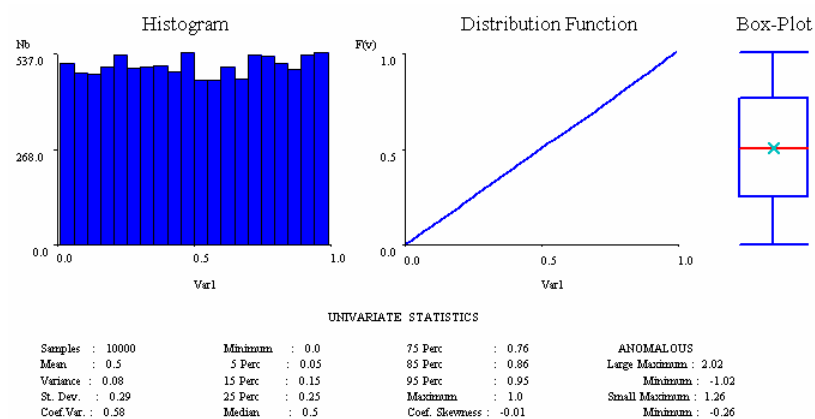


Fig. 23 Histograma da população de 100 elementos com lei uniforme

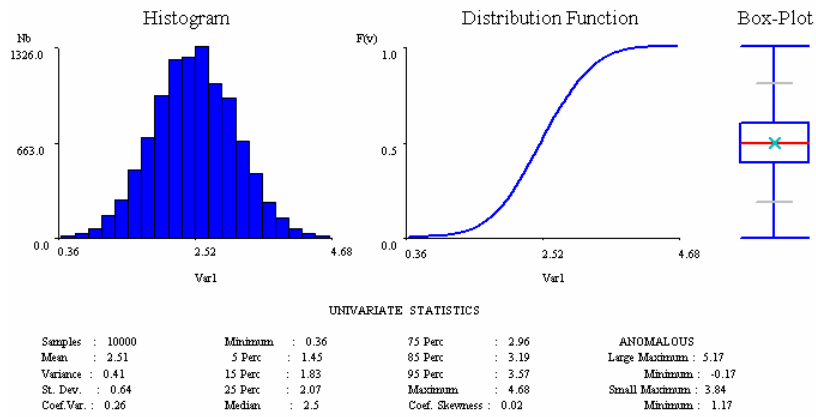


Fig. 24 Histograma da população de 100 elementos com lei gaussiana

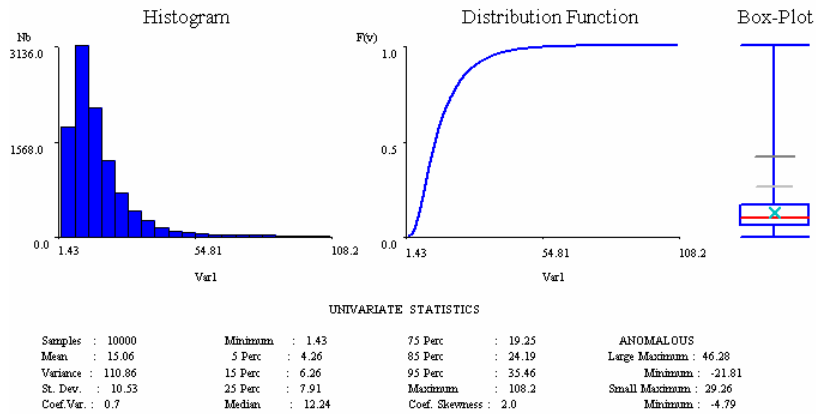


Fig. 25 Histograma da população de 100 elementos com lei lognormal

Foram ajustados manualmente modelos teóricos de variograma, cujos resultados se apresentam seguidamente (figuras 26, 27 e 28). Foram sempre considerados variogramas omnidirecionais porque as amostras foram geradas pressupondo um fenómeno anisótropo.

Modelo = SPH (C = 0,065; a = 18 m)

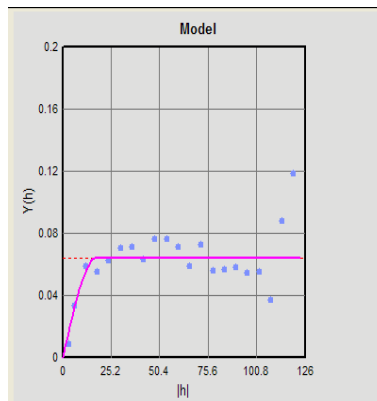


Fig. 26 Variograma experimental e modelo teórico ajustado à população com lei uniforme

Modelo = SPH (C = 0,28; a = 20 m)

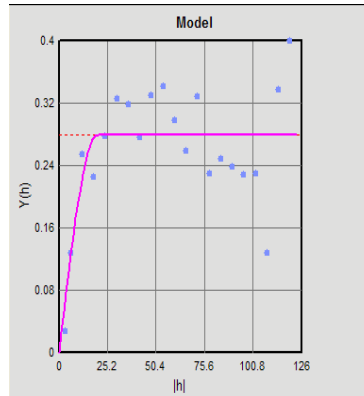


Fig. 27 Variograma experimental e modelo teórico ajustado à população com lei gaussiana

Modelo = SPH (C = 84.82; a = 22 m)

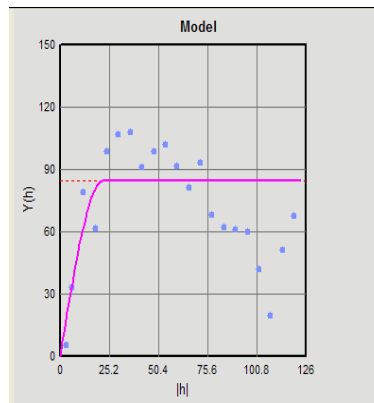


Fig. 28 Variograma experimental e modelo teórico ajustado à população com lei lognormal

Seguidamente procedeu-se a um conjunto de testes de validação cruzada (tabelas 1 a 3) para cada lei de distribuição, considerando o estimador de *krigagem* normal e mantendo sempre a vizinhança de *krigagem* em 10 amostras.

Nos testes de validação cruzada, fez-se sempre a estimação com efeito de pepita puro para obter um valor de referência para o erro que seria obtido, por exemplo, com um estimador de médias móveis. Testaram-se várias amplitudes (sempre para o tipo modelo escolhido no ajustamento) e para a amplitude onde se obtiveram os melhores resultados testaram-se efeitos de pepita com 10, 20 e 30%.

| Modelo variograma | m_z | m_{z^*} | EM | EQM |
|---|----------------|----------------|-----------------|----------------|
| Efeito de pepita puro | 0,54010 | 0,53578 | 0,00432 | 0,05561 |
| SPH ($C_0=0\%$; $a=5$) | 0,54010 | 0,53664 | 0,00346 | 0,04820 |
| SPH ($C_0=0\%$; $a=7,5$) | 0,54010 | 0,53910 | 0,00100 | 0,04375 |
| SPH ($C_0=0\%$; $a=8,2$) | 0,54010 | 0,54013 | -0,00003 | 0,04236 |
| SPH ($C_0=0\%$; $a=10$) | 0,54010 | 0,54178 | -0,00168 | 0,03945 |
| SPH ($C_0=0\%$; $a=12,5$) | 0,54010 | 0,54301 | -0,00291 | 0,03749 |
| SPH ($C_0=0\%$; $a=15$) | 0,54010 | 0,54451 | -0,00441 | 0,03720 |
| SPH ($C_0=0\%$; $a=18$)⁶ | 0,54010 | 0,54531 | -0,00521 | 0,03748 |
| SPH ($C_0=0\%$; $a=20$) | 0,54010 | 0,54527 | -0,00517 | 0,03787 |
| SPH ($C_0=0\%$; $a=22$) | 0,54010 | 0,54479 | -0,00469 | 0,03841 |
| SPH ($C_0=0\%$; $a=24$) | 0,54010 | 0,54463 | -0,00453 | 0,03899 |
| SPH ($C_0=0\%$; $a=30$) | 0,54010 | 0,54425 | -0,00415 | 0,04005 |
| SPH ($C_0=0\%$; $a=40$) | 0,54010 | 0,54441 | -0,00431 | 0,04035 |
| SPH ($C_0=10\%$; $a=8,2$) | 0,54010 | 0,53954 | 0,00056 | 0,04301 |
| SPH ($C_0=20\%$; $a=8,2$) | 0,54010 | 0,53913 | 0,00097 | 0,04377 |
| SPH ($C_0=30\%$; $a=8,2$) | 0,54010 | 0,53880 | 0,00030 | 0,04462 |
| | | | | |
| Modelo automático (1): POW ($C_0=0,001477$; $s=0,02944$; $p=0,18241$) | 0.54010 | 0.53756 | 0.00254 | 0.04629 |
| Modelo automático (2): SPH ($C_0=0,00340$; $C_1=0,04832$; $a=17,6$) | 0.54010 | 0.54434 | -0.00424 | 0.03716 |

Tabela 1. Resultados da validação cruzada para vários parâmetros do variograma e população com lei uniforme

⁶ Parâmetros do modelo ajustado manualmente.

| Modelo variograma | m_z | m_{z^*} | EM | EQM |
|--|----------------|----------------|-----------------|----------------|
| Efeito de pepita puro | 2,58077 | 2,57135 | 0,00942 | 0,24080 |
| SPH ($C_0 = 0\%$; $a = 5$) | 2,58077 | 2,57336 | 0,00741 | 0,21447 |
| SPH ($C_0 = 0\%$; $a = 7,5$) | 2,58077 | 2,57787 | 0,00290 | 0,19374 |
| SPH ($C_0 = 0\%$; $a = 8,5$) | 2,58077 | 2,58085 | -0,00008 | 0,18262 |
| SPH ($C_0 = 0\%$; $a = 10$) | 2,58077 | 2,58335 | -0,00258 | 0,16834 |
| SPH ($C_0 = 0\%$; $a = 12,5$) | 2,58077 | 2,58570 | -0,00493 | 0,15415 |
| SPH ($C_0 = 0\%$; $a = 15$) | 2,58077 | 2,58852 | -0,00775 | 0,14842 |
| SPH ($C_0 = 0\%$; $a = 17$) | 2,58077 | 2,59001 | -0,00924 | 0,14641 |
| SPH ($C_0 = 0\%$; $a = 20$)⁷ | 2,58077 | 2,59030 | -0,00953 | 0,14828 |
| SPH ($C_0 = 0\%$; $a = 22$) | 2,58077 | 2,58936 | -0,00859 | 0,15101 |
| SPH ($C_0 = 0\%$; $a = 24$) | 2,58077 | 2,58858 | -0,00781 | 0,15334 |
| SPH ($C_0 = 0\%$; $a = 30$) | 2,58077 | 2,58676 | -0,00599 | 0,15734 |
| SPH ($C_0 = 0\%$; $a = 40$) | 2,58077 | 2,58690 | -0,00613 | 0,15873 |
| SPH ($C_0 = 10\%$; $a = 8,5$) | 2,58077 | 2,57968 | 0,00109 | 0,18594 |
| SPH ($C_0 = 20\%$; $a = 8,5$) | 2,58077 | 2,57881 | 0,00196 | 0,18960 |
| SPH ($C_0 = 30\%$; $a = 8,5$) | 2,58077 | 2,57799 | 0,00278 | 0,19367 |
| | | | | |
| Modelo automático (1): GAU ($C_0 = 0,015956$; $C_1 = 0,20266$; $a = 28,985$) | 2,58077 | 2,59548 | -0.01471 | 0.13814 |
| Modelo automático (2): SPH ($C_0 = 0,015956$; $C_1 = 0,20266$; $a = 20,385$) | 2,58077 | 2.58944 | -0.00867 | 0.14522 |

Tabela 2. Resultados da validação cruzada para vários parâmetros do variograma e população com lei gaussiana

⁷ Parâmetros do modelo ajustado manualmente.

| Modelo variograma | m_z | m_{z^*} | EM | EQM |
|---|-----------------|-----------------|-----------------|------------------|
| Efeito de pepita puro | 15,26958 | 15,03919 | 0,23039 | 74,233534 |
| SPH ($C_0 = 0\%$; $a = 5$) | 15,26958 | 15,07012 | 0,19946 | 69,794235 |
| SPH ($C_0 = 0\%$; $a = 7,5$) | 15,26958 | 15,12596 | 0,14362 | 64,526086 |
| SPH ($C_0 = 0\%$; $a = 10$) | 15,26958 | 15,21462 | 0,05496 | 55,239728 |
| SPH ($C_0 = 0\%$; $a = 12,5$) | 15,26958 | 15,24294 | 0,02664 | 49,927845 |
| SPH ($C_0 = 0\%$; $a = 14,9$) | 15,26958 | 15,26965 | -0,00007 | 47,632371 |
| SPH ($C_0 = 0\%$; $a = 15$) | 15,26958 | 15,27113 | -0,00155 | 47,543863 |
| SPH ($C_0 = 0\%$; $a = 17,5$) | 15,26958 | 15,29614 | -0,02656 | 45,942889 |
| SPH ($C_0 = 0\%$; $a = 20$) | 15,26958 | 15,31022 | -0,04064 | 45,938383 |
| SPH ($C_0 = 0\%$; $a = 22$)⁸ | 15,26958 | 15,29700 | -0,02742 | 46,404061 |
| SPH ($C_0 = 0\%$; $a = 24$) | 15,26958 | 15,27745 | -0,00787 | 46,803844 |
| SPH ($C_0 = 0\%$; $a = 30$) | 15,26958 | 15,24279 | 0,02679 | 47,505769 |
| SPH ($C_0 = 0\%$; $a = 40$) | 15,26958 | 15,24341 | 0,02617 | 47,855826 |
| SPH ($C_0 = 0\%$; $a = 60$) | 15,26958 | 15,23729 | 0,03229 | 48,255001 |
| SPH ($C_0 = 10\%$; $a = 8,5$) | 15,26958 | 15,27176 | -0,00218 | 48,121715 |
| SPH ($C_0 = 20\%$; $a = 8,5$) | 15,26958 | 15,24400 | 0,02558 | 50,314570 |
| SPH ($C_0 = 30\%$; $a = 8,5$) | 15,26958 | 15,21956 | 0,05002 | 52,550688 |
| | | | | |
| Modelo automático (1): GAU ($C_0 = 9,95137$; $C_1 = 47,221$; $a = 24,2141$) | 15,26958 | 15,26667 | 0,00291 | 47,39360 |
| Modelo automático (2): GAU ($C_0 = 2,7225$; $C_1 = 57,456$; $a = 19,0645$) | 15,26958 | 15,28047 | -0,01089 | 47,65314 |

Tabela 3. Resultados da validação cruzada para vários parâmetros do variograma e população com lei lognormal

Apresentam-se, de seguida (figuras 29 a 31), tabelas de representação dos valores reais *versus* valores estimados para as populações com as diferentes leis consideradas e com os modelos de variograma referentes ao modelo de ajustamento manual, modelo a que corresponde o menor EM, modelo a que corresponde o menor EQM e o modelo ajustado pelo algoritmo

⁸ Parâmetros do modelo ajustado manualmente.

genético. Finalmente, nas figuras 32 a 34, apresentam-se os resultados do ajustamento com o algoritmo genético referentes às mesmas populações.

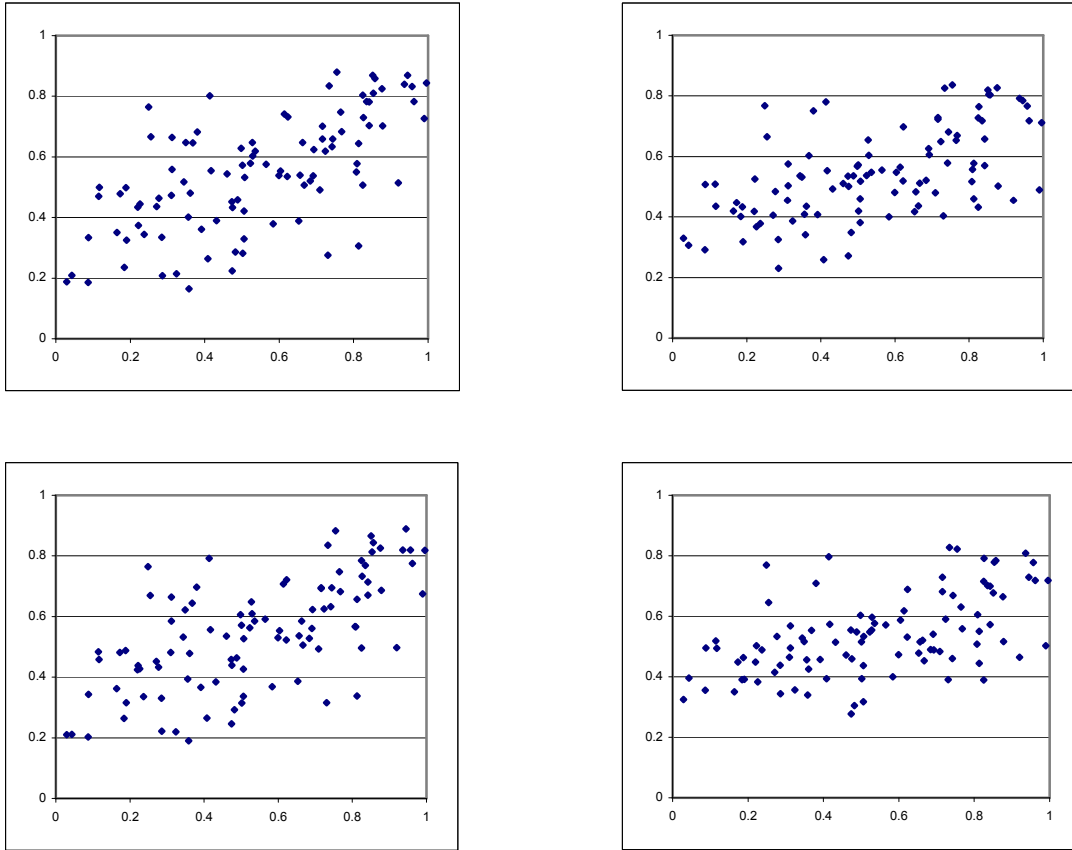
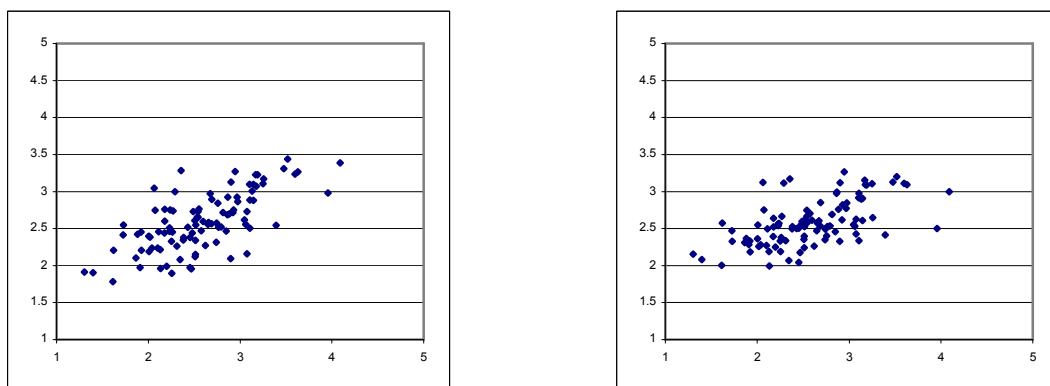


Fig. 29 População com lei uniforme: representação dos valores reais vs valores estimados para os modelos de variograma com $a = 18$ (modelo do ajustamento manual), $a = 8,2$ (modelo a que corresponde o menor EM) e $a = 15$ (modelo a que corresponde o menor EQM) e o modelo ajustado pelo algoritmo genético



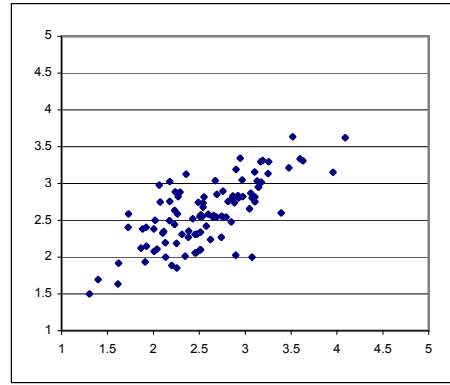
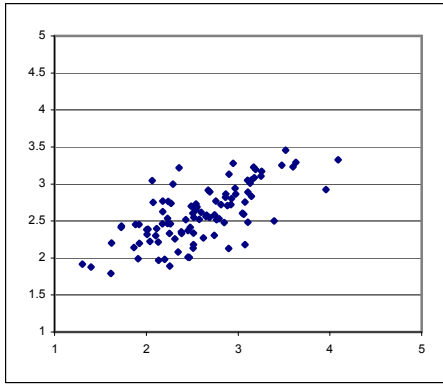


Fig. 30 População com lei gaussiana: representação dos valores reais vs valores estimados para os modelos de variograma com $a = 20$ (modelo do ajustamento manual), $a = 8,5$ (modelo a que corresponde o menor EM) e $a = 17$ (modelo a que corresponde o menor EQM) e o modelo ajustado pelo algoritmo genético

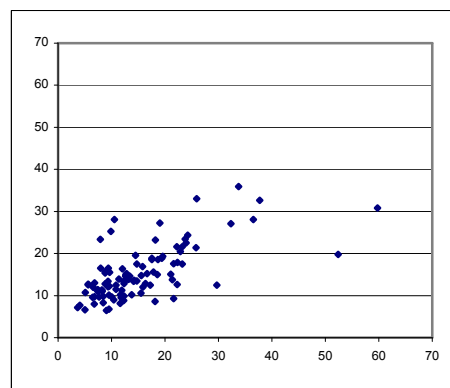
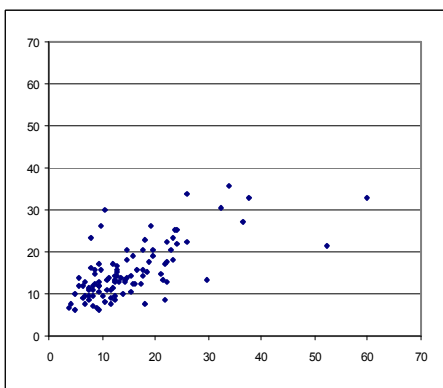
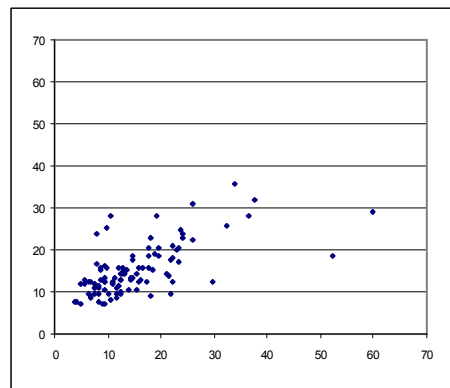
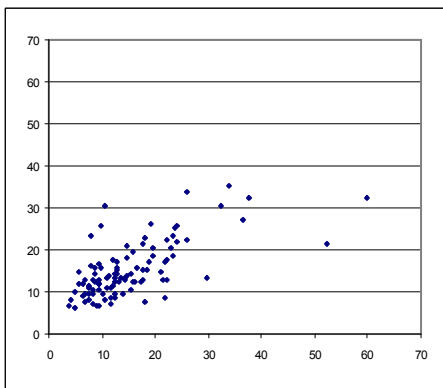


Fig. 31 População com lei logonormal: representação dos valores reais vs valores estimados para os modelos de variograma com $a = 22$ (modelo do ajustamento manual), $a = 14,9$ (modelo a que corresponde o menor EM) e $a = 20$ (modelo a que corresponde o menor EQM) e o modelo ajustado pelo algoritmo genético

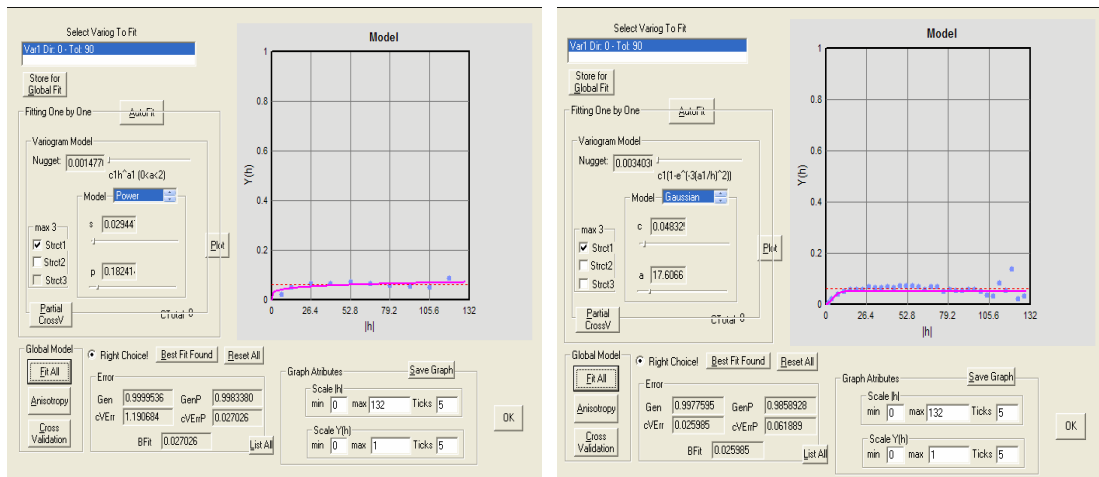


Fig. 32 Resultados do ajustamento com o algoritmo genético para a população com lei uniforme (esquerda: 10 lags; direita 33 lags)

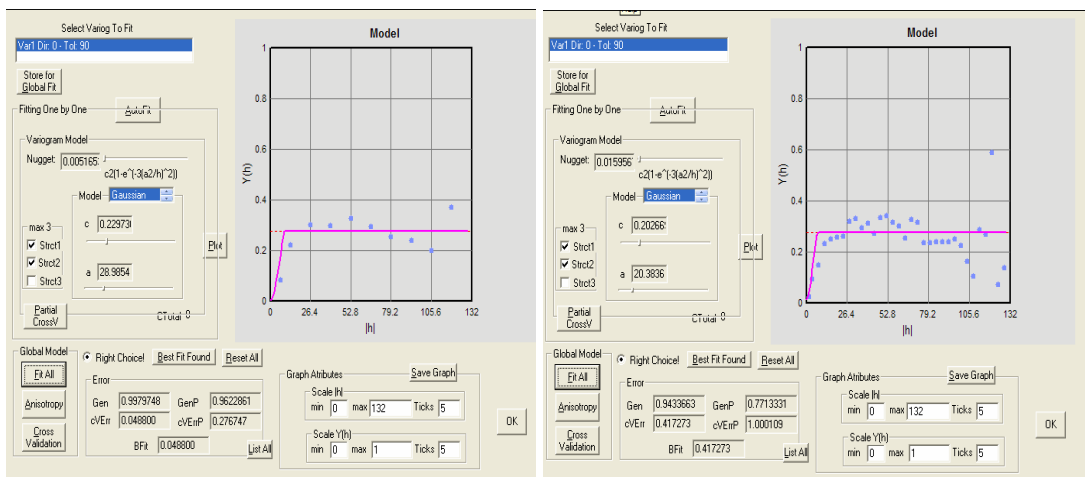


Fig. 33 Resultados do ajustamento com o algoritmo genético para a população com lei gaussiana (esquerda: 10 lags; direita 33 lags)

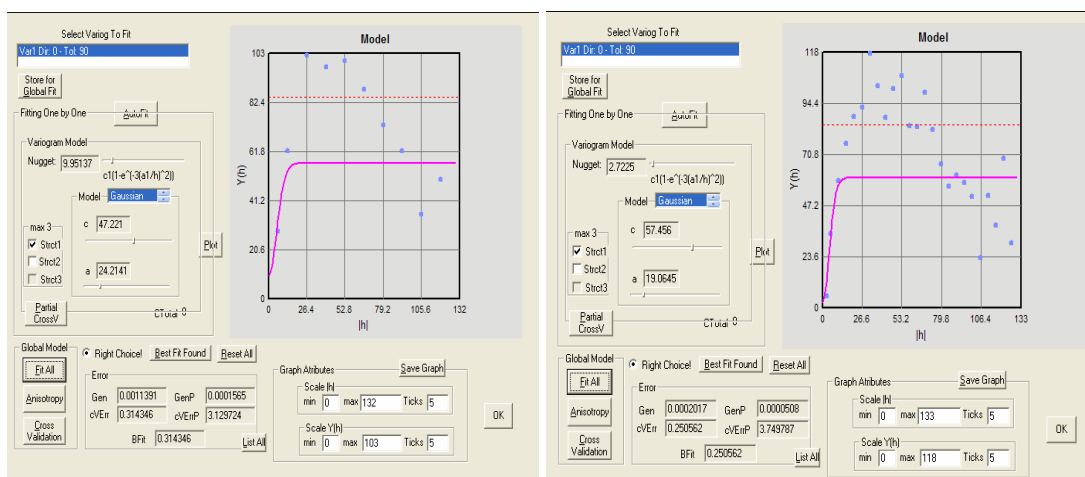


Fig. 34 Resultados do ajustamento com o algoritmo genético para a população com lei lognormal (esquerda: 10 lags; direita 33 lags)

3.5.9. Estudo comparativo com outros *softwares* disponíveis

Para avaliar as potencialidades deste *software*, em face dos programas disponíveis de maior utilização nesta área, efectuou-se um estudo comparativo, cujas conclusões se reportam em seguida.

Para a escolha dos *softwares* de comparação usaram-se os seguintes critérios:

- a) *Softwares* para ambiente Windows® (podendo eventualmente correr em ambiente MSDOS®, pois este existe subjacente ao Windows) pois, este é o ambiente usado na construção do sistema desenvolvido e é também o ambiente mais utilizado pela comunidade científica desta área.
- b) *Softwares* que contenham ferramentas para o estudo da variografia;
- c) *Softwares* de uso gratuito ou comercial de ampla utilização nesta área.

Assim, com base nas premissas anteriores, escolheram-se, dentre os gratuitos de utilização livre, os seguintes *softwares*:

- VARWIN ([PA 96]) bastante referenciado neste trabalho por se pensar ser um padrão no que respeita ao ambiente de janelas facilmente manuseável pelo utilizador;
- GSLIB ([DJ 98]) que, apesar de correr em ambiente MSDOS®, é ainda hoje um *software* muito utilizado e referenciado na comunidade científica. Este *software* disponibiliza, inclusive, as suas fontes, potenciando assim a construção de outros *softwares*, com ambientes mais “amigáveis” para os utilizadores (por exemplo, o WINGSLIB® [HPWGSLIB]).

Dentre os *softwares* comerciais escolheram-se

- SURFER ([HP SURF]) por ser um dos *softwares* comerciais mais usados pelo utilizador deste tipo de dados;
- GEOMS ([HP GEOMS]) por ser um *software* desenvolvido no Departamento de Minas e Georecursos do IST.

Quanto às características que se consideram inovadoras no sistema desenvolvido destacam-se:

- Processamento “directo” de variografia e estimação/simulação sobre imagens matriciais sem saídas de dados por canal ou por *bmp* estimado e/ou simulado;
- Modelação automática 2D utilizando algoritmos genéticos com teste de diferença de quadrados seguido, de teste de validação cruzada;

Assim, no que respeita ao processamento directo de variografia e estimação/simulação sobre imagens matriciais não há referência na literatura de nenhum *software* que possibilite esta funcionalidade.

Em todos os *softwares* citados, e em todos os que se consultaram através da *internet* (veja-se a listagem de *softwares* contida em [HP AIGEOS]), apenas é possível a análise de imagens se for processado cada canal do RGB, *per se*. A operação de separação dos dados do ficheiro *bmp* por canal e da integração dos resultados para reconstituição da imagem terão que ser efectuadas através de *softwares* específicos para processamento de imagem.

Esta limitação resultará, naturalmente, numa maior complexidade de processos para o utilizador e numa maior ineficiência do processo global pois, para além dos erros humanos inerentes a processos repetitivos, o peso dos dados obrigará a efectuar três processamentos independentes (se admitirmos canais RGB) para a variografia (a variografia para os três canais tenderá, na maior parte das situações, a ser calculada para as mesmas direcções e os mesmos passos/número de passos).

Quanto à modelação automática apenas o SURFER a permite usando, tal como em outros *softwares* que efectuem modelação automática encontrados na *internet* (por exemplo, GSOFFICE ([HP GSOF]), métodos de optimização determinísticos com erros baseados no critério de Cressie.

Os resultados obtidos na modelação automática efectuada pelo *Surfer* são muito incipientes (tal como, é admitido na sua documentação⁹) e a modelação automática (modelação 2D, tal

⁹ “... Please note that the *AutoFit* tool is meant to hone and polish a set of user specified parameters. ***AutoFit*** does not select the appropriate model form. It should not be used as an exploratory tool in the search for an appropriate model. It merely fits the parameters **once you select a reasonable model**. ...” in *Surfer Help*

como no sistema desenvolvido nesta dissertação) não selecciona o modelo mais apropriado, apenas ajusta os parâmetros após selecção prévia do modelo apropriado.

Quanto aos grandes volumes de dados a processar e à integração dos diferentes módulos num único *software* com as vantagens que daí advêm:

- o VARWIN apenas permite o estudo da variografia e a modelação manual e suporta no máximo 500.000 pares [PA 96] (o volume de dados mínimo em imagens é da ordem dos 100x100 e a formação de pares revela-se bastante ineficiente para os dados em malha regular);
- o GSLIB é bastante eficiente em tempos de execução mas não utilização de memória dinâmica (o que obriga a alterações nos programas fonte quando se pretende executar grandes volumes de dados);
- o Surfer processa, de forma eficiente, grandes volumes de dados mas apenas possui módulos integrados para a *krigagem*;
- o GEOMS permite o processamento de grandes volumes de dados e, tal como o GSLIB, possui módulos integrados de estimação e simulação, mas não realiza modelação automática.

Conclui-se assim que as grandes diferenças entre os *softwares* citados e o *software* desenvolvido centram-se nas funções existentes no *MODIMAV* que possibilitam o processamento conjunto dos canais de uma imagem aceite como entrada e a modelação automática através de algoritmos genéticos dos variogramas experimentais.

4. Descrição do Sistema e Manual do Utilizador

Neste capítulo detalham-se as principais funcionalidades do sistema desenvolvido e, de seguida, apresenta-se o manual do utilizador, descrevendo as entradas e saídas de dados mostrando alguns exemplos típicos de ficheiros de dados de entrada que ilustram as duas estruturas de dados, referentes a amostragem regular ou irregular, que podem ser processados em termos de variografia e modelação, por estimação e/ou simulação. Descrevem-se ainda, na óptica do utilizador, os ecrãs referentes ao processamento dos dados no que se refere à variografia e modelação, estimação e/ou simulação.

Finalmente, descreve-se o sistema na óptica da implementação, apresentando-se o ambiente em que foi concebido, o diagrama de classes obtido através do *Rational Rose Enterprise Edition* (*Rational Software Corporation*®), detalham-se alguns aspectos considerados relevantes nas diferentes funcionalidades implementadas, tipificam-se as estruturas de maior peso utilizadas para armazenamento de dados em memória e os ficheiros de ligação entre módulos executáveis e referenciam-se as bibliotecas e os módulos que foram importados e utilizados no seu desenvolvimento.

4.1. Síntese das principais funcionalidades

Em termos genéricos a sequência de tarefas para processamento de dados no sistema *MODIMAV* é ilustrada na figura seguinte (fig. 35):

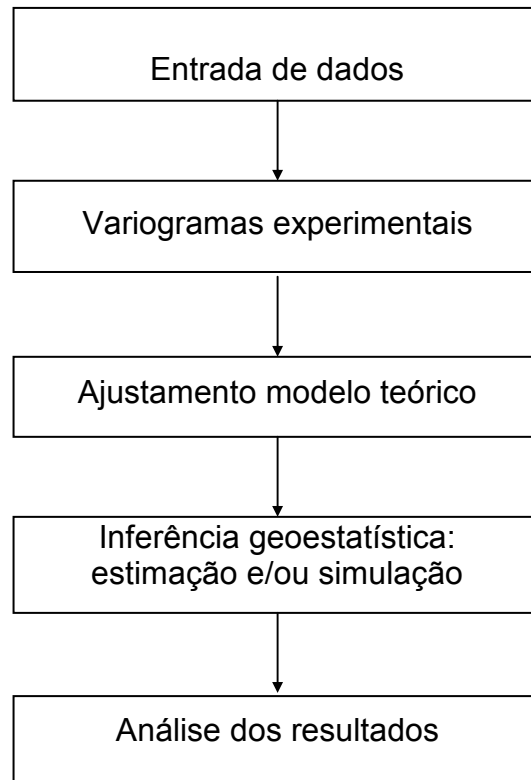


Fig. 35 Diagrama de funções

I) ENTRADA DE DADOS

- No início de cada projecto é, em primeiro lugar, seleccionado o ficheiro de dados a processar que contém a(s) variável(is). Seguidamente é identificado o tipo de dados (imagem, plano de amostragem de dados segundo malha regular ou irregular dispersos) e as características da variável em estudo (nome, limites, etc.). No final desta fase o sistema está apto para fazer o processamento pretendido.
- Observação de dados: visualização de diagramas de superfície e de dispersão dos dados/variável em estudo.

II) VARIOGRAMAS EXPERIMENTAIS

- Pré-cálculo dos variogramas experimentais da variável em estudo segundo várias direcções. Se o ficheiro for referido a dados dispersos, são calculadas a distância e a

orientação entre todos os pares de pontos para permitir a posterior visualização segundo vários agrupamentos por classes. No caso do ficheiro ser uma imagem, esta será aproveitada, tal como no GSLIB [DJ 98], para o cálculo do variograma experimental nas direcções principais da rede regular de pontos (assumindo-se que a imagem está correctamente orientada, NS e EW) e nas direcções diagonais.

- Cálculo das seguintes medidas experimentais de continuidade espacial: variograma, variograma normalizado, covariância e madograma para as direcções e classes de distância pretendidas. A escolha do passo “mais adequado” para as classes de distância é auxiliada pela representação inicial de todos os pontos do variograma experimental (não considerando classes de distância).
- No caso dos dados referirem-se a uma imagem com várias bandas ou canais de cor (por exemplo, do tipo R-red, G-green, B-blue) serão calculados, simultaneamente, os variogramas para todas as bandas ou canais de cor e ajustados independentemente;
- Visualização dos variogramas experimentais sob a forma de um gráfico direccional, de superfície ou de dispersão.

III) AJUSTAMENTO MODELO TEÓRICO

- Numa primeira fase é proposto o ajustamento automático de um modelo teórico global para todas as direcções. Este modelo pode ser validado com testes de validação cruzada e sofrer pequenos ajustes manualmente. Seguidamente, com base no modelo global é ajustado, automaticamente, um modelo teórico para cada direcção em que foi calculado o variograma experimental, que também fica sujeito a testes de validação cruzada e que também poderá ser “afinado” manualmente.
- Em todos os casos de ajustamento, é detectada, automaticamente, a anisotropia geométrica para todos os casos em que a maior amplitude seja superior em 50% da menor;
- O modelo teórico final é memorizado para permitir a ligação com as opções de estimação e/ou simulação.

IV) INFERÊNCIA GEOESTATÍSTICA: ESTIMAÇÃO E/OU SIMULAÇÃO

No que respeita a resultados baseados no modelo ajustado na etapa anterior, o sistema desenvolvido permite utilizar, com utilização do modelo teórico do variograma e anisotropia, os seguintes modelos de inferência geoestatística:

- *Krigagem normal* – Obtém-se uma nova imagem estimada respeitante a cada variável trabalhada. Por exemplo, no caso do estudo da variável cor a partir de uma imagem inicial, obtém-se uma nova imagem estimada referente a uma área adjacente não amostrada ou do mesmo local mas com outra resolução espacial, respectivamente, codificada com o mesmo número de canais da imagem original.
- *Cokrigagem colocalizada (colocated cokriging)* – Para utilizar este modelo de estimação considera-se uma imagem de entrada que constitui a variável secundária e um conjunto de amostras pontuais com a variável primária. Obtém-se o coeficiente de correlação entre as duas variáveis (primária e secundária) nas localizações comuns e uma nova imagem estimada por cada variável trabalhada. Por exemplo, no caso do estudo da variável precipitação medida num número limitado de estações meteorológicas espaçadas, pode ser considerado o Modelo Digital de Terreno como informação secundária condicionante, obtendo-se uma nova imagem estimada, contínua, da variável precipitação.
- *Simulação geoestatística* – No caso da simulação obtêm-se várias realizações ou imagens equiprováveis da variável trabalhada e a variância local do conjunto das realizações, que é indicador da incerteza local.

V) VALIDAÇÃO DOS RESULTADOS

- Visualização dos resultados da estimação e/ou simulação sob a forma de mapas ou imagens.

Deste sistema realçam-se as seguintes características:

- São processados dados georeferenciados a 2/3D bem como imagens *matriciais* (2D) sem limitações de tamanho salvo as que resultem da memória disponível;
- No caso de ficheiros armazenando estruturas de dados irregulares ou regulares mas onde são fornecidas as coordenadas (ficheiros com coordenadas explícitas) podem realizar-se validações ao conjunto de dados e a variável em estudo é escolhida pelo utilizador devendo os valores referentes às coordenadas X, Y e Z constituir as duas/três primeiras colunas do ficheiro;
- Os resultados quantitativos obtidos para os diferentes estatísticos após a realização de uma operação escolhida podem ser gravados num ficheiro *.txt* (codificação ASCII);
- Todos os gráficos visualizados podem ser gravados em ficheiros *.bmp* Windows®;

No que diz respeito à construção de variogramas, deve notar-se que:

- O utilizador fornece todos os parâmetros necessários para o processamento requerido através de um menu;
- Para a construção de qualquer tipo de representação gráfica são sempre efectuados os cálculos de todos os estatísticos bem como a variância dos dados, a distância média, médias e variâncias esquerdas e direitas das variáveis (secção **2.2.1.** do cap. **2**);

No ajustamento do modelo teórico do variograma há que atender a:

- No modo manual, é requerido ao utilizador a introdução de todos os parâmetros necessários para a construção de modelos teóricos direccionais;
- A modelação teórica direccional é sempre sujeita a validação cruzada, podendo o utilizador escolher o modelo que obtiver menor erro após um número considerado suficiente em termos de tentativas de modelação;
- Para construir qualquer representação gráfica são sempre efectuados os cálculos de todos os estatísticos, bem como a variância dos dados, a distancia média, médias e variâncias esquerda e direita das variáveis;

- O modelo teórico global é obtido automaticamente e pode sempre ser testado através da validação cruzada;
- A anisotropia é calculada de modo automático sempre que a razão entre a maior e a menor amplitude for superior a 1.5;

Na *krigagem* simples ou *colocalizada* e na simulação, deve referir-se que:

- O utilizador dispõe de janelas de interface com o GSLIB ([DJ 98]) onde, por defeito, são preenchidos todos os campos relativos a informação já adquirida pelo sistema e, no caso da *krigagem colocalizada*, após a introdução do nome do ficheiro que contém a informação secundária, é calculada a correlação global de forma automática;
- Na simulação é possível calcular a variância das simulações efectuadas;
- É sempre possível visualizar graficamente os resultados obtidos.

4.2. Manual do utilizador

Nesta secção descrevem-se, na perspectiva do utilizador, os principais componentes deste sistema ilustrados por pequenos exemplos que demonstram as potencialidades do *software* desenvolvido.

Este programa foi desenvolvido para ambiente Windows (Microsoft®) em linguagem de programação Visual C++ (Microsoft®) na versão 6. Já foi testado nos ambientes Windows 95/98/ME/2000/NT, não tendo sido até à data detectados quaisquer problemas de memória ou gráficos, devendo apenas notar-se a degradação dos tempos de execução em computadores com memória RAM inferior a 256MB.

Os menus e submenus principais são ilustrados na figura seguinte (fig. 36). Para maior divulgação do programa, todos os menus foram escritos em língua inglesa.

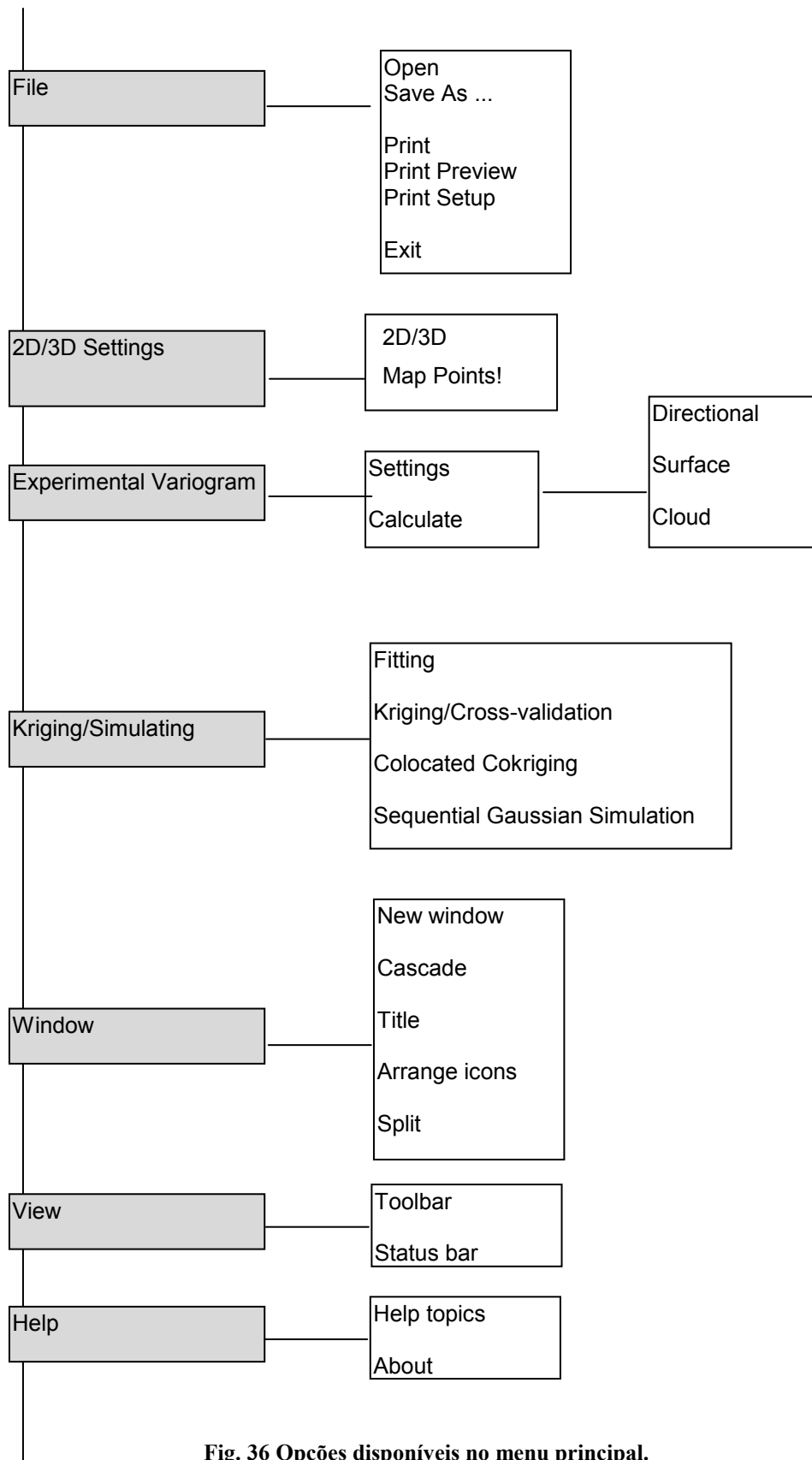


Fig. 36 Opções disponíveis no menu principal.

4.2.1. Entradas e saídas de dados

São consideradas entradas do sistema um ficheiro de dados bem como as entradas fornecidas pelo utilizador via teclado referentes a parâmetros ou opções solicitadas. Constituem saídas do sistema, as que são apresentadas no ecrã bem como as que são gravadas em ficheiros.

4.2.1.1. Ficheiros de dados admissíveis

A análise geoestatística só pode ser realizada para estruturas de dados georeferenciados. A georeferenciação pode ser explícita pelas coordenadas de todas as amostras (ficheiro onde as primeiras colunas são as coordenadas, a duas ou três dimensões) ou implícita, no caso de uma imagem com malha regular de dados.

A primeira caixa de diálogo (figura 37) permite definir as características de georeferenciação do ficheiro de dados.

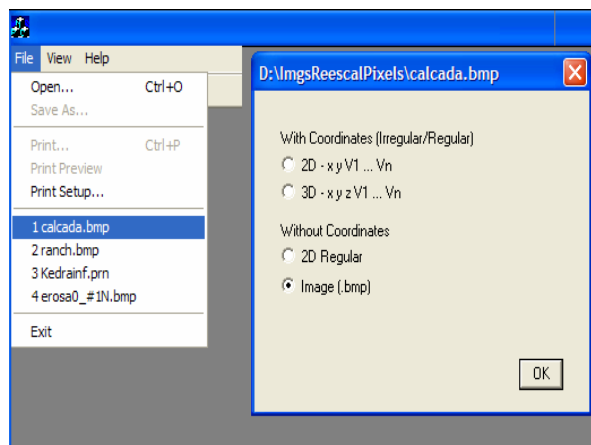


Fig. 37 Caixas de diálogo para a escolha do ficheiro

No caso das coordenadas serem definidas explicitamente em todas as linhas do ficheiro de dados, deve assinalar-se a primeira opção da caixa de diálogo (*With coordinates [Irregular/Regular]*) exibida na fig. 37. Dentro desta opção pode seleccionar-se que os dados são 2D ou 3D.

No caso do ficheiro não ter coordenadas, foram consideradas duas situações admissíveis: ficheiros de texto representando estruturas de dados espacialmente regulares (opção *2D Regular* na caixa de diálogo já referida) e ficheiros de imagens com formato *bitmap* (extensão *bmp*) a que corresponde a opção *Image* na caixa de diálogo. Este é o caso típico de ficheiros

de imagem onde apenas constam, sob forma de matriz ($m \times n$), os valores de cor dada por valores do tipo RGB (*Red, Green e Blue*).

4.2.1.2. Descrição dos ficheiros de entrada

Foram considerados três tipos de estruturas presentes em ficheiros de entrada: ficheiros que contêm dados com coordenadas explícitas ou implícitas e ficheiros que contêm imagens RGB.

A) Estruturas de dados com coordenadas explícitas

O ficheiro de entrada é composto por uma lista de colunas de dados, separados por um espaço ou tabulação, em formato ASCII; cada linha corresponde a um dado ponto coordenado ou amostra. Neste tipo de ficheiro os valores das primeiras duas ou três colunas têm de corresponder às coordenadas, respectivamente X Y ou X Y e Z (sempre por esta ordem).

No extracto de um ficheiro deste tipo, apresentado na fig. 38A, a 1ª linha (amostra) corresponde à localização X=288, Y= 311, sendo representada por três variáveis var1=850, var2=11.5, var3=18.25 .

| | | | | |
|-------|-------|------|------|-------|
| 288.0 | 311.0 | .850 | 11.5 | 18.25 |
| 285.6 | 288.0 | .630 | 8.50 | 30.25 |
| 273.6 | 269.0 | 1.02 | 7.00 | 20.00 |

Fig. 38A Malha irregular 2D

B) Estruturas de dados sem coordenadas explícitas

O ficheiro de entrada é composto por colunas dados separados por um espaço ou tabulação em formato ASCII, devendo cada linha corresponder aos valores das variáveis para um dado ponto. Neste tipo de ficheiro, subentende-se que as coordenadas serão obtidas de forma explícita, efectuando primeiramente um ciclo para X, depois para Y e finalmente para Z.

Por exemplo, considerando que a coordenada X varia entre 1 a 5 e a coordenada Y entre 1 e 2, um extracto do ficheiro (fig. 38B) conteria apenas uma coluna de dados com a variável (*Var*).

| X | Y | Var |
|---|---|-----|
| 1 | 1 | 129 |
| 2 | 1 | 131 |
| 3 | 1 | 55 |
| 4 | 1 | 56 |
| 5 | 1 | 155 |
| 1 | 2 | 167 |

Fig. 38B Malha regular 2D

C) Imagens RGB

Caso de imagens em formato *bitmap*, 24 bits para Windows® (por exemplo, imagem de calçada visualizada com o Adobe Photoshop® (fig. 39)).

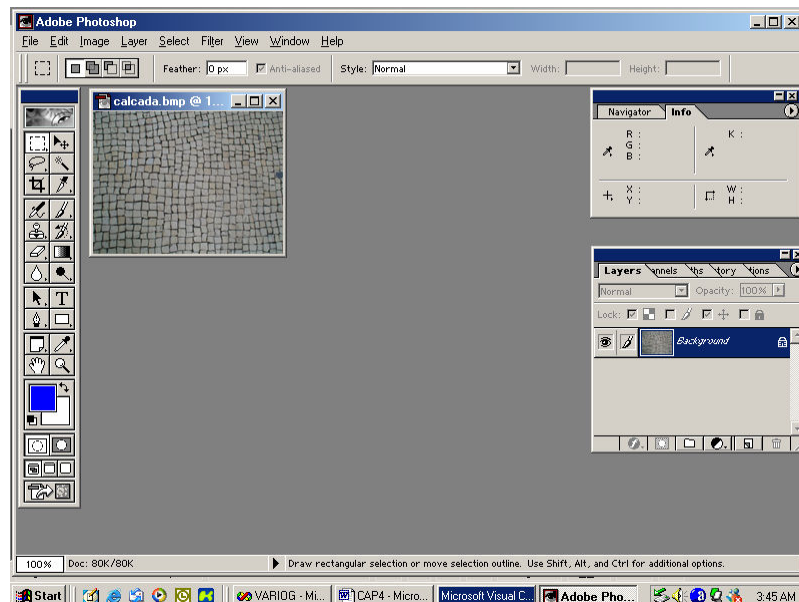


Fig. 39 Exemplo de uma imagem matricial em formato bitmap com 190x142 pixels

4.2.1.3. Ficheiros de saída

As diferentes opções disponibilizadas através de menus permitem a visualização das saídas obtidas no ecrã, bem como a sua gravação em ficheiros *bitmap*, com a extensão *bmp*.

No caso quantitativo:

- Para os quadros de síntese de estatísticos são gerados ficheiros de tipo ASCII, onde na primeira linha é identificado o diagrama onde foram obtidos e as linhas seguintes são do tipo das que se encontram nas janelas que fornecem os estatísticos obtidos;
- Para os ficheiros de resultados da *krigagem* e *cokrigagem colocalizada* estes ficheiros são saídas dos módulos do GSLIB [DJ 98] e contêm (fig. 40) na primeira linha o nome do ficheiro, na segunda o número de variáveis, de seguida o nome de cada variável, um por linha e, finalmente, linhas constituídas pelos valores de X, Y, Z, valor real, valor estimado, variância de estimação e erro. Aos valores não estimados é atribuído o decimal -999.0.

```
KT3D ESTIMATES WITH: G:\TESTES\erosaTR.dat
5
X
Y
Z
Estimate
EstimationVariance
0.000 0.000 0.000 179.103 36.196
1.000 0.000 0.000 180.601 36.195
2.000 0.000 0.000 174.000 0.000
3.000 0.000 0.000 178.000 0.000
4.000 0.000 0.000 180.000 0.000
```

Fig. 40 Extracto do ficheiro de resultados da *krigagem* (*.out)

Os ficheiros correspondentes às janelas de *krigagem/cokrigagem colocalizada* e *simulação* são gravados com o nome do ficheiro a tratar sufixado com *.par* no directório onde existe o ficheiro de dados.

Por exemplo, se o ficheiro de dados for *kedrainf.prn* existente na pasta *g:\testes* o ficheiro de parâmetros para a *krigagem* será o ficheiro *kedrainf.par* e será gravado na mesma pasta.

No caso da *krigagem* e da *simulação* a partir de ficheiros de imagem a sequência de *krigagem* para os diferentes canais redundará numa sobreposição dos ficheiros de parâmetros.

4.2.1.4. Validações e escolha de variáveis/unidades de medida no ficheiro de dados

A validação e escolha de variáveis foi descrita com base no tipo de estrutura de dados em presença. Assim distinguir-se-ão os procedimentos a adoptar para os caso de estruturas de dados com ou sem coordenadas explícitas.

A) Malha regular/irregular com coordenadas explícitas

Após a escolha do tipo de estrutura de dados, deve ser desencadeada a opção *2D/3D* do submenu *2D/3DSettings*.

No caso de ficheiros com coordenadas, as opções da caixa de diálogo são as seguintes (fig. 41):

- tipo de validação pretendida quando existem dados com as mesmas coordenadas, às designações
- unidades em que as variáveis estão expressas
- delimitação dos seus valores num intervalo fixado pelo utilizador.

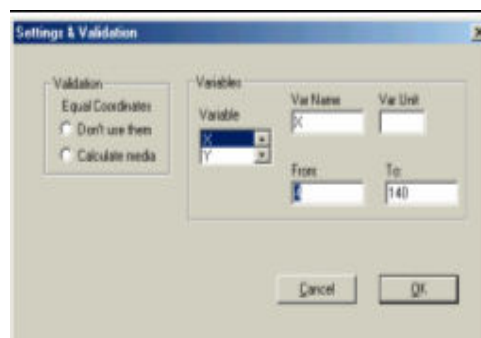


Fig. 41 Caixa de diálogo 2D/3D (malha com coordenadas explícitas)

Na sequência do processamento aparecerá no ecrã a estatística correspondente ao número de variáveis, dados e o número previsível de pares de pontos que será constituído (fig. 42).

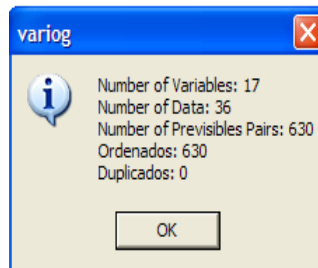


Fig. 42 Estatísticas iniciais (malha com coordenadas explícitas)

B) Estrutura de dados regular sem coordenadas explícitas

Tal como no caso anterior, após a escolha do tipo de ficheiro, deve ser desencadeada a opção *2D/3D* do submenu *2D/3DSettings* cuja caixa de diálogo (fig. 43) permite definir os limites possíveis para o valor da variável, definir o número de linhas e colunas a utilizar nos cálculos, o espaçamento entre nós (em linha e em coluna) (respectivamente *UnitSizeY* e *UnitSizeX*) bem como as designações e as unidades em que as variáveis serão expressas. Note-se que, neste caso, apenas são processados dados regulares a 2D.

C) Ficheiro imagem *bitmap* (extensão *bmp*)

Neste caso, a opção *2D/3D* do submenu *2D/3DSettings* permitirá (fig. 43), visualizar a imagem escolhida numa janela fixa, definir (através de percentagem) e visualizar a região da imagem que irá ser trabalhada, definir os limites possíveis para o valor da variável (0..255), definir o número de linhas e colunas a utilizar nos cálculos, o espaçamento entre nós (em linha e em coluna) (respectivamente *UnitSizeY* e *UnitSizeX*) bem como as designações e as unidades em que as variáveis serão expressas.

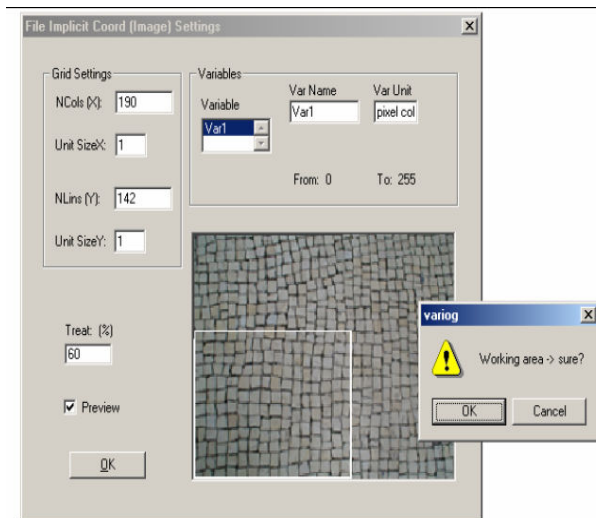


Fig. 43 Caixa de diálogo 2D/3D (caso de ficheiro de imagem ou regular 2D)

4.2.1.5. Visualização dos dados de entrada

Para todas as estruturas de dados de entrada, com ou sem coordenadas explícitas, é possível observar os dados numa visualização gráfica 2D através da opção *Map Points* do submenu (*2D/3DSettings*).

Esta representação gráfica (fig. 44), em que se optou pela utilização de um reticulado que permite a leitura mais exacta dos valores representados, na sua totalidade ou parcelarmente, poderá, eventualmente, ser utilizada para a detecção de possíveis erros em coordenadas.

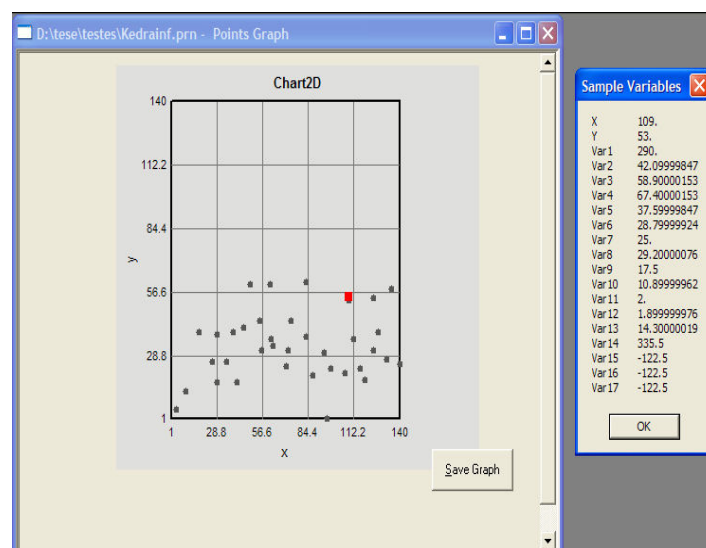


Fig. 44 Representação gráfica de uma estrutura de dados 2D com coordenadas explícitas

Se for premida a tecla esquerda do rato sobre um dos pontos representado no gráfico obter-se-á, noutra janela, as suas coordenadas e os valores das variáveis associadas (fig. 44) possibilitando ainda o armazenamento do gráfico obtido num ficheiro *bmp* através do botão *Save Graph* exibido na janela gráfica (fig. 44).

No caso de estruturas de dados 3D é possível visualizar a projecção das localizações das amostras nos planos XOY, XOZ ou YOZ de acordo com a escolha do utilizador (fig. 45).

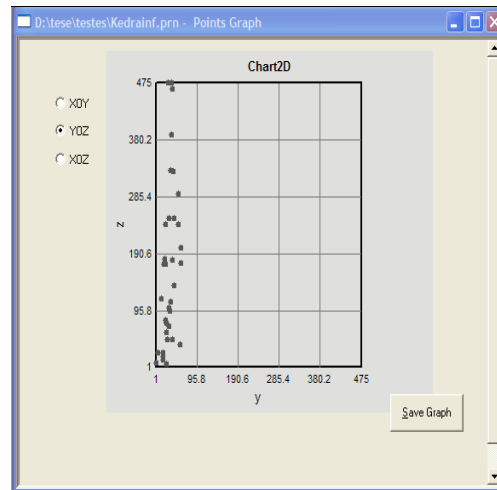


Fig. 45 Representação gráfica de uma estrutura de dados 3D no plano YOZ

No caso da ausência de coordenadas a representação gráfica não é relevante, pois a estrutura de dados regular, por questões de densidade gráfica, reduzir-se-á a um rectângulo como podemos ver na figura seguinte (fig. 46). Não obstante, para preservar o paralelismo em relação aos procedimentos relativos aos dois tipos de estruturas de dados, esta foi considerada permitindo, no caso de imagens, visualizar o valor RGB de um determinado *pixel*.

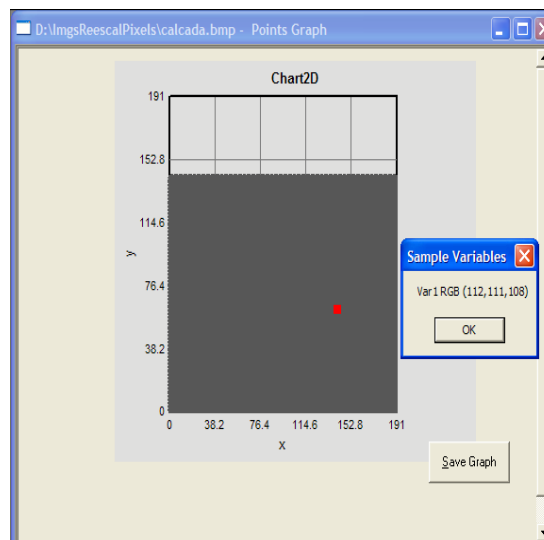


Fig. 46 Representação gráfica de uma estrutura de dados 2D com coordenadas implícitas

4.2.2. Variografia e ajustamento do modelo teórico

4.2.2.1. Análise estrutural

Para a realização da análise estrutural é seleccionada, em primeiro lugar, uma função estatística e um valor para assinalar a impossibilidade de cálculo. Estas opções são disponibilizadas na caixa de diálogo desencadeada pela opção *Settings* (fig. 47) do submenu *Experimental Variograms*.

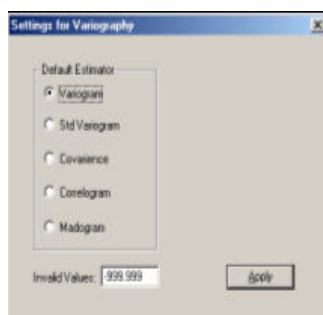


Fig. 47 Caixa de diálogo *Settings for Variography*

4.2.2.2. Gráficos de superfície

Os gráficos de superfície são seleccionados, para os dois tipos de estruturas de dados considerados, via submenus: *Experimental Variograms* -> *Calculate* cuja opção *Surface* desencadeia janelas semelhantes à representada na fig. 48.

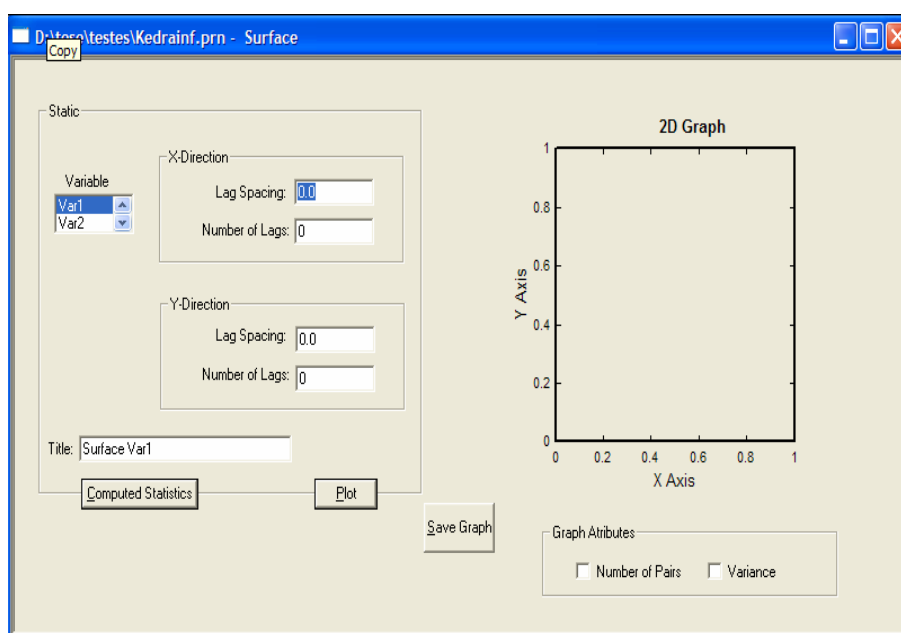


Fig. 48 Janela para estruturas de dados com coordenadas explícitas

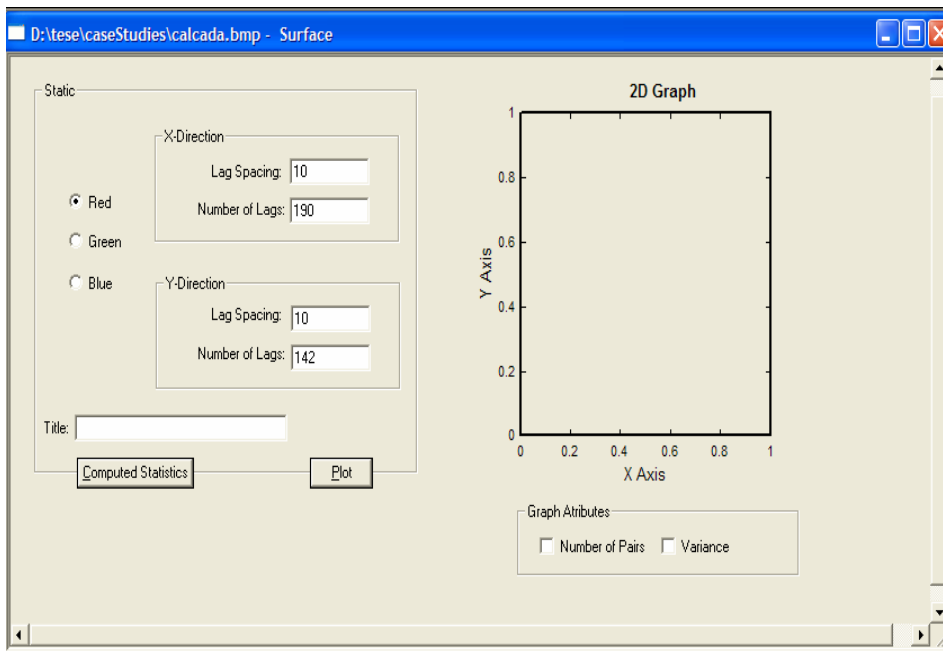


Fig. 49 Janela para estruturas de dados com coordenadas explícitas

Para calcular estes diagramas são pedidos, ao utilizador, os seguintes parâmetros (figuras 49 e 50):

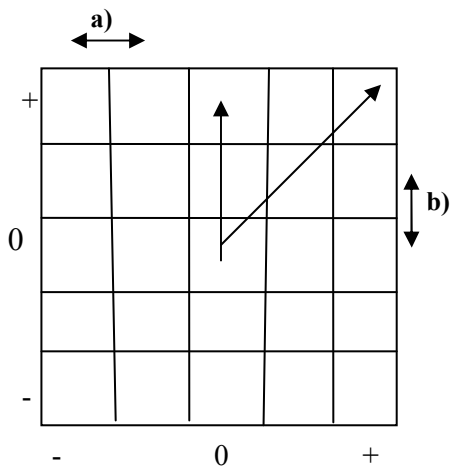


Fig. 50 Parâmetros: Superfície

- passo para X (**Fig. 50 a)**);
- passo para Y (**Fig. 50 b)**);
- número de passos em X (número de passos = 2 - **Fig. 50**);
- número de passos em Y (número de passos = 2 - **Fig. 50**).

Se premir a tecla esquerda do rato sobre o botão *Plot* desta janela será desenhado o diagrama de superfície para o caso das estruturas de dados com coordenadas (fig. 51). No caso das estruturas de dados sem coordenadas (fig. 52), o gráfico será desenhado após ter escolhido o

conjunto de dados pretendidos na visualização (relativos aos valores de R ou G ou B), podendo, em ambos os casos (figuras 51 e 52) visualizar:

- o número de pontos que foram considerados para cada passo (opção *Number of Pairs*);
- a indicação do valor da variância que é assinalada na régua de cores (opção *Variance*).

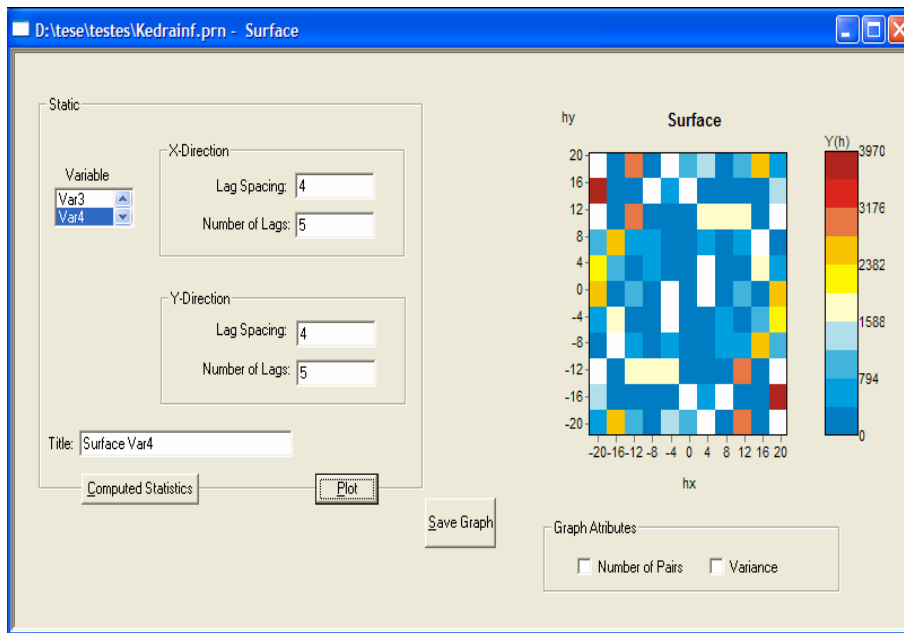


Fig. 51 Janela *Surface* com os resultados do variograma para dados com coordenadas explícitas

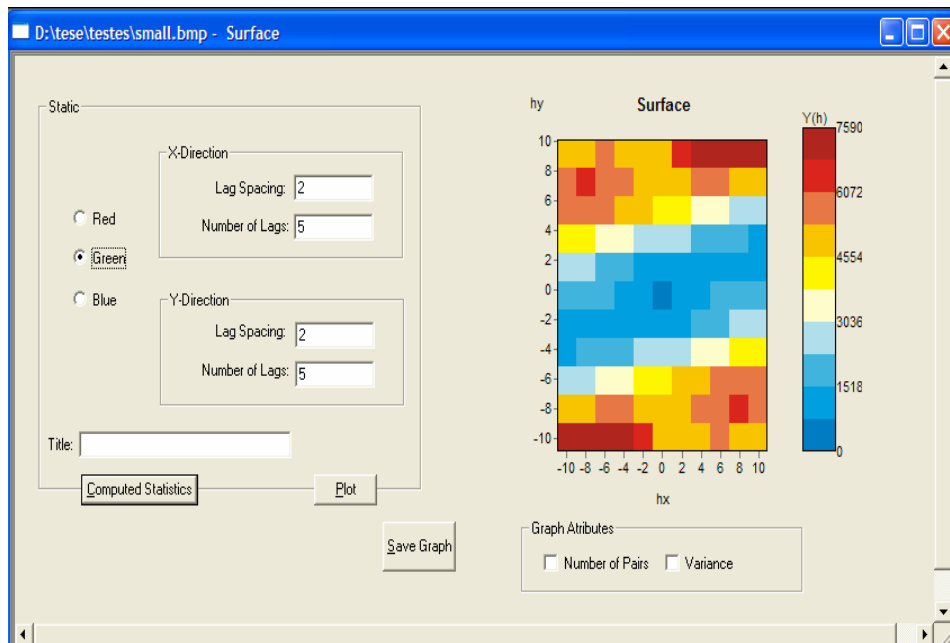


Fig. 52 Janela *Surface* com os resultados do variograma para dados sem coordenadas explícitas

O quadro resumo dos estatísticos calculados (por exemplo, fig. 53), é semelhante para os dois tipos de malha e também para o caso dos gráficos direccionais, e pode ser visualizado pressionando com o rato no botão *Computed Statistics* da janela gráfica.

D:\teste\testes\Kedrainf.prn - Surface Var4 Statistics

Surface Var4 Statistics

Number of lags along X: 5 Number of lags along Y: 5
Lag distance along X: 4 Lag distance along Y: 4
Variance: 1133.13

| DeltaX | DeltaY | NPairs | Mean[h] | Variogram | Std Variogram | Covariance | Correlogram | Madogram | Mean(-h) |
|--------|--------|--------|---------|-----------|---------------|------------|-------------|----------|----------|
| -20 | -20 | 3 | 27.378 | 596.103 | 6.7759 | 4.14007 | 0.0470601 | 14.7 | 85.4 |
| -16 | -20 | 1 | 24.8395 | 2464.02 | -999.999 | 0 | -999.999 | 35.1 | 79.1 |
| -12 | -20 | 2 | 22.2294 | 903.633 | 2.36321 | -382.375 | -1 | 21.025 | 100.25 |
| -8 | -20 | 2 | 21.5439 | 256.805 | 10.3875 | 24.7225 | 1 | 8.75 | 71.35 |
| -4 | -20 | 2 | 19.8201 | 1576.03 | 8.42062 | -187.163 | -1 | 23.1 | 78.95 |
| 0 | -20 | 1 | 21 | 1012.5 | -999.999 | 0 | -999.999 | 22.5 | 122.4 |
| 4 | -20 | 0 | N/A | N/A | N/A | N/A | N/A | N/A | N/A |
| 8 | -20 | 2 | 20.1685 | 385.09 | 916.873 | -0.420004 | -1 | 13.85 | 119.5 |
| 12 | -20 | 1 | 24.1868 | 3042 | -999.999 | 0 | -999.999 | 39 | 122.4 |
| 16 | -20 | 1 | 25.6125 | 184.32 | -999.999 | 0 | -999.999 | 9.6 | 81.8 |
| 20 | -20 | 0 | N/A | N/A | N/A | N/A | N/A | N/A | N/A |
| -20 | -16 | 1 | 24.2074 | 1573.61 | -999.999 | 0 | -999.999 | 28.05 | 149.3 |
| -16 | -16 | 1 | 20.5183 | 340.605 | -999.999 | 0 | -999.999 | 13.05 | 70.5 |
| -12 | -16 | 1 | 19.4165 | 222.605 | -999.999 | 0 | -999.999 | 10.55 | 87.3 |
| -8 | -16 | 1 | 19.2354 | 109.52 | -999.999 | 0 | -999.999 | 7.4 | 71.9 |
| -4 | -16 | 1 | 15.8114 | 42.32 | -999.999 | 0 | -999.999 | 4.6 | 119.8 |

Fig. 53 Exemplo dos resultados sob a forma de tabela das superfícies do variograma para amostras com coordenadas explícitas

Na figura seguinte (fig. 54) pode ver-se a janela resumo de estatísticos para o caso de estruturas de dados sem coordenadas específicas, devendo notar-se que, neste caso, para cada estatístico são simultaneamente calculados os valores para cada variável do terno RGB.

D:\teste\testes\small.bmp - Statistics

Statistics

Number of lags along X: 5 Number of lags along Y: 5
Lag distance along X: 2 Lag distance along Y: 2
Variance: RGB(4447.04,3026.48,2479.98)

| DeltaX | DeltaY | NPairs | Color | Mean[h] | Variogram | Std Variogram | Covariance | Correlogram | Madogram | Mean(-h) |
|--------|--------|--------|-------|---------|-----------|---------------|------------|-------------|----------|----------|
| -10 | -10 | 60 | Red | 13.6831 | 11166.8 | 14.2706 | -227.343 | -0.290533 | 70.2333 | 158.117 |
| | | | Green | 13.6831 | 7596.63 | 12.2847 | -311.168 | -0.503198 | 57.1583 | 139.367 |
| | | | Blue | 13.6831 | 6447.52 | 14.6397 | -225.33 | -0.511635 | 53.4833 | 127.7 |
| -8 | -10 | 66 | Red | 12.2965 | 10644.4 | 12.4743 | -443.051 | -0.519217 | 67.5 | 155.258 |
| | | | Green | 12.2965 | 7077.42 | 9.98944 | -410.495 | -0.579395 | 54.0682 | 135.894 |
| | | | Blue | 12.2965 | 5948.58 | 11.3458 | -267.132 | -0.509503 | 50.5455 | 124.879 |
| -6 | -10 | 72 | Red | 11.0984 | 10705.3 | 5.20098 | -926.113 | -0.449937 | 68.3819 | 157.944 |
| | | | Green | 11.0984 | 7110.44 | 4.0045 | -604.455 | -0.340421 | 55.4861 | 137.764 |
| | | | Blue | 11.0984 | 6088.38 | 4.12907 | -462.332 | -0.313548 | 52.1458 | 126.208 |
| -4 | -10 | 78 | Red | 10.156 | 10586 | 5.1479 | -1135.09 | -0.55199 | 67.3013 | 159.756 |
| | | | Green | 10.156 | 6869.62 | 3.90297 | -702.762 | -0.399274 | 54.3333 | 138.487 |
| | | | Blue | 10.156 | 5783.51 | 3.87077 | -486.817 | -0.325816 | 50.2756 | 126.756 |

Fig. 54 Exemplo dos resultados sob a forma de tabela das superfícies do variograma para amostras sem coordenadas explícitas

Em todos os tipos de gráfico o botão *Save Graph* permite o armazenamento do gráfico num ficheiro *bmp*.

4.2.2.3. Gráficos direccionais

Os gráficos direccionais são seleccionados via submenus: *Experimental Variograms* -> *Calculate* cuja opção *Directional* desencadeia as seguintes janelas:

A) Estruturas de dados com coordenadas explícitas

Coordenadas a 2D:

A janela relativa aos diagramas direccionais (dados 2D) é ilustrada na figura seguinte (fig. 55).

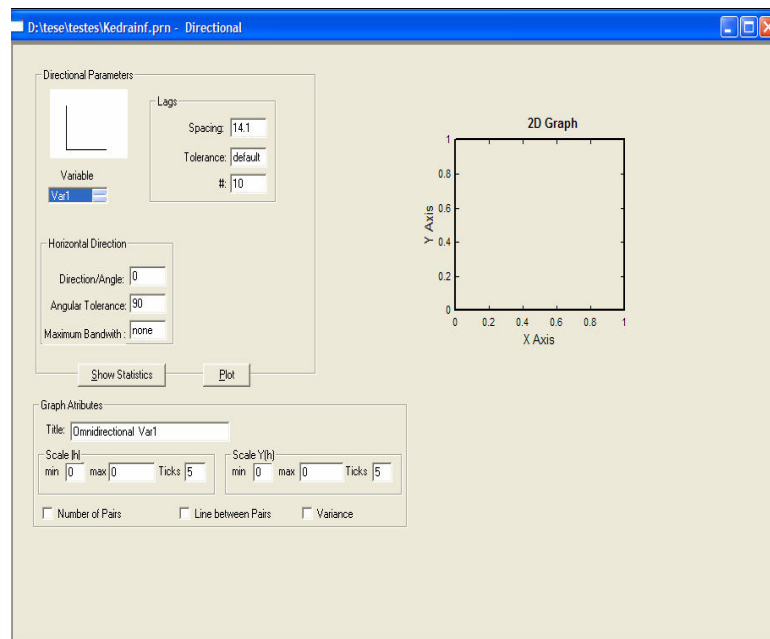


Fig. 55 Janela *Directional* (caso malha 2D com coordenadas explícitas)

Coordenadas a 3D: adiciona-se *dip*.

Na fig. 56 ilustra-se a janela relativa aos diagramas direccionais (dados 3D).

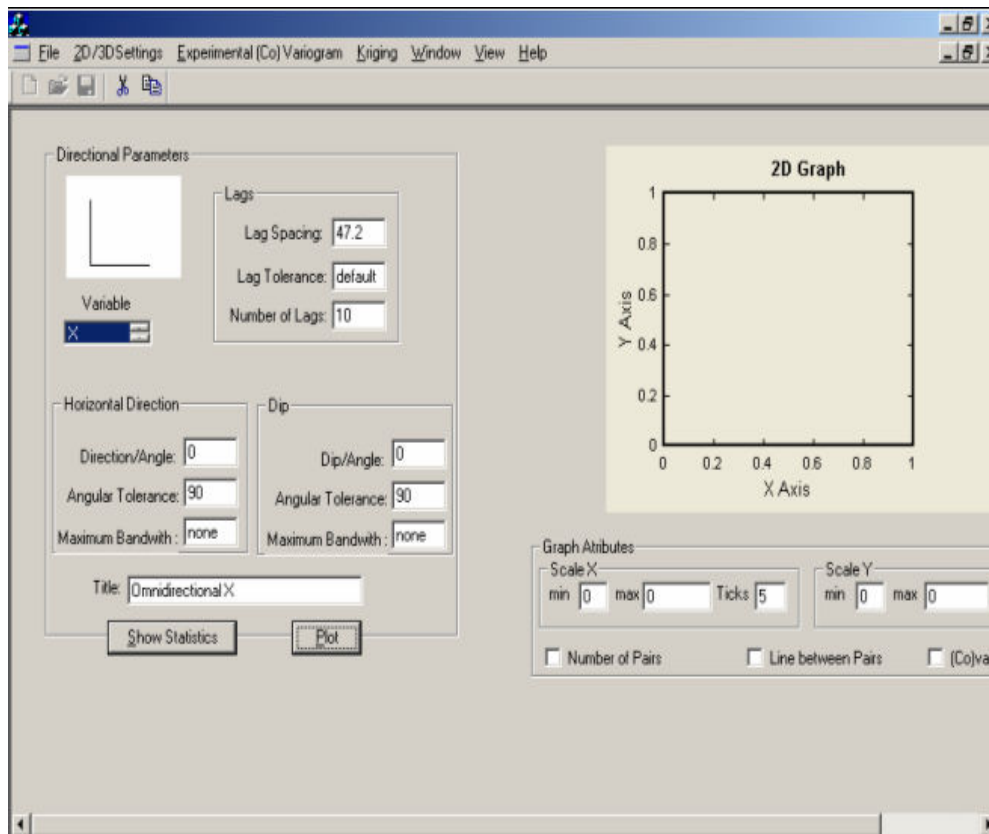


Fig. 56 Janela *Directional* (caso malha 3D sem coordenadas explícitas)

Neste caso são pedidos ao utilizador os seguintes parâmetros:

- passo (Fig. 57 - a))
- tolerância do passo (Fig. 57 - b))
- número de passos (na Fig. 57 número de passos = 4));
- direcção horizontal (Fig. 57 - α em graus) que é medida como um ângulo trigonométrico (origem em Este e em sentido contrário aos ponteiros do relógio) correspondendo à direcção horizontal no caso 3D e a direcção vertical é medida como um ângulo trigonométrico (origem em Norte e no sentido dos ponteiros do relógio);
- tolerância angular (Fig. 57 - β em graus);
- máxima largura de banda (*bandwidth*) (Fig. 57 - c)).

Se for premido a tecla esquerda do rato sobre o botão *Plot* desta janela será desenhado o gráfico direccional onde se encontra (fig. 58):

- o número de pontos que foram considerados para cada passo (opção *Number of Pairs*);
- uma linha a unir os diferentes pontos (opção *Line between Pairs*);
- uma linha horizontal correspondente ao valor da variância dos dados (opção *Variance*).

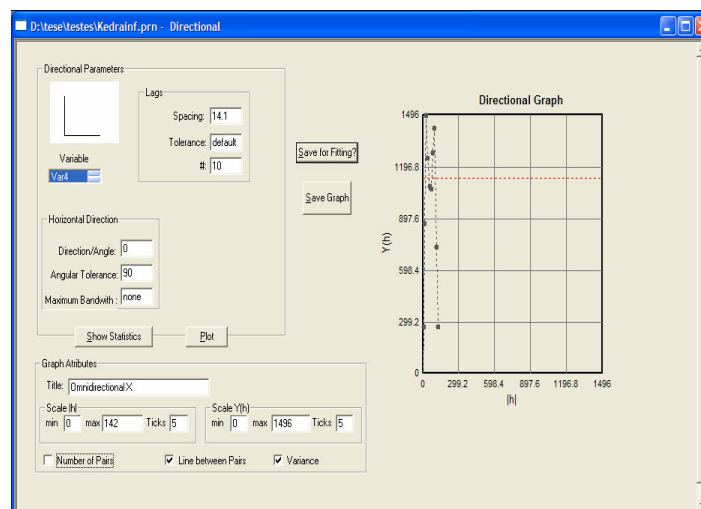


Fig. 58 Exemplo dos resultados sob a forma de gráfico direccional do variograma para amostras com coordenadas explícitas

Nesta janela é ainda possível alterar a escala de valores dos eixos, da distância e do variograma, bem como o número de subdivisões a considerar em cada eixo. Também admite-se um passo zero ao qual pertencem todos os pares cuja distância seja menor do que a tolerância do passo.

B) Estruturas de dados em malha regular ou com coordenadas explícitas

No caso de estruturas de dados sem coordenadas explícitas é aberta a janela constante da figura seguinte (fig. 59).

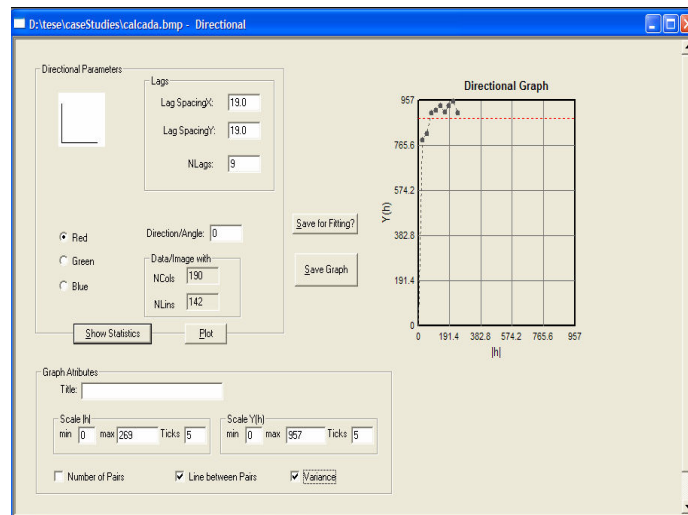


Fig. 59 Exemplo dos resultados sob a forma de gráfico direccional do variograma para amostras sem coordenadas explícitas

Neste caso são pedidos ao utilizador os seguintes parâmetros:

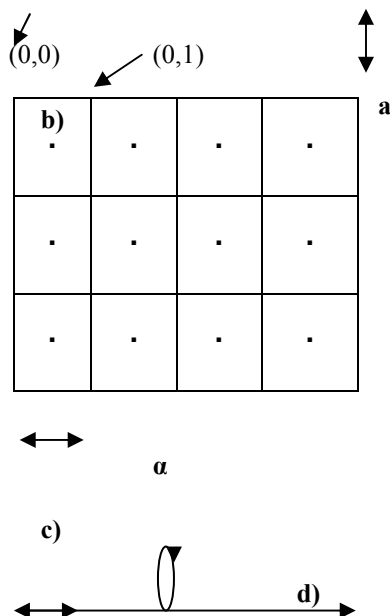


Fig. 60 Parâmetros – Direccional malha (6x5) sem coordenadas explícitas

- espaçamento para X (*Lag Spacing X* – Fig. 60)
- distância entre colunas (Fig. 60 - b)
- espaçamento para Y (*Lag Spacing Y* Fig. 60)
- distância entre linhas (Fig. 60 - a)
- direcção (Fig. 60 – α em graus correspondente à direcção d)) que é medida como um ângulo trigonométrico (origem em Este e em sentido contrário aos ponteiros do relógio). Por defeito. Neste exemplo $\alpha=0$;
- número de passos pretendidos na direcção considerada (na Fig. 60 o passo (c) é 1)

Neste exemplo, $UnitSizeX=2$ e $UnitSizeY=2$, pois cada nó é o ponto de intersecção entre uma linha e uma coluna (opções da caixa de diálogo *2D/3D* constante da Fig. 60), $lagSpacingX = 1$ e $lagSpacingY = 1$ e o par de pontos $((0,0),(0,1))$ pertence ao 1º passo (pertencem ao primeiro passo todos os pontos cuja distância pertence ao intervalo $]0, \sqrt{2}]$). Neste caso não existe passo zero pois não é considerada a tolerância do passo.

Por defeito, são assumidos os seguintes valores para os parâmetros:

- a direcção horizontal é 0° ;
- espaçamento para X é 0;
- espaçamento para Y é 0;

Se for premido a tecla esquerda do rato sobre o botão *Plot* desta janela, o gráfico será desenhado após a escolha do conjunto de dados pretendido (relativo aos valores de R ou G ou B) podendo ainda visualizar-se, tal como no caso anterior, o número de pontos que foram considerados para cada passo, uma linha a unir os diferentes pontos e/ou uma linha horizontal correspondente ao valor da variância (covariância) dos dados.

Nos dois casos, a janela resumo dos estatísticos calculados (por exemplo, fig. 61), tal como no caso dos gráficos de superfície (fig. 53), é visualizada pressionando com o rato o botão *Show Statistics* da janela gráfica *Statistics*

| Lag | NPairs | Mean h | Variogram | Std Variogram | Covariance | Correlogram | Madogram | Mean{-h} | Mean{+h} |
|-----|--------|---------|-----------|---------------|------------|-------------|----------|----------|----------|
| 0 | 6 | 5.10593 | 268.088 | 0.218905 | 956.59 | 0.781095 | 10.65 | 119.583 | 119.583 |
| 1 | 188 | 14.9918 | 869.458 | 0.80238 | 214.141 | 0.19762 | 16.1383 | 103.31 | 103.31 |
| 2 | 244 | 28.3361 | 1495.8 | 1.15028 | -195.424 | -0.150283 | 22.2119 | 98.9324 | 98.9324 |
| 3 | 232 | 42.1511 | 1247.45 | 1.16567 | -177.296 | -0.165673 | 19.3737 | 98.0461 | 98.0461 |
| 4 | 182 | 56.0959 | 1084.37 | 1.01195 | -12.8021 | -0.011947 | 17.9995 | 96.0445 | 96.0445 |
| 5 | 158 | 70.0841 | 1067.91 | 1.01074 | -11.351 | -0.0107433 | 18.3911 | 93.419 | 93.419 |
| 6 | 114 | 84.3099 | 1280.02 | 1.07765 | -92.229 | -0.0776472 | 20.0079 | 94.1202 | 94.1202 |
| 7 | 72 | 98.2716 | 1421.78 | 1.29518 | -324.033 | -0.295181 | 19.7528 | 90.4472 | 90.4472 |
| 8 | 40 | 111.423 | 730.211 | 1.48817 | -239.535 | -0.488173 | 15.9825 | 77.7775 | 77.7775 |
| 9 | 20 | 125.99 | 271 | 1.20561 | -46.2181 | -0.205613 | 8.365 | 70.855 | 70.855 |

Fig. 61 Exemplo dos resultados sob a forma de tabela do variograma para amostras com coordenadas explícitas

No caso de malha sem coordenadas explícitas, a janela resumo dos estatísticos calculados (por exemplo, fig. 62), tal como no caso dos gráficos de superfície (fig. 54) apresenta os valores calculados para os dados R, G e B.

| Lag | NPairs | Mean h | Variogram | Std Variogram | Covariance | Correlogram | Madogram | Mean{h} | Mean{h} |
|-----|--------|---------|-----------|---------------|------------|-------------|----------|---------|---------|
| 0 | 6 | 5.10593 | 268.088 | 0.218905 | 956.59 | 0.781095 | 10.65 | 119.583 | 119.583 |
| 1 | 188 | 14.9918 | 869.458 | 0.80238 | 214.141 | 0.19762 | 16.1383 | 103.31 | 103.31 |
| 2 | 244 | 28.3361 | 1495.8 | 1.15028 | -195.424 | -0.150283 | 22.2119 | 98.9324 | 98.9324 |
| 3 | 232 | 42.1511 | 1247.45 | 1.16567 | -177.296 | -0.165673 | 19.3737 | 98.0461 | 98.0461 |
| 4 | 182 | 56.0959 | 1084.37 | 1.01195 | -12.8021 | -0.011947 | 17.9995 | 96.0445 | 96.0445 |
| 5 | 158 | 70.0841 | 1067.91 | 1.01074 | -11.351 | -0.0107433 | 18.3911 | 93.419 | 93.419 |
| 6 | 114 | 84.3099 | 1280.02 | 1.07765 | -92.229 | -0.0776472 | 20.0079 | 94.1202 | 94.1202 |
| 7 | 72 | 98.2716 | 1421.78 | 1.29518 | -324.033 | -0.295181 | 19.7528 | 90.4472 | 90.4472 |
| 8 | 40 | 111.423 | 730.211 | 1.48817 | -239.535 | -0.488173 | 15.9825 | 77.7775 | 77.7775 |
| 9 | 20 | 125.99 | 271 | 1.20561 | -46.2181 | -0.205613 | 8.365 | 70.855 | 70.855 |

Fig. 62 Exemplo dos resultados sob a forma de tabela do variograma para amostras sem coordenadas explícitas

No caso da malha ter coordenadas explícitas, se for pressionada a tecla esquerda do rato sobre um ponto do gráfico, poderá visualizar-se o diagrama de dispersão correspondente à construção do ponto, permitindo a avaliação do contributo de cada ponto para o cálculo dos estatísticos (fig. 63).

No caso da malha não ter coordenadas explícitas, não foi introduzida esta opção, dado que o custo em cálculo e a complexidade prática da sua utilização pela densidade de elementos em questão, não o justifica.

Em ambos os casos, a opção *Save for Fitting* permitirá armazenar os dados do referido gráfico em memória para posterior modelação. Note-se que, por uma questão de eficiência, no caso das imagens, são armazenados os dados referentes aos três variogramas obtidos no que respeita aos canais R G B.

4.2.2.4. Diagramas de dispersão

Na construção do diagrama de dispersão, é contraposto ao valor de uma variável na posição x o valor da mesma variável na posição $x + \vec{h}$, onde \vec{h} é o vector distância (caso de diagramas de dispersão simples).

Segundo Isaaks e Srivastava [IS 89] e Pannatier [PA 96] estes diagramas são usados para identificar pares de localização que influenciem fortemente uma medida de continuidade espacial. Assim, neste sistema, é permitida a identificação dos pontos e a sua eliminação temporária ou definitiva, sendo automaticamente actualizados o gráfico direccional e os estatísticos.

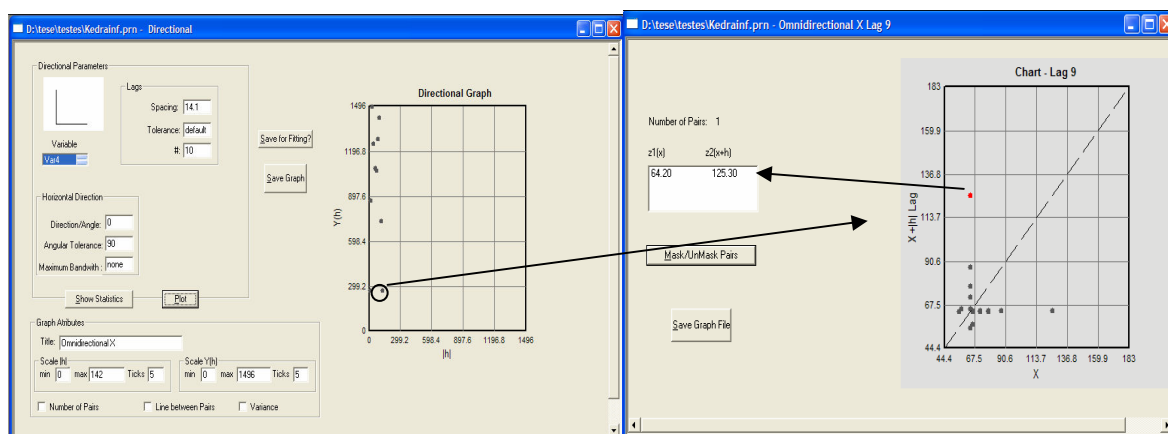


Fig. 63 Diagrama de dispersão para o Lag 2 (estrutura com coordenadas explícitas)

A construção do diagrama de dispersão (fig. 64) é desencadeada pressionando a tecla esquerda do rato sobre um ponto do gráfico direccional, permitindo visionar os dados resumidos no ponto (correspondente a um determinado passo) e seleccionar pares de amostras que, após simulação da sua eliminação, poderão ser efectivamente eliminados.

Os diagramas de dispersão são seleccionados via submenus: *Experimental Variograms* -> *Calculate* cujo botão *Plot* da opção *Cloud* origina:

A) no caso de estruturas de dados com coordenadas explícitas, a janela constante da Fig. 64;

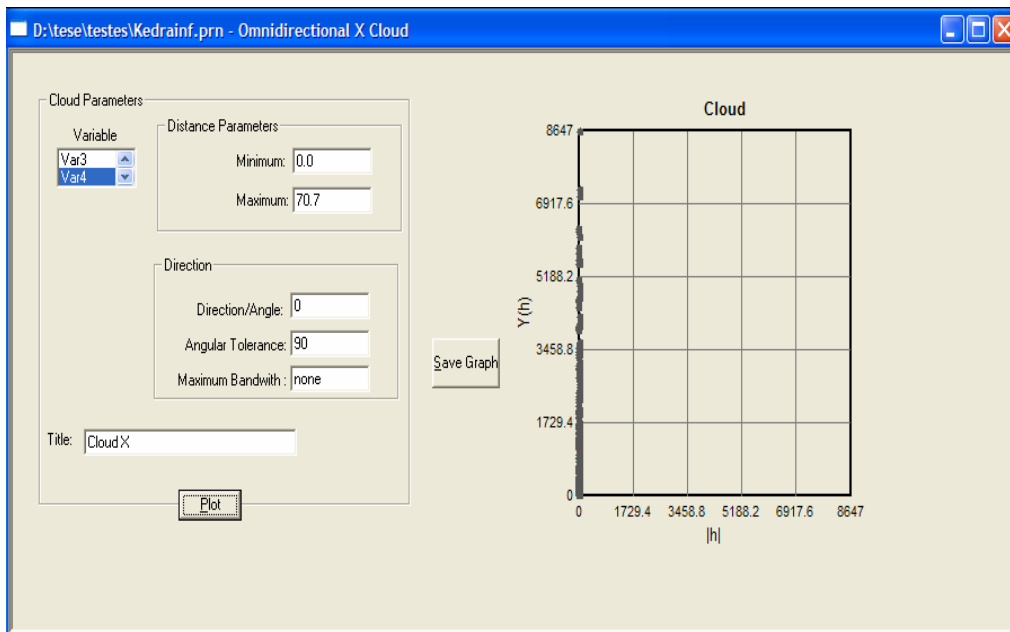


Fig. 64 Exemplo dos resultados sob a forma de gráfico de dispersão do variograma para amostras com coordenadas explícitas

B) no caso de estruturas de dados sem coordenadas explícitas, a janela constante da fig. 65;

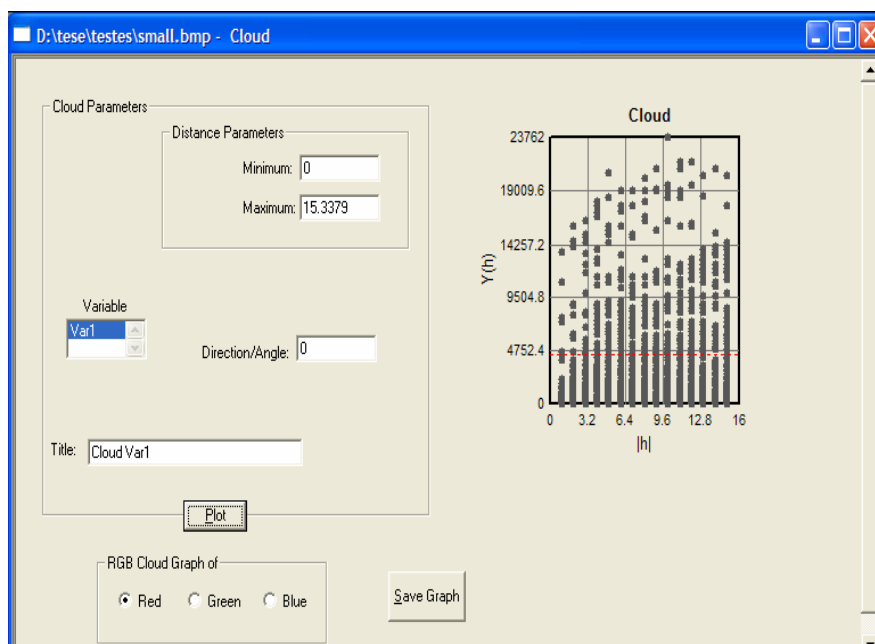


Fig. 65 Exemplo dos resultados sob a forma de gráfico de dispersão do variograma para amostras sem coordenadas explícitas

Para conjunto de amostras com coordenadas explícitas são pedidos ao utilizador os seguintes parâmetros:

- distância mínima;
- distância máxima ou distância de corte;
- direcção angular (azimute e inclinação) considerando a direcção angular (em graus) medida como um ângulo trigonométrico (origem em Este e em sentido contrário aos ponteiros do relógio);
- tolerância angular;
- largura de banda máxima.

Para o conjunto de amostras sem coordenadas explícitas apenas são pedidos ao utilizador os seguintes parâmetros:

- distância mínima;
- distância máxima;
- direcção angular.

Por defeito, são assumidos os seguintes valores para os seguintes parâmetros:

- A distância mínima é zero;
- A distância máxima é metade da distância máxima calculada entre pares de pontos;
- A direcção angular é 0° ;
- A tolerância angular é 90° (em graus). Se a direcção angular for 0 e a tolerância 90° é apresentado um diagrama omnidireccional;
- Não é considerado qualquer valor para a largura de banda máxima.

Neste tipo de diagrama não é fornecido qualquer janela resumo de estatísticos dado que a sua extensão dificultaria a leitura. Todavia, a pedido do utilizador, é possível gravar o quadro resumo num ficheiro.

4.2.2.5. Ajustamento da função teórica ao variograma experimental

A modelação teórica é accionada via submenu *Kriging* cuja opção *Fitting Model* desencadeia a janela constante da Fig. 66.

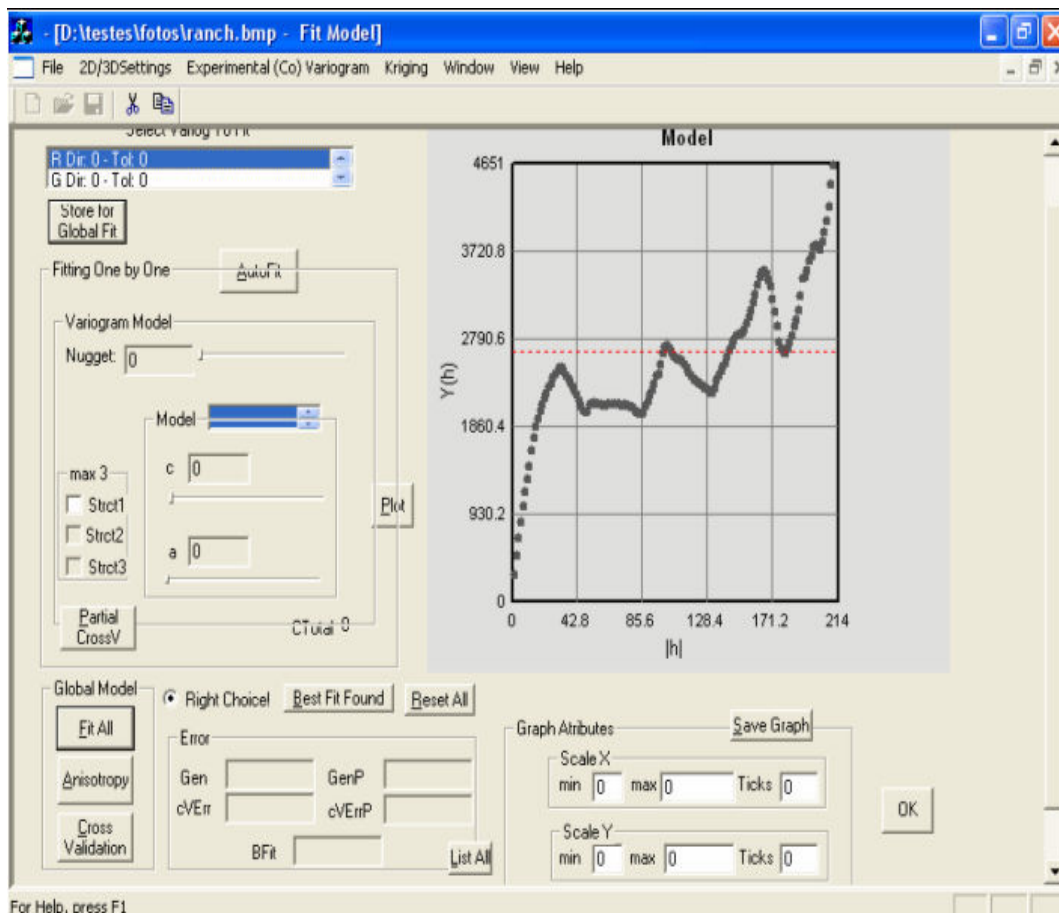


Fig. 66 Janela *Fitting Model*

A modelação dos variogramas teóricos é realizada de acordo com a seguinte sequência de procedimentos:

- 1 - Armazenam-se os variogramas direccionais que se pretende que contribuam para o modelo global;
- 2 - Calcula-se um modelo global que é válido para qualquer direcção (são gerados dois modelos (um referente a potência e outro a funções não potência e, automaticamente, é escolhido o que tiver menor erro após a validação cruzada);
- 3 - Modela-se cada variograma condicionado pelos valores do modelo global para obter modelos direccionais que diferem no parâmetro amplitude (a) ou potência (s) para possibilitar a detecção automática de anisotropia;
- 4 - Calcula-se a anisotropia.

O modelo final só pode ser obtido de forma automática através do botão *Fit All* contribuindo para o seu cálculo todos os variogramas direccionais armazenados através do botão *Store to Fit* podendo ser visualizados graficamente (fig. 67) os pontos relativos aos variogramas nas direcções consideradas, bem como o gráfico da função relativa ao modelo teórico obtido.

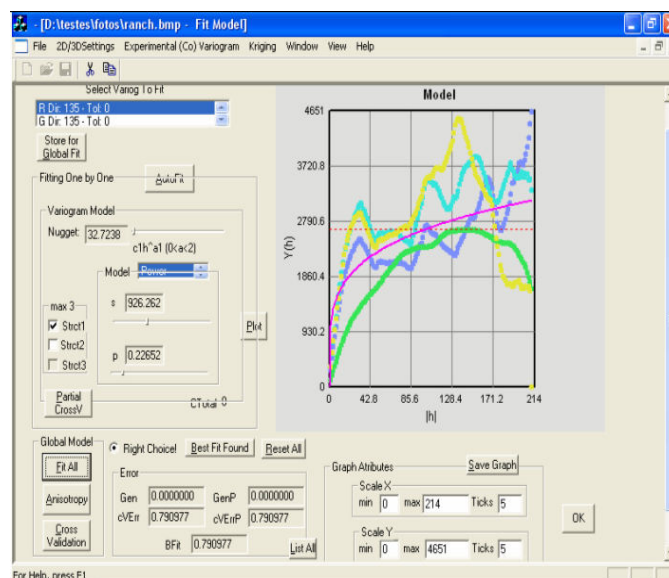


Fig. 67 Janela *Fitting* após o *Fit All* (imagem)

Na geração do modelo final o teste de validação cruzada é desencadeado automaticamente e obtém-se os erros, respectivamente, o erro obtido na geração automática do modelo relativo a modelos não potência (*Gen - Error*), o erro obtido na geração automática do modelo relativo ao modelo de potência (*GenP - Error*), o erro obtido após o teste de validação cruzada para modelos não potência (*Cross Val - Error*), o erro obtido após o teste de validação cruzada para modelos potência (*Cross ValP - Error*), e o menor erro cometido até ao momento (*Best Fit - Error*).

Note-se que, se não pretender aceitar o modelo escolhido automaticamente tem duas opções: ou escolhe outro modelo através da geração de novos modelos ou clica na opção *Right Choice!* Permitindo-lhe assim optar pelo outro modelo gerado

A modelação de cada variograma permite visualizar gráfica e analiticamente a função teórica ajustada, bem como o erro associado, ajustando as amplitudes ou potências de forma a calcular posteriormente a anisotropia. A modelação dos variogramas direccionais (fig. 68) pode efectuar-se das seguintes formas:

- Automaticamente, através do botão *AutoFit*. Desta forma a função teórica é gerada por um algoritmo genético e é submetida, automaticamente, a um teste de validação cruzada.
- Manualmente através do botão *Plot*, após a introdução dos seguintes parâmetros:
 - Função teórica que poderá ser escolhida entre esférica, exponencial, gaussiana e potência;
 - Efeito de pepita (*nugget effect*);
 - Patamar (*c*) ou, no caso da função potência, factor de escala (*s*);
 - Amplitude (*a*) ou, no caso da função potência, um expoente entre 0 e 2.

O erro cometido no ajuste do modelo efectuado manualmente é obtido através da opção *Partial CrossV* constante do submenu *Fitting Model*.

O ajuste dos ‘modelos direccionais’ apenas é efectuado para a amplitude ou potência (a) pois quer o tipo de modelo, o efeito de pepita ($c0$) e o patamar ou factor de escala (c) são propagados do modelo global para todos os modelos direccionais.

Em qualquer dos casos, ajuste automático ou manual, após o teste de validação cruzada, obtêm-se as saídas (fig. 68):

- o gráfico, onde podemos analisar o variograma e a função modelada;
- o C_{total} que é a soma do patamar com o efeito de pepita e que se deve aproximar da variância;
- os erros, respectivamente, o erro obtido na geração automática do modelo (*Genetic - Error*), o erro obtido após o teste de validação cruzada (*Cross Val - Error*) e o menor erro cometido até ao momento (*Best Fit - Error*).

Também em qualquer dos casos é construído um quadro resumo com todos os modelos e erros respectivos calculados para um dado variograma, que poderá ser visualizado através do botão *List All*.

O modelo teórico com menor erro é sempre armazenado e poderá ser recuperado através do botão *Best Fit Saved*.

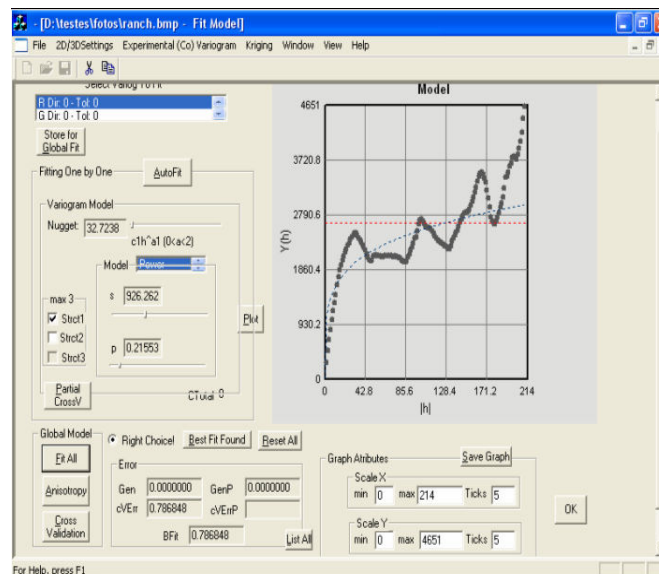


Fig. 68 Janela *Fitting Model* após o *Auto Fit*

Nesta janela gráfica, tal como na janela dos diagramas direccionais, é possível alterar a escala de valores dos eixos horizontal e vertical bem como o número de subdivisões a considerar em cada eixo.

4.2.3. Estimação

Nesta secção descrevem os procedimentos a adoptar para realizar a estimação ou simulação a partir do modelo teórico encontrado, bem como as saídas possíveis para os resultados obtidos após essas operações.

4.2.3.1. Ordinary *Krigagem*

A *krigagem* é desencadeada pela opção Ordinary *Kriging*/*CrossValidation* do submenu *Kriging* apresentando a janela constante da Fig. 69.

Nesta janela são introduzidos todos os parâmetros necessários para efectuar a *krigagem* usando o módulo *kt3d* do GSLIB [DJ 98].

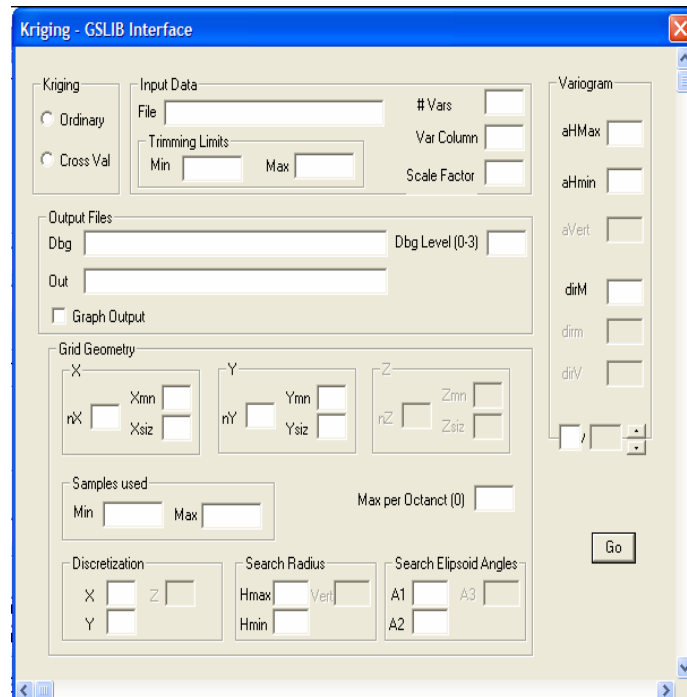


Fig. 69 Janela de interface da opção Ordinary *Kriging*

Os parâmetros a introduzir são os seguintes:

- *kriging*: escolha de uma malha de pontos ou blocos ou teste de validação cruzada (por defeito, *grid estimation*);
- *file*: o ficheiro que contém os dados (por defeito, o último ficheiro utilizado);
- *var column*: uma coluna da variável a estimar (as coordenadas também contam). Por defeito, a variável sobre a qual se efectuou a modelação;
- *trimming limits*: os valores máximo e mínimo a considerar na estimação (mínimo e máximo inclusivé);
- *factor scale*: percentagem de aumento ou redução (apenas para efeitos de reescalonamento dos valores);
- *output files*:
 - *dbg*: nome do ficheiro de debug onde será escrito um relatório do programa de *krigagem* (por defeito, o nome do ficheiro de entrada acoplado a *krig* e com sufixo *dbg*). Por exemplo: nome do ficheiro de entrada *d:\testes\kedrainf.prn* - > nome do ficheiro de debug *d:\testes\kedrainfkrig.dbg*;
 - *dbg*: um inteiro de 0 a 3 representando o nível de *debug* pretendido;
 - *out*: nome do ficheiro com os resultados estimados (por defeito, o nome do ficheiro de entrada acoplado a *krig* e com sufixo *krig*). No exemplo anterior, o ficheiro de saída designar-se-ia por *d:\testes\kedrainfkrig.krig*;
 - *graph output*: opção para obter os resultados sob a forma gráfica;
- *grid geometry*:
 - *nx, xmn, xsiz*: definição da malha regular para o eixo do X, respectivamente, número de pontos, ponto inicial e distância entre pontos;
 - *ny, ymn, ysiz*: o mesmo que o anterior para Y;
 - *nz, zmn, zsiz*: o mesmo que o anterior para Z

- *samples used*: o mínimo e o máximo número de pontos usado para estimar um bloco;
- *discretization*: número de pontos de discretização para um bloco (por defeito é assumido 1 o que corresponde à estimação pontual);
- *search radius*: raios de pesquisa máximo e mínimo horizontal e na direcção vertical (por defeito, é utilizada a dimensão da área de estudo);
- *search ellipsoid angles*: ângulos que definem a orientação angular de pesquisa num elipsoide (por defeito é utilizada a direcção principal);
- *variogram*: as amplitudes máxima e mínima horizontais e a vertical bem como os ângulos que definem a anisotropia geométrica (por defeito usam-se os parâmetros calculados automaticamente).

Apenas se considera anisotropia se o quociente entre a maior amplitude e a menor amplitude dos modelos teóricos calculados para as diferentes direcções for superior a 1.5.

Após efectuada a estimação, para além da criação do ficheiro *.krg*, podem visualizar-se graficamente os resultados obtidos, tendo sido considerados dois tipos de representação:

- o gráfico de superfície colorido se os dados de entrada não resultarem de uma imagem (fig. 70)

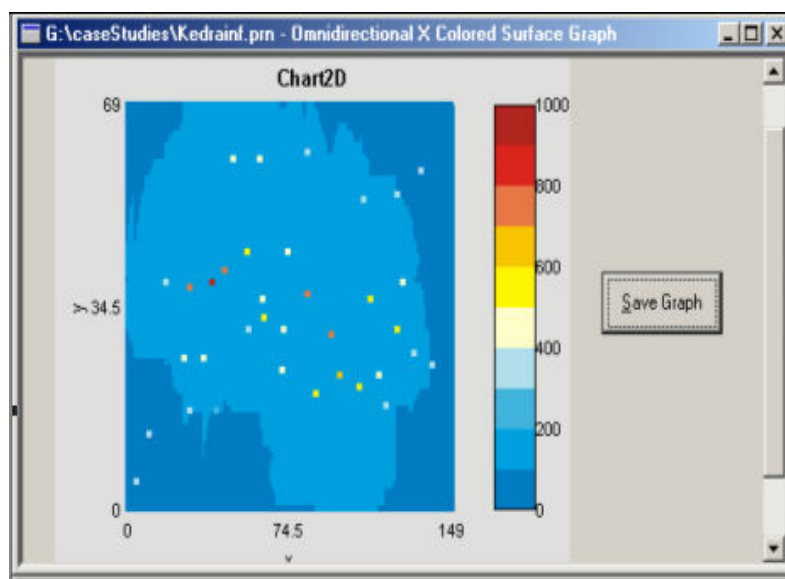


Fig. 70 Janela de resultados de estimação (casos não imagem)

- a imagem original e a final, para os casos em que foram estimados canais RGB (fig. 71).



Fig. 71 Janela de resultados de estimação (caso de imagem)

4.2.3.2. Cokrigagem colocalizada

A *cokrigagem colocalizada* é desencadeada pela opção *Co-located/CoKriging* do submenu *Kriging*, apresentando a janela constante da fig. 72.

Nesta janela são introduzidos os parâmetros necessários para efectuar a *cokrigagem colocalizada* usando o módulo *cokb3d* (adaptado) do GSLIB [DJ 98].

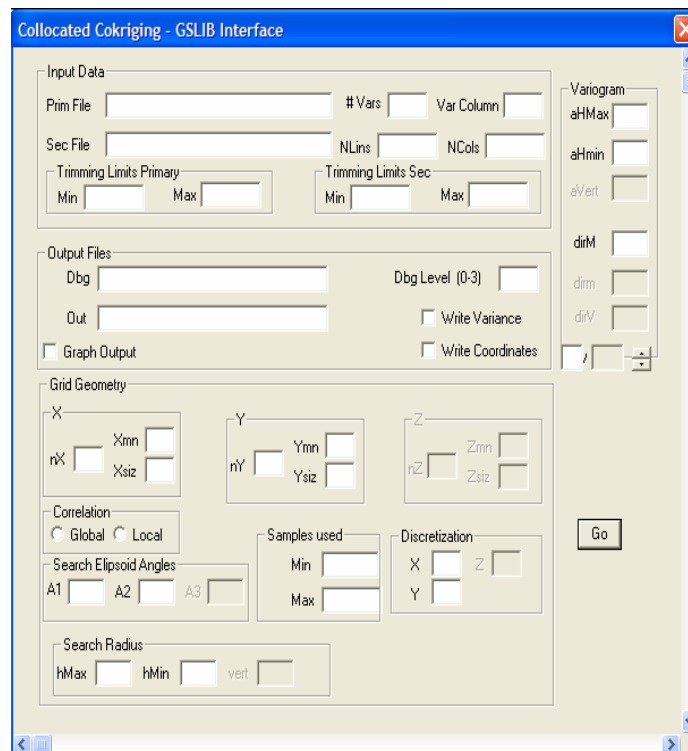


Fig. 72 Janela interface *Cokrigagem colocalizada*

Os parâmetros a introduzir são os seguintes:

- *Prim file*: o ficheiro principal (por defeito, o último ficheiro utilizado);
- *Sec file*: o ficheiro com os dados secundários;
- *Var columl*: uma coluna da variável a estimar (as coordenadas também contam). Por defeito, a variável sobre a qual se efectuou a modelação;
- *trimming limits*: os valores máximo e mínimo a considerar na estimação (mínimo e máximo inclusive);
- *output files*:
 - *dbg*: nome do ficheiro de *debug* onde será escrito um relatório do programa de *krigagem* (por defeito, o nome do ficheiro de entrada acoplado a *krig* e com sufixo *dbg*). Por exemplo: nome do ficheiro de entrada *d:\testes\kedrainf.prn* - > nome do ficheiro de *debug* *d:\testes\kedrainfcol.dbg*;

- *dbg*: um inteiro de 0 a 3 representando o nível de *debug* pretendido;
- *out*: nome do ficheiro com os dados estimados (por defeito, o nome é o nome do ficheiro de entrada acoplado a *col* mas com sufixo *kr*). No exemplo anterior, o ficheiro de saída designar-se-ia por *d:\testes\kedrainfcol.krg*.

No exemplo anterior, o ficheiro de saída designar-se-ia por *d:\testes\kedrainfcol.krg*;

- *dbg*: um inteiro de 0 a 3 representando o nível de *debug* pretendido;
 - *variance*: se se pretende apresentar no ficheiro de saída a variância de estimação (por defeito (sim));
 - *coordinates*: se se pretende apresentar no ficheiro de saída as coordenadas dos pontos (por defeito (sim));
- *output files*:
 - *graph output*: opção para obter os resultados sob a forma gráfica;
 - *grid geometry*:
 - *nx, xmn, xsiz*: definição da malha regular para o eixo do X, respectivamente, número de pontos, ponto inicial e distância entre pontos;
 - *ny, ymn, ysiz*: o mesmo que o anterior para Y;
 - *nz, zmn, zsiz*: o mesmo que o anterior para Z
 - *correlation*: assinalar se pretende correlação global ou local;
 - *samples used*: o mínimo e o máximo número de pontos usado para estimar um bloco;
 - *discretization*: número de pontos de discretização para um bloco (por defeito é assumido 1 o que corresponde à estimação pontual);
 - *search radius*: raios de pesquisa máximo e mínimo horizontal e na direcção vertical (por defeito, é utilizada a dimensão da área de estudo);

- *search ellipsoid angles*: ângulos que definem a orientação angular de pesquisa num elipsoide (por defeito é utilizada a direcção principal);
- *variogram*: as amplitudes máxima e mínima horizontais e a vertical bem como os ângulos que definem a anisotropia geométrica (por defeito usam-se os parâmetros calculados automaticamente).

Apenas se considera existir a existência de anisotropia se o quociente entre a maior amplitude e a menor amplitude dos modelos teóricos calculados para as diferentes direcções for superior a 1.5.

Após efectuada a estimação *colocalizada*, para além da criação do ficheiro *.krig*, pode visualizar-se graficamente os resultados obtidos através de um gráfico colorido de superfície (fig. 73).

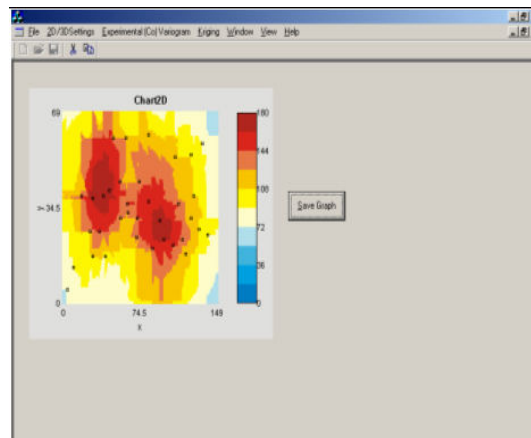


Fig. 73 Janela Resultados de *Cokrigagem* colocalizada

4.2.4. Simulação Sequencial Gaussiana

A simulação é desencadeada pela opção *SGS* do submenu *Kriging/Simulating*, apresentando a janela constante da fig. 74.

Nesta janela são introduzidos os parâmetros necessários para efectuar a simulação usando o módulo *sgsim* do GSLIB [DJ 98] após as necessárias adaptações de integração neste sistema.

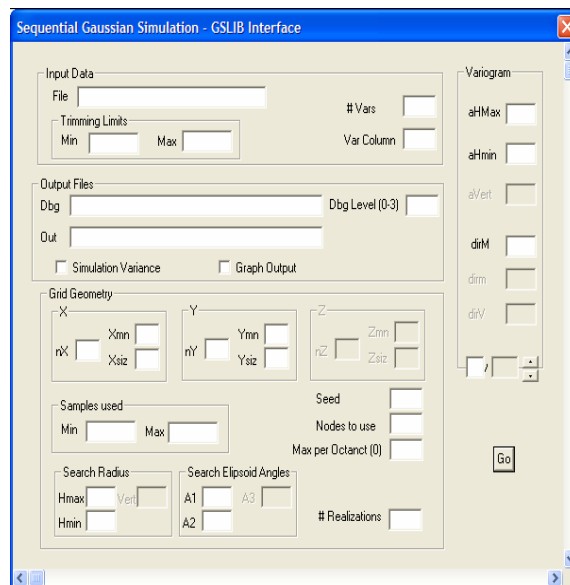


Fig. 74 Janela interface simulação sequencial gaussiana

Os parâmetros a introduzir são os seguintes:

- *file*: o ficheiro que contém os dados (por defeito, o último ficheiro utilizado);
- *var column*: uma coluna da variável a estimar (as coordenadas também contam). Por defeito, a variável sobre a qual se efectuou a modelação;
- *trimming limits*: os valores máximo e mínimo a considerar na estimação (mínimo e máximo inclusive);
- *output files*:
 - *dbg*: nome do ficheiro de *debug* onde será escrito um relatório do programa de simulação (por defeito, o nome do ficheiro de entrada acoplado a *sim* e com sufixo *dbg*). Por exemplo: nome do ficheiro de entrada *d:\testes\kedrainf.prn* - > nome do ficheiro de *debug* *d:\testes\kedrainfsim.dbg*;
 - *dbg*: um inteiro de 0 a 3 representando o nível de *debug* pretendido;
 - *out*: nome do ficheiro com os dados estimados (por defeito, o nome é o mesmo do ficheiro de entrada acoplado a *sim* e com sufixo *krj*).

No ex. anterior, o ficheiro de saída designar-se-ia por *d:\testes\kedrainfsim.krg*;

- *graph output*: opção para obter os resultados sob a forma gráfica;
- *simulation var*: opção para ser calculada a variância entre simulações;
- *grid geometry*:
 - *nx, xmn, xsiz*: definição da malha regular para o eixo do X, respectivamente, número de pontos, ponto inicial e distância entre pontos;
 - *ny, ymn, ysiz*: o mesmo que o anterior para Y;
 - *nz, zmn, zsiz*: o mesmo que o anterior para Z
- *samples used*: o mínimo e o máximo número de pontos usado para estimar um bloco;
- *random seed*: semente usada para gerar números aleatórios;
- *nodes to use*: número de pontos já simulados usados na simulação de um outro;
- *# realizations*: número de realizações a efectuar;
- *search radius*: raios de pesquisa máximo e mínimo horizontal e na direcção vertical (por defeito, é utilizada a dimensão da área de estudo);
- *search ellipsoid angles*: ângulos que definem a orientação angular de pesquisa num elipsoide (por defeito é utilizada a direcção principal);
- *variogram*: as amplitudes máxima e mínima horizontais e a vertical bem como os ângulos que definem a anisotropia geométrica (por defeito usam-se os parâmetros calculados automaticamente).

Apenas se considera existir anisotropia se o quociente entre a maior amplitude e a menor amplitude dos modelos teóricos calculados para as diferentes direcções for superior a 1.5.

Após efectuada a simulação, para além da criação do ficheiro *.krg*, tal como na *krigagem* é possível visualizar-se graficamente os resultados obtidos de acordo com o tipo de representações já anteriormente referidos, ou seja:

- o gráfico de superfície colorido, no caso de dados de entrada estarem armazenados em ficheiros ASCII
- a imagem original e a final, para os casos em que foram simulados canais RGB

4.3. Alguns aspectos particulares da implementação

Consideram-se aspectos particulares da implementação o ambiente de desenvolvimento, a descrição dos algoritmos foram considerados relevantes, a descrição das estruturas intermédias de maior peso, os ficheiros necessários como interface entre módulos executáveis e todos os módulos e bibliotecas que facilitam aspectos pontuais da implementação.

4.3.1. Ambiente de desenvolvimento

Para o desenvolvimento do sistema utilizaram-se as potencialidades do Visual C++ (Microsoft®) no que se refere a projectos *MDI (Multi Document Interface)* que são caracterizados por 4 componentes principais:

- objecto *Application* – o primeiro objecto a ser executado pelo Windows e responsável pela coordenação global do sistema;
- objecto *Main Window* - responsável pela gestão das janelas que permitem a visualização de saídas;
- objecto *Document* onde deve ser efectuado o processamento principal dos dados;
- objecto *View* - que manuseia a área cliente da janela, ou seja, onde se formatam e geram as saídas.

Para o desenho deste sistema foi respeitada a estrutura supra-referida, criando:

- um objecto *Document* onde são efectuados todos os tratamentos e desencadeadas todas as caixas de diálogo relativas ao processamento do ficheiro de dados;

- vários objectos do tipo *View* que preparam os vários tipos de saídas proporcionadas (janelas para os diferentes tipos de gráficos, janelas descritivas contendo resumo de estatísticos, janelas permitindo a visualização de imagens, etc.).

O *design* e a funcionalidade das janelas são semelhantes aos propostos pelo Windows (Microsoft®), numa tentativa de construção de *software* facilmente manuseável pelo utilizador.

4.3.2. Diagrama de classes

Na página seguinte é apresentado o diagrama de classes correspondente ao sistema desenvolvido onde as classes construídas no sistema (rectângulos coloridos a amarelo) herdam atributos e métodos das classes existentes no Visual C++ (rectângulos cinzentos).

Note-se que estão representados ligações com dois tipos de setas:

- as setas com a extremidade triangular que ligam classes a super-classes e que são direccionadas no sentido da herança, ou seja, existem métodos e atributos da super-classe que são herdados pela classe, por ex., a classe *CNTGraph* herda da classe *CWnd* do Visual C++ os métodos e atributos que necessita para produzir uma janela do tipo das janelas do Microsoft® Windows;
- as outras setas ligam classes independentes mas que estão associadas, ou seja, relacionam-se entre si sendo a seta dirigida para a classe que é invocada, por ex., a classe *CModel* é independente da classe *CGenetic* mas utiliza-a para realizar a modulação automática.

Não são visualizadas no diagrama de classes todas as relações associativas existentes pois a maior parte são realizadas para cada classe do tipo *View* associada a uma classe do tipo *Doc* e, se necessário, a uma classe do tipo *Dialog* e a sua representação tornaria o diagrama de classes muito pouco legível.

4.3.3. Descrição dos algoritmos mais relevantes

Seguidamente descrevem-se os algoritmos mais relevantes e alguns detalhes de implementação das funcionalidades para entrada de dados, estudo variográfico, modelação e interface para *krigagem/cokrigagem colocalizada* e *simulação* de imagens.

4.3.3.1. Entrada de dados

O armazenamento dos dados de entrada em memória e o pré-processamento dessa informação (validação e delimitação) foram implementados de acordo com as diferentes características dos tipos de ficheiros de entrada.

Para abordar a problemática do pré-processamento dos dados de entrada adoptaram-se duas metodologias: uma respeitante aos dados com coordenadas explícitas (em geral irregulares) e outra referente aos dados com coordenadas implícitas (em geral regulares). Note-se que as imagens podem ser consideradas como uma generalização de um conjunto de dados com coordenadas implícitas.

A) Dados com coordenadas explícitas

Como foi referido anteriormente, este tipo de dados engloba amostras distribuídas irregularmente a 2D ou 3D, tendo sido convencionado, por uma questão de facilidade de implementação, que os valores atribuídos às coordenadas, designadamente X, Y ou Z devem corresponder às primeiras duas ou três colunas do ficheiro (cap. 5. secção 5.3.3.1.).

Aos dados contidos nestes ficheiros, armazenados em colunas e linhas sem qualquer cabeçalho, corresponderão variáveis cujas designações podem ser introduzidas através de menus.

Dado que todos os estatísticos de análise estrutural utilizam a distância euclidiana entre amostras, entendeu-se que, para aumentar a eficiência no cálculo dos estatísticos, se poderia usar um procedimento semelhante ao adoptado na construção do Varwin [PA 96] ou seja,

criar uma estrutura em memória que armazene sequencialmente todos os pares de pontos e as respectivas distâncias calculadas previamente.

Note-se que esta opção tem custos do ponto de vista de tempo de processamento, principalmente quando o conjunto das amostras é grande. Contudo, na maior parte das aplicações com amostras irregulares, esta situação raramente é verificada, donde o procedimento em questão é, em geral, compensador no que respeita à sua eficiência.

Para a ordenação por distâncias implementou-se o algoritmo *quicksort* que, dos algoritmos conhecidos para grandes conjuntos de dados e segundo [PTVF 02], é considerado o mais rápido. Apesar do Visual C++ possuir uma função (*quicksort()*), esta não responde aos requisitos necessários para o tipo de dados a ordenar uma vez que apenas permite os tipos pré-construídos e não objectos do tipo *CArray*¹⁰. Optou-se, assim, por criar uma estrutura ordenada apenas com identificadores dos elementos sujeitos à ordenação.

Principais etapas do pré-processamento

As etapas de pré-processamento, após a caracterização do ficheiro (nome, tipo de dados e, neste caso, número de variáveis), a identificação pelo utilizador da variável com que pretende trabalhar, as designações e limites das variáveis e a expressão do seu requisito no que respeita a validação de dados com coordenadas iguais serão as seguintes:

1º Armazenar os dados de entrada em memória;

2º Realizar as validações e restringir as variáveis de acordo com o solicitado pelo utilizador;

3º Calcular as distâncias entre dados e ordenar por ordem crescente os pares obtidos, em função das suas distâncias.

¹⁰ Classe de objectos do Visual C++ que suporta dados do tipo vector implementando gerindo dinamicamente a memória a alocar

B) Dados sem coordenadas explícitas

No caso da inexistência de coordenadas explícitas, tal como em GSLIB ([DJ 98]), cada dado terá como coordenadas (x, y) no caso 2D ou (x, y, z) no caso 3D devendo percorrer-se primeiro X, depois Y e finalmente Z.

Os vectores nas linguagens de programação C/ C++ usam como primeiro índice o zero e apenas podem ser definidos a uma dimensão.

Assim, para a implementação de estruturas do tipo matriz, ou se usa o tipo normalizado em C/C++ (unidimensional), ou se constrói um novo tipo multidimensional com as funcionalidades de linha e coluna. A opção neste *software* foi usar a classe *CArray* que corresponde a um vector unidimensional com alocação dinâmica.

Face à escolha da classe *CArray* para armazenar uma matriz sob a forma de um vector unidimensional será necessário converter as coordenadas x e y num único índice (i) através da fórmula:

$$i = y * nx + x \quad \text{sendo } nx - n^\circ \text{ de colunas de } x.$$

A conversão inversa, ou seja, a obtenção das coordenadas x e y dado um índice (i), far-se-á pelas fórmulas seguintes:

$$y = \text{int}\left(\frac{i}{nx}\right) \quad x = i - y * nx \quad \text{sendo } \text{int}() \text{ a função que devolve o maior inteiro contido no cociente.}$$

Finalmente, no caso particular das imagens, deve referir-se que foi usada a biblioteca *CxImage* (descrita na secção 3.4. deste cap.) para facilitar o acesso directo aos valores dos canais que são agrupados em cada *pixel* situado num par (*linha, coluna*).

Principais etapas do pré-processamento

Neste caso, após a caracterização do ficheiro (nome, tipo de dados e o nº de variáveis para o caso de dados não correspondendo a imagem) e a identificação pelo utilizador da variável que pretende processar (apenas para o caso de dados não correspondendo a imagem), bem como as suas designações e os limites, as principais etapas serão:

1º Armazenar os dados de entrada em memória;

2º Realizar as validações e delimitação das variáveis de acordo com o solicitado pelo utilizador.

4.3.3.2. Cálculos de Estatísticos

Para o cálculo de estatísticos, tal como para o pré-processamento dos dados de entrada, adoptaram-se dois tipos de metodologias: uma para dados com coordenadas explícitas e outra para dados com coordenadas implícitas.

Quanto aos parâmetros necessários para o cálculo dos estatísticos deve notar-se que:

- Não é necessário considerar qualquer tolerância para os dados contidos em estruturas de dados regulares;
- As distâncias angulares devem ser introduzidas em graus (serão convertidas internamente em radianos) e, tal como é definido em Matemática e em [PA 96] mas ao contrário do GSLIB ([DJ 98]), consideram-se os ângulos com o sentido contrário aos ponteiros do relógio como ângulos positivos (ver cap. 5);
- Como este *software* permite, em cada instante, saber o valor de todos os estatísticos referidos no cap. 2, os resultados obtidos terão que permanecer em memória para poderem ser reutilizados sem necessidade de efectuar novos cálculos.

A) Particularidades para o caso de dados com coordenadas explícitas

No caso de dados com coordenadas explícitas, como já foi referido anteriormente, estes foram armazenados em pares ordenados pela distância entre eles. Assim, para o cálculo do gráfico direccional apenas será necessário testar um dos limites das classes de distância (h) donde, a estrutura de dados ordenados será percorrida sequencialmente sujeita ao teste do limite superior da distância para cada classe.

B) Particularidades para o caso de dados sem coordenadas explícitas

Neste caso, que pressupõe a existência de uma malha regular, os cálculos das distâncias terão, tal como GSLIB ([DJ 98]), que ter em conta o tamanho da unidade da malha quer para X, quer para Y. Assim, a distância entre pontos será dada por:

$$h_{xy} = \sqrt{((x_2 - x_1) * unit_x)^2 + ((y_2 - y_1) * unit_y)^2} \quad \text{onde } unit_x \text{ e } unit_y \text{ representam, respectivamente, a unidade da malha para X e para Y.}$$

Sendo a malha regular caracterizada pelo o grande volume de dados (ordem mínima testada de 100x100) a processar, usou-se um artifício de cálculo que possibilita a substituição por uma simples soma do teste angular necessário para obter os pares que pertencem a uma dada direcção.

Assim, note-se que, para um dado com localização (lin, col),

- a) O par para a direcção **0** terá localização ($lin, col+1$)
- b) O par para a direcção **45** terá localização ($lin+1, col+1$)
- c) O par para a direcção **90** terá localização ($lin-1, col$)
- d) O par para a direcção **135** terá localização ($lin-1, col-1$)

...

Nos casos de dados resultantes de imagens, ainda numa perspectiva de optimização, foram simultaneamente calculados os estatísticos para os 3 canais de cor (pesquisando apenas uma

vez o valor da distância) e, nos gráficos, se for detectado o caso de imagem, será visualizado, em cada momento, apenas um gráfico (referente a um canal).

Ainda na fase de cálculo de estatísticos, são criados ficheiros com o *design* de entrada necessário para efectuarem a ligação com os módulos executáveis do GSLIB ([DJ 98]) (secção 4.4.4.).

Principais etapas do cálculo de estatísticos

Após a escolha do estatístico para o qual se pretende efectuar os cálculos e do tipo de gráfico pretendido:

1º São escolhidas as opções para os diferentes parâmetros bem como a variável que se pretende estudar (no caso em que exista mais que uma variável);

2º São efectuados os cálculos relativos aos estatísticos e o gráfico é visualizado numa janela gráfica;

3º Através de um botão é possibilitado ao utilizador o armazenamento dos pontos referentes ao cálculo do estatístico para utilização futura na modelação;

4º Repete-se o ciclo de 1 a 3 enquanto o utilizador pretender o cálculo do estatístico noutra direcção e o seu posterior armazenamento.

4.3.3.3. Representações gráficas

Para as representações gráficas foi utilizado como base um controlo importado de uso público (**2D Graph ActiveX Control** - secção 4.4.5. deste cap.).

Este controlo tem na sua base a construção de gráficos como conjuntos de pontos ou linhas, bem como a implementação das opções necessárias às funcionalidades normalmente usadas pelo utilizador na alteração do *design* de um gráfico, tais como, a limitação dos valores presentes nos gráficos, a cor da apresentação, etc.

Foi necessário complementar essa arquitectura com funções que permitissem a representação gráfica em gráficos de superfície e mapas de cor, bem como a alteração de algumas funções para permitir, por exemplo, que no caso de gráficos direccionais seja também visualizada a variância (se for solicitado pelo utilizador).

4.3.3.4. Modelação

Foram implementadas dois tipos de modelação. A modelação manual em que o utilizador ajusta um modelo teórico e a modelação automática em que o utilizador apenas intervém, se o pretender, para corrigir manualmente o ajuste automático efectuado.

Destacam-se nos parágrafos seguintes os aspectos mais relevantes em cada um dos tipos de modelação.

A) Modelação Automática

A modelação automática é realizada com um algoritmo genético que, como foi fundamentado na secção 3.1.2.1. do capítulo 3, é o mais adequado para a resolução deste tipo de problemas.

As principais etapas da implementação da modelação automática são as seguintes:

1º São escolhidos os variogramas direccionais da lista de variogramas memorizados que intervirão no modelo global;

2º Através de um botão desencadeia-se a modelação automática implicando, sob ponto de vista da implementação, a invocação do algoritmo genético por duas etapas:

- na primeira etapa gerando o modelo que melhor se ajuste dentre todos os modelos à excepção do modelo de potência;
- na segunda etapa gerando um modelo de potência;

sendo escolhido, por defeito; o modelo que tiver menor erro que será sujeito à validação cruzada;

3º O modelo global e os variogramas direccionais seleccionados, bem como, os erros calculados após a validação cruzada são visualizados;

4º Com o modelo global ajustado proceder-se-á à modelação automática de cada variograma direccionais. Esta modelação será condicionada pelo modelo global obtido, pelo patamar ou efeito de escala obtidos e pelo efeito de pepita, ou seja, apenas possibilitará o ajustamento da amplitude ou potência em cada ‘modelo direccionais’ para possibilitar a posterior detecção de anisotropias. Os variogramas e a respectiva variância, bem como o modelo ajustado são visualizados numa janela gráfica;

6º É calculada a anisotropia, caso exista, se solicitado pelo utilizador;

7º Em qualquer momento é possível ajustar manualmente os parâmetros do modelo teórico;

8º A validação cruzada é desencadeada através do botão respectivo e, dever-se-á repetir a modelação automática, quer no caso do modelo global quer no caso do modelo de variogramas direccionais, até que o utilizador considere o modelo ajustado;

9º O ultimo modelo em memória é o modelo memorizado e será o modelo usado na *krigagem/cokrigagem colocalizada* e na simulação.

A validação cruzada direccionais é apenas efectuada para os pontos que pertencem a essa direcção que são guardados num ficheiro.

No caso de ficheiros de imagem as validações cruzadas são efectuadas apenas para um conjunto de 30% dos dados retirados aleatoriamente do ficheiro de imagens.

O *erro* cometido na validação cruzada é calculado de acordo com a seguinte formula

$$\text{erro} = \frac{\text{variancia}(e - r)}{\text{media}(\text{variancia}_e)}$$
 onde variancia_e é a variância de estimação e $\text{variancia}(e - r)$ é a

variância da diferença entre os estimados e os reais (prova-se que este cociente tende para 1 [IS 89]).

Cálculo automático da anisotropia

Considerar-se-á a existência de anisotropia quando o cociente entre a maior amplitude e a menor amplitude (razão entre distâncias no caso do modelo de potência) dos modelos direccionais for superior a 1.5.

Em caso de existência de anisotropia, a direcção correspondente à maior amplitude será a direcção principal e a maior amplitude (maior distância) corresponderá ao semi-eixo principal da elipse correspondente à anisotropia.

A direcção secundária será a direcção perpendicular à direcção principal e, se o variograma experimental para essa direcção existir, considerar-se-á a sua amplitude (distância) como o eixo menor da elipse, caso contrário dever-se-á calcular o semi-eixo menor da elipse sabendo o semi-eixo principal e o valor da amplitude para outra direcção (por exemplo, o ponto da elipse de coordenadas $(0, a)$ correspondente à direcção de 90°).

A anisotropia só será visualizada nas janelas de preparação para a *krigagem/cokrigagem* e simulação e é sempre passível de aí ser introduzida manualmente.

B) Modelação manual

A modelação manual, que apenas pode ser efectuada para modelar variogramas direccionais, é implementada segundo o procedimento seguinte:

1º É escolhido um variograma direccional da lista de variogramas memorizados;

2º O variograma e a variância são visualizados numa janela gráfica onde serão também visualizados o gráfico das funções escolhidas para as diferentes estruturas para os valores introduzidos nos parâmetros;

3º É possível ao utilizador (através de um botão) desencadear a validação cruzada para obter o erro relativo ao modelo teórico concebido;

4º Se o utilizador considerar que o modelo está ajustado, deve ser requerida a sua memorização para, posteriormente, produzir automaticamente o modelo global;

5º Após ter sido efectuada a modelação de todas as direcções o utilizador pode solicitar o ajuste global dos modelos direccionais que foram memorizados;

6º É sempre possível recomeçar o processo de ajuste reiniciando o processo de modelação;

7º É sempre possível visualizar todos os modelos direccionais já calculados e os respectivos erros bem como recuperar o modelo com menor erro já encontrado.

4.3.3.5. Validação cruzada, *Krigagem* e *cokrigagem colocalizada*, simulação

Para a validação cruzada, *krigagem*, *cokrigagem colocalizada* e simulação utilizaram-se os módulos executáveis *kt3d.exe*, *cokb3d.exe*, *sgsim.exe* (secção 4.3.5.) do GSLIB ([DJ 98]) com as devidas adaptações para permitirem processar um maior volume de dados e realizar a *cokrigagem colocalizada*.

Estando fora do âmbito desta tese a construção de módulos específicos para *krigagem*, *cokrigagem* e simulação, optou-se pela construção de uma interface que possibilite a criação do ficheiro de parâmetros necessário para as execuções e pela invocação dos módulos referidos através duma chamada ao MSDOS®.

Não foi considerada viável a opção de se considerar estes módulos programados em Fortran como sub-programas pois, tendo este sistema como objectivo o processamento de grande volume de dados, tal opção obrigaria à reestruturação total dos módulos para possibilitar alocações dinâmicas.

A validação cruzada é realizada sem interacção do utilizador pois todos os elementos necessários são, ao momento, conhecidos pelo sistema.

Na *krigagem*, que também é desencadeada através do mesmo módulo, é, no mínimo, necessária a introdução do nº de linhas e colunas que se pretende *krigar* (o mesmo acontece na *cokrigagem colocalizada* e na simulação). Assim, optou-se por fornecer todos os parâmetros preenchidos por defeito ao utilizador para que este possa decidir sobre a sua manutenção ou reconfiguração.

Quanto à *cokrigagem colocalizada* para além dos parâmetros já referidos anteriormente é necessário, pelo menos, indicar os parâmetros relativos ao ficheiro secundário como por exemplo o nome do ficheiro. Para este tipo de *krigagem* é necessário o coeficiente de correlação que, no caso do coeficiente de correlação global é automaticamente calculado e gravado em ficheiro.

No caso particular da simulação é possível, se solicitado, obter a imagem após uma simulação e a variância das imagens obtidas para n realizações. Essa variância, por canal de imagem, vai sendo guardada num ficheiro que contém variâncias.

Deve ainda notar-se que, para a validação cruzada em canais de imagem, foram apenas utilizadas 30% das amostras pertencentes ao ficheiro relativo a essa direcção ou canal sendo retiradas sequencialmente 3 em cada 10.

No que se refere aos resultados obtidos pela *krigagem*, *cokrigagem colocalizada* e simulação estes podem ser vistos através dos ficheiros de dados fornecidos à saída dos módulos ou, por opção, podem ser vistos graficamente no ambiente do sistema desenvolvido.

Os resultados obtidos podem ser visualizados de acordo com o tipo de dados de entrada:

- Imagem se os dados de entrada eram uma imagem
- Mapa de cores nas outras situações

4.3.3.6. *Krigagem de imagens*

Os três canais RGB de uma imagem terão naturalmente modelos associados a cada canal. Assim, para o caso das imagens, optou-se por permitir apenas a modelação de cada canal seguida da *krigagem* do canal respectivo, sendo os resultados finais obtidos através da combinação destas *krigagens*.

Assim, no caso de imagens, o utilizador terá que realizar a seguinte sequência de operações:

1º Obter o modelo para um canal escolhido;

2º *Krigar* os dados para o canal referido;

3º Proceder da mesma forma para o segundo e para o terceiro canal escolhido;

4º Após ser processada a terceira *krigagem*, o sistema, se solicitada uma apresentação gráfica, recompõe a imagem que, após a *krigagem*, resultou em valores para os três canais e grava a imagem em disco.

Para a resolução de problemas de reescalonamento foi necessário acrescentar um módulo que reconvertisse a matriz correspondente à imagem que se pretende *krigar* num ficheiro onde os *pixels* fossem redistribuídos pela matriz ampliada.

4.3.3.7. Simulação de imagens

Tal como na *krigagem* deve ser feita a modelação de cada canal seguida da simulação do canal respectivo, sendo os resultados finais obtidos através da combinação destas simulações.

No caso da simulação de imagens, o utilizador deverá realizar a seguinte sequência de operações:

1º Obter o modelo para um canal escolhido;

2º Simular os dados para o canal referido;

3º Proceder da mesma forma para o segundo e para o terceiro canal escolhido;

4º Após ser processada a terceira simulação, o sistema, se solicitada uma apresentação gráfica, recompõe a imagem que, após a simulação, resultou em valores para os três canais e grava a imagem em disco

5º Se for indicado pelo utilizador que pretende a variância da simulação, após serem efectuadas as realizações solicitadas pelo módulo *sgsim* do GSLIB [DJ 98], o sistema procederá da seguinte forma:

Para cada canal:

- É calculada a variância de todos os valores obtidos nas realizações para cada localização (x,y);

- Cada variância é armazenados na mesma localização (x,y) numa estrutura intermédia;
- A variância da amostra para o canal será a média das variâncias de todas as localizações (x,y).

4.3.4. Principais estruturas intermédias

As principais estruturas intermédias de dados utilizadas são de dois tipos:

- Estruturas onde são armazenados os dados de entrada e todos os cálculos efectuados possibilitando a sua reutilização durante uma sessão. Estas estruturas são de alocação dinâmica do tipo *CArray* (vector) e têm grande peso na memória.
- Ficheiros de parâmetros, utilizados como transferência de informação entre módulos executáveis não concebidos no âmbito deste sistema (*krigagem* e simulação).

De seguida descrevem-se, de forma genérica, as estruturas utilizadas, quer do tipo vector, quer do tipo ficheiro.

A) Vectores

Considerando os tipos de ficheiros de entrada foi necessário, para além do armazenamento dos dados de entrada na RAM, utilizar vectores para o armazenamento dos cálculos estatísticos efectuados, designadamente, nos casos de gráficos direccionais, de superfície e de dispersão.

Os vectores para o armazenamento dos cálculos estatísticos efectuados terão dois tipos de estruturas distintas correspondentes ao caso das coordenadas implícitas e explícitas.

B) Ficheiros

Na fase de cálculo de estatísticos, são criados ficheiros com o *design* de entrada necessário para o GSLIB ([DJ 98]), ou seja, cujo cabeçalho contenha o nome do ficheiro, o nº de

variáveis e as designações das variáveis (no caso das coordenadas não serem explícitas, é ainda necessário gerá-las).

Também, durante a fase de cálculo dos estatísticos, são criados ficheiros contendo os dados que pertencem a cada direcção para se permitir a realização posterior do teste de validação cruzada. No caso das imagens, estes ficheiros contêm os dados para cada canal onde os nomes principais são sufixados, por um caracter (R, G ou B).

Por exemplo, se o ficheiro que contem a imagem se chamar *arrabida.bmp* os ficheiros que contêm cada um dos canais chamar-se-ão “*arrabidaR.dat*”, “*arrabidaB.dat*” e “*arrabidaG.dat*”. Os ficheiros que serão usados nas validações cruzadas, por exemplo, o ficheiro respeitante ao canal *Blue* na direcção **0°** será “*arrabidaCrossB0.dat*”.

Na fase da *krigagem* e, quando o objectivo for a resolução de problemas de reescalonamento, os 3 ficheiros com formato adequado ao GSLIB ([DJ 98]) serão reconvertidos pois os *pixels* de entrada terão que ser redistribuídos para o reescalonamento pretendido.

4.3.5. Bibliotecas e módulos importados

Para a construção deste sistema foram utilizadas as seguintes bibliotecas e importados os seguintes objectos listados seguidamente:

CxImage - CxImage é uma biblioteca publica para processamento e conversão de imagem ([HP CXIMG]).

2D Graph ActiveX Control - Um controlo ActiveX, frequentemente referenciado como OCX, é um *software* (componente) que não pode ser executado por si próprio podendo apenas ser integrado em ambientes Microsoft® Visual ou Microsoft® Internet Explorer.

Este OCX ([HP OCX]), de uso livre, possibilita as representações gráficas a 2D através de linhas e pontos. Foi necessário adaptá-lo às necessidades do sistema, não só na produção de gráficos por linhas e pontos que melhor se adaptassem aos requisitos do sistema desenvolvido, como também para possibilitar a produção de gráficos de superfície, mapas de cor e histogramas.

TrandomMersenne - Classe para uso livre que possibilita a geração de números aleatórios do tipo Mersenne Twister [HP RAND].

QuickSort() - QuickSort() é um *template* que implementa o algoritmo de ordenação *quicksort* para a classe CArray do Visual C++.

Módulos executáveis utilizados

KT3D COKB3D e SGSIM

Foram utilizados os módulos executáveis do GSLIB [DJ 98] respeitantes ao *kriging* (módulo *kt3d*), *cokrigagem colocalizada* (módulo *cokb3d*) e simulação (módulo *sgsim*). Estes programas escritos em Fortran foram alterados para permitir estruturas de dados com 80000 elementos bem como para efectuar a *cokrigagem colocalizada* (alterações efectuadas no GEOMS [HP GEOMS]).

5. Estudo de casos

Para ilustrar a abordagem metodológica apresentada nos capítulos anteriores e a utilização do programa informático *MODIMAV*, apresentam-se três exemplos práticos e distintos onde são abordados o reescalonamento de imagens, a inferência espacial usando imagem como informação secundária e a captura de padrões para a sua posterior reprodução por simulação geoestatística.

O primeiro exemplo diz respeito ao reescalonamento de imagens por ampliação. A utilização da estimação por *krigagem*, com um variograma ajustado automaticamente, para o reescalonamento da imagem constitui uma alternativa relativamente aos processos de interpolação clássicos.

No segundo exemplo faz-se a estimação e simulação de uma variável principal (pluviosidade) com medições pontuais dessa variável e considerando uma imagem secundária relacionada com a variável principal (modelo digital do terreno). Trata-se de um problema de inferência espacial, onde se passa de amostras pontuais para uma imagem da área em estudo com uma resolução espacial que pode ser previamente estabelecida e alterada. Neste exemplo comparam-se as metodologias de estimação por *krigagem* e simulação geoestatística.

No terceiro exemplo faz-se a reprodução de padrões extraídos de imagens de um determinado objecto (neste exemplo foram consideradas chapas de rocha ornamental) por simulação geoestatística.

5.1. Reescalonamento de imagens

Para um problema de reescalonamento isotrópico, ou seja, quando a taxa de ampliação a aplicar na imagem é da mesma grandeza em todas as direcções, deve proceder-se da seguinte forma:

- desagregar a imagem nos canais de cor que a constituem;

- calcular o variograma experimental de cada canal de cor e ajustar um modelo teórico a cada canal;
- aumentar a imagem original na percentagem requerida e estimar por *krigagem* normal os valores que ficam por preencher após o reescalonamento;
- aglutinar novamente os canais de cor e representar a imagem.

Para ilustrar esta metodologia de reescalonamento foram usados três exemplos de ampliação de 30% imagens para imagens que se consideram representativas do tipo de problemas que é habitual encontrar na prática: uma fotografia em tons de cinzento (três canais de cor - RGB, respectivamente, vermelho, verde e azul todos iguais), uma fotografia a cores (três canais de cor - RGB, respectivamente, vermelho, verde e azul) e a composição RGB de uma imagem de satélite Landsat TM (também três canais de cor – RGB).

A metodologia aplicada encontra-se sintetizada nos diagramas das figuras 75 e 76, respectivamente para imagens com uma banda (visualização em tons de cinzento) e três bandas (visualização como uma composição colorida RGB).

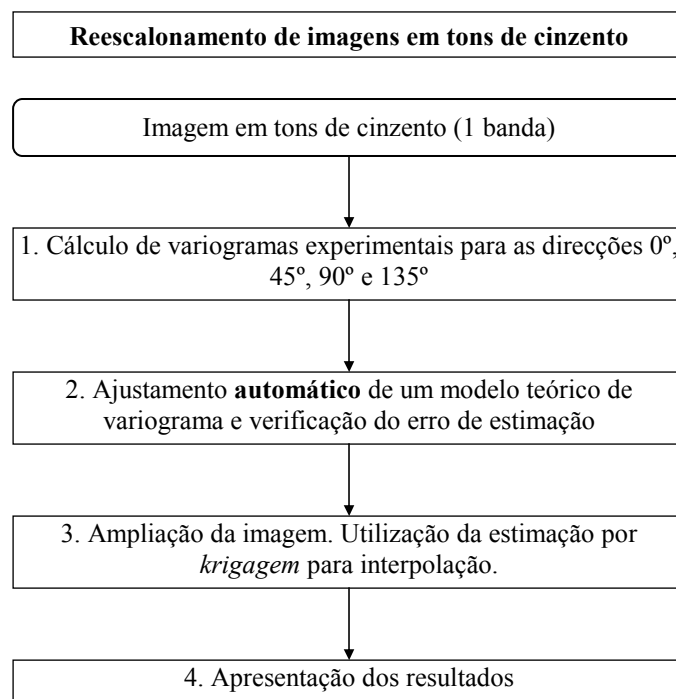


Fig. 75 Diagrama da metodologia proposta para a ampliação de imagens com uma banda

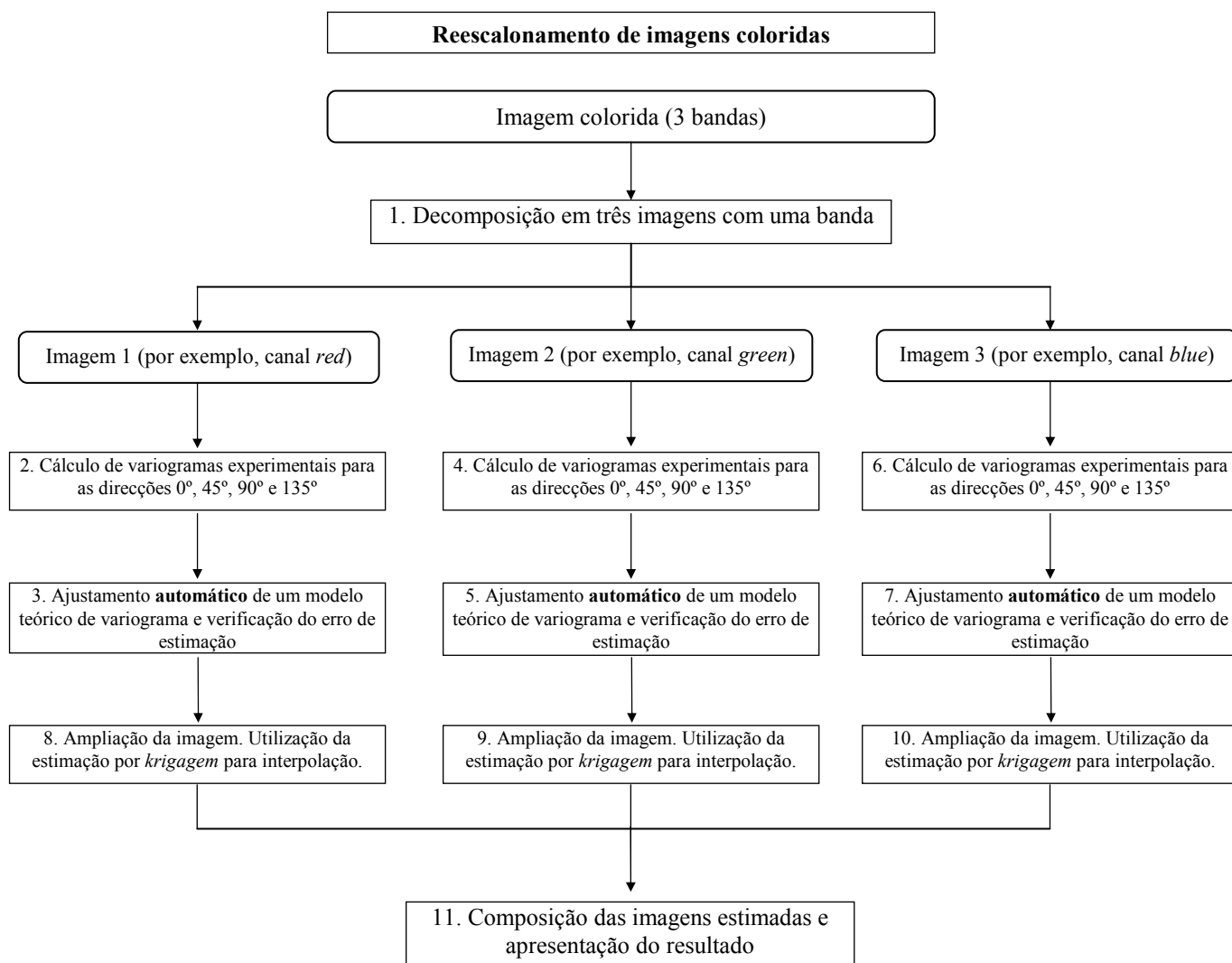


Fig. 76 Diagrama da metodologia proposta para a ampliação de imagens com três bandas

5.1.1. Fotografia em tons de cinzento

Neste caso pretende-se ampliar uma fotografia digitalizada (fig. 77) escolhida de modo a conter diferentes texturas e formas com formato *bitmap-grayscale* com a dimensão inicial 150x176 que, sendo caracterizada pela existência do mesmo valor nos três canais, poder-se-á ampliar usando apenas um modelo teórico que será replicado para os três canais.



Fig. 77 Fotografia em tons de cinzento

O primeiro passo consistiu no cálculo dos variogramas experimentais para as direcções de 0°, 45°, 90° e 135° e ajustamento do respectivo modelo teórico.

Os variogramas foram calculados com 150 passos de um *pixel* para as direcções de azimute 0° e 90°. Os resultados do variograma experimental e o modelo teórico do ajustamento automático estão reportados nas figuras 78 e 79. Na figura 78 mostra-se o modelo considerando a imagem isotrópica (variograma médio omnidireccional) e na figura 79 os variogramas para 4 direcções, de 45 em 45 graus.

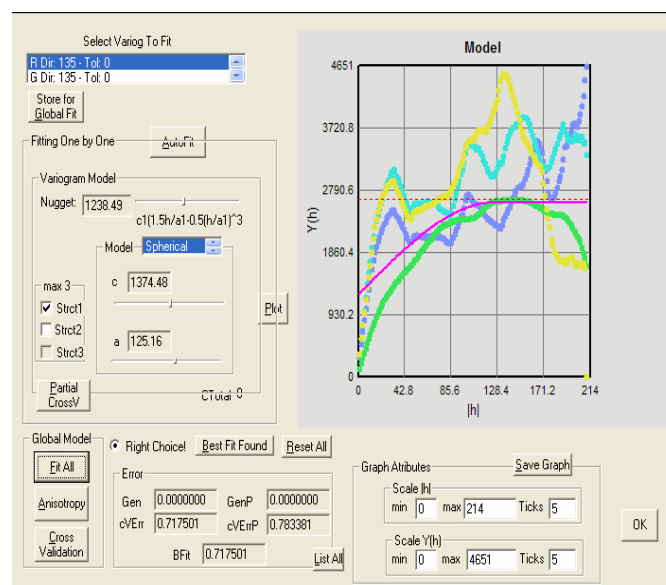


Fig. 78 Variogramas experimentais e modelo teórico omnidireccional

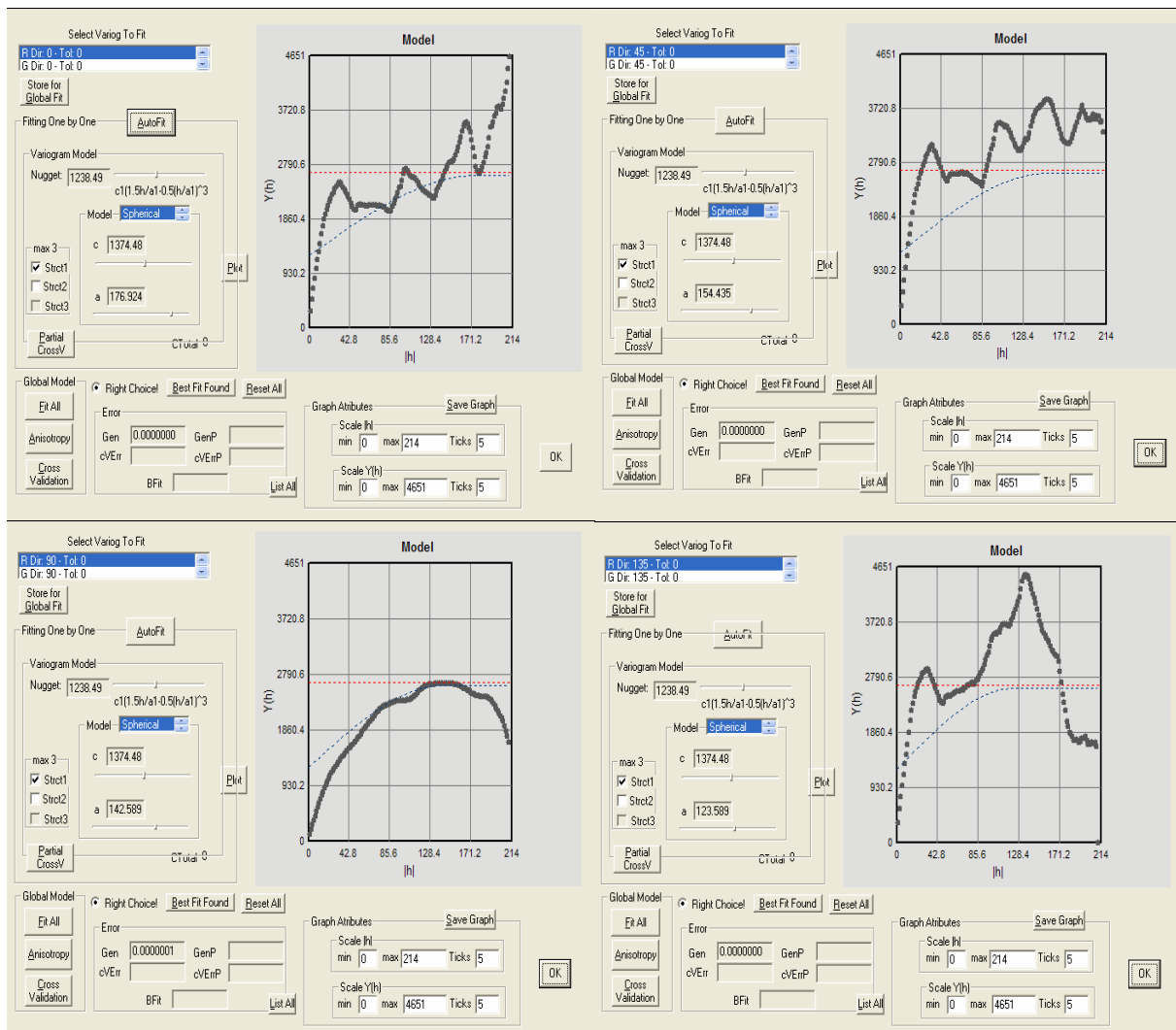


Fig. 79 Variogramas experimentais e modelo teórico, direcção 0°, 45°, 90° e 135°

Para a estimação por *krigagem* utilizaram-se os seguintes parâmetros:

- $n_x = 195$; $n_y = 229$;
- n° amostras (mínimo/máximo): 3; 16;
- raio de pesquisa máximo: 195 *pixels*;
- direcção principal de pesquisa: 0°;
- variograma: modelo teórico com uma estrutura, do tipo esférico: $c_0=1238,49$; $c_1=1374,48$; amplitudes por direcção: $a_1(90^\circ)=142,59$; $a_2(0^\circ)=176,94$

A imagem inicial, a imagem ampliada a 30% com o valor dos *pixels* ainda não estimados e a imagem resultante da estimação por *krigagem* dos *pixels* obtida usando o software aqui desenvolvido podem ser visualizadas na figura 80.

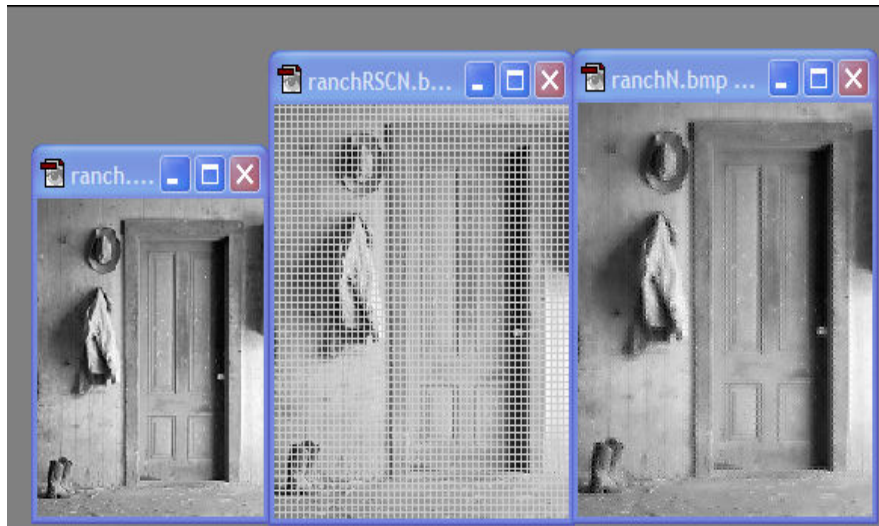


Fig. 80 Foto inicial, ampliação a 30% com *pixels* “sem cor” e resultado da *krigagem*

A figura seguinte (fig. 81) ilustra os resultados equivalentes obtidos no *Adobe Photoshop®* utilizando três métodos descritos na secção 2.4.6, designadamente, método da vizinhança mais próxima, interpolação bilinear e interpolação bicúbica.

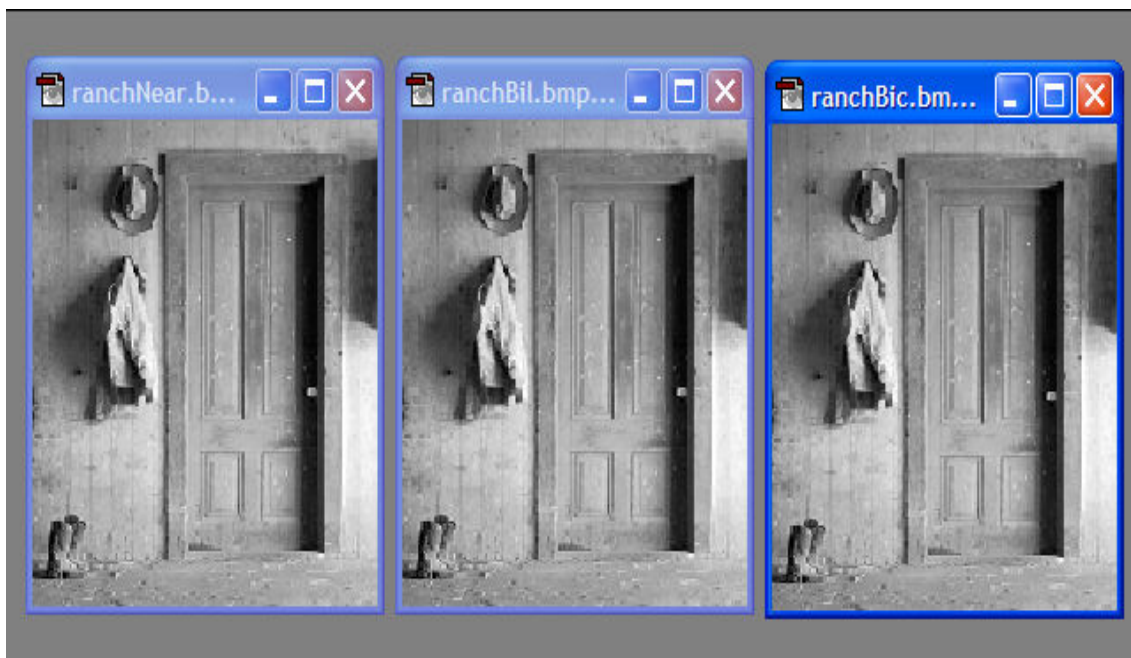


Fig. 81 (da esquerda para a direita) Ampliações a 30% obtidos pelo método da vizinhança mais próxima, interpolação bilinear e bicúbica

5.1.2. Fotografia a cores

Neste caso pretende-se ampliar uma fotografia digitalizada (fig. 82) de padrão mais regular do que o considerado anteriormente com formato *bitmap* – 8 bits por canal e com dimensão de 190x142.

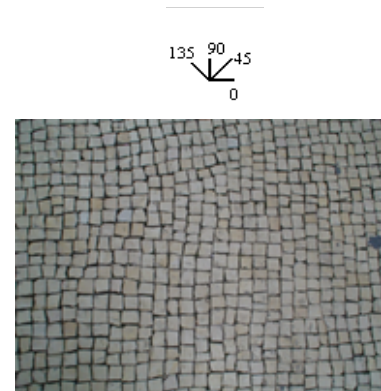


Fig. 82 Fotografia a cores

Para cada canal, e de forma independente, foram calculados os variogramas experimentais para as direcções 0°, 45°, 90° e 135° e ajustados automaticamente modelos teóricos de variograma.

Os variogramas experimentais foram calculados com 142 passos de 1 *pixel* para as direcções de azimute 0° e 90°. Os resultados obtidos pelo ajustamento automático para o três canais de cor (*red*, *green* e *blue*) estão reportados nas figuras seguintes (figuras 83 a 88).

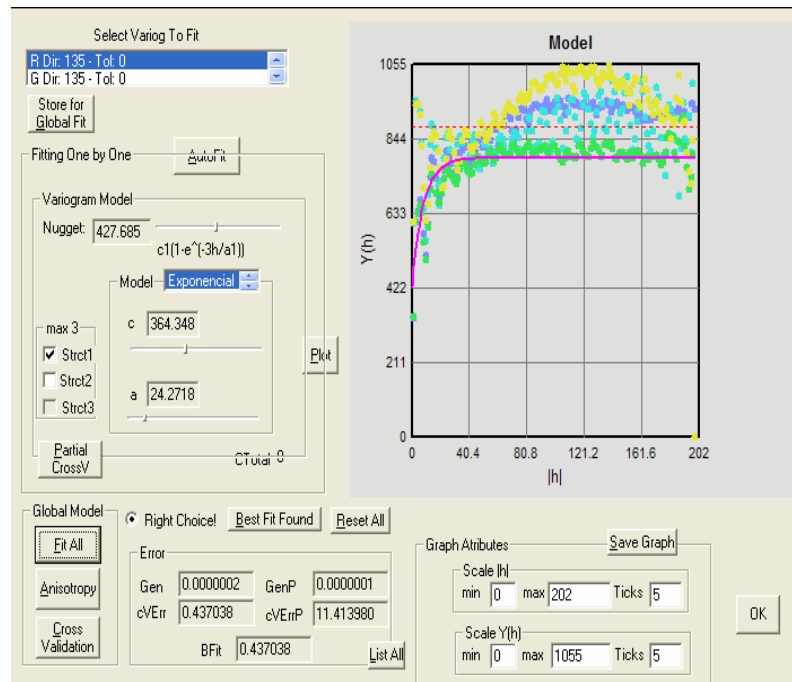


Fig. 83 Variogramas experimentais e modelo teórico, canal *Red*, modelo omnidireccional

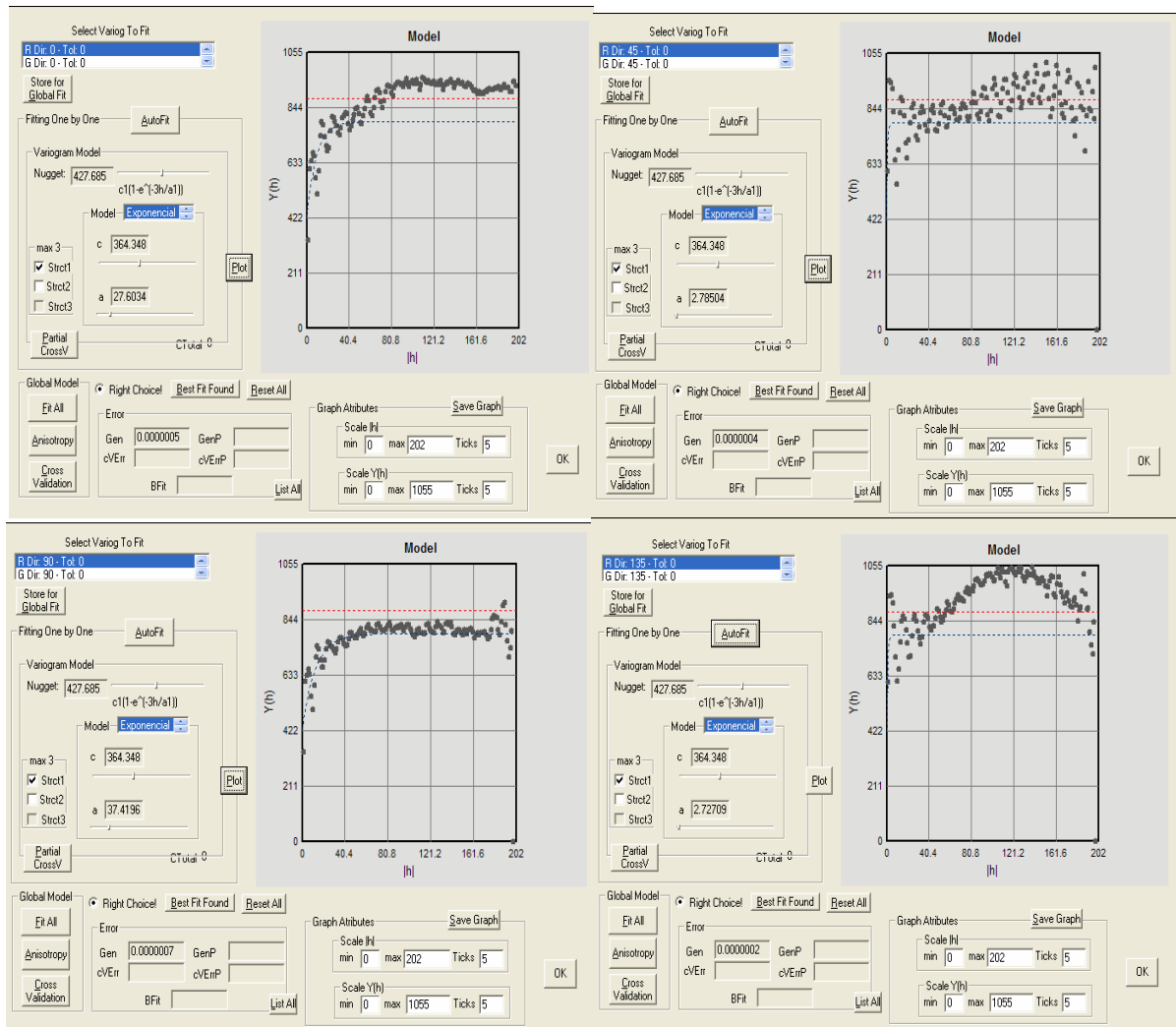


Fig. 84 Variogramas experimentais e modelo teórico, canal Red, direcção 0°, 45°, 90° e 135°

Para a estimação por *krigagem* do canal de cor Red utilizaram-se os seguintes parâmetros:

- $n_x = 247$; $n_y = 185$;
- n° amostras (mínimo/máximo): 3; 16;
- raio de pesquisa máximo: 247 pixels;
- direcção principal de pesquisa: 90°;
- variograma: modelo teórico com uma estrutura, do tipo exponencial: $c_0=427,69$;
 $c_1=364,35$; amplitudes por direcção: $a_1(0^\circ)=27,60$; $a_2(90^\circ)=37,42$

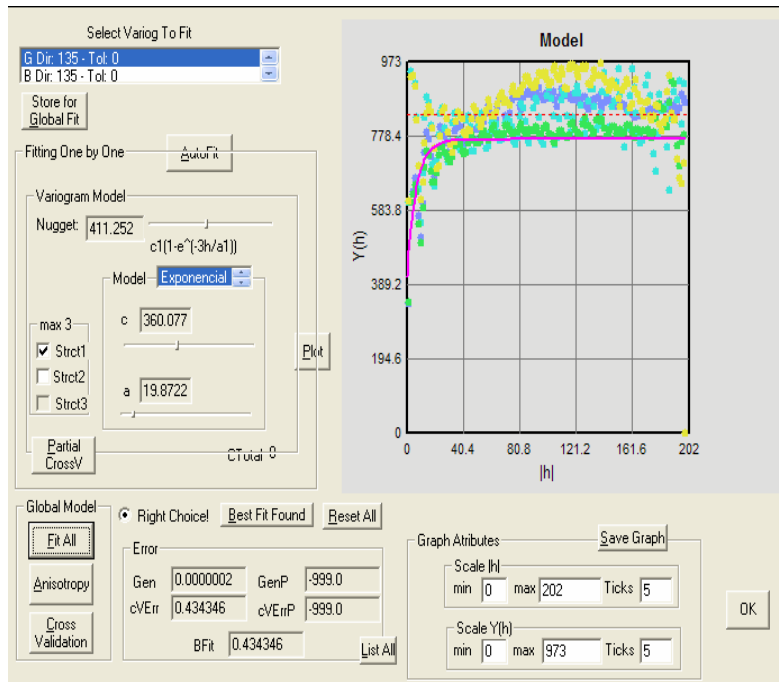


Fig. 85 Variogramas experimentais e modelo teórico, canal *Green*, modelo omnidireccional

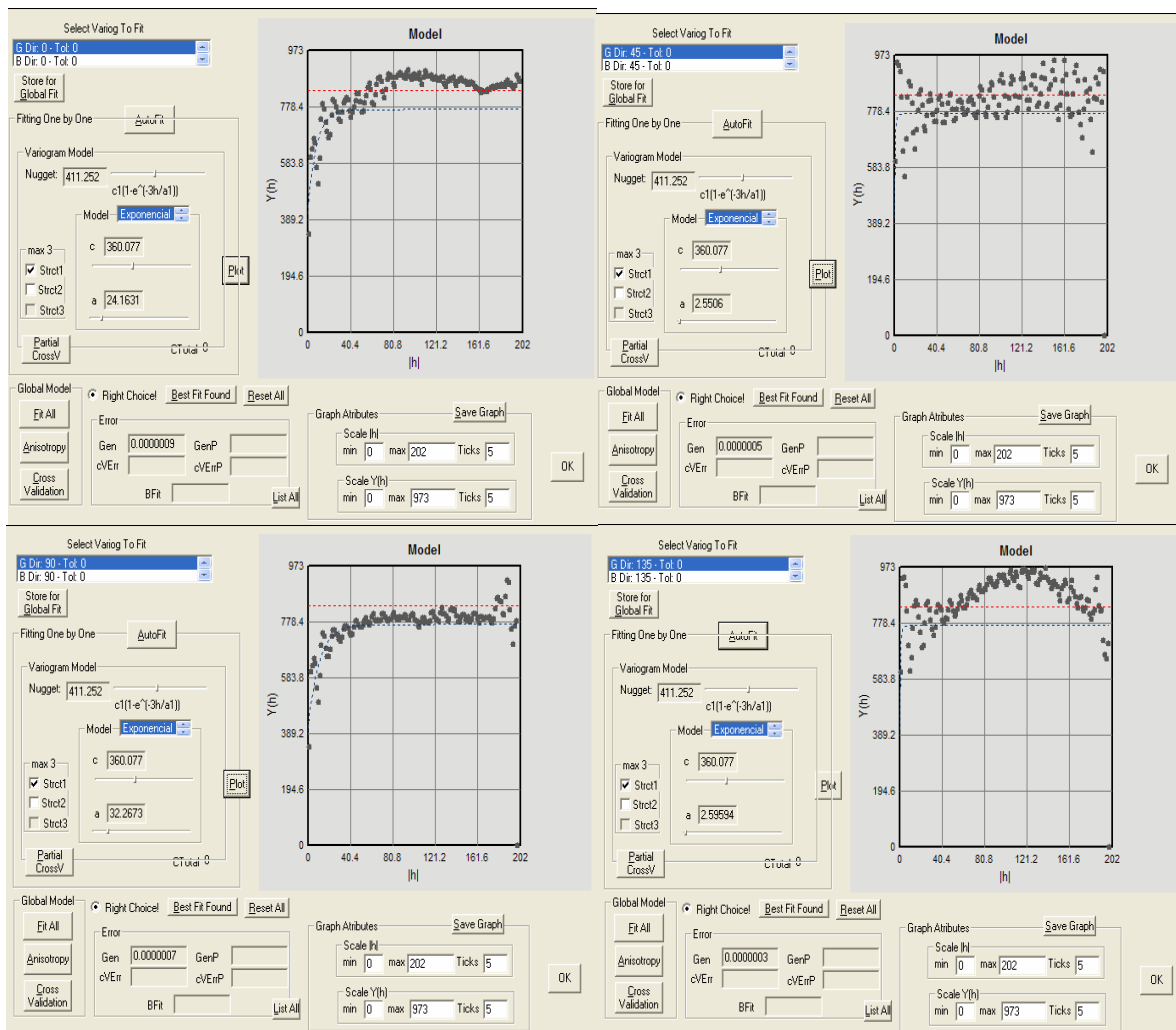


Fig. 86 Variogramas experimentais e modelo teórico, canal *Green*, direcção 0°, 45°, 90° e 135°

Para a estimação por *krigagem* do canal de cor *Green* utilizaram-se os seguintes parâmetros:

- $n_x = 247$; $n_y = 185$;
- nº amostras (mínimo/máximo): 3; 16;
- raio de pesquisa máximo: 247 *pixels*;
- direcção principal de pesquisa: 90°;
- variograma: modelo teórico com uma estrutura, do tipo exponencial: $c_0=411,25$;
 $c_1=360,08$; amplitudes por direcção: $a_1(0^\circ)=24,16$; $a_2(90^\circ)=32,27$

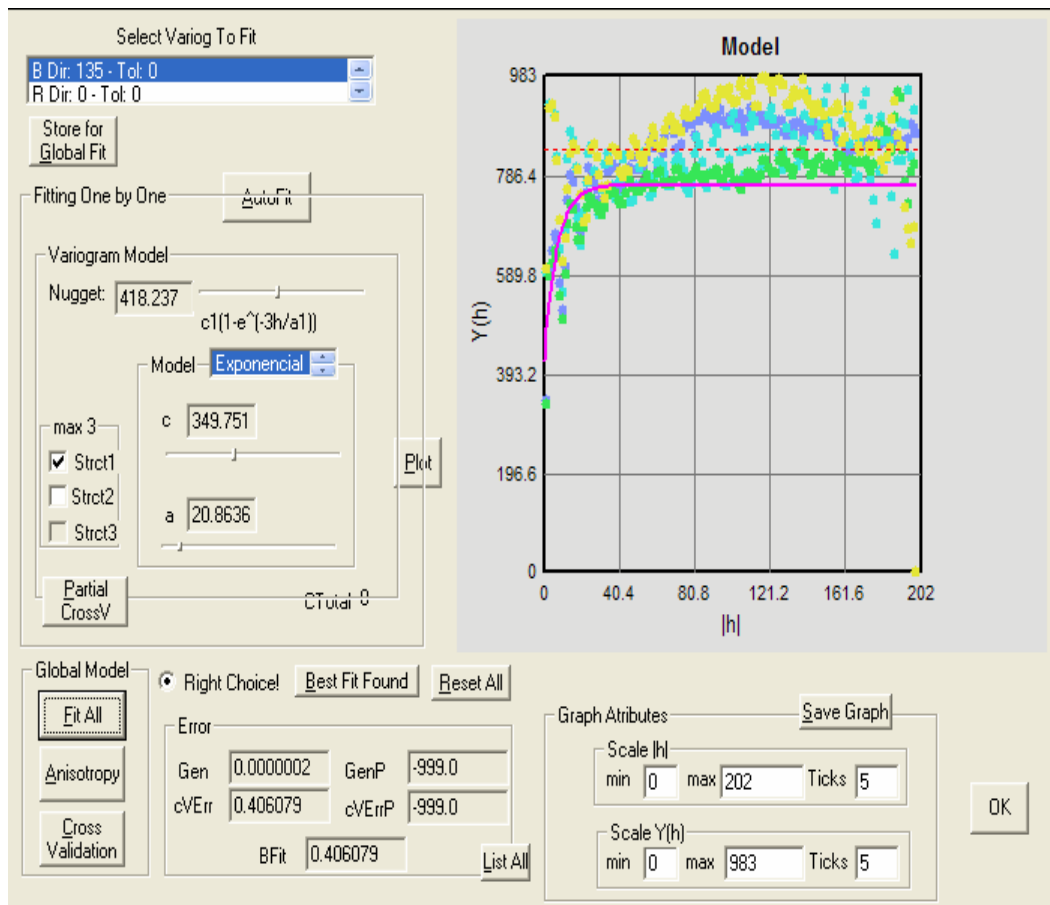


Fig. 87 Variogramas experimentais e modelo teórico, canal *Blue*, modelo omnidireccional

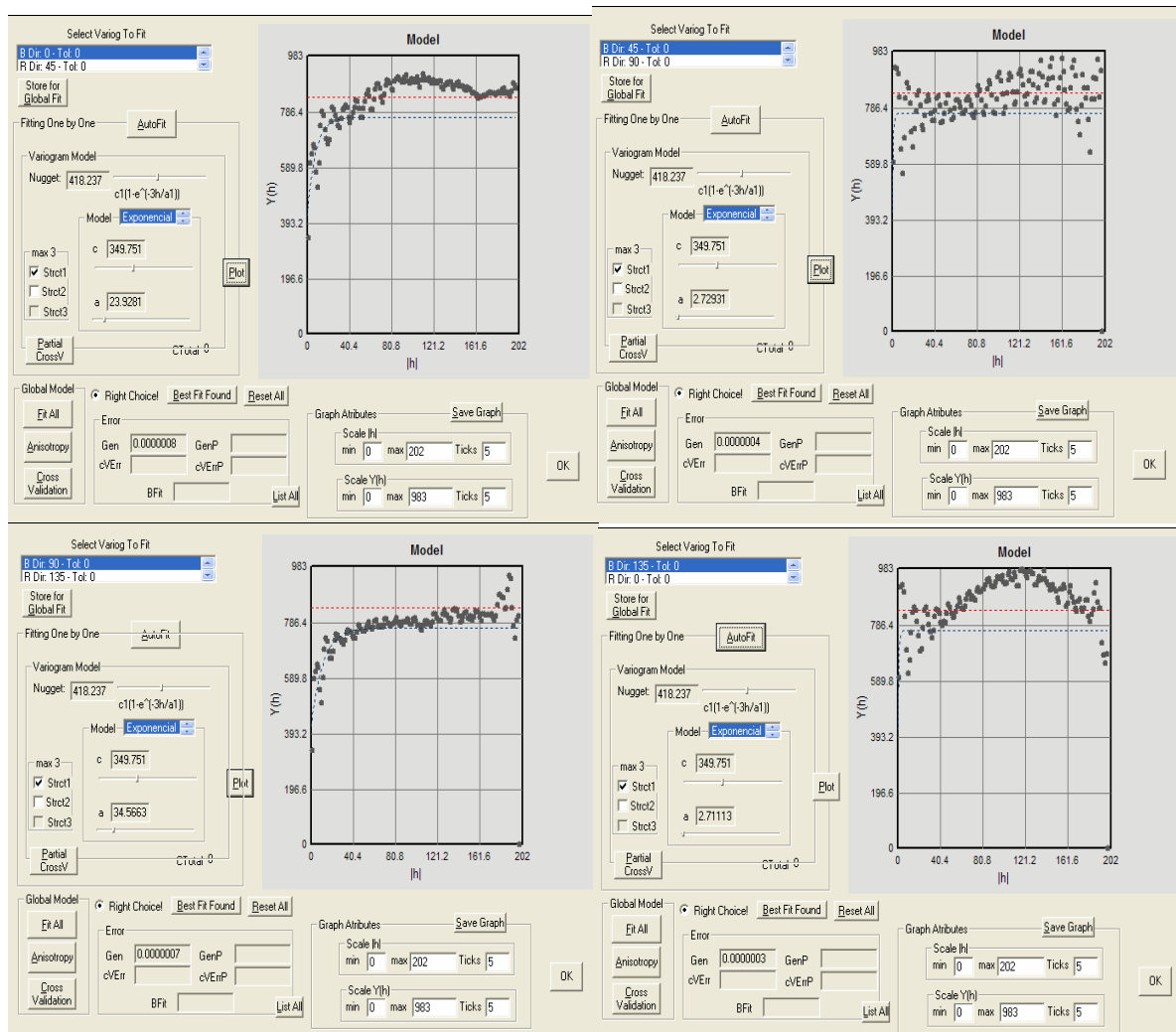


Fig. 88 Variogramas experimentais e modelo teórico, canal Blue, direcção 0°, 45°, 90° e 135°

Para a estimação por *krigagem* do canal de cor Blue utilizaram-se os seguintes parâmetros:

- $n_x = 247$; $n_y = 185$;
- nº amostras (mínimo/máximo): 3; 16;
- raio de pesquisa máximo: 247 pixels;
- direcção principal de pesquisa: 90°;
- variograma: modelo teórico com uma estrutura, do tipo exponencial: $c_0=418,24$; $c_1=349,75$; amplitudes por direcção: $a_1(0^\circ)=23,93$; $a_2(90^\circ)=34,57$

A imagem inicial, a imagem ampliada a 30% com o valor dos *pixels* ainda não estimados e a imagem resultante da estimação por *krigagem* dos *pixels* obtida usando o software aqui desenvolvido podem ser visualizadas na figura 89.

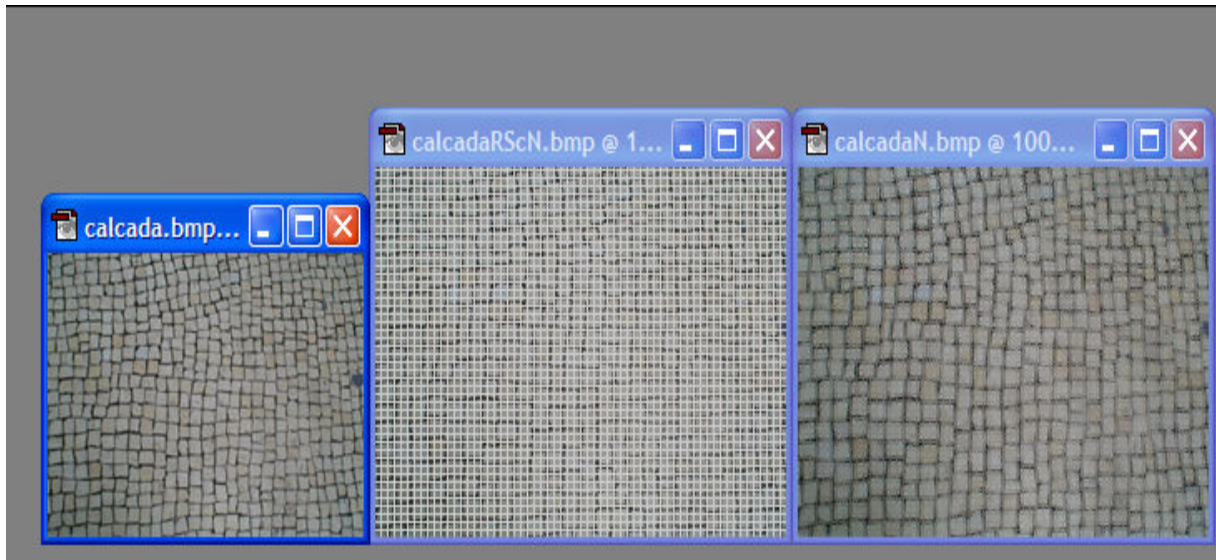


Fig. 89 Foto inicial, ampliação a 30% com *pixels* “sem cor” e resultado da *krigagem*

A figura seguinte (fig. 90) ilustra os resultados equivalentes obtidos no *Adobe Photoshop®* utilizando três métodos descritos na secção 2.4.6, designadamente, método da vizinhança mais próxima, interpolação bilinear e interpolação bicúbica.



Fig. 90 (da esquerda para a direita) Ampliações a 30% com o método da vizinhança mais próxima, interpolação bilinear e interpolação bicúbica

5.1.3. Imagem de satélite

Neste caso, pretende-se ampliar uma zona de uma imagem de satélite LANDSAT TM (fig. 91) fornecida com formato *bitmap* – RGB com 8 bits por canal e com dimensão de 142x143. Esta imagem é típica das que se obtêm por detecção remota e também aqui é necessário calcular os variogramas para as direcções de 0°, 45°, 90° e 135° e os respectivos modelos teóricos para cada canal.

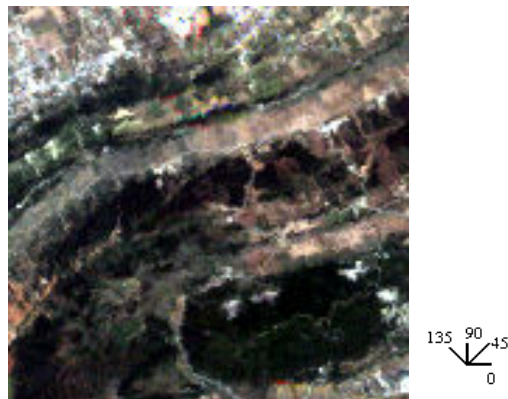


Fig. 91 Composição RGB de 3 bandas de uma imagem de satélite Landsat TM

Os variogramas experimentais foram calculados com 142 passos de 1 *pixel* para as direcções de azimute 0° e 90°. Os resultados obtidos pelo ajustamento automático para o três canais de cor (*red*, *green* e *blue*) estão reportados nas figuras seguintes (figuras 92 a 97).

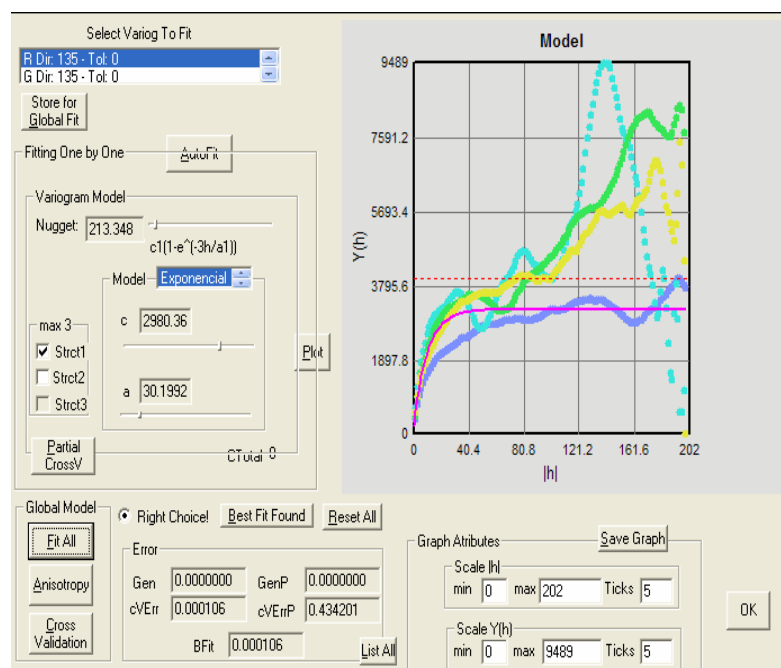


Fig. 92 Variogramas experimentais e modelo teórico, canal *Red*, modelo omnidireccional

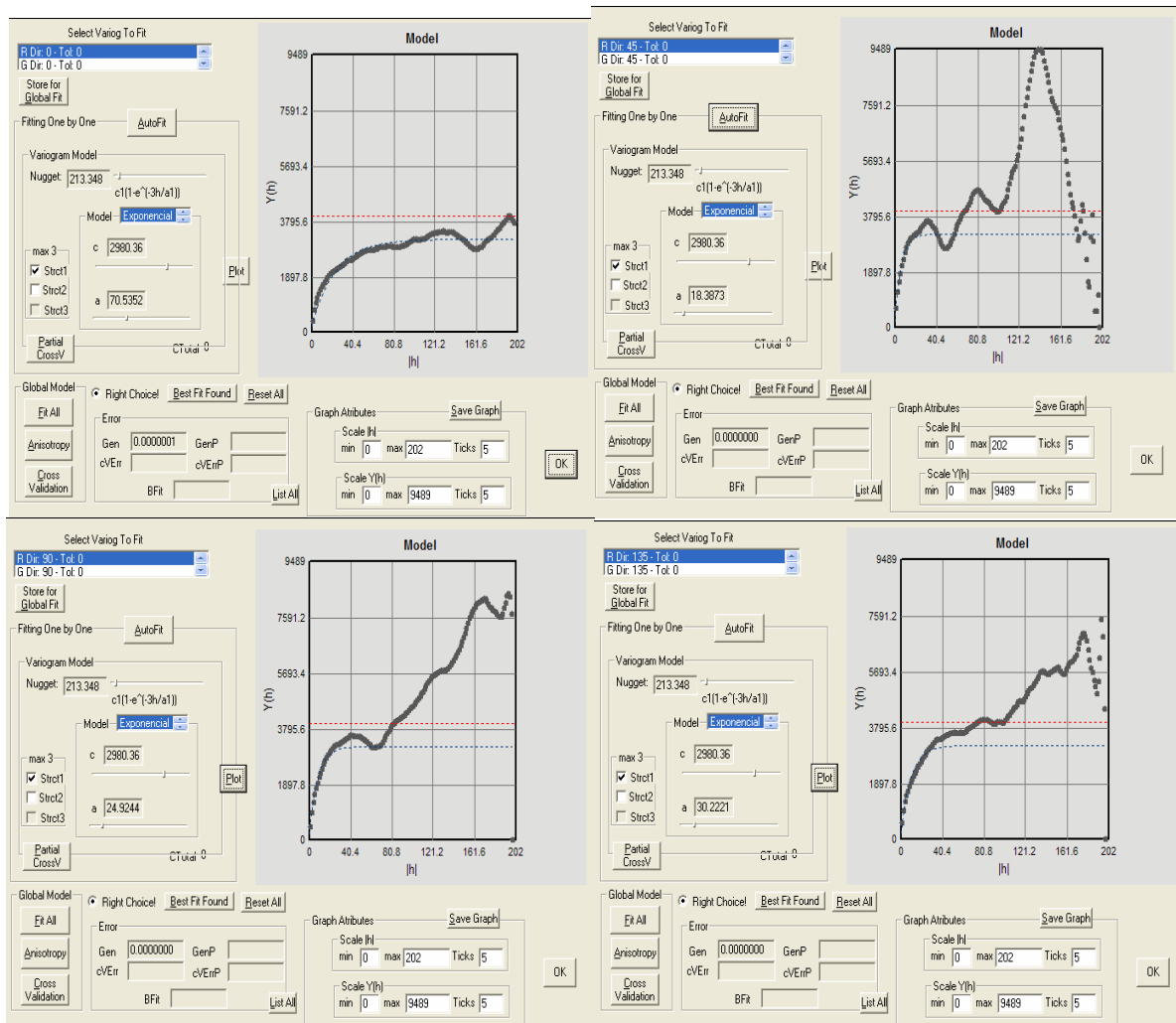


Fig. 93 Variogramas experimentais e modelo teórico, canal Red, direcção 0°, 45°, 90° e 135°

Para a estimação por *krigagem* do canal de cor Red utilizaram-se os seguintes parâmetros:

- $n_x = 185$; $n_y = 186$;
- nº amostras (mínimo/máximo): 3; 16;
- raio de pesquisa máximo: 185 pixels;
- direcção principal de pesquisa: 90°;
- variograma: modelo teórico com uma estrutura, do tipo exponencial: $c_0=213,35$; $c_1=2980.36$; amplitudes por direcção: $a_1(90^\circ)=24.92$; $a_2(0^\circ)=70.54$

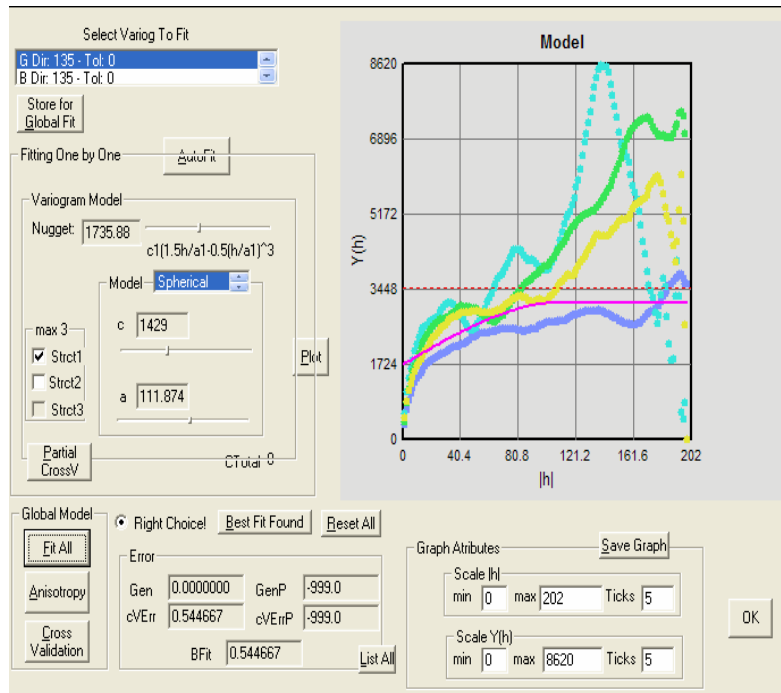


Fig. 94 Variogramas experimentais e modelo teórico, canal *Green*, modelo omnidireccional

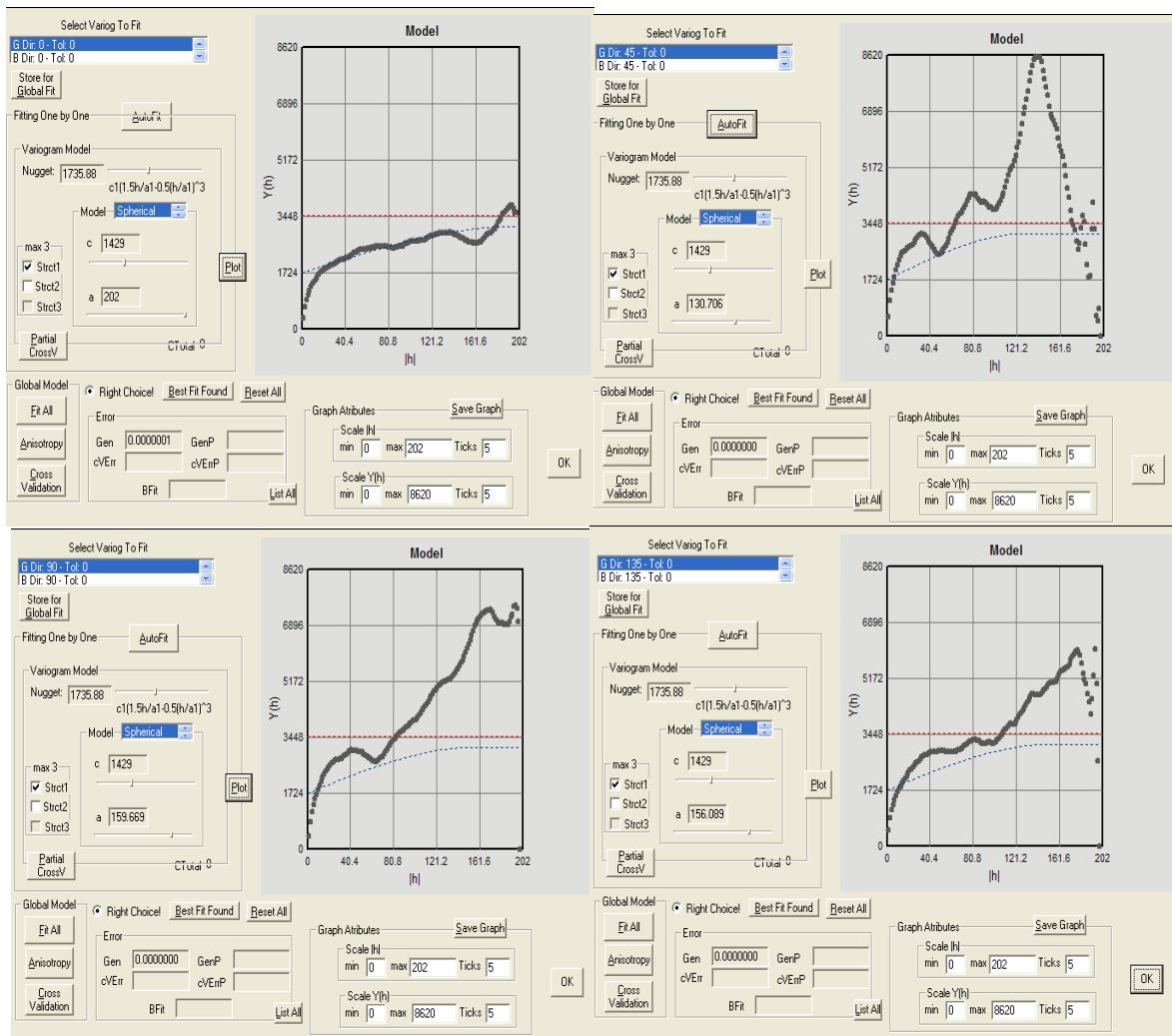


Fig. 95 Variogramas experimentais e modelo teórico, canal *Green*, direcção 0°, 45°, 90° e 135°

Para a estimação por *krigagem* do canal de cor *Green* utilizaram-se os seguintes parâmetros:

- $n_x = 185$; $n_y = 186$;
- nº amostras (mínimo/máximo): 3; 16;
- raio de pesquisa máximo: 185 *pixels*;
- direcção principal de pesquisa: 0° ;
- variograma: modelo teórico com uma estrutura, do tipo exponencial: $c_0=1735,88$; $c_1=1429,00$; amplitudes por direcção: $a_1(0^\circ)=202,00$; $a_1(90^\circ)=159,67$

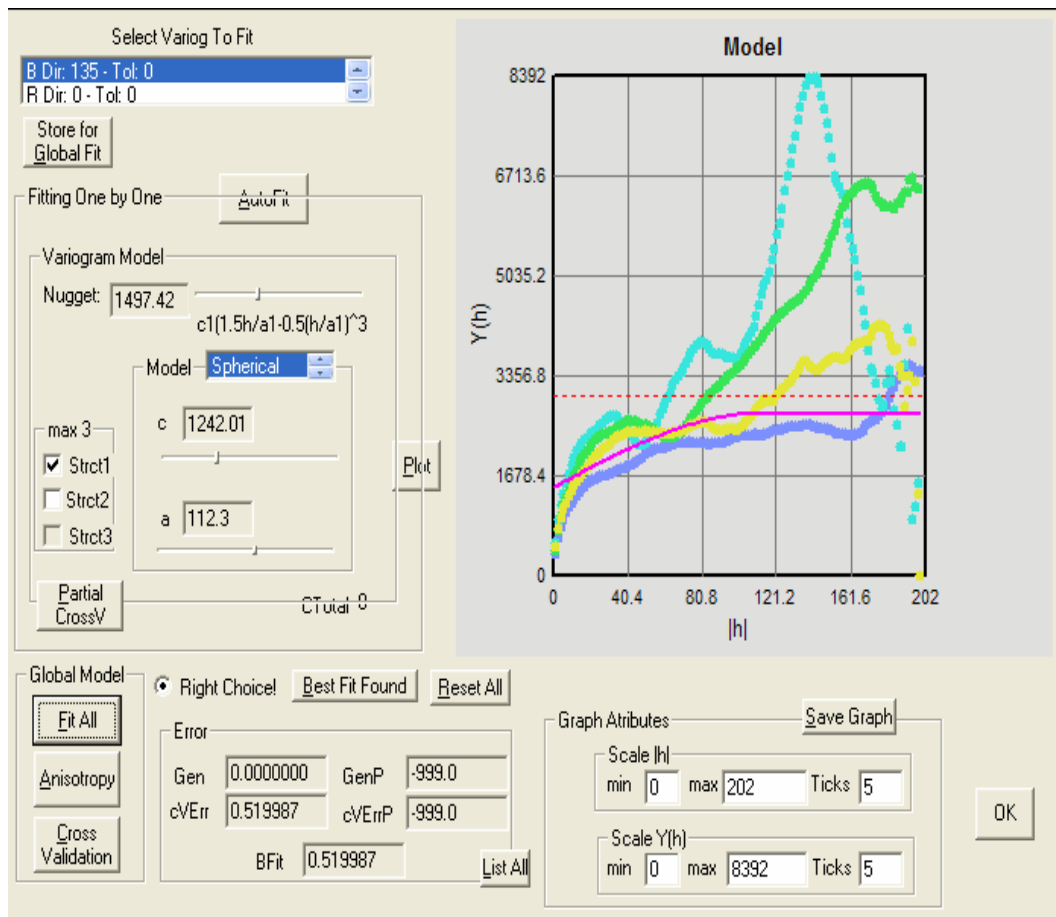


Fig. 96 Variogramas experimentais e modelo teórico, canal *Blue*, modelo omnidireccional

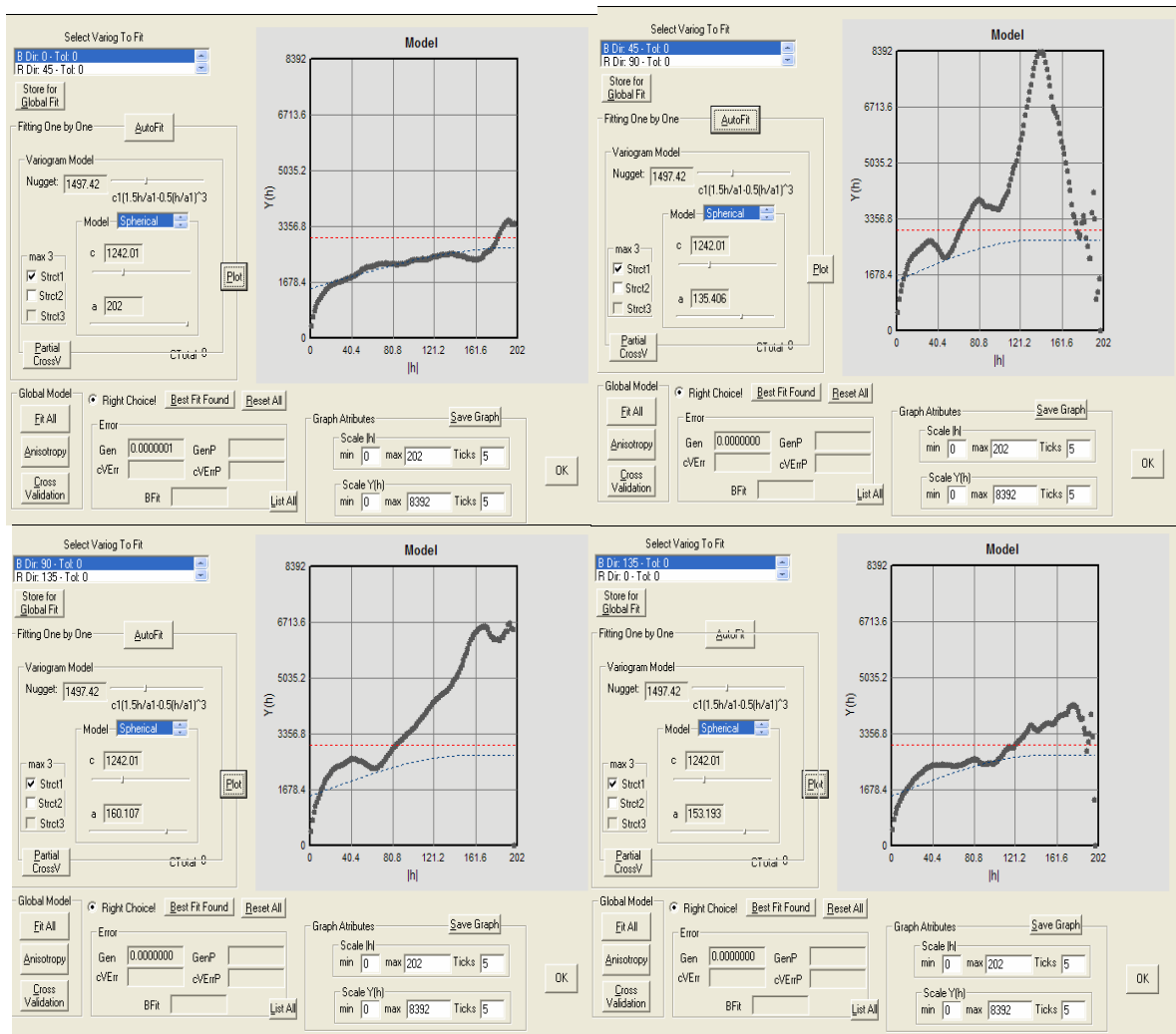


Fig. 97 Variogramas experimentais e modelo teórico, canal Blue, direcção 0°, 45°, 90° e 135°

Para a estimação por *krigagem* do canal de cor Blue utilizaram-se os seguintes parâmetros:

- $n_x = 185$; $n_y = 186$;
- nº amostras (mínimo/máximo): 3; 16;
- raio de pesquisa máximo: 185 pixels;
- direcção principal de pesquisa: 0°;
- variograma: modelo teórico com uma estrutura, do tipo esférico: $c_0=1497,42$;
 $c_1=1242,00$; amplitudes por direcção: $a_1(90^\circ)= 160,11$; $a_2(0^\circ)=202,00$

A imagem inicial, a imagem ampliada a 30% com o valor dos *pixels* ainda não estimados e a imagem resultante da estimação por *krigagem* dos *pixels* obtida usando o software aqui desenvolvido podem ser visualizadas na figura 98.

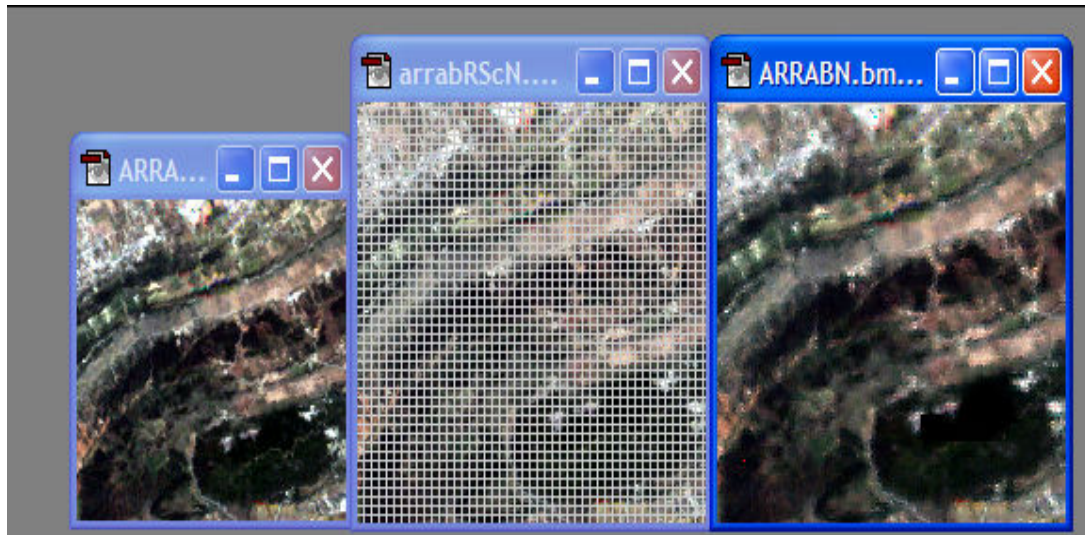


Fig. 98 Imagem inicial, ampliação a 30% com *pixels* “sem cor” e resultado da *krigagem*

A figura seguinte (fig. 99) ilustra as ampliações realizadas no *Adobe Photoshop®* utilizando os três métodos descritos na secção 2.4.6, designadamente, métodos da vizinhança mais próxima, interpolação bilinear e interpolação bicúbica.

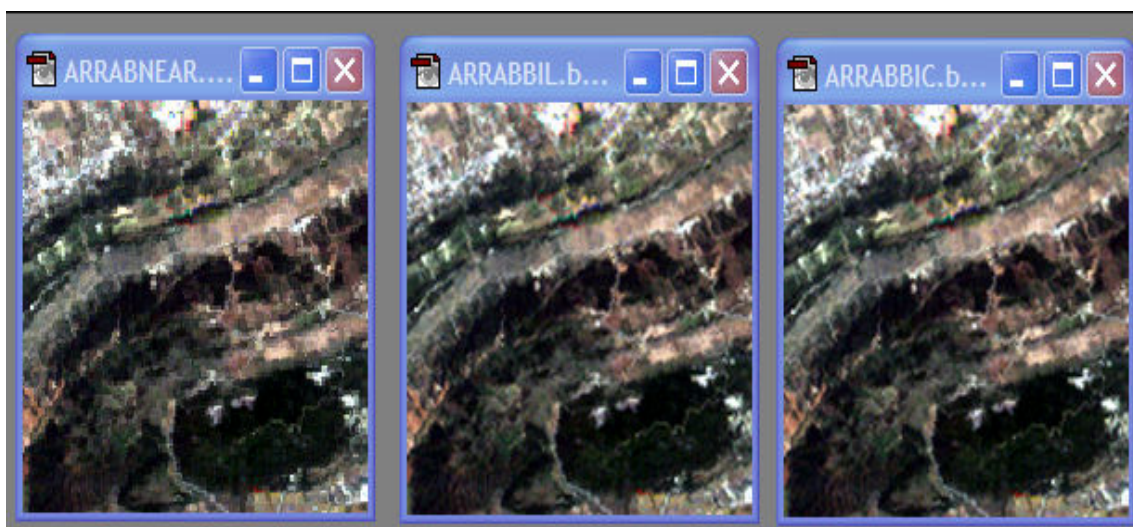


Fig. 99 (da esquerda para a direita) ampliações a 30% com método da vizinhança mais próxima, interpolação bilinear e interpolação bicúbica

5.1.4. Análise dos resultados

Este exemplo permite ilustrar a aplicação da estimação geoestatística no domínio do reescalonamento de imagens como alternativa aos processos de interpolação tradicionais (não geoestatísticos) tal como é sugerido em [HP MICRO]¹¹.

Assim, escolheram-se imagens referentes a dois domínios aplicativos: o domínio do design gráfico na relação texto/foto onde, no efeito visual dominante, devem prevalecer as vertentes subjectivas perceptuais e os conceitos estéticos do *designer* (caso das duas primeiras imagens) e a utilização do *zoom* para fornecer um maior detalhe em zonas de imagem (caso da terceira imagem, cujos canais RGB foram extraídos de uma imagem de satélite).

Após a obtenção de imagens ampliadas estimadas por *krigagem*, com um variograma ajustado automaticamente, as imagens originais foram ampliadas utilizando os métodos de interpolação não geoestatísticos presentes no *Adobe Photoshop*® para possibilitar a comparação do método proposto com os métodos “tradicionais”.

Ressalvando que as imagens impressas apenas têm uma função ilustrativa na comparação das imagens obtidas observa-se o seguinte:

- no caso da imagem a preto e branco (figuras 80 e 81 da página 135) é notório na imagem estimada por *krigagem* um menor contraste e recorte (fig. 80 *versus* fig. 81), inclusive, nas zonas de sombra;
- no caso da imagem a cores (figuras 89 e 90 da página 141) o tema padrão “calçada à portuguesa” aparece mais evidente nas imagens mais contrastadas que são, respectivamente, as resultantes de interpolação pelo método da vizinhança mais próxima e bicúbica;
- no caso da imagem de satélite (figuras 98 e 99 da página 147) as imagens obtidas têm um aspecto visual semelhante observando uma maior exactidão/nitidez de pormenor na imagem obtida por *krigagem*.

¹¹ “...*The additional detail is not real, however, so fractal interpolation is not appropriate for measurement purposes. When accurate measurements are the goal, more precise algorithms such as the **kriging** technique should be employed. ...*”

No sentido de tentar quantificar de alguma forma os aspectos subjectivos atrás enunciados, realizou-se um estudo comparativo entre a *krigagem* e os estimadores “tradicionais” citados. Assim, para realizar a comparação dos vários estimadores, procedeu-se da seguinte forma:

- retirou-se a linha 100 da imagem constante na figura 77 (pag. 133). Os *pixels* desta linha encontram-se assinalados a negro na fig. 100;
- reconstruiu-se a referida linha utilizando os interpoladores designados como “tradicionais”¹², bem como o estimador geoestatístico;
- calculou-se o erro quadrático médio (EQM) entre os valores retirados (mas conhecidos) e os estimados para cada tipo de interpolação realizada.



Fig. 100 Fotografia com realce dos *pixels* da linha 100

Como pode verificar-se no resumo seguinte (que se apresenta na totalidade no Apêndice I) onde o menor erro de estimação é obtido por estimação geoestatística o que indicia, não obstante a necessidade futura de estudos posteriores mais aprofundados, ser o interpolador mais preciso dos quatro estimadores considerados.

| NearErr | BilinErr | BicErr | KrigErr |
|----------------|-----------------|---------------|----------------|
| 4847 | 4600 | 4613 | 604 |

Neste resumo NearErr representa o EQM relativo à interpolação por vizinhança mais próxima, BilinErr representa o EQM relativo à interpolação bilinear, BicErr representa o EQM relativo à interpolação bicúbica e KrigErr representa o EQM relativo ao método aqui desenvolvido.

¹² Usou-se como linguagem de programação, a linguagem interpretada *noobeed* [HP NOO] que possui métodos numa classe designada por *Image* que implementam os interpoladores: vizinhança mais próxima, bilinear e bicúbico.

5.2. Inferência espacial

Nos processos de inferência espacial utilizando métodos geoestatísticos podem ser utilizados métodos de estimação e de simulação. Os métodos de estimação têm como objectivo a construção de um mapa médio baseado em medidas de continuidade espacial (covariâncias espaciais e variogramas); os métodos de simulação têm como objectivo a construção de várias imagens equiprováveis de um determinado fenómeno, condicionais ao conjunto de dados experimentais, aos estatísticos básicos e às medidas de continuidade espacial observadas pelas amostras.

As imagens obtidas por simulação constituem uma ferramenta essencial na análise estocástica espacial, como, por exemplo, o cálculo da probabilidade de ocorrência de cenários extremos e a avaliação da incerteza espacial de um determinado fenómeno.

Este exemplo é ilustrativo da aplicação das técnicas de inferência por estimação e simulação geoestatística para a geração de imagens aplicadas às ciências da terra e do ambiente. Acresce ainda que a metodologia de geração de imagens, quer por estimação quer por simulação, poderá ser enriquecida sempre que existam outras imagens, ditas secundárias, que estejam correlacionadas com a variável a caracterizar.

O exemplo escolhido diz respeito à caracterização espacial da variável pluviosidade, referido à região do Algarve. A caracterização espacial desta variável é fundamental em inúmeros estudos, de que se destacam a elaboração de cartas de erosão. De todas as variáveis que contribuem para a construção de uma carta de erosão, nomeadamente o índice RUSLE (Revised Universal Soil Loss Equation) [LO 01], a precipitação é a que verdadeiramente apresenta as características de uma variável regionalizada (variável aleatória mas simultaneamente estruturada espacialmente). Assim, para a sua modelação, as técnicas geoestatísticas são as mais adequadas.

Com base em informação meteorológica referente ao período de um ano (precipitação mensal acumulada) proveniente de 36 estações meteorológicas no Algarve (informação principal), o objectivo deste exemplo é gerar uma imagem estimada média e várias imagens simuladas equiprováveis dos quantitativos anuais da precipitação.

Na geração da imagem estimada e das imagens simuladas consideram-se os valores do modelo digital de terreno como informação secundária, correlacionável com a precipitação média anual. A contribuição da informação secundária é quantificada pelo coeficiente de correlação global, estimado a partir do conjunto das estações.

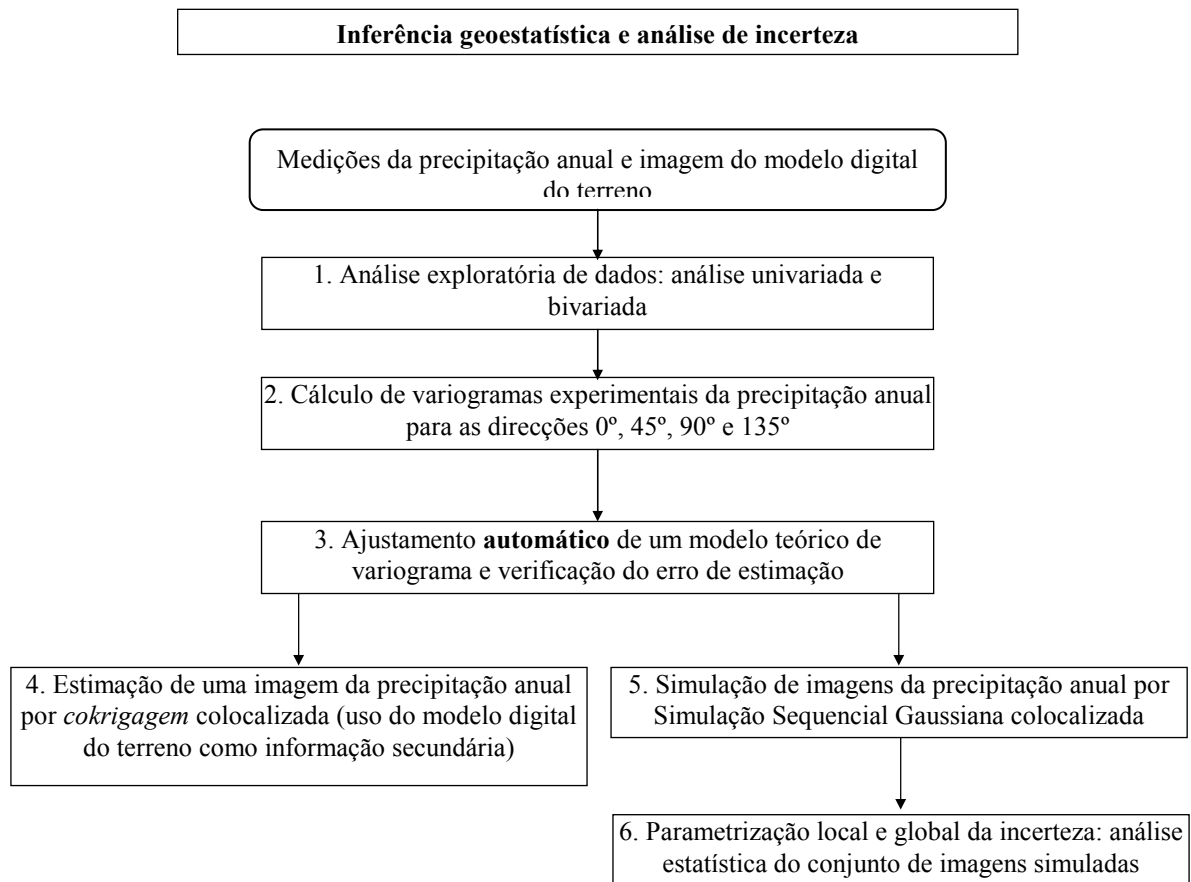


Fig. 101 Diagrama síntese da metodologia proposta para a estimação de uma imagem média da precipitação e parametrização da incerteza por simulação

A sequência de etapas proposta para este exemplo é resumida nas seguintes alíneas (fig. 101):

- a) Determinação do coeficiente de correlação global entre os valores de precipitação anual medidos nas estações e o modelo digital de terreno;
- b) Análise da continuidade espacial dos valores de precipitação: cálculo de variogramas experimentais, pesquisa de anisotropias e ajustamento automático de um modelo teórico;

- c) Construção de uma imagem média através da estimação por *cokrigagem* colocalizada, onde o modelo digital do terreno é considerado como informação secundária. A título de exemplo apresenta-se a imagem equivalente que seria obtida com a simples estimação por *krigagem* normal.
- d) Construção de imagens equiprováveis da precipitação por simulação pelo método de Simulação Sequencial Gaussiana colocalizada, onde o modelo digital do terreno é também considerado como informação secundária.
- e) Análise estatística do conjunto de imagens simuladas.

A informação de partida está distribuída por dois ficheiros:

I) *kedrainf.prn* – ficheiro com informação sobre precipitação obtidas nas estações pluviométricas, designadamente, x, y, cota, precipitação total mensal (mm) medida durante 12 meses em 36 estações e precipitação total (mm). Esta informação é disponibilizada pelo Instituto de Meteorologia.

II) *mdt.out* – ficheiro com informação do modelo digital de terreno, referente à região do Algarve, com resolução espacial de 1 km numa malha de 150 colunas por 70 linhas (fig. 102).

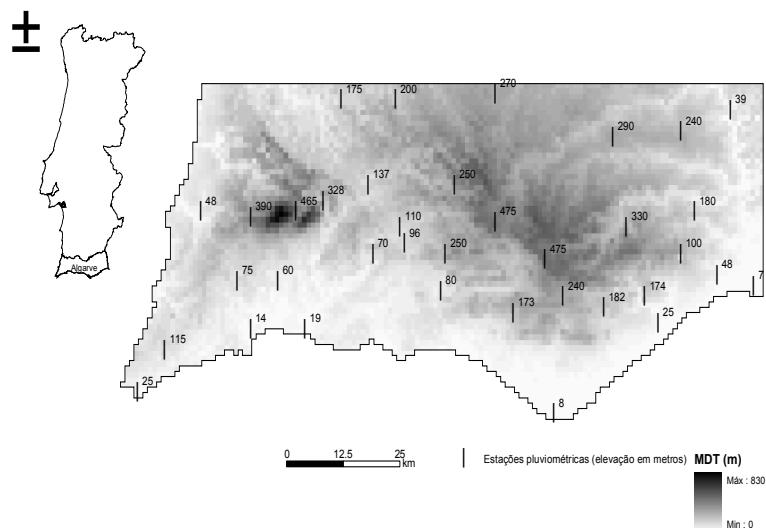


Fig. 102 Modelo digital de terreno do Algarve

5.2.1. Análise Estatística

Foram determinados, em primeiro lugar, estatísticas univariadas e bivariadas ao conjunto de dados cujos resultados se apresentam seguidamente (figuras 103 a 105).

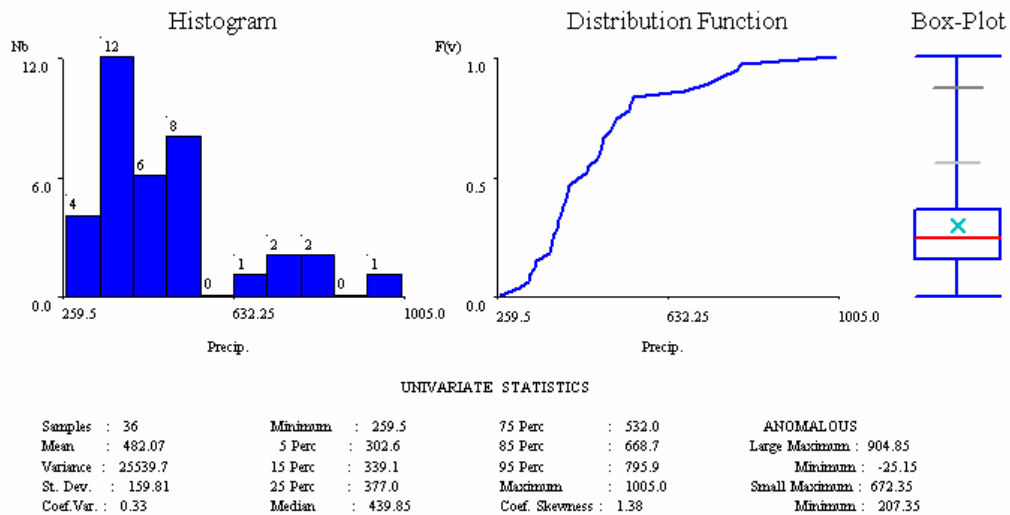


Fig. 103 Estatísticas univariadas da precipitação

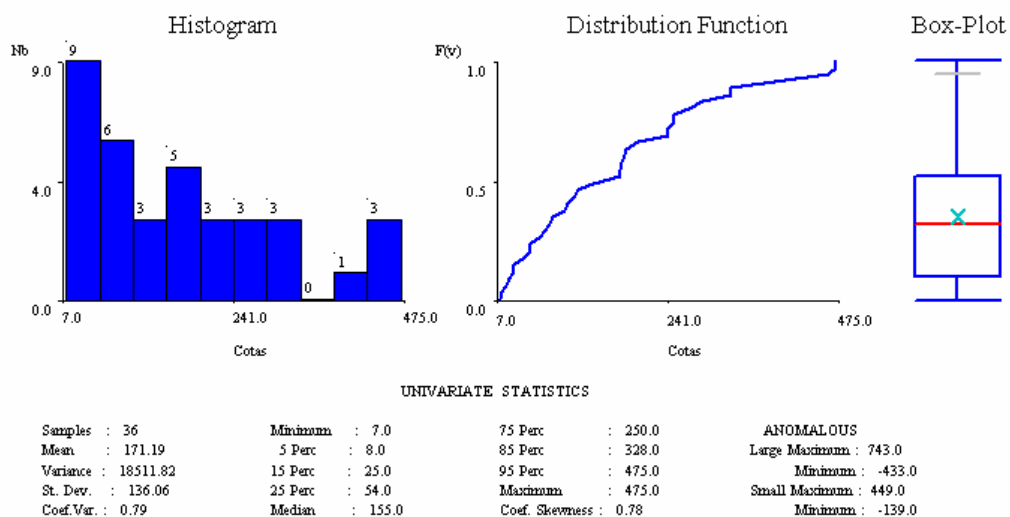


Fig. 104 Estatísticas univariadas das cotas para as 36 localizações das estações meteorológicas

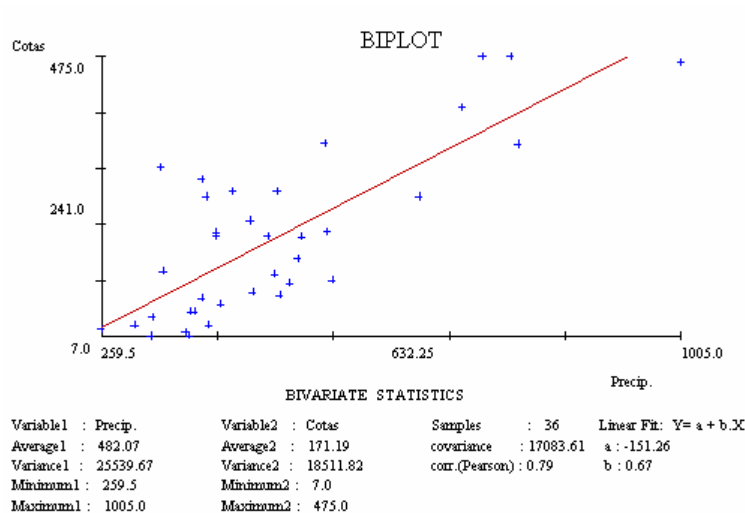


Fig. 105 Estatística bivariada entre a precipitação e as cotas

O histograma da variável precipitação mostra existir uma bi-distribuição separando os valores altos e os valores mais baixos. Para o modelo digital do terreno verifica-se também que existem duas populações. A correlação entre as duas variáveis é boa e quantificada por um coeficiente de correlação de 0,79.

5.2.2. Análise variográfica

No que diz respeito à variografia da variável pluviometria, foram calculados os variogramas para as direcções de 0°, 45°, 90° e 135° com uma tolerância de 30° e um passo de 10.5 km (16 passos) e também o variograma omnidireccional. Verificou-se a inexistência de anisotropia, talvez porque o número de amostras era insuficiente para esta análise.

O variograma omnidireccional foi modelado de forma automática, obtendo-se o respectivo modelo teórico (fig. 106) através do procedimento descrito anteriormente. O efeito de pepita representa cerca de 18% da variância total.

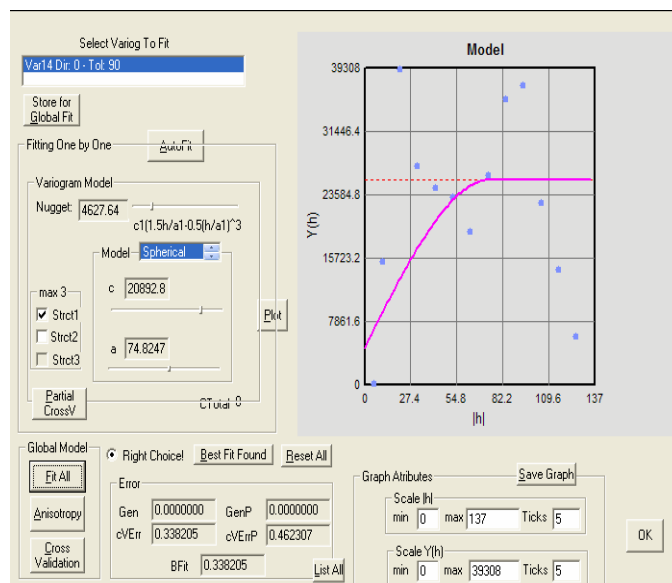


Fig. 106 Modelo teórico de variograma da precipitação total

5.2.3. Estimação por *krigagem*

Para a estimação, foram utilizados os métodos de estimação por *krigagem* normal e por *cokrigagem* colocalizada, esta última considerando o modelo digital do terreno como informação secundária.

A *cokrigagem* colocalizada foi realizada através do módulo do GSLIB (*cokb3d*), com as alterações introduzidas no geoMS [HP GEOMS]. Este modelo é desencadeado através da respectiva janela de interface do *MODIMAV*¹³.

Para a estimação utilizaram-se os seguintes parâmetros:

- $n_x=150$; $n_y=70$;
- nº amostras vizinhas: min=1; máx=20;
- raio máximo de pesquisa: 150;
- modelo teórico do tipo isotrópico com uma estrutura do tipo esférico:
 $c_0=4627,64$; $c_1 = 20892,80$; $a_1 = 74,83$ (conforme fig. 106).

¹³ Software desenvolvido nesta dissertação

Os mapas de precipitação estimados por *krigagem* normal e por *cokrigagem* colocalizada encontram-se na fig. 107. Os mapas com a variância de *krigagem* encontram-se na fig. 108.

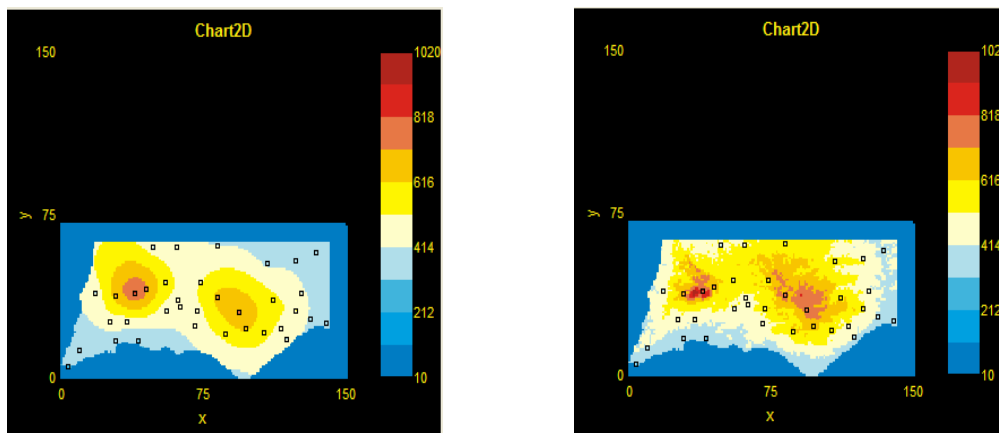


Fig. 107 Estimação da pluviosidade média na região do Algarve. Esq) *Krigagem* normal; dir) *cokrigagem* colocalizada

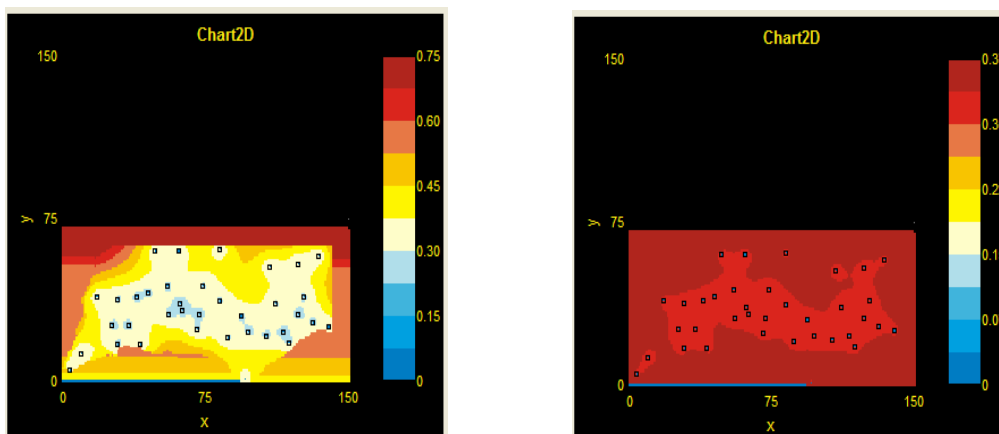


Fig. 108 Variância de *krigagem* da pluviosidade estimada na região do Algarve. Esq) *Krigagem* normal; dir) *cokrigagem* colocalizada

5.2.4. Simulação de imagens

A simulação sem imagem secundária e com imagem secundária foi realizada através do módulo do GSLIB (*sgsim*) que é desencadeado após a selecção dos parâmetros na janela respectiva de interface do *MODIMAV*.

Nas simulações com e sem imagem secundária, para além do modelo teórico já referido anteriormente, utilizaram-se os seguintes parâmetros:

- semente para a geração de números aleatórios: 23118;
- nº máximo de nós vizinhos seleccionados em cada iteração na simulação: 20

Na figura 109 são apresentados o gráfico resultante dos valores obtidos para três realizações da simulação da pluviosidade média sem variável secundária bem como a representação gráfica da variância calculada para cada ponto ao longo de 100 simulações.

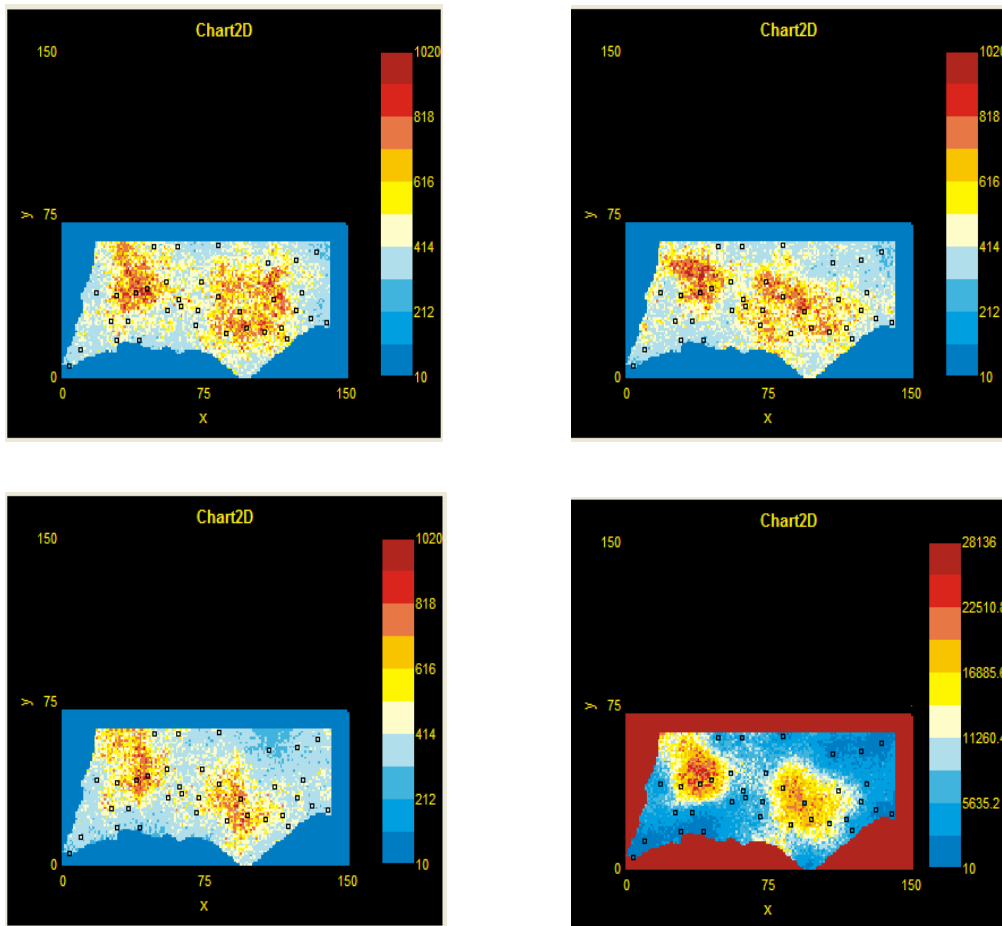


Fig. 109 Simulação da pluviosidade na região do Algarve. a) b) e c) 3 imagens simuladas; d) variância local das 100 realizações

Na figura 110 são apresentados o gráfico resultante dos valores obtidos para três realizações da simulação da pluviosidade média com imagem secundária bem como a representação gráfica da variância calculada para cada ponto ao longo de 100 simulações.

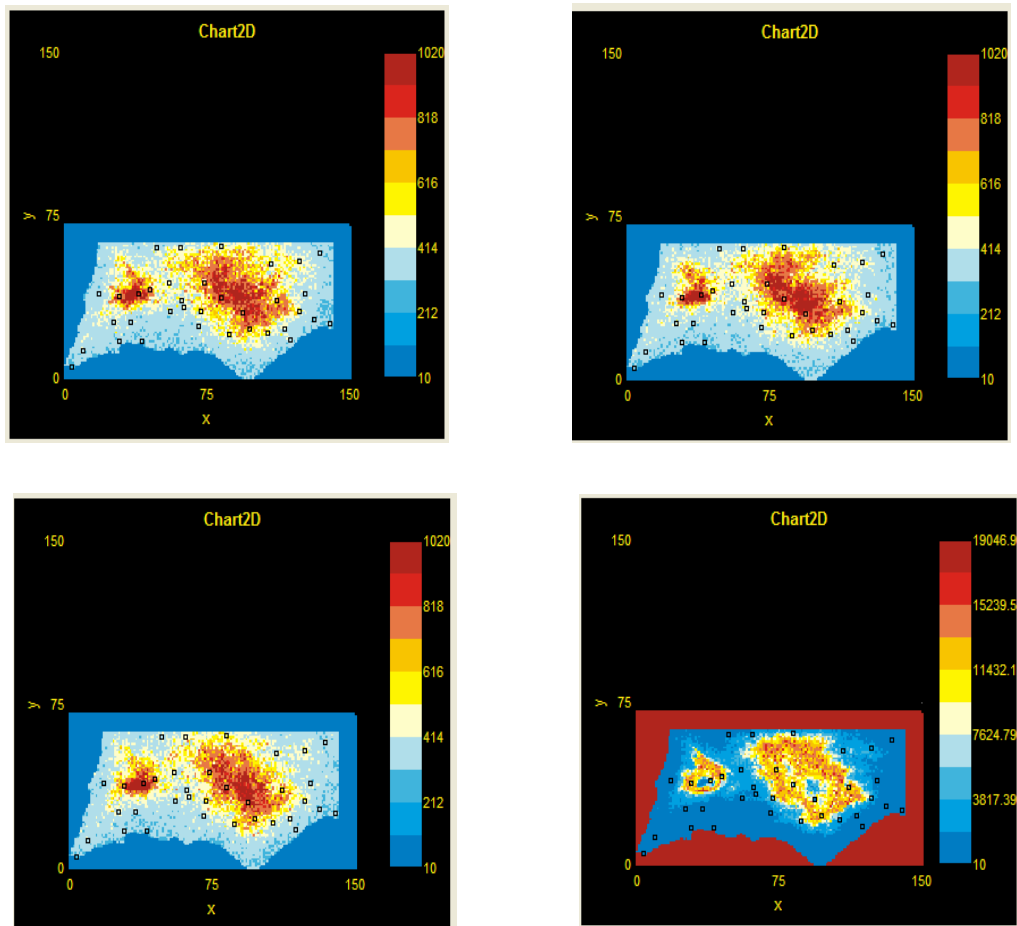


Fig. 110 Simulação da pluviosidade na região do Algarve. a) b) e c) 3 imagens co-simuladas com o modelo digital do terreno; d) variância local das 100 realizações

5.2.5. Análise dos resultados

Este exemplo comparativo de estimação e simulação geoestatística permite demonstrar a capacidade das técnicas de simulação estocástica na caracterização espacial de uma variável em estudo e no estabelecimento e validação de zonas de incerteza.

Em primeiro lugar foi realizada a estimação da quantidade de precipitação sem e com uma imagem secundária, neste caso de estudo o modelo digital do terreno. Nesta situação particular, Goovaerts [GO 99] com estes mesmos dados, mostra que o modelo que considera a variável secundária reproduz melhor os dados de partida num teste de validação cruzada, pelo que os modelos de co-estimação são os mais adequados para gerar imagens estimadas.

Comparando as imagens obtidas, a imagem estimada por *krigagem* normal é uma imagem muito mais atenuada do que a imagem obtida por *cokrigagem* colocada, onde é perfeitamente visível a morfologia do terreno (elevações).

A título meramente ilustrativo, apresentam-se os mapas da variância de estimação de cada procedimento, respectivamente, *krigagem* normal e *cokrigagem* colocada. Estes mapas reflectem apenas a “distância” estatística de cada ponto da imagem à amostra mais próxima, o que os torna semelhantes, a menos de uma constante. Todavia os mapas da variância de *krigagem* são pobres no que se refere à representação da incerteza, pois não tomam em conta a variabilidade local dos valores no plano de amostragem nem a sua interacção com o modelo de variograma.

A simulação de imagens permite interpretar a variabilidade local e global relativa a um dado conjunto amostral. Em termos de imagens simuladas, também as que consideram o modelo digital do terreno como informação secundária são muito influenciadas por esta informação.

A variância local das 100 realizações constitui um mapa de incerteza local do conjunto de amostras. Dado que se trata de uma variância calculada nos dois casos sobre 100 realizações, os mapas podem ser comparados localmente e em termos de grandeza absoluta.

Nos dois modelos de simulação considerados (sem e com imagem secundária) verifica-se que a maior variabilidade de valores é observada nos locais onde existe maior heterogeneidade ou “contradição” de medições, que constituem as zonas mais elevadas. Também, todos os valores são globalmente mais baixos, o que significa que a inclusão de uma imagem secundária num processo de simulação (passando a co-simulação) reduz o espaço das soluções de um processo de simulação condicional.

5.3. Reprodução de padrões de cor

Neste estudo, onde pretende reproduzir-se padrões de cor de amostras de mármore polido através da simulação condicional, os exemplos apresentados são referentes a duas amostras de mármore rosa de Estremoz, concretamente, uma amostra muito homogénea em tons de rosa e uma amostra com veios escuros, esbatidos.

Assim, em primeiro lugar, calcula-se o modelo teórico de cada canal de cor (fig. 111) que fornecerá a estrutura espacial da cor. Posteriormente são simuladas imagens de chapas de mármore, condicionadas aos histogramas dos canais de cor e aos variogramas experimentais.

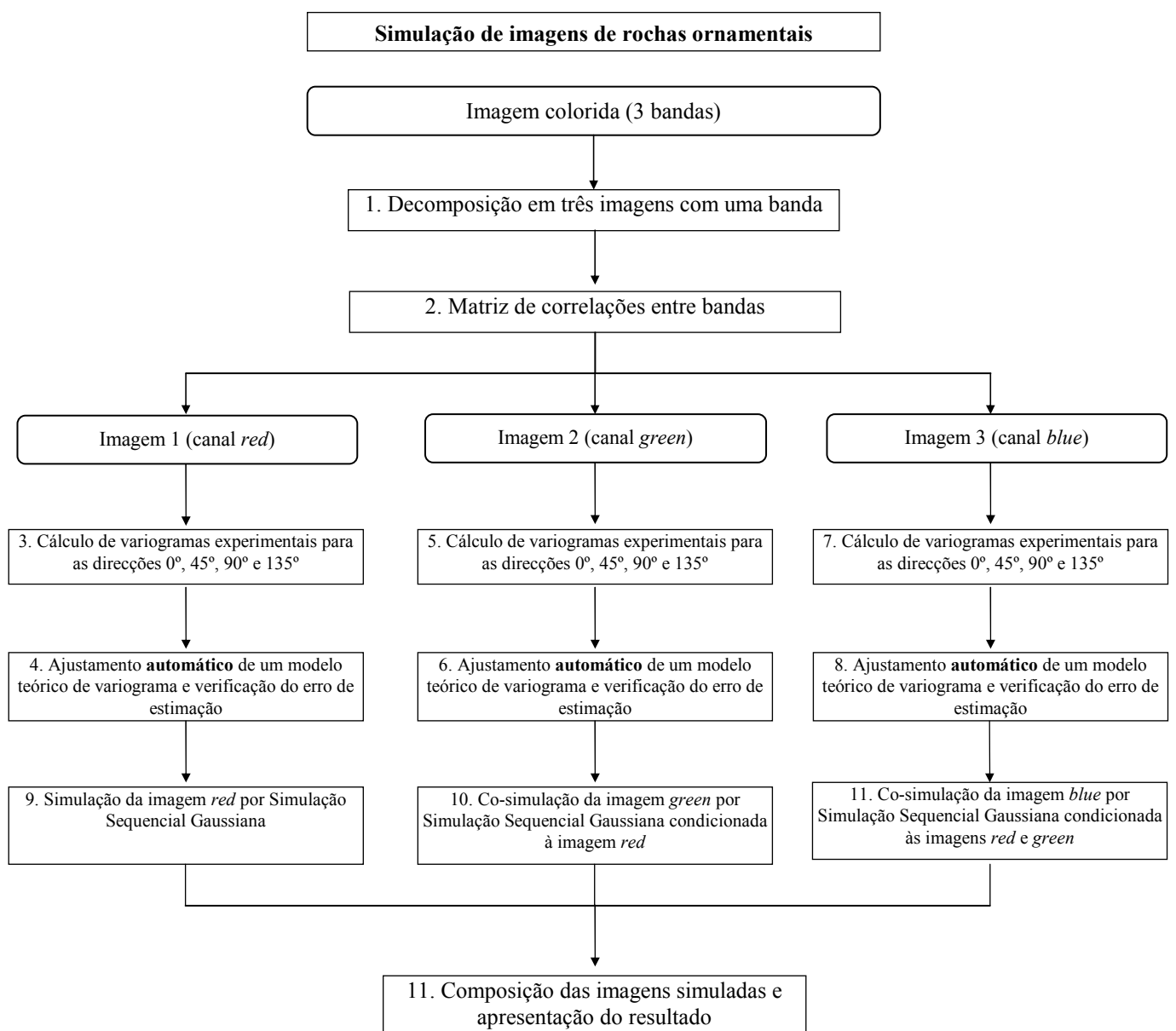


Fig. 111 Diagrama da metodologia proposta para a simulação de imagens de padrões de rochas ornamentais

5.3.1. Amostral

Neste caso, pretende-se obter a variabilidade de um padrão de cor através da simulação condicional da imagem digitalizada (fig.112) fornecida com formato *bitmap* – RGB com 8 bits por canal e com resolução de 146 por 150.



Fig. 112 Amostra de mármore

Em primeiro lugar, calcularam-se os variogramas experimentais para as direcções de 0°, 45°, 90° e 135° e ajustaram-se automaticamente os respectivos modelos teóricos para cada canal.

Os variogramas foram calculados com 40 passos e um passo de 4 *pixels* para X e Y. e os resultados obtidos pelo ajustamento automático estão reportados nas figuras 113 a 118.

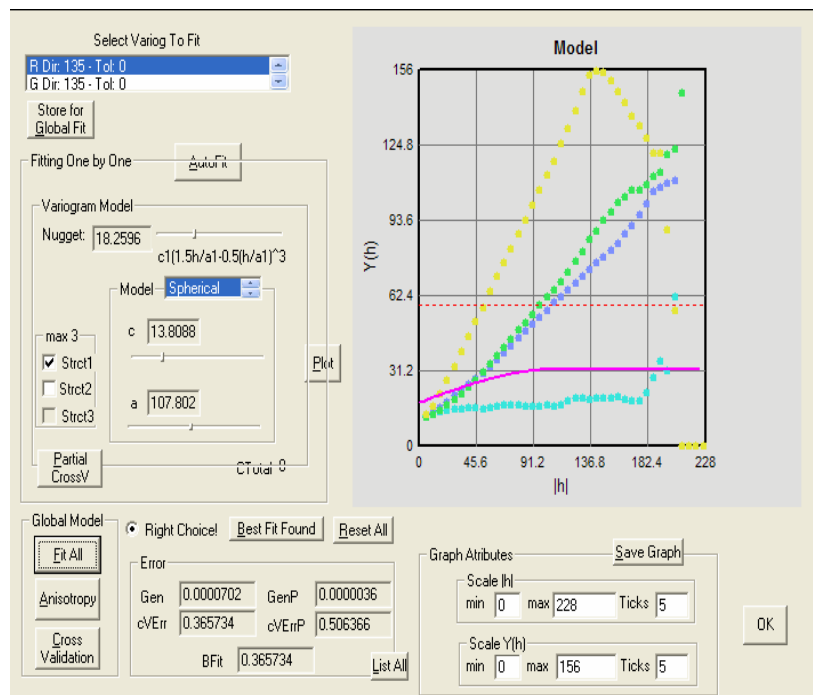


Fig. 113 Variogramas experimentais e modelo teórico, canal Red, modelo omnidireccional

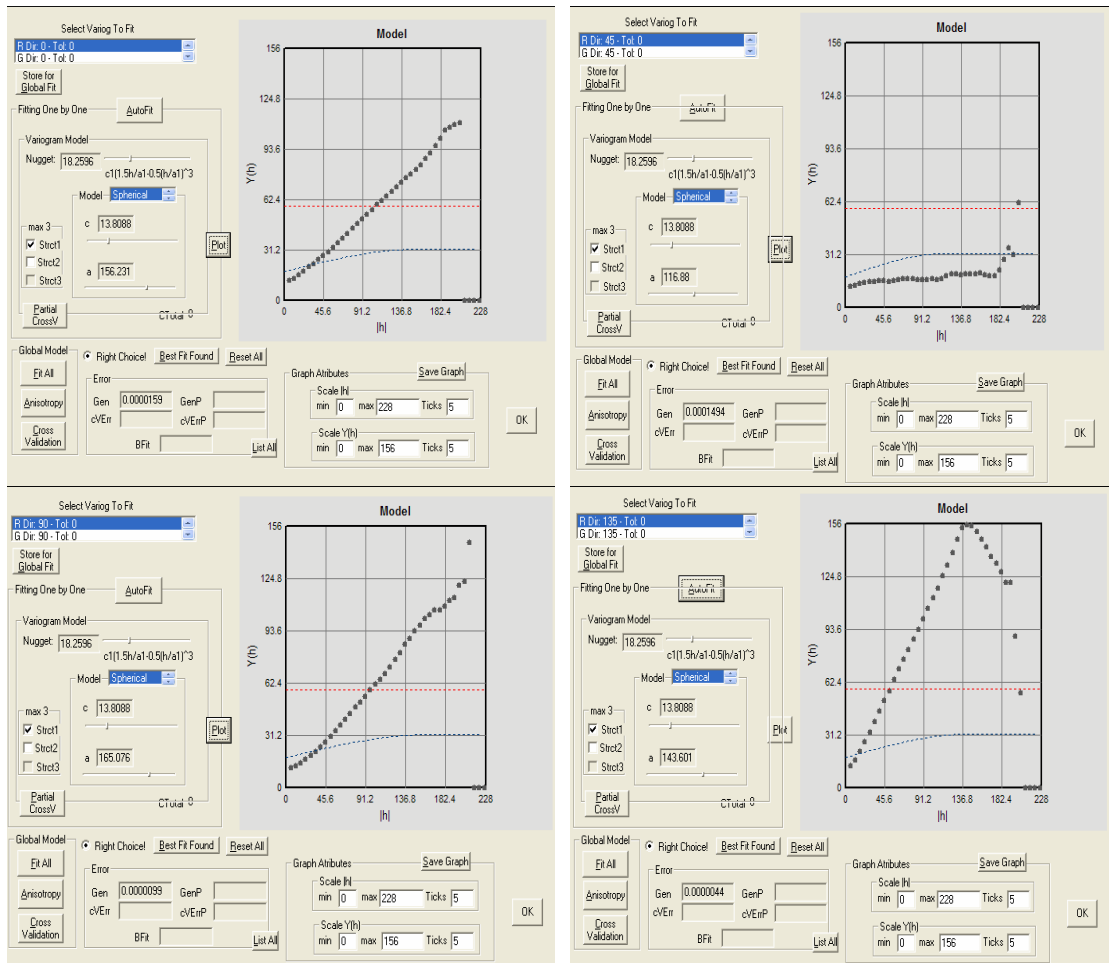


Fig. 114 Variogramas experimentais e modelo teórico, canal Red, direcção 0°, 45°, 90° e 135°

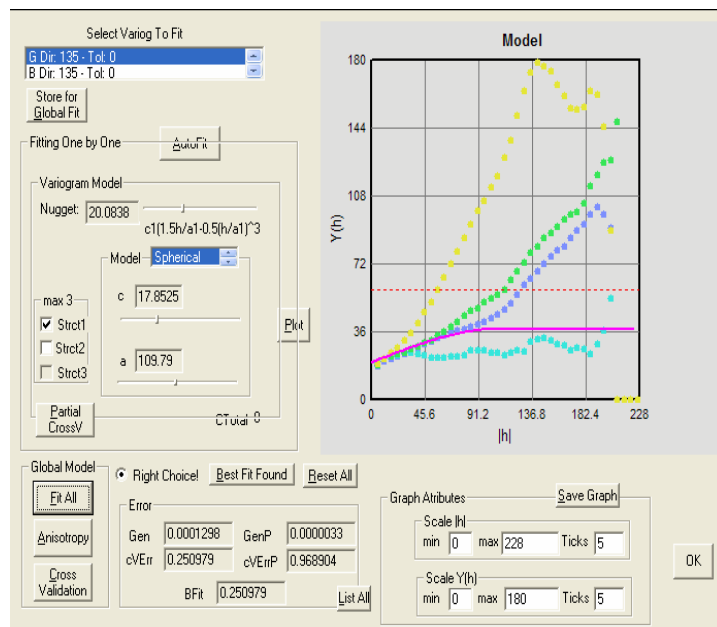


Fig. 115 Variogramas experimentais e modelo teórico, canal Green, modelo omnidireccional

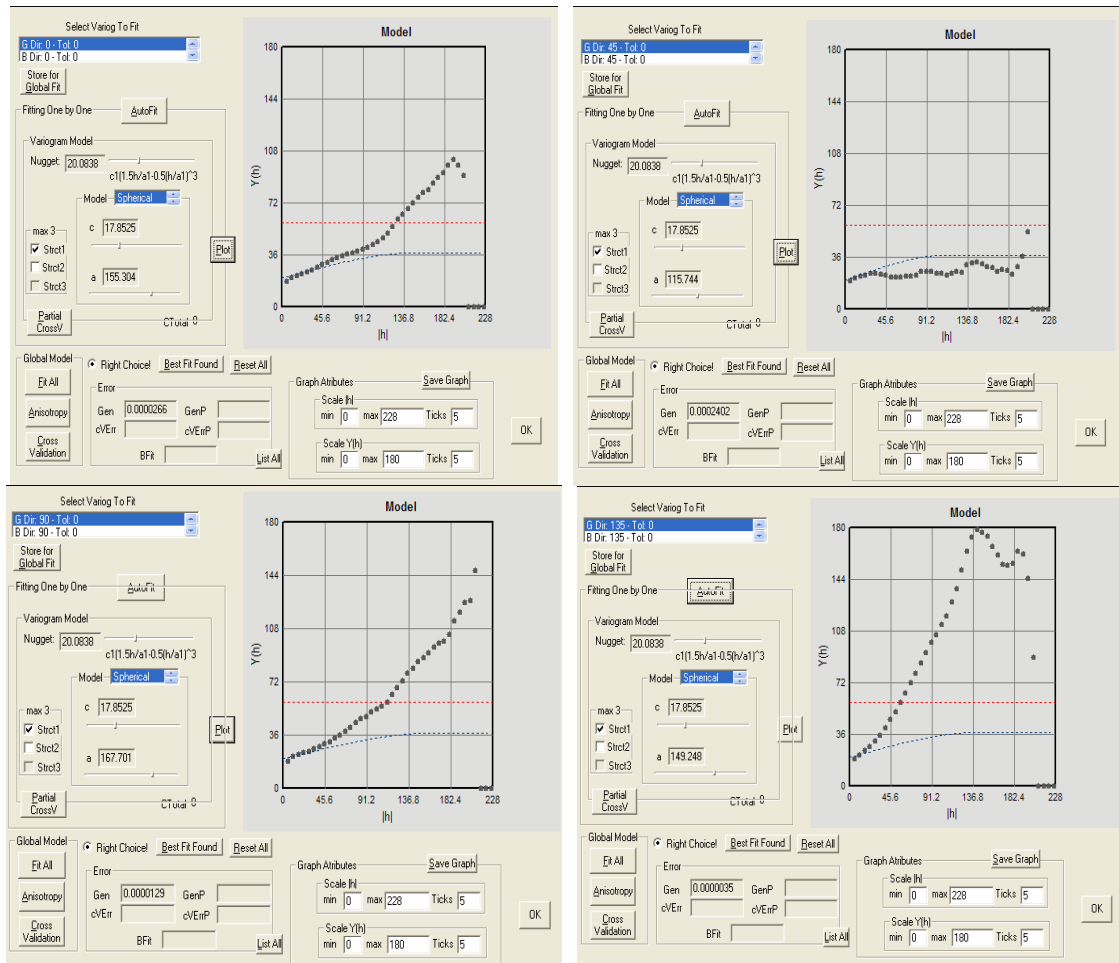


Fig. 116 Variogramas experimentais e modelo teórico, canal *Green*, direcção 0°, 45°, 90° e , 135°

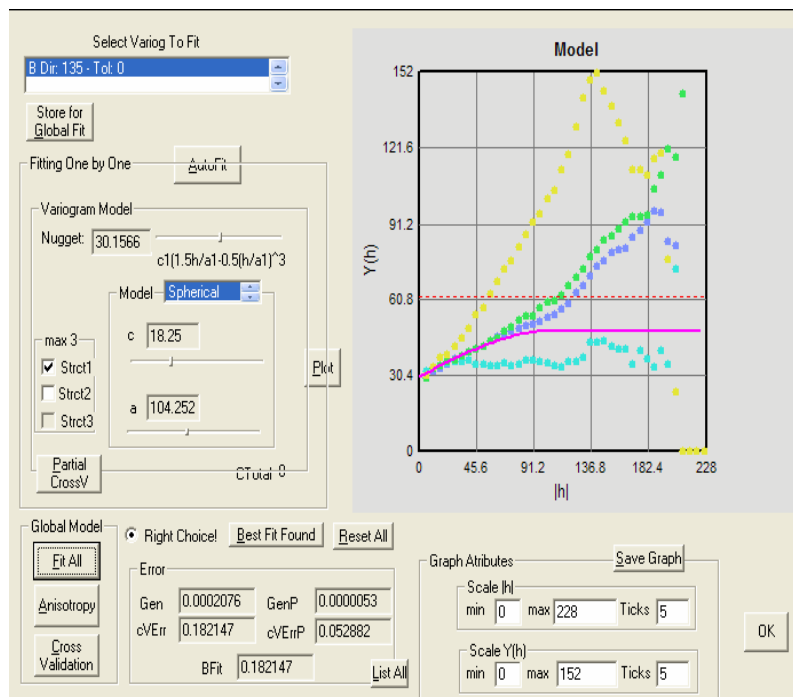


Fig. 117 Variogramas experimentais e modelo teórico, canal *Blue*, modelo omnidireccional

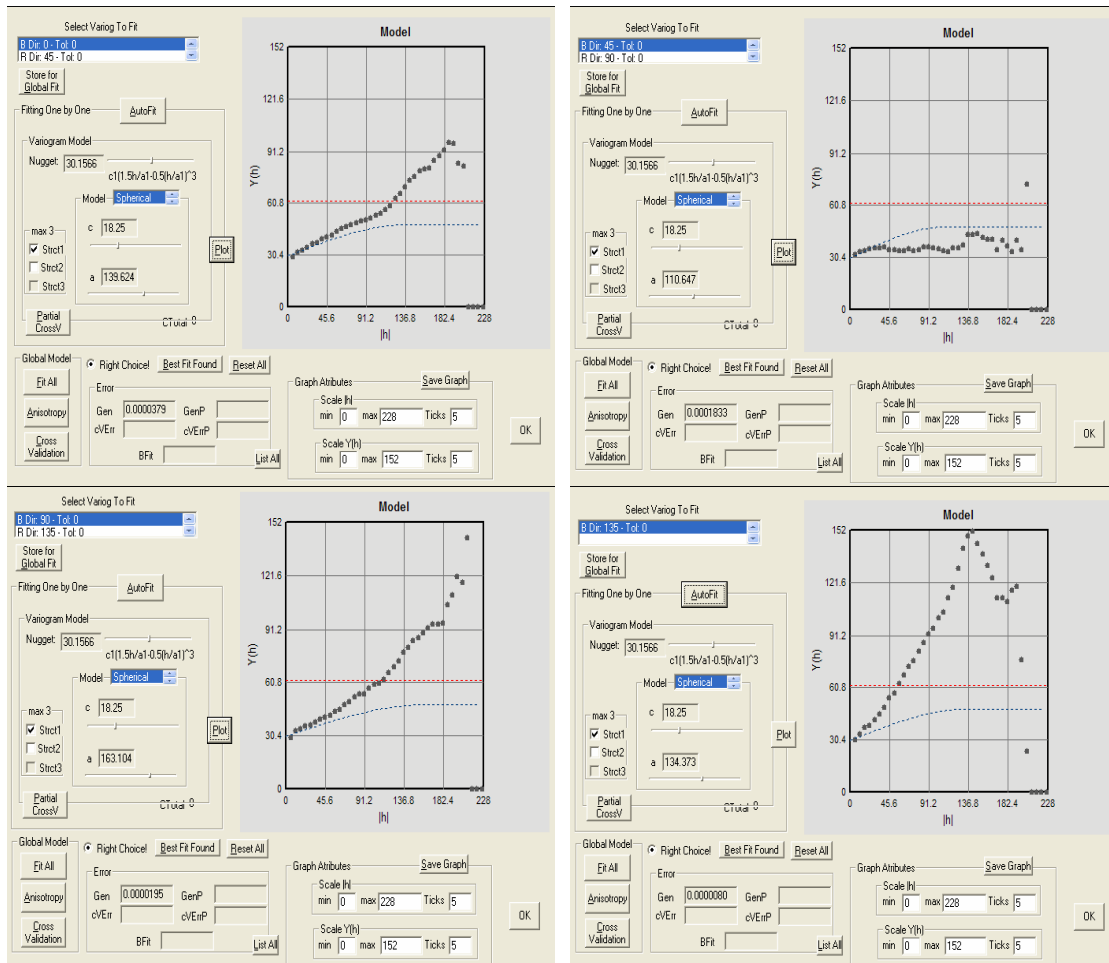


Fig. 118 Variogramas experimentais e modelo teórico, canal *Blue*, direcção 0°, 45°, 90° e , 135°

Observa-se que a correlação entre os três canais de cor é sempre bastante elevada (tabela 5), variando o coeficiente de correlação entre 0,7919 (vermelho e azul) e 0,9071 (verde e azul). Esta elevada correlação entre os três canais implica a utilização de um modelo de coregionalização para a construção de imagens simuladas dos padrões de mármore.

| | <i>Red</i> | <i>Green</i> | <i>Blue</i> |
|--------------|------------|--------------|-------------|
| <i>Red</i> | 1,000 | 0,8261 | 0,7919 |
| <i>Green</i> | 0,8261 | 1,0000 | 0,9071 |
| <i>Blue</i> | 0,7919 | 0,9071 | 1,0000 |

Tabela 4. Tabela de correlações entre os canais de cor *Red*, *Green* e *Blue*.

Foram simuladas três imagens do canal *Red* por Simulação Sequencial Gaussiana e 3 imagens dos canais *Green* e *Blue* por Co-Simulação Sequencial Gaussiana segundo a metodologia proposta na figura 111. Os resultados das imagens para os canais são apresentados nas figuras 119 a 121.

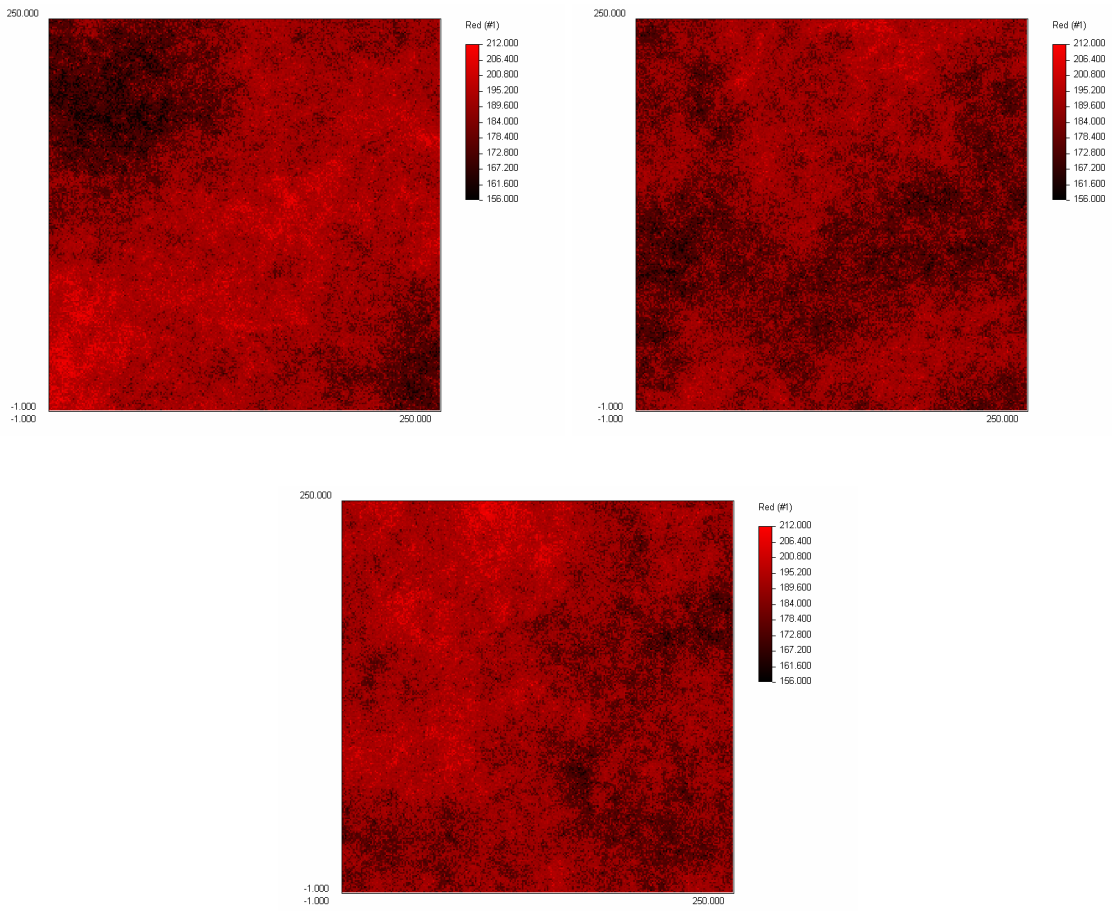
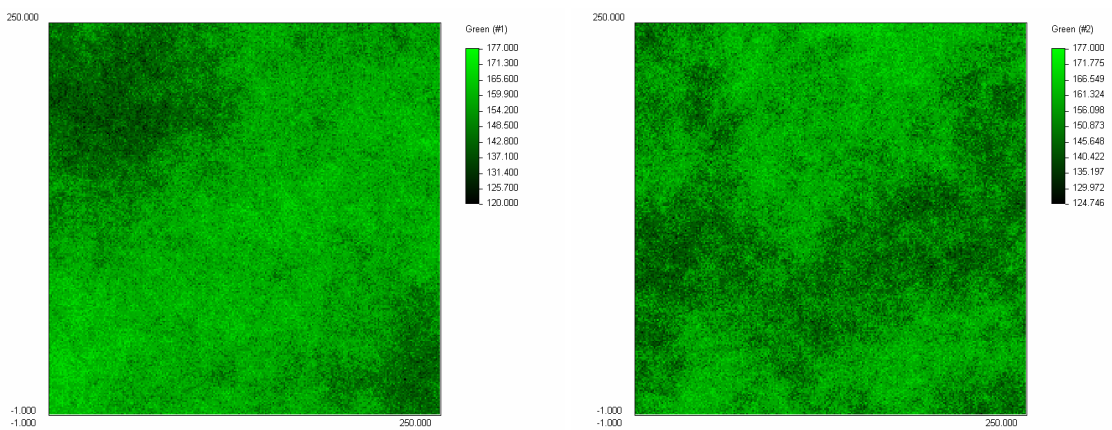


Fig. 119 3 imagens simuladas, canal *Red*



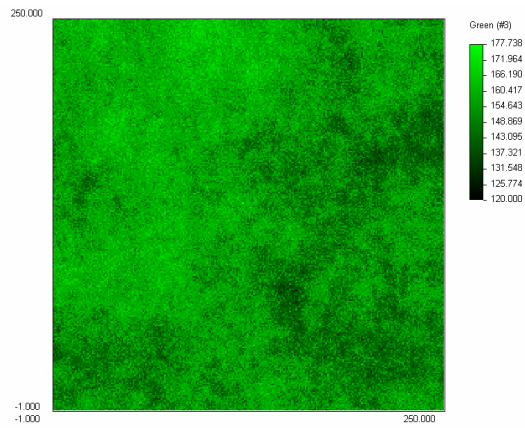


Fig. 120 3 imagens simuladas, canal *Green*

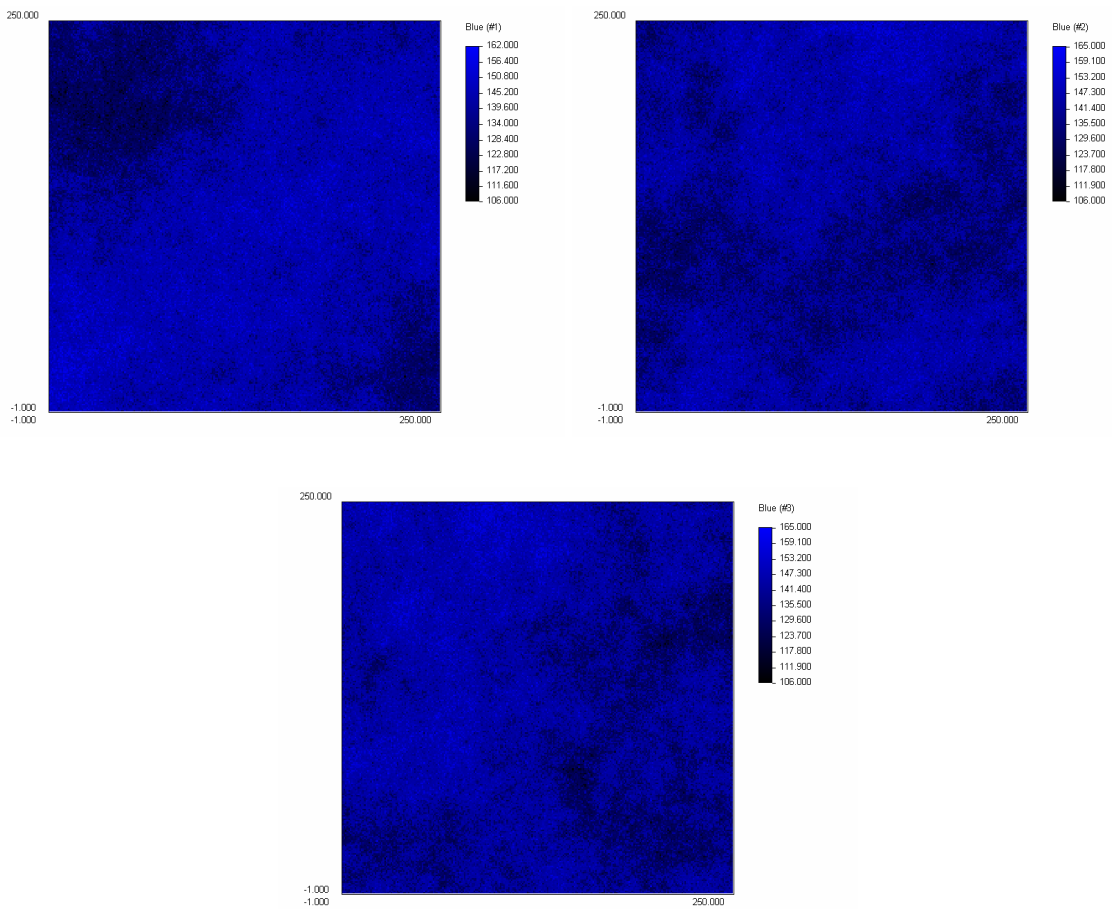


Fig. 121 – 3 imagens simuladas, canal *Blue*

A combinação dos canais de cor das imagens homólogas permite apresentar três imagens simuladas da placa de mármore rosa onde assenta este estudo (fig. 122).

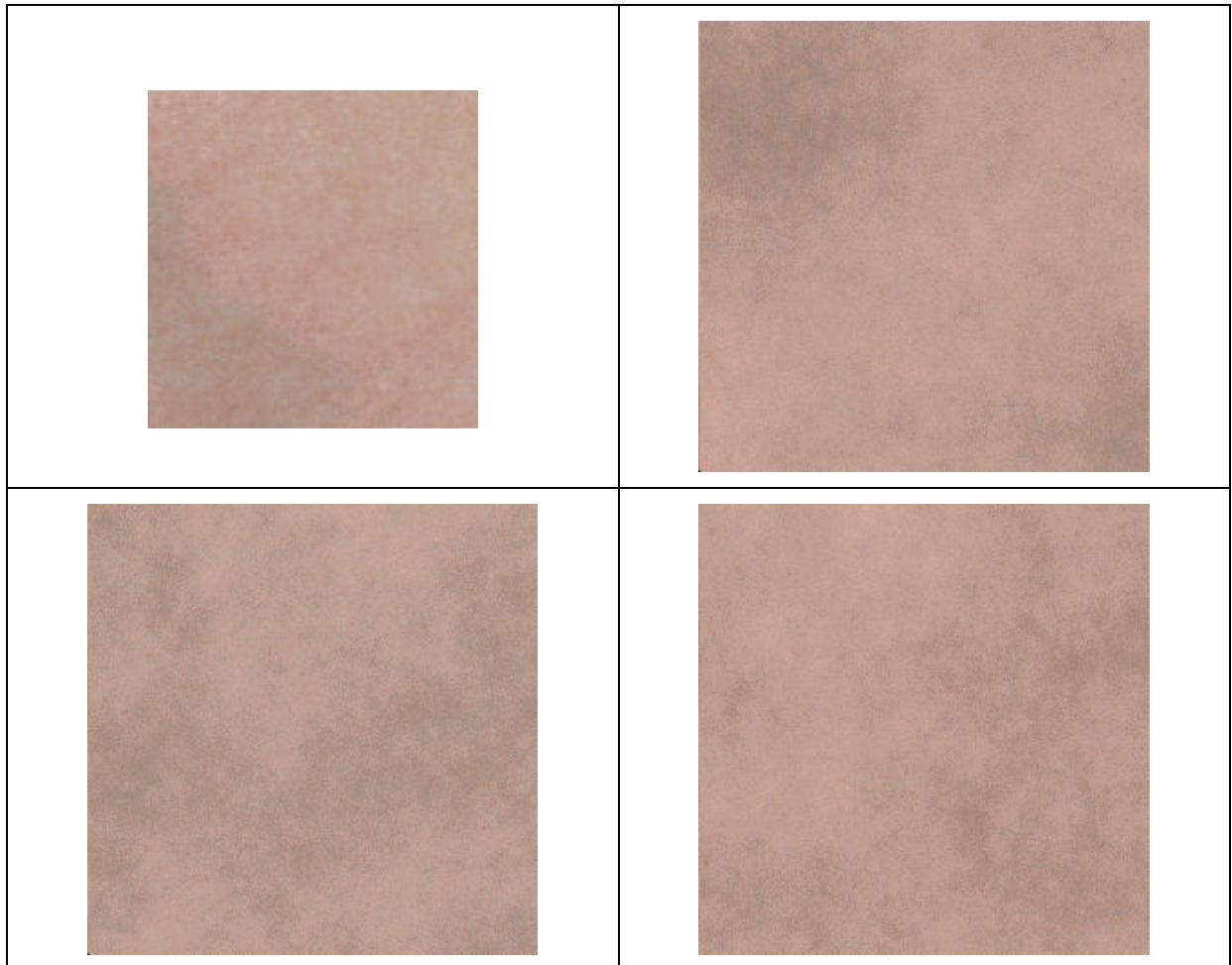


Fig. 122 Amostra inicial e 3 amostras simuladas

Para validar os resultados obtidos, foram calculados os variogramas das imagens simuladas (figuras 123 a 125) e comparados, para a imagem real *versus* as simulações #1, #2 e #3, os respectivos estatísticos de cor (tabela 5 a tabela 7).

Nestes modelos pode notar-se a influência do conjunto de pontos de cada variograma experimental usados no ajuste automático do modelo teórico (como foi referido no capítulo 3 apenas foram considerados para o ajuste automático os pontos situados abaixo da variância).

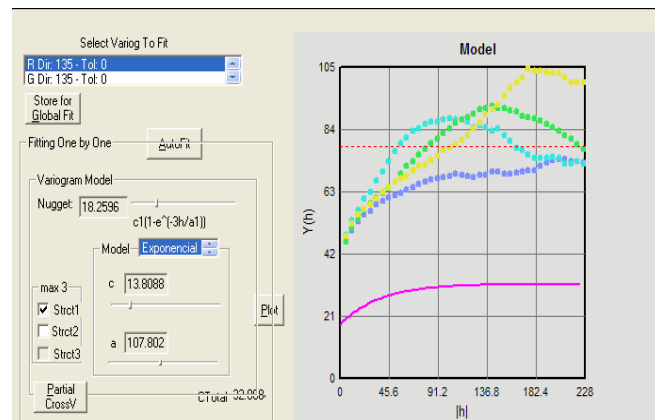
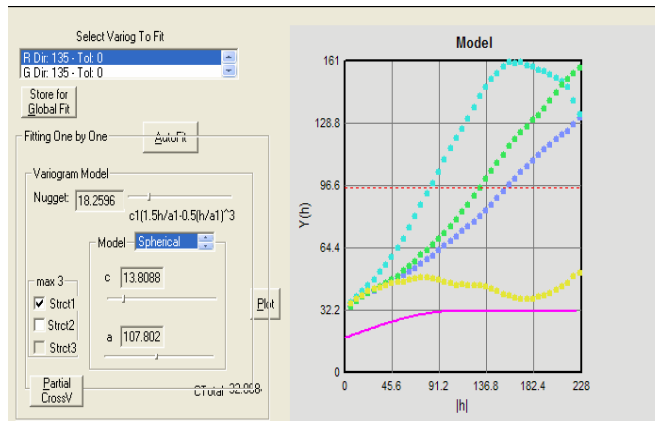
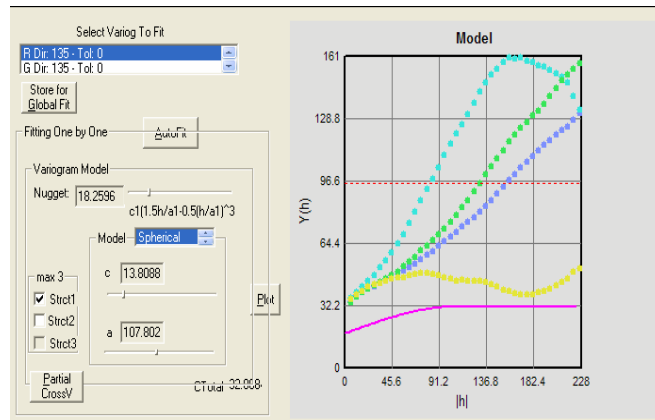


Fig. 123 Variogramas experimentais do canal *Red* sobrepostos ao modelo teórico dos dados

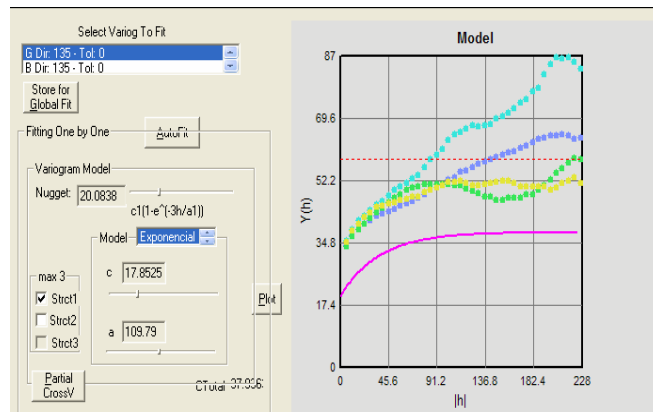
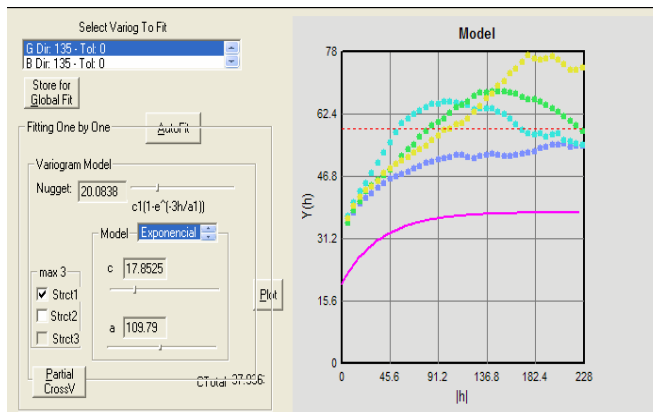
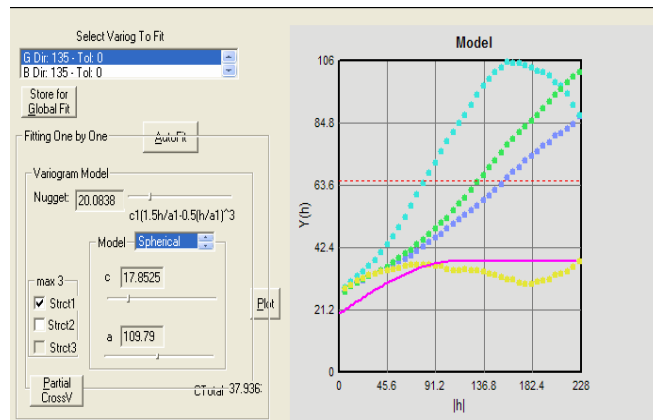


Fig. 124 Variogramas experimentais do canal *Green* sobrepostos ao modelo teórico dos dados

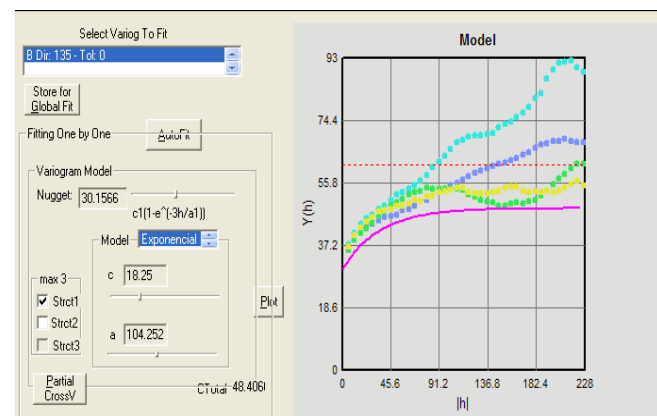
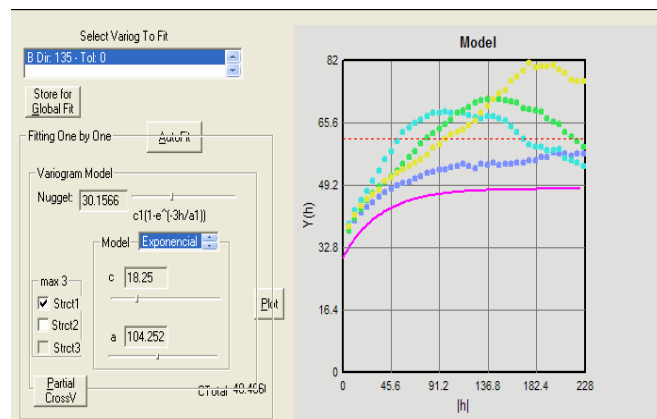
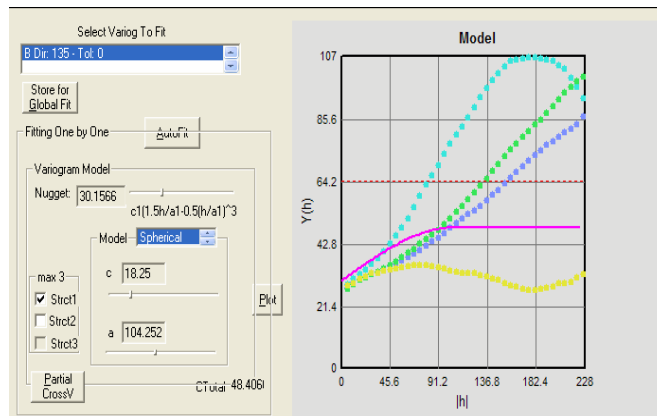


Fig. 125 Variogramas experimentais do canal *Blue* sobrepostos ao modelo teórico dos dados

| Intervalos dos canais de cor | | | Frequências da imagem original em valores | | Frequências da imagem simulada em valores | | Diferenças | | |
|------------------------------|-----------|-----------|---|----------------|---|----------------|----------------|----------------|-----------------|
| R | G | B | absolutos | relativos | absolutos | relativos | totais | positivos | Negativos |
| [129;144] | [97;112] | [81;96] | 0 | 0.00000 | 10 | 0.00016 | -0.00016 | 0.00000 | -0.00016 |
| [145;160] | [97;112] | [81;96] | 2 | 0.00009 | 9 | 0.00014 | -0.00005 | 0.00000 | -0.00005 |
| [129;144] | [113;128] | [81;96] | 0 | 0.00000 | 12 | 0.00019 | -0.00019 | 0.00000 | -0.00019 |
| [145;160] | [113;128] | [81;96] | 4 | 0.00018 | 19 | 0.00030 | -0.00012 | 0.00000 | -0.00012 |
| [129;144] | [97;112] | [97;112] | 0 | 0.00000 | 12 | 0.00019 | -0.00019 | 0.00000 | -0.00019 |
| [145;160] | [97;112] | [97;112] | 2 | 0.00009 | 26 | 0.00042 | -0.00032 | 0.00000 | -0.00032 |
| [129;144] | [113;128] | [97;112] | 7 | 0.00032 | 330 | 0.00528 | -0.00496 | 0.00000 | -0.00496 |
| [145;160] | [113;128] | [97;112] | 437 | 0.01995 | 3389 | 0.05422 | -0.03427 | 0.00000 | -0.03427 |
| [161;176] | [113;128] | [97;112] | 168 | 0.00767 | 494 | 0.00790 | -0.00023 | 0.00000 | -0.00023 |
| [177;192] | [113;128] | [97;112] | 0 | 0.00000 | 8 | 0.00013 | -0.00013 | 0.00000 | -0.00013 |
| [145;160] | [129;144] | [97;112] | 216 | 0.00986 | 301 | 0.00482 | 0.00505 | 0.00505 | 0.00000 |
| [161;176] | [129;144] | [97;112] | 86 | 0.00393 | 242 | 0.00387 | 0.00005 | 0.00005 | 0.00000 |
| [177;192] | [129;144] | [97;112] | 1 | 0.00005 | 11 | 0.00018 | -0.00013 | 0.00000 | -0.00013 |
| [129;144] | [113;128] | [113;128] | 0 | 0.00000 | 46 | 0.00074 | -0.00074 | 0.00000 | -0.00074 |
| [145;160] | [113;128] | [113;128] | 756 | 0.03452 | 3104 | 0.04966 | -0.01514 | 0.00000 | -0.01514 |
| [161;176] | [113;128] | [113;128] | 626 | 0.02858 | 1217 | 0.01947 | 0.00911 | 0.00911 | 0.00000 |
| [177;192] | [113;128] | [113;128] | 0 | 0.00000 | 23 | 0.00037 | -0.00037 | 0.00000 | -0.00037 |
| [145;160] | [129;144] | [113;128] | 1106 | 0.05050 | 3135 | 0.05016 | 0.00034 | 0.00034 | 0.00000 |
| [161;176] | [129;144] | [113;128] | 5754 | 0.26274 | 11323 | 0.18117 | 0.08157 | 0.08157 | 0.00000 |
| [177;192] | [129;144] | [113;128] | 1199 | 0.05475 | 5833 | 0.09333 | -0.03858 | 0.00000 | -0.03858 |
| [145;160] | [145;160] | [113;128] | 0 | 0.00000 | 2 | 0.00003 | -0.00003 | 0.00000 | -0.00003 |
| [161;176] | [145;160] | [113;128] | 735 | 0.03356 | 408 | 0.00653 | 0.02703 | 0.02703 | 0.00000 |
| [177;192] | [145;160] | [113;128] | 408 | 0.01863 | 987 | 0.01579 | 0.00284 | 0.00284 | 0.00000 |
| [193;208] | [145;160] | [113;128] | 0 | 0.00000 | 3 | 0.00005 | -0.00005 | 0.00000 | -0.00005 |
| [145;160] | [113;128] | [129;144] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [161;176] | [113;128] | [129;144] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [145;160] | [129;144] | [129;144] | 13 | 0.00059 | 200 | 0.00320 | -0.00261 | 0.00000 | -0.00261 |
| [161;176] | [129;144] | [129;144] | 2402 | 0.10968 | 5902 | 0.09443 | 0.01525 | 0.01525 | 0.00000 |
| [177;192] | [129;144] | [129;144] | 2314 | 0.10566 | 7940 | 0.12704 | -0.02138 | 0.00000 | -0.02138 |
| [129;144] | [145;160] | [129;144] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [145;160] | [145;160] | [129;144] | 0 | 0.00000 | 7 | 0.00011 | -0.00011 | 0.00000 | -0.00011 |
| [161;176] | [145;160] | [129;144] | 1323 | 0.06041 | 3354 | 0.05366 | 0.00675 | 0.00675 | 0.00000 |
| [177;192] | [145;160] | [129;144] | 4072 | 0.18594 | 13824 | 0.22118 | -0.03525 | 0.00000 | -0.03525 |
| [193;208] | [145;160] | [129;144] | 6 | 0.00027 | 247 | 0.00395 | -0.00368 | 0.00000 | -0.00368 |
| [177;192] | [161;176] | [129;144] | 4 | 0.00018 | 1 | 0.00002 | 0.00017 | 0.00017 | 0.00000 |
| [193;208] | [161;176] | [129;144] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [161;176] | [129;144] | [145;160] | 4 | 0.00018 | 0 | 0.00000 | 0.00018 | 0.00018 | 0.00000 |
| [161;176] | [145;160] | [145;160] | 3 | 0.00014 | 4 | 0.00006 | 0.00007 | 0.00007 | 0.00000 |
| [177;192] | [145;160] | [145;160] | 232 | 0.01059 | 67 | 0.00107 | 0.00952 | 0.00952 | 0.00000 |
| [193;208] | [145;160] | [145;160] | 9 | 0.00041 | 5 | 0.00008 | 0.00033 | 0.00033 | 0.00000 |
| [177;192] | [161;176] | [145;160] | 10 | 0.00046 | 0 | 0.00000 | 0.00046 | 0.00046 | 0.00000 |
| [193;208] | [161;176] | [145;160] | 1 | 0.00005 | 1 | 0.00002 | 0.00003 | 0.00003 | 0.00000 |
| Totais | | | 21900 | 1.00000 | 62500 | 1.00000 | 0.00000 | 0.15876 | -0.15876 |

Tabela 5. Comparação entre frequências absolutas e relativas para intervalos de cor relativamente à imagem original e imagem simulada (#1)

| Intervalos dos canais de cor | | | Frequências da imagem original em valores | | Frequências da imagem simulada em valores | | Diferenças | | |
|------------------------------|-----------|-----------|---|----------------|---|----------------|----------------|----------------|-----------------|
| R | G | B | absolutos | relativos | absolutos | relativos | totais | positivos | Negativos |
| [129;144] | [97;112] | [81;96] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [145;160] | [97;112] | [81;96] | 2 | 0.00009 | 2 | 0.00003 | 0.00006 | 0.00006 | 0.00000 |
| [129;144] | [113;128] | [81;96] | 0 | 0.00000 | 2 | 0.00003 | -0.00003 | 0.00000 | -0.00003 |
| [145;160] | [113;128] | [81;96] | 4 | 0.00018 | 2 | 0.00003 | 0.00015 | 0.00015 | 0.00000 |
| [129;144] | [97;112] | [97;112] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [145;160] | [97;112] | [97;112] | 2 | 0.00009 | 4 | 0.00006 | 0.00003 | 0.00003 | 0.00000 |
| [129;144] | [113;128] | [97;112] | 7 | 0.00032 | 62 | 0.00099 | -0.00067 | 0.00000 | -0.00067 |
| [145;160] | [113;128] | [97;112] | 437 | 0.01995 | 2186 | 0.03498 | -0.01502 | 0.00000 | -0.01502 |
| [161;176] | [113;128] | [97;112] | 168 | 0.00767 | 398 | 0.00637 | 0.00130 | 0.00130 | 0.00000 |
| [129;144] | [129;144] | [97;112] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [145;160] | [129;144] | [97;112] | 216 | 0.00986 | 476 | 0.00762 | 0.00225 | 0.00225 | 0.00000 |
| [161;176] | [129;144] | [97;112] | 86 | 0.00393 | 274 | 0.00438 | -0.00046 | 0.00000 | -0.00046 |
| [177;192] | [129;144] | [97;112] | 1 | 0.00005 | 3 | 0.00005 | 0.00000 | 0.00000 | 0.00000 |
| [129;144] | [113;128] | [113;128] | 0 | 0.00000 | 26 | 0.00042 | -0.00042 | 0.00000 | -0.00042 |
| [145;160] | [113;128] | [113;128] | 756 | 0.03452 | 3569 | 0.05710 | -0.02258 | 0.00000 | -0.02258 |
| [161;176] | [113;128] | [113;128] | 626 | 0.02858 | 1117 | 0.01787 | 0.01071 | 0.01071 | 0.00000 |
| [177;192] | [113;128] | [113;128] | 0 | 0.00000 | 2 | 0.00003 | -0.00003 | 0.00000 | -0.00003 |
| [129;144] | [129;144] | [113;128] | 0 | 0.00000 | 2 | 0.00003 | -0.00003 | 0.00000 | -0.00003 |
| [145;160] | [129;144] | [113;128] | 1106 | 0.05050 | 8043 | 0.12869 | -0.07819 | 0.00000 | -0.07819 |
| [161;176] | [129;144] | [113;128] | 5754 | 0.26274 | 13954 | 0.22326 | 0.03948 | 0.03948 | 0.00000 |
| [177;192] | [129;144] | [113;128] | 1199 | 0.05475 | 1636 | 0.02618 | 0.02857 | 0.02857 | 0.00000 |
| [145;160] | [145;160] | [113;128] | 0 | 0.00000 | 6 | 0.00010 | -0.00010 | 0.00000 | -0.00010 |
| [161;176] | [145;160] | [113;128] | 735 | 0.03356 | 716 | 0.01146 | 0.02211 | 0.02211 | 0.00000 |
| [177;192] | [145;160] | [113;128] | 408 | 0.01863 | 488 | 0.00781 | 0.01082 | 0.01082 | 0.00000 |
| [161;176] | [113;128] | [129;144] | 0 | 0.00000 | 3 | 0.00005 | -0.00005 | 0.00000 | -0.00005 |
| [145;160] | [129;144] | [129;144] | 13 | 0.00059 | 923 | 0.01477 | -0.01417 | 0.00000 | -0.01417 |
| [161;176] | [129;144] | [129;144] | 2402 | 0.10968 | 8259 | 0.13214 | -0.02246 | 0.00000 | -0.02246 |
| [177;192] | [129;144] | [129;144] | 2314 | 0.10566 | 3089 | 0.04942 | 0.05624 | 0.05624 | 0.00000 |
| [145;160] | [145;160] | [129;144] | 0 | 0.00000 | 54 | 0.00086 | -0.00086 | 0.00000 | -0.00086 |
| [161;176] | [145;160] | [129;144] | 1323 | 0.06041 | 7602 | 0.12163 | -0.06122 | 0.00000 | -0.06122 |
| [177;192] | [145;160] | [129;144] | 4072 | 0.18594 | 9446 | 0.15114 | 0.03480 | 0.03480 | 0.00000 |
| [193;208] | [145;160] | [129;144] | 6 | 0.00027 | 20 | 0.00032 | -0.00005 | 0.00000 | -0.00005 |
| [177;192] | [161;176] | [129;144] | 4 | 0.00018 | 0 | 0.00000 | 0.00018 | 0.00018 | 0.00000 |
| [161;176] | [129;144] | [145;160] | 4 | 0.00018 | 0 | 0.00000 | 0.00018 | 0.00018 | 0.00000 |
| [177;192] | [129;144] | [145;160] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [161;176] | [145;160] | [145;160] | 3 | 0.00014 | 20 | 0.00032 | -0.00018 | 0.00000 | -0.00018 |
| [177;192] | [145;160] | [145;160] | 232 | 0.01059 | 107 | 0.00171 | 0.00888 | 0.00888 | 0.00000 |
| [193;208] | [145;160] | [145;160] | 9 | 0.00041 | 3 | 0.00005 | 0.00036 | 0.00036 | 0.00000 |
| [177;192] | [161;176] | [145;160] | 10 | 0.00046 | 2 | 0.00003 | 0.00042 | 0.00042 | 0.00000 |
| [193;208] | [161;176] | [145;160] | 1 | 0.00005 | 0 | 0.00000 | 0.00005 | 0.00005 | 0.00000 |
| Totais | | | 21900 | 1.00000 | 62500 | 1.00000 | 0.00000 | 0.21659 | -0.21659 |

Tabela 6. Comparação entre frequências absolutas e relativas para intervalos de cor relativamente à imagem original e imagem simulada (#2)

| Intervalos dos canais de cor | | | Frequências da imagem original em valores | | Frequências da imagem simulada em valores | | Diferenças | | |
|------------------------------|-----------|-----------|---|----------------|---|----------------|----------------|----------------|-----------------|
| R | G | B | absolutos | relativos | absolutos | relativos | totais | positivos | Negativos |
| [129;144] | [97;112] | [81;96] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [145;160] | [97;112] | [81;96] | 2 | 0.00009 | 5 | 0.00008 | 0.00001 | 0.00001 | 0.00000 |
| [145;160] | [113;128] | [81;96] | 4 | 0.00018 | 11 | 0.00018 | 0.00001 | 0.00001 | 0.00000 |
| [129;144] | [97;112] | [97;112] | 0 | 0.00000 | 2 | 0.00003 | -0.00003 | 0.00000 | -0.00003 |
| [145;160] | [97;112] | [97;112] | 2 | 0.00009 | 7 | 0.00011 | -0.00002 | 0.00000 | -0.00002 |
| [129;144] | [113;128] | [97;112] | 7 | 0.00032 | 27 | 0.00043 | -0.00011 | 0.00000 | -0.00011 |
| [145;160] | [113;128] | [97;112] | 437 | 0.01995 | 2201 | 0.03522 | -0.01526 | 0.00000 | -0.01526 |
| [161;176] | [113;128] | [97;112] | 168 | 0.00767 | 873 | 0.01397 | -0.00630 | 0.00000 | -0.00630 |
| [177;192] | [113;128] | [97;112] | 0 | 0.00000 | 5 | 0.00008 | -0.00008 | 0.00000 | -0.00008 |
| [145;160] | [129;144] | [97;112] | 216 | 0.00986 | 277 | 0.00443 | 0.00543 | 0.00543 | 0.00000 |
| [161;176] | [129;144] | [97;112] | 86 | 0.00393 | 402 | 0.00643 | -0.00251 | 0.00000 | -0.00251 |
| [177;192] | [129;144] | [97;112] | 1 | 0.00005 | 17 | 0.00027 | -0.00023 | 0.00000 | -0.00023 |
| [129;144] | [113;128] | [113;128] | 0 | 0.00000 | 3 | 0.00005 | -0.00005 | 0.00000 | -0.00005 |
| [145;160] | [113;128] | [113;128] | 756 | 0.03452 | 2585 | 0.04136 | -0.00684 | 0.00000 | -0.00684 |
| [161;176] | [113;128] | [113;128] | 626 | 0.02858 | 1854 | 0.02966 | -0.00108 | 0.00000 | -0.00108 |
| [177;192] | [113;128] | [113;128] | 0 | 0.00000 | 29 | 0.00046 | -0.00046 | 0.00000 | -0.00046 |
| [145;160] | [129;144] | [113;128] | 1106 | 0.05050 | 3020 | 0.04832 | 0.00218 | 0.00218 | 0.00000 |
| [161;176] | [129;144] | [113;128] | 5754 | 0.26274 | 14890 | 0.23824 | 0.02450 | 0.02450 | 0.00000 |
| [177;192] | [129;144] | [113;128] | 1199 | 0.05475 | 4789 | 0.07662 | -0.02188 | 0.00000 | -0.02188 |
| [161;176] | [145;160] | [113;128] | 735 | 0.03356 | 523 | 0.00837 | 0.02519 | 0.02519 | 0.00000 |
| [177;192] | [145;160] | [113;128] | 408 | 0.01863 | 720 | 0.01152 | 0.00711 | 0.00711 | 0.00000 |
| [145;160] | [113;128] | [129;144] | 0 | 0.00000 | 2 | 0.00003 | -0.00003 | 0.00000 | -0.00003 |
| [161;176] | [113;128] | [129;144] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [145;160] | [129;144] | [129;144] | 13 | 0.00059 | 187 | 0.00299 | -0.00240 | 0.00000 | -0.00240 |
| [161;176] | [129;144] | [129;144] | 2402 | 0.10968 | 6904 | 0.11046 | -0.00078 | 0.00000 | -0.00078 |
| [177;192] | [129;144] | [129;144] | 2314 | 0.10566 | 6429 | 0.10286 | 0.00280 | 0.00280 | 0.00000 |
| [145;160] | [145;160] | [129;144] | 0 | 0.00000 | 2 | 0.00003 | -0.00003 | 0.00000 | -0.00003 |
| [161;176] | [145;160] | [129;144] | 1323 | 0.06041 | 4012 | 0.06419 | -0.00378 | 0.00000 | -0.00378 |
| [177;192] | [145;160] | [129;144] | 4072 | 0.18594 | 12411 | 0.19858 | -0.01264 | 0.00000 | -0.01264 |
| [193;208] | [145;160] | [129;144] | 6 | 0.00027 | 153 | 0.00245 | -0.00217 | 0.00000 | -0.00217 |
| [177;192] | [161;176] | [129;144] | 4 | 0.00018 | 1 | 0.00002 | 0.00017 | 0.00017 | 0.00000 |
| [193;208] | [161;176] | [129;144] | 0 | 0.00000 | 3 | 0.00005 | -0.00005 | 0.00000 | -0.00005 |
| [161;176] | [129;144] | [145;160] | 4 | 0.00018 | 0 | 0.00000 | 0.00018 | 0.00018 | 0.00000 |
| [161;176] | [145;160] | [145;160] | 3 | 0.00014 | 3 | 0.00005 | 0.00009 | 0.00009 | 0.00000 |
| [177;192] | [145;160] | [145;160] | 232 | 0.01059 | 136 | 0.00218 | 0.00842 | 0.00842 | 0.00000 |
| [193;208] | [145;160] | [145;160] | 9 | 0.00041 | 11 | 0.00018 | 0.00023 | 0.00023 | 0.00000 |
| [177;192] | [161;176] | [145;160] | 10 | 0.00046 | 1 | 0.00002 | 0.00044 | 0.00044 | 0.00000 |
| [193;208] | [161;176] | [145;160] | 1 | 0.00005 | 3 | 0.00005 | 0.00000 | 0.00000 | 0.00000 |
| Totais | | | 21900 | 1.00000 | 62500 | 1.00000 | 0.00000 | 0.07676 | -0.07676 |

Tabela 7. Comparação entre frequências absolutas e relativas para intervalos de cor relativamente à imagem original e imagem simulada (#3).

Na tabela 8 apresenta-se um quadro resumo da percentagem de *pixels* que alteraram o seu intervalo de cor bem como a percentagem de *pixels* existentes em intervalos de cor que desaparecem nas imagens simuladas podendo constatar-se as percentagens de alterações de intervalo de cor e de *pixels* que desaparecem não apresentam valores significativos.

| Imagem simulada | #1 | #2 | #3 |
|---|-----------|-----------|-----------|
| % de <i>pixels</i> em novos intervalos de cor | 0,2032 | 0,1584 | 0,0768 |
| % de <i>pixels</i> em intervalos de cor que desaparecem | 0,0639 | 0,0411 | 0,0183 |

Tabela 8. Frequência de *pixels* (em %) convertidos em novos intervalos de cor e intervalos de cor que desaparecem nas imagens simuladas

Finalmente, na tabela 9 apresenta-se um quadro resumo dos modelos teóricos e respectivos parâmetros obtidos para a imagem original e para as três imagens simuladas.

| Canais | Imagens | Modelo | C₀ | C₁ | a |
|---------------|------------------------|-----------------|----------------------|----------------------|----------------|
| Red | Imagem original | Esférico | 18,8596 | 13,8028 | 107,802 |
| | Imagem simulada 1 | Esférico | 37,0817 | 26,7949 | 124,561 |
| | Imagem simulada 2 | Esférico | 39,0122 | 33,7474 | 85,1623 |
| | Imagem simulada 3 | Esférico | 34,6455 | 27,0308 | 88,9598 |
| Green | Imagem original | Esférico | 20,0838 | 17,8525 | 109,79 |
| | Imagem simulada 1 | Esférico | 28,6701 | 17,9309 | 126,1279 |
| | Imagem simulada 2 | Esférico | 29,3696 | 25,2773 | 83,8252 |
| | Imagem simulada 3 | Esférico | 29,1191 | 22,3841 | 83,9007 |
| Blue | Imagem original | Esférico | 30,1566 | 18,25 | 104,252 |
| | Imagem simulada 1 | Esférico | 27,3895 | 17,9050 | 120,6222 |
| | Imagem simulada 2 | Esférico | 30,7747 | 25,8596 | 74,3365 |
| | Imagem simulada 3 | Esférico | 30,6084 | 23,8335 | 80,8055 |

Tabela 9. Comparação dos modelos teóricos da imagem original e das imagens simuladas

5.3.2. Amostra 2

Neste caso, também se pretende obter a variabilidade de um padrão de cor através da simulação condicional da imagem digitalizada (fig. 126) fornecida com formato *bitmap* – RGB com 8 bits por canal e com uma resolução de 146 x 138.



Fig. 126 Amostra de mármore

Assim, é necessário calcular os variogramas para as direcções de 0°, 45°, 90° e 135° e os respectivos modelos teóricos para cada canal. Os variogramas foram calculados com 40 passos e um passo de 4 *pixels* para X e Y, e os resultados obtidos pelo ajustamento automático estão reportados nas figuras 127 a 132.

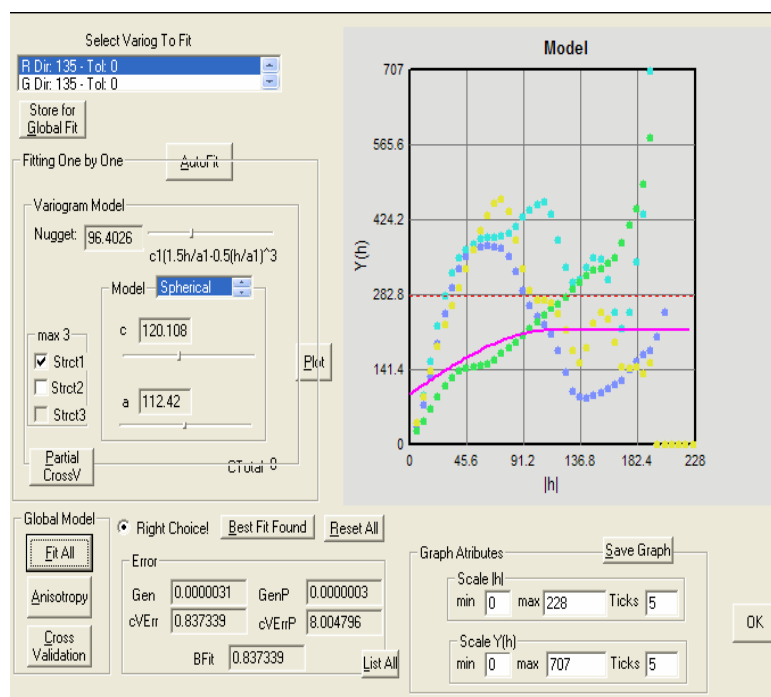


Fig. 127 Variogramas experimentais e modelo teórico, canal *Red*, modelo omnidireccional

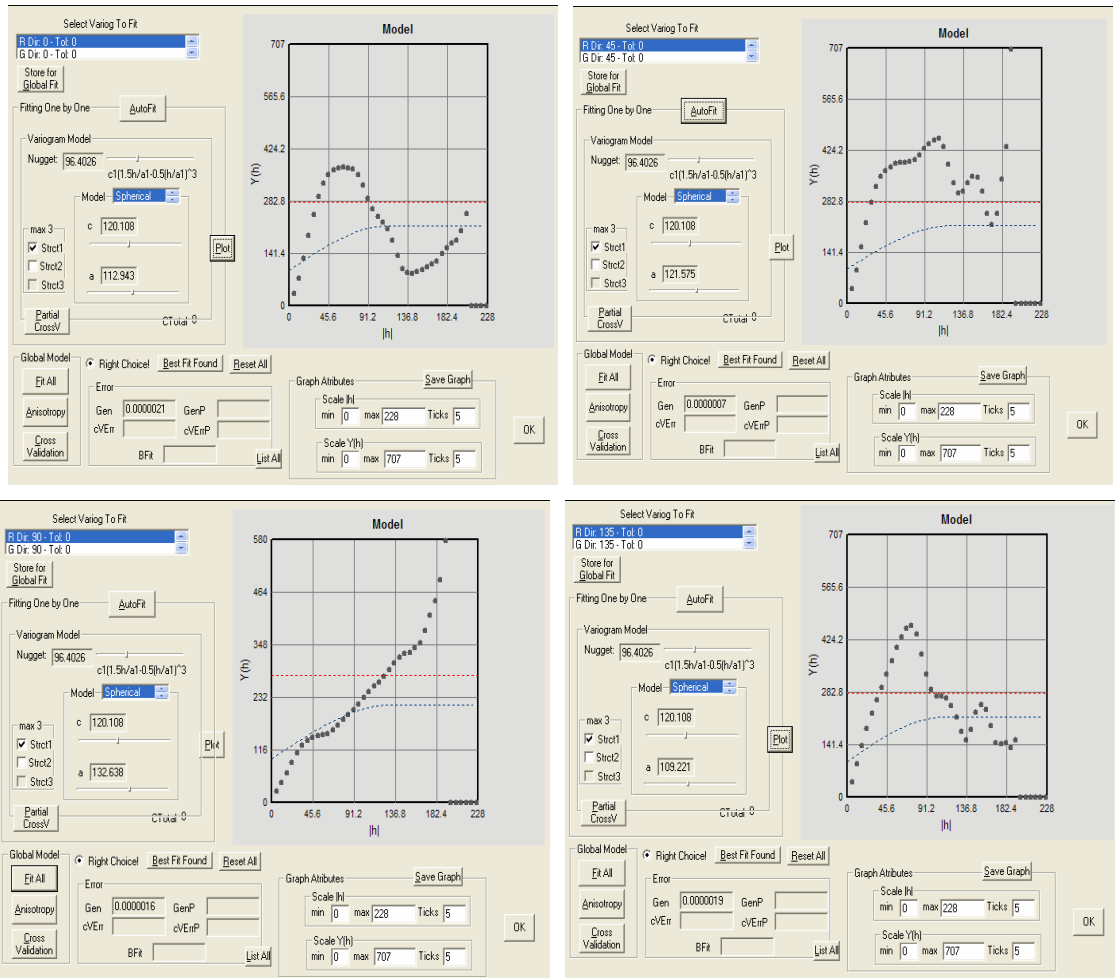


Fig. 128 Variogramas experimentais e modelo teórico, canal Red, direcção 0°, 45°, 90° e , 135°

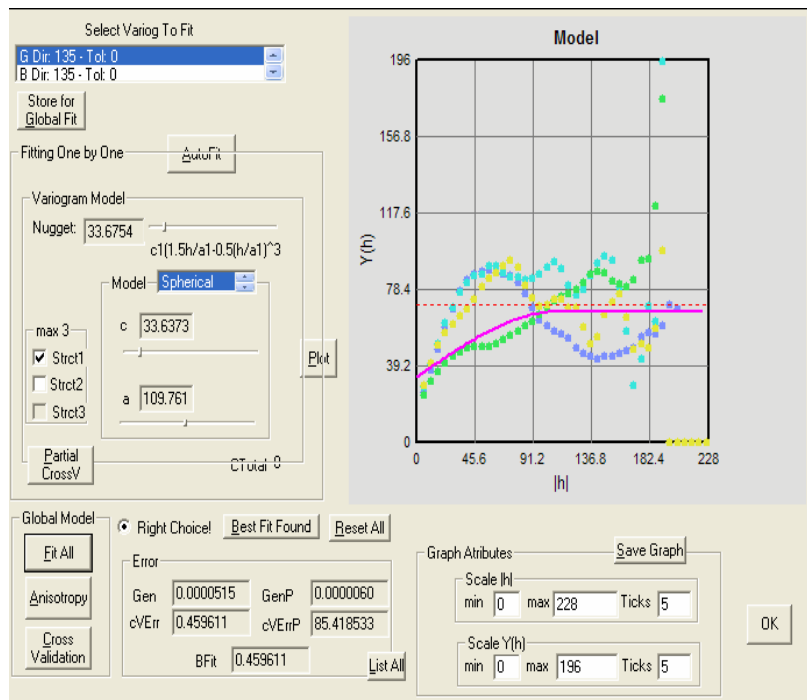


Fig. 129 Variogramas experimentais e modelo teórico, canal Green, modelo omnidireccional

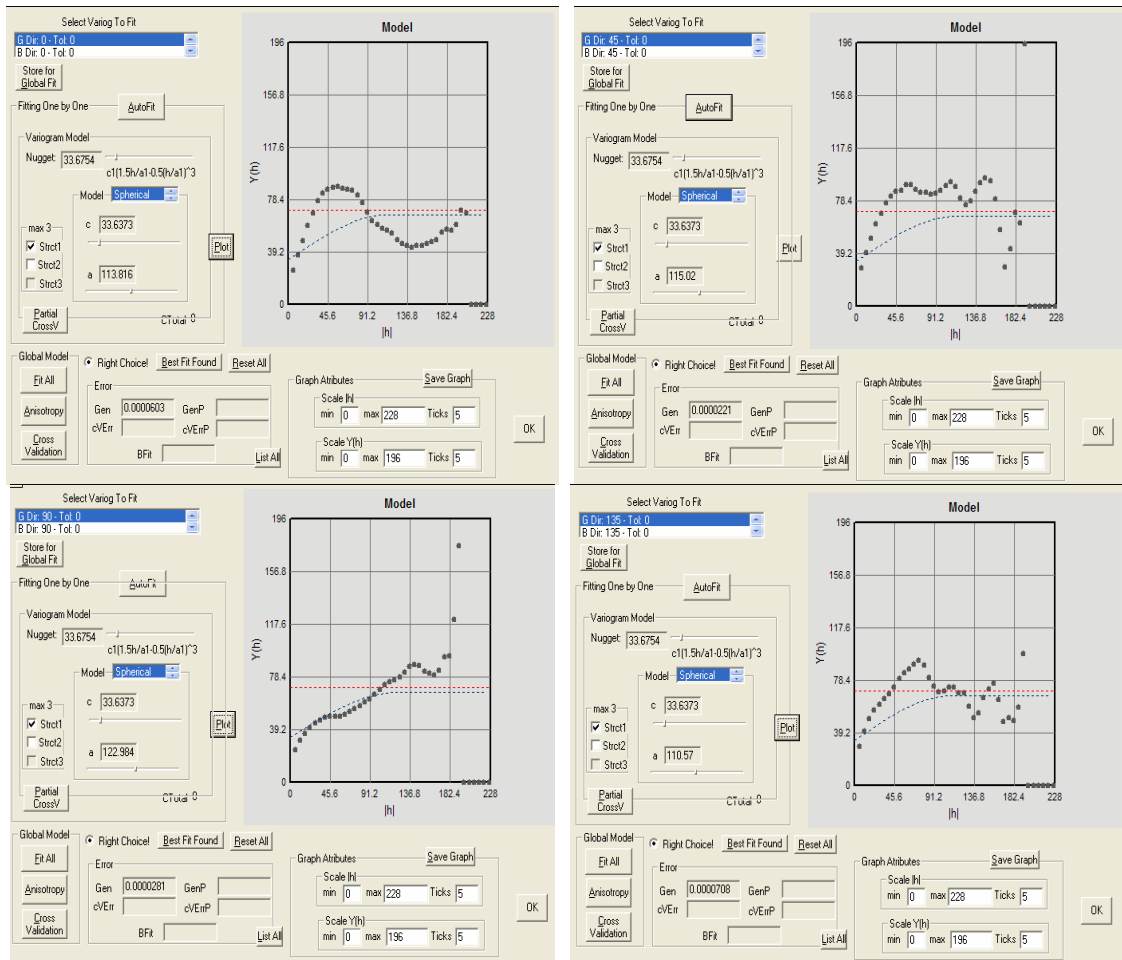


Fig. 130 Variogramas experimentais e modelo teórico, canal *Green*, direcção 0°, 45°, 90° e 135°

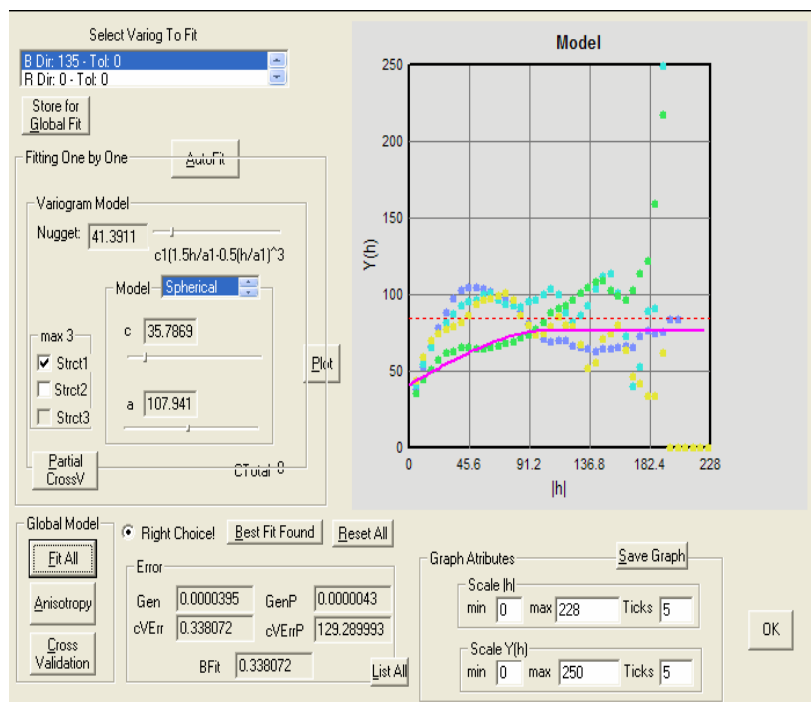


Fig. 131 Variogramas experimentais e modelo teórico, canal *Blue*, modelo omnidireccional

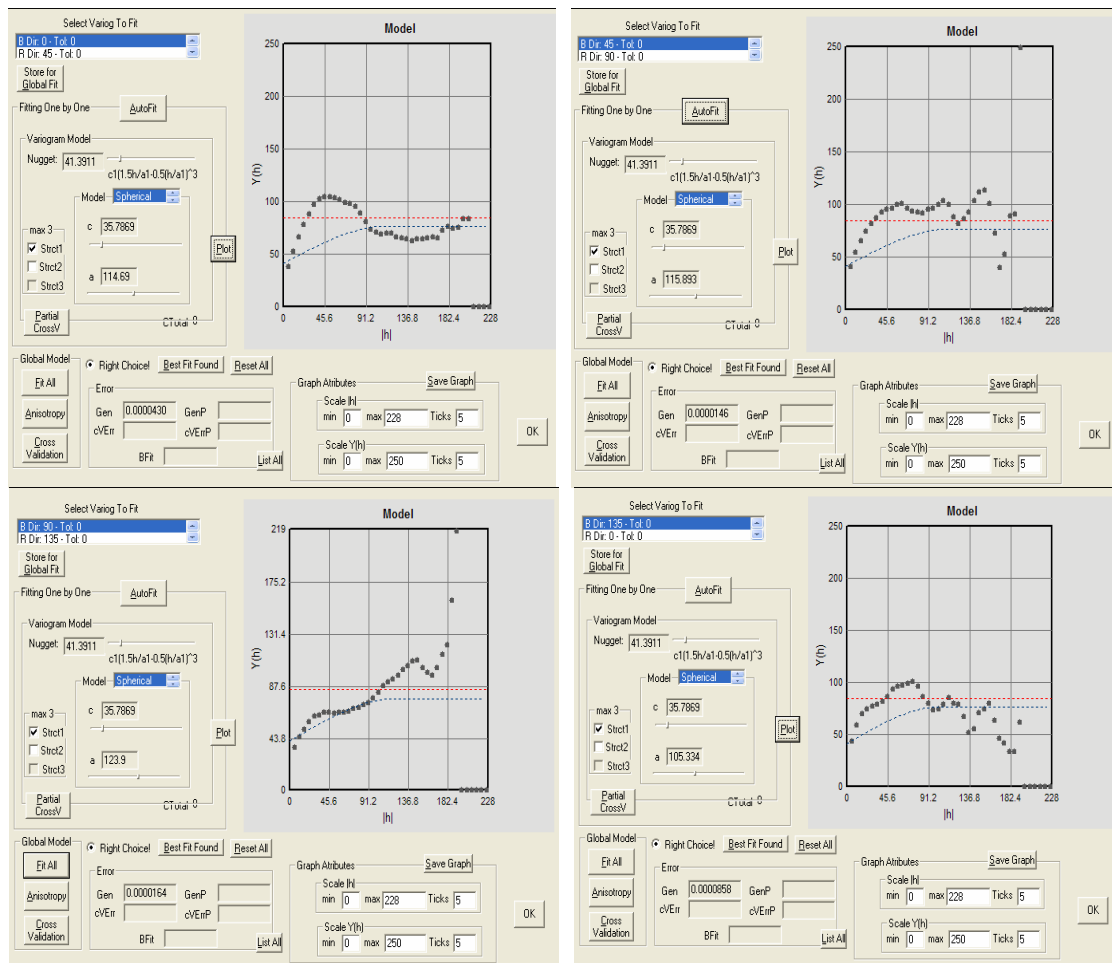


Fig. 132 Variogramas experimentais e modelo teórico, canal Blue, direcção 0°, 45°, 90° e 135°

Observa-se que a correlação entre os três canais de cor é sempre bastante elevada (tabela 10), variando o coeficiente de correlação entre 0,7443 (vermelho e azul) e 0,93 (verde e azul). Esta elevada correlação entre os três canais de cor mostra, mais uma vez, a necessidade de utilizar um modelo de coregionalização para a construção de imagens simuladas dos padrões de mármore.

| | <i>Red</i> | <i>Green</i> | <i>Blue</i> |
|--------------|------------|--------------|-------------|
| <i>Red</i> | 1,0000 | 0,8100 | 0,7443 |
| <i>Green</i> | 0,8100 | 1,0000 | 0,9300 |
| <i>Blue</i> | 0,7443 | 0,9300 | 1,0000 |

Tabela 10. Tabela de correlações entre os canais de cor Red, Green e Blue

Foram simuladas por Simulação Sequencial Gaussiana 3 imagens do canal *Red* e por Co-Simulação Sequencial Gaussiana 3 imagens dos canais *Green* e *Blue* segundo a metodologia proposta na figura 111. Os resultados são apresentados nas figuras 133 a 135.

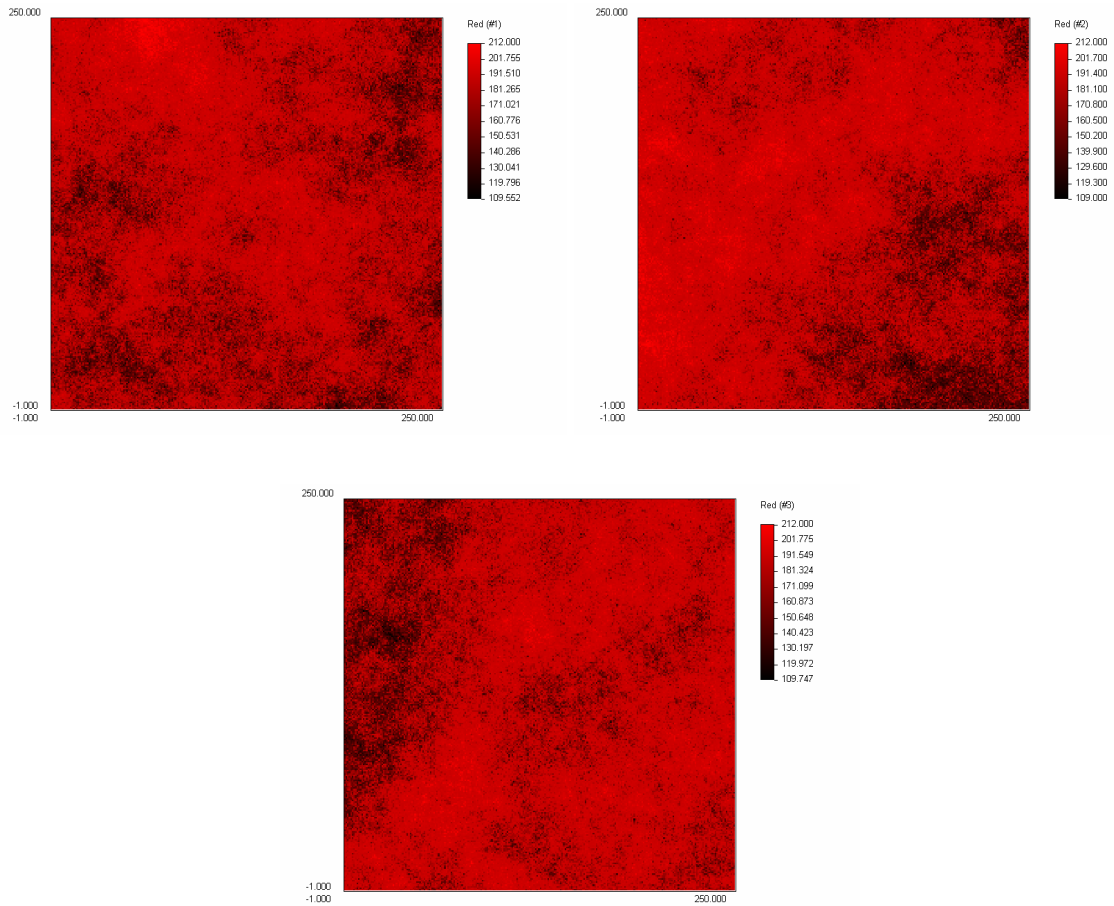
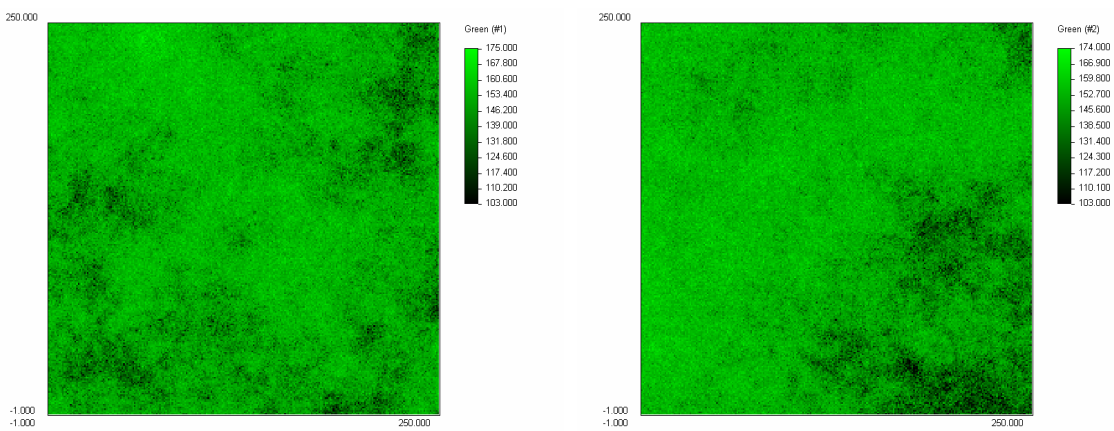


Fig. 133 3 imagens simuladas, canal *Red*



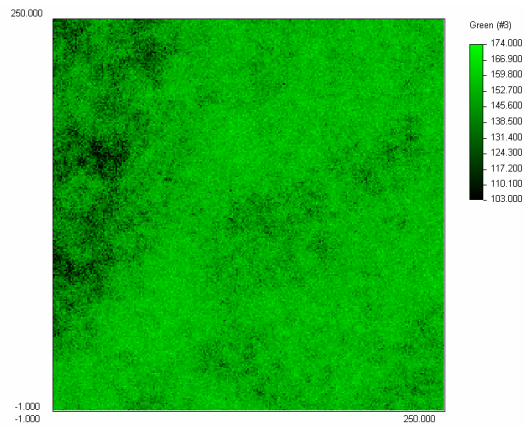


Fig. 134 3 imagens simuladas, canal *Green*

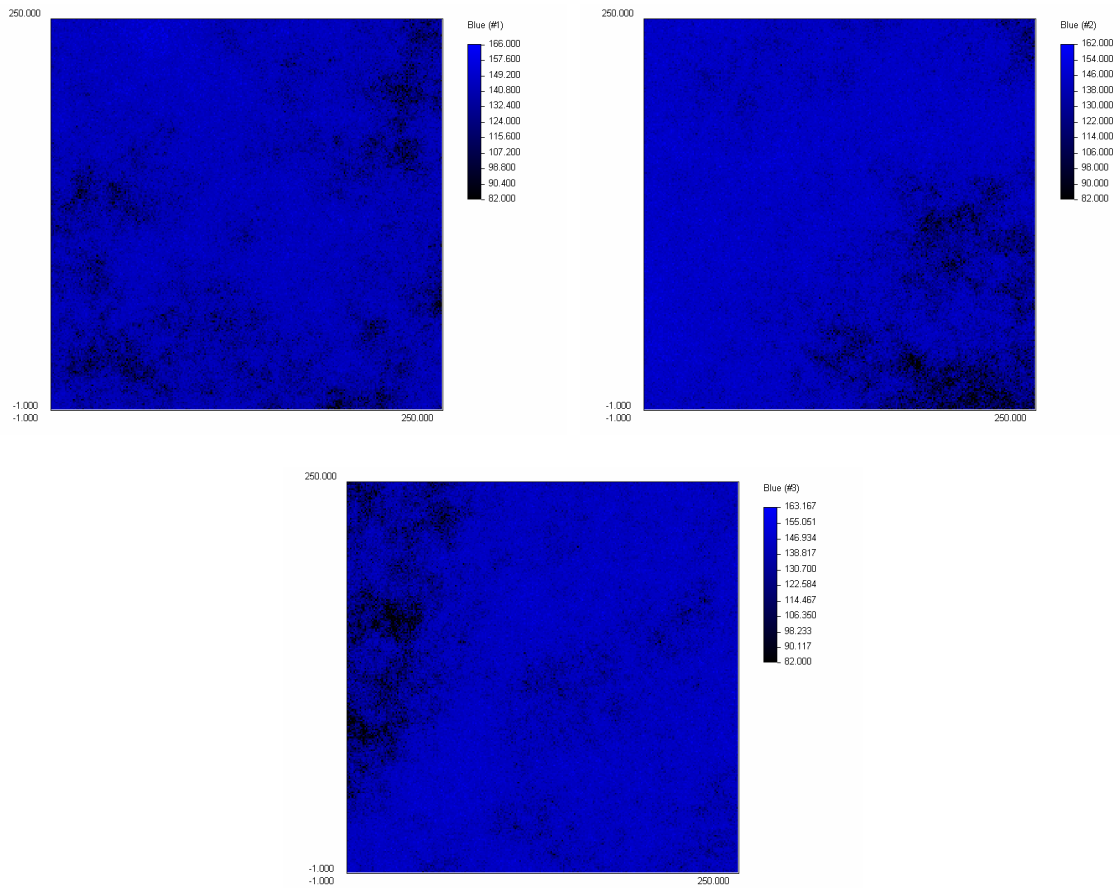


Fig. 135 3 imagens simuladas, canal *Blue*

A combinação dos canais de cor das imagens homólogas permite apresentar 3 imagens simuladas de mármore rosa (fig. 136).

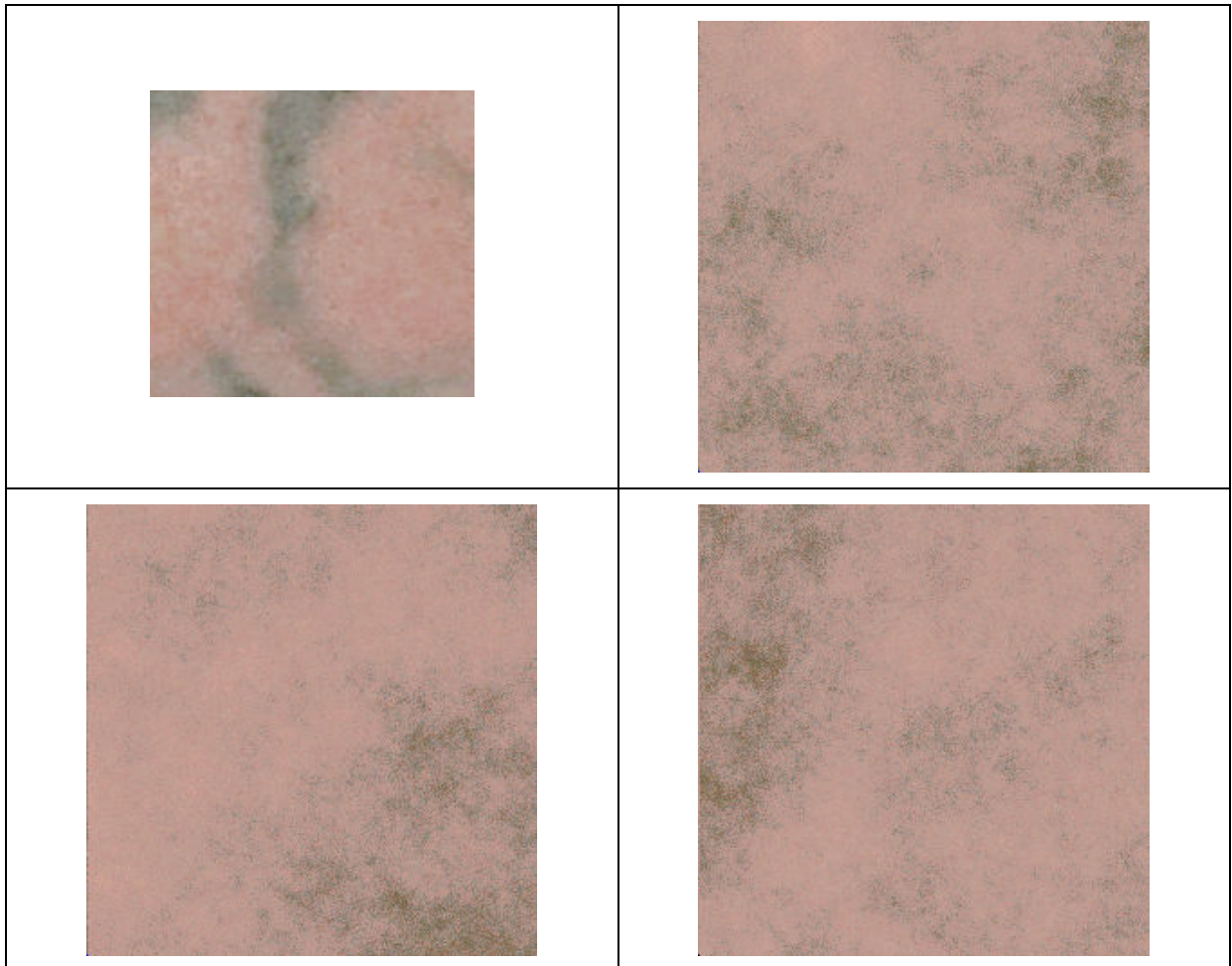


Fig. 136 Amostra inicial e 3 amostras simuladas

Para validar os resultados obtidos foram calculados os variogramas das imagens simuladas (figuras 137 a 139) e comparados os respectivos estatísticos de cor relativos à imagem real e às três simulações (#1, #2 e #3) (tabelas 11 a 16).

Como já foi referido na secção 5.3.1. em alguns canais é notória a influência do conjunto de ponto usados no ajuste automático do modelo teórico (como por exemplo no modelo teórico do canal Blue ilustrado na figura 139).

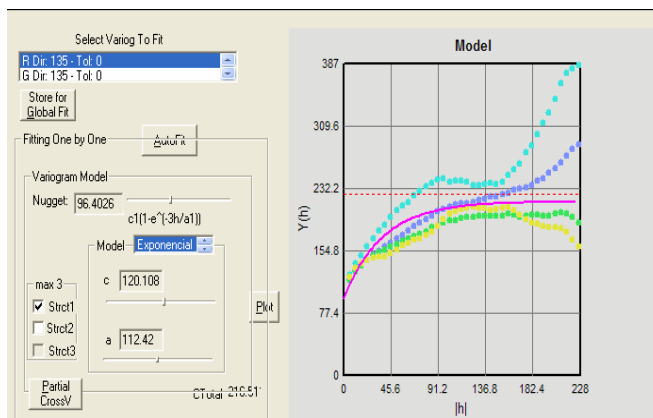
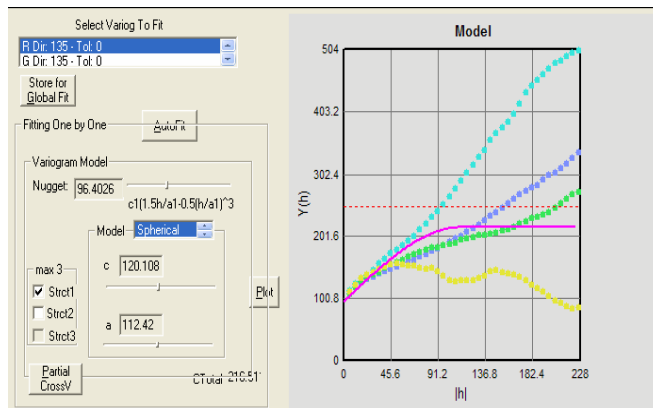
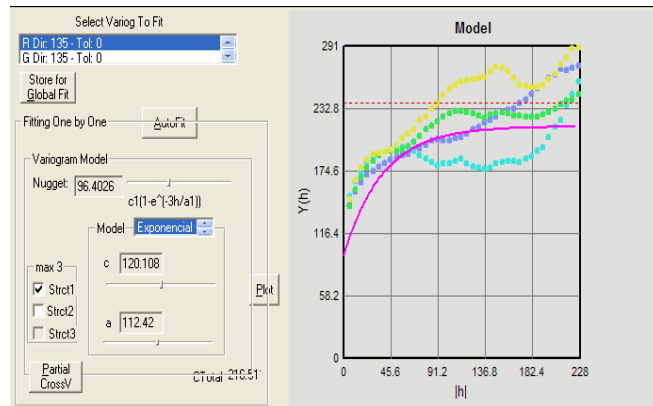


Fig. 137 Variogramas experimentais do canal Red sobrepostos ao modelo teórico dos dados

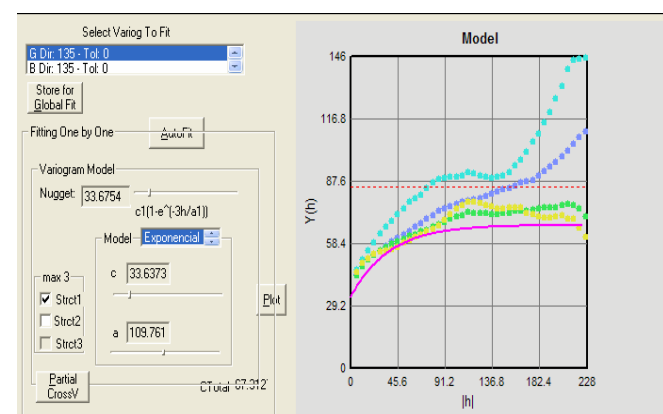
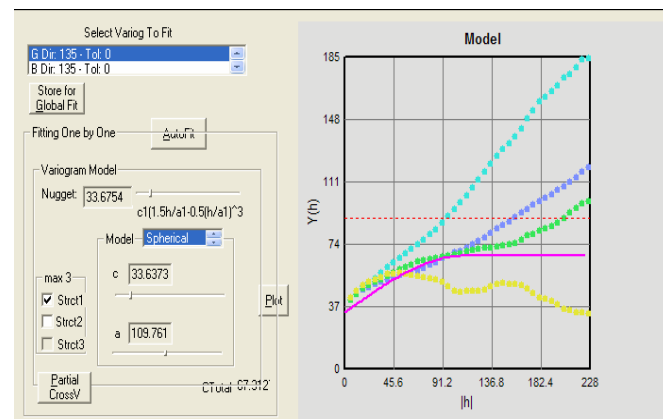
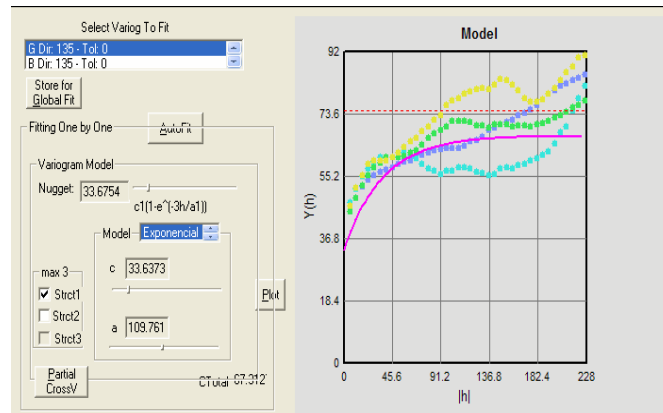


Fig. 138 Variogramas experimentais do canal *Green* sobrepostos ao modelo teórico dos dados

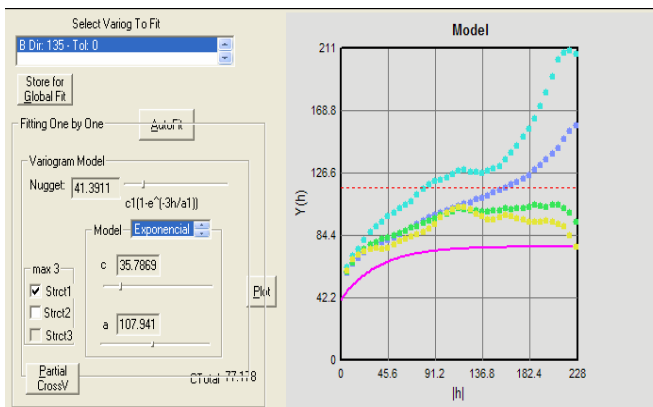
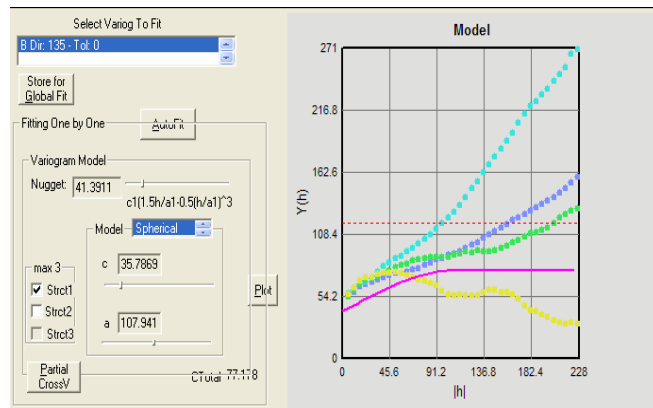
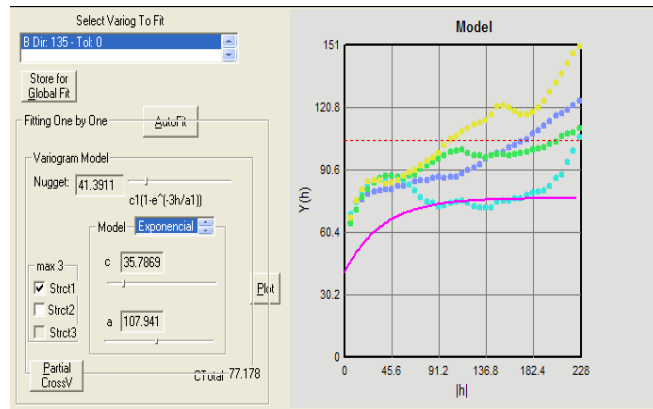


Fig. 139 Variogramas experimentais do canal *Blue* sobrepostos ao modelo teórico dos dados

| Intervalos dos canais de cor | | | Frequências da imagem original em valores | | Frequências da imagem simulada em valores | | Diferenças | | |
|------------------------------|-----------|----------|---|-----------|---|-----------|------------|-----------|-----------|
| R | G | B | absolutos | relativos | absolutos | relativos | totais | positivos | negativos |
| [81;96] | [81;96] | [65;80] | 6 | 0.00030 | 5 | 0.00008 | 0.00022 | 0.00022 | 0.00000 |
| [97;112] | [81;96] | [65;80] | 6 | 0.00030 | 26 | 0.00042 | -0.00012 | 0.00000 | -0.00012 |
| [113;128] | [81;96] | [65;80] | 0 | 0.00000 | 71 | 0.00114 | -0.00114 | 0.00000 | -0.00114 |
| [129;144] | [81;96] | [65;80] | 0 | 0.00000 | 21 | 0.00034 | -0.00034 | 0.00000 | -0.00034 |
| [145;160] | [81;96] | [65;80] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [97;112] | [97;112] | [65;80] | 5 | 0.00025 | 35 | 0.00056 | -0.00031 | 0.00000 | -0.00031 |
| [113;128] | [97;112] | [65;80] | 0 | 0.00000 | 245 | 0.00392 | -0.00392 | 0.00000 | -0.00392 |
| [129;144] | [97;112] | [65;80] | 0 | 0.00000 | 127 | 0.00203 | -0.00203 | 0.00000 | -0.00203 |
| [145;160] | [97;112] | [65;80] | 0 | 0.00000 | 4 | 0.00006 | -0.00006 | 0.00000 | -0.00006 |
| [97;112] | [113;128] | [65;80] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [113;128] | [113;128] | [65;80] | 0 | 0.00000 | 5 | 0.00008 | -0.00008 | 0.00000 | -0.00008 |
| [129;144] | [113;128] | [65;80] | 0 | 0.00000 | 9 | 0.00014 | -0.00014 | 0.00000 | -0.00014 |
| [145;160] | [113;128] | [65;80] | 0 | 0.00000 | 2 | 0.00003 | -0.00003 | 0.00000 | -0.00003 |
| [97;112] | [81;96] | [81;96] | 2 | 0.00010 | 0 | 0.00000 | 0.00010 | 0.00010 | 0.00000 |
| [81;96] | [97;112] | [81;96] | 1 | 0.00005 | 0 | 0.00000 | 0.00005 | 0.00005 | 0.00000 |
| [97;112] | [97;112] | [81;96] | 53 | 0.00263 | 24 | 0.00038 | 0.00225 | 0.00225 | 0.00000 |
| [113;128] | [97;112] | [81;96] | 53 | 0.00263 | 297 | 0.00475 | -0.00212 | 0.00000 | -0.00212 |
| [129;144] | [97;112] | [81;96] | 2 | 0.00010 | 244 | 0.00390 | -0.00380 | 0.00000 | -0.00380 |
| [145;160] | [97;112] | [81;96] | 0 | 0.00000 | 33 | 0.00053 | -0.00053 | 0.00000 | -0.00053 |
| [161;176] | [97;112] | [81;96] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [97;112] | [113;128] | [81;96] | 1 | 0.00005 | 8 | 0.00013 | -0.00008 | 0.00000 | -0.00008 |
| [113;128] | [113;128] | [81;96] | 32 | 0.00159 | 167 | 0.00267 | -0.00108 | 0.00000 | -0.00108 |
| [129;144] | [113;128] | [81;96] | 7 | 0.00035 | 271 | 0.00434 | -0.00399 | 0.00000 | -0.00399 |
| [145;160] | [113;128] | [81;96] | 0 | 0.00000 | 72 | 0.00115 | -0.00115 | 0.00000 | -0.00115 |
| [161;176] | [113;128] | [81;96] | 0 | 0.00000 | 8 | 0.00013 | -0.00013 | 0.00000 | -0.00013 |
| [97;112] | [97;112] | [97;112] | 8 | 0.00040 | 4 | 0.00006 | 0.00033 | 0.00033 | 0.00000 |
| [113;128] | [97;112] | [97;112] | 69 | 0.00342 | 93 | 0.00149 | 0.00194 | 0.00194 | 0.00000 |
| [129;144] | [97;112] | [97;112] | 5 | 0.00025 | 94 | 0.00150 | -0.00126 | 0.00000 | -0.00126 |
| [145;160] | [97;112] | [97;112] | 0 | 0.00000 | 13 | 0.00021 | -0.00021 | 0.00000 | -0.00021 |
| [161;176] | [97;112] | [97;112] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [97;112] | [113;128] | [97;112] | 14 | 0.00069 | 7 | 0.00011 | 0.00058 | 0.00058 | 0.00000 |
| [113;128] | [113;128] | [97;112] | 624 | 0.03097 | 909 | 0.01454 | 0.01643 | 0.01643 | 0.00000 |
| [129;144] | [113;128] | [97;112] | 789 | 0.03916 | 2191 | 0.03506 | 0.00410 | 0.00410 | 0.00000 |
| [145;160] | [113;128] | [97;112] | 260 | 0.01290 | 1569 | 0.02510 | -0.01220 | 0.00000 | -0.01220 |
| [161;176] | [113;128] | [97;112] | 120 | 0.00596 | 504 | 0.00806 | -0.00211 | 0.00000 | -0.00211 |
| [177;192] | [113;128] | [97;112] | 4 | 0.00020 | 15 | 0.00024 | -0.00004 | 0.00000 | -0.00004 |
| [113;128] | [129;144] | [97;112] | 4 | 0.00020 | 5 | 0.00008 | 0.00012 | 0.00012 | 0.00000 |
| [129;144] | [129;144] | [97;112] | 22 | 0.00109 | 38 | 0.00061 | 0.00048 | 0.00048 | 0.00000 |
| [145;160] | [129;144] | [97;112] | 37 | 0.00184 | 126 | 0.00202 | -0.00018 | 0.00000 | -0.00018 |
| [161;176] | [129;144] | [97;112] | 68 | 0.00338 | 136 | 0.00218 | 0.00120 | 0.00120 | 0.00000 |

Tabela 11. Comparação entre frequências absolutas e relativas para intervalos de cor relativamente à imagem original e imagem simulada (#1).

| Intervalos dos canais de cor | | | Frequências da imagem original em valores | | Frequências da imagem simulada em valores | | Diferenças | | |
|------------------------------|-----------|-----------|---|----------------|---|----------------|----------------|----------------|-----------------|
| R | G | B | absolutos | relativos | absolutos | relativos | totais | positivos | negativos |
| [177;192] | [129;144] | [97;112] | 12 | 0.00060 | 11 | 0.00018 | 0.00042 | 0.00042 | 0.00000 |
| [129;144] | [97;112] | [113;128] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [97;112] | [113;128] | [113;128] | 1 | 0.00005 | 1 | 0.00002 | 0.00003 | 0.00003 | 0.00000 |
| [113;128] | [113;128] | [113;128] | 195 | 0.00968 | 269 | 0.00430 | 0.00537 | 0.00537 | 0.00000 |
| [129;144] | [113;128] | [113;128] | 650 | 0.03226 | 1380 | 0.02208 | 0.01018 | 0.01018 | 0.00000 |
| [145;160] | [113;128] | [113;128] | 606 | 0.03008 | 1786 | 0.02858 | 0.00150 | 0.00150 | 0.00000 |
| [161;176] | [113;128] | [113;128] | 304 | 0.01509 | 1180 | 0.01888 | -0.00379 | 0.00000 | -0.00379 |
| [177;192] | [113;128] | [113;128] | 9 | 0.00045 | 63 | 0.00101 | -0.00056 | 0.00000 | -0.00056 |
| [113;128] | [129;144] | [113;128] | 29 | 0.00144 | 80 | 0.00128 | 0.00016 | 0.00016 | 0.00000 |
| [129;144] | [129;144] | [113;128] | 429 | 0.02129 | 1069 | 0.01710 | 0.00419 | 0.00419 | 0.00000 |
| [145;160] | [129;144] | [113;128] | 1809 | 0.08979 | 5558 | 0.08893 | 0.00086 | 0.00086 | 0.00000 |
| [161;176] | [129;144] | [113;128] | 4966 | 0.24648 | 15313 | 0.24501 | 0.00147 | 0.00147 | 0.00000 |
| [177;192] | [129;144] | [113;128] | 1447 | 0.07182 | 4830 | 0.07728 | -0.00546 | 0.00000 | -0.00546 |
| [193;208] | [129;144] | [113;128] | 0 | 0.00000 | 2 | 0.00003 | -0.00003 | 0.00000 | -0.00003 |
| [145;160] | [145;160] | [113;128] | 1 | 0.00005 | 1 | 0.00002 | 0.00003 | 0.00003 | 0.00000 |
| [161;176] | [145;160] | [113;128] | 28 | 0.00139 | 55 | 0.00088 | 0.00051 | 0.00051 | 0.00000 |
| [177;192] | [145;160] | [113;128] | 36 | 0.00179 | 88 | 0.00141 | 0.00038 | 0.00038 | 0.00000 |
| [193;208] | [145;160] | [113;128] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [113;128] | [113;128] | [129;144] | 1 | 0.00005 | 1 | 0.00002 | 0.00003 | 0.00003 | 0.00000 |
| [129;144] | [113;128] | [129;144] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [145;160] | [113;128] | [129;144] | 2 | 0.00010 | 4 | 0.00006 | 0.00004 | 0.00004 | 0.00000 |
| [161;176] | [113;128] | [129;144] | 1 | 0.00005 | 0 | 0.00000 | 0.00005 | 0.00005 | 0.00000 |
| [177;192] | [113;128] | [129;144] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [113;128] | [129;144] | [129;144] | 0 | 0.00000 | 3 | 0.00005 | -0.00005 | 0.00000 | -0.00005 |
| [129;144] | [129;144] | [129;144] | 20 | 0.00099 | 51 | 0.00082 | 0.00018 | 0.00018 | 0.00000 |
| [145;160] | [129;144] | [129;144] | 433 | 0.02149 | 1125 | 0.01800 | 0.00349 | 0.00349 | 0.00000 |
| [161;176] | [129;144] | [129;144] | 2572 | 0.12766 | 9804 | 0.15686 | -0.02921 | 0.00000 | -0.02921 |
| [177;192] | [129;144] | [129;144] | 1982 | 0.09837 | 7356 | 0.11770 | -0.01932 | 0.00000 | -0.01932 |
| [193;208] | [129;144] | [129;144] | 0 | 0.00000 | 29 | 0.00046 | -0.00046 | 0.00000 | -0.00046 |
| [129;144] | [145;160] | [129;144] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [145;160] | [145;160] | [129;144] | 21 | 0.00104 | 26 | 0.00042 | 0.00063 | 0.00063 | 0.00000 |
| [161;176] | [145;160] | [129;144] | 668 | 0.03315 | 1622 | 0.02595 | 0.00720 | 0.00720 | 0.00000 |
| [177;192] | [145;160] | [129;144] | 1525 | 0.07569 | 3238 | 0.05181 | 0.02388 | 0.02388 | 0.00000 |
| [193;208] | [145;160] | [129;144] | 2 | 0.00010 | 117 | 0.00187 | -0.00177 | 0.00000 | -0.00177 |
| [193;208] | [161;176] | [129;144] | 1 | 0.00005 | 0 | 0.00000 | 0.00005 | 0.00005 | 0.00000 |
| [161;176] | [129;144] | [145;160] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [145;160] | [145;160] | [145;160] | 1 | 0.00005 | 0 | 0.00000 | 0.00005 | 0.00005 | 0.00000 |
| [161;176] | [145;160] | [145;160] | 16 | 0.00079 | 4 | 0.00006 | 0.00073 | 0.00073 | 0.00000 |
| [177;192] | [145;160] | [145;160] | 169 | 0.00839 | 41 | 0.00066 | 0.00773 | 0.00773 | 0.00000 |
| [193;208] | [145;160] | [145;160] | 13 | 0.00065 | 5 | 0.00008 | 0.00057 | 0.00057 | 0.00000 |
| [177;192] | [161;176] | [145;160] | 2 | 0.00010 | 0 | 0.00000 | 0.00010 | 0.00010 | 0.00000 |
| [193;208] | [161;176] | [145;160] | 5 | 0.00025 | 0 | 0.00000 | 0.00025 | 0.00025 | 0.00000 |
| Totais | | | 20148 | 1.00000 | 62500 | 1.00000 | 0.00000 | 0.09788 | -0.09788 |

Tabela 12. Comparação entre frequências absolutas e relativas para intervalos de cor relativamente à imagem original e imagem simulada (#1) (continuação)

| Intervalos dos canais de cor | | | Frequências da imagem original em valores | | Frequências da imagem simulada em valores | | Diferenças | | |
|------------------------------|-----------|----------|---|-----------|---|-----------|------------|-----------|-----------|
| R | G | B | absolutos | relativos | absolutos | relativos | totais | positivos | negativos |
| [81;96] | [81;96] | [65;80] | 6 | 0.00030 | 7 | 0.00011 | 0.00019 | 0.00019 | 0.00000 |
| [97;112] | [81;96] | [65;80] | 6 | 0.00030 | 67 | 0.00107 | -0.00077 | 0.00000 | -0.00077 |
| [113;128] | [81;96] | [65;80] | 0 | 0.00000 | 220 | 0.00352 | -0.00352 | 0.00000 | -0.00352 |
| [129;144] | [81;96] | [65;80] | 0 | 0.00000 | 58 | 0.00093 | -0.00093 | 0.00000 | -0.00093 |
| [145;160] | [81;96] | [65;80] | 0 | 0.00000 | 4 | 0.00006 | -0.00006 | 0.00000 | -0.00006 |
| [97;112] | [97;112] | [65;80] | 5 | 0.00025 | 46 | 0.00074 | -0.00049 | 0.00000 | -0.00049 |
| [113;128] | [97;112] | [65;80] | 0 | 0.00000 | 289 | 0.00462 | -0.00462 | 0.00000 | -0.00462 |
| [129;144] | [97;112] | [65;80] | 0 | 0.00000 | 168 | 0.00269 | -0.00269 | 0.00000 | -0.00269 |
| [145;160] | [97;112] | [65;80] | 0 | 0.00000 | 11 | 0.00018 | -0.00018 | 0.00000 | -0.00018 |
| [113;128] | [113;128] | [65;80] | 0 | 0.00000 | 3 | 0.00005 | -0.00005 | 0.00000 | -0.00005 |
| [129;144] | [113;128] | [65;80] | 0 | 0.00000 | 2 | 0.00003 | -0.00003 | 0.00000 | -0.00003 |
| [145;160] | [113;128] | [65;80] | 0 | 0.00000 | 2 | 0.00003 | -0.00003 | 0.00000 | -0.00003 |
| [97;112] | [81;96] | [81;96] | 2 | 0.00010 | 0 | 0.00000 | 0.00010 | 0.00010 | 0.00000 |
| [113;128] | [81;96] | [81;96] | 0 | 0.00000 | 2 | 0.00003 | -0.00003 | 0.00000 | -0.00003 |
| [129;144] | [81;96] | [81;96] | 0 | 0.00000 | 2 | 0.00003 | -0.00003 | 0.00000 | -0.00003 |
| [81;96] | [97;112] | [81;96] | 1 | 0.00005 | 1 | 0.00002 | 0.00003 | 0.00003 | 0.00000 |
| [97;112] | [97;112] | [81;96] | 53 | 0.00263 | 30 | 0.00048 | 0.00215 | 0.00215 | 0.00000 |
| [113;128] | [97;112] | [81;96] | 53 | 0.00263 | 433 | 0.00693 | -0.00430 | 0.00000 | -0.00430 |
| [129;144] | [97;112] | [81;96] | 2 | 0.00010 | 356 | 0.00570 | -0.00560 | 0.00000 | -0.00560 |
| [145;160] | [97;112] | [81;96] | 0 | 0.00000 | 42 | 0.00067 | -0.00067 | 0.00000 | -0.00067 |
| [161;176] | [97;112] | [81;96] | 0 | 0.00000 | 6 | 0.00010 | -0.00010 | 0.00000 | -0.00010 |
| [97;112] | [113;128] | [81;96] | 1 | 0.00005 | 4 | 0.00006 | -0.00001 | 0.00000 | -0.00001 |
| [113;128] | [113;128] | [81;96] | 32 | 0.00159 | 135 | 0.00216 | -0.00057 | 0.00000 | -0.00057 |
| [129;144] | [113;128] | [81;96] | 7 | 0.00035 | 186 | 0.00298 | -0.00263 | 0.00000 | -0.00263 |
| [145;160] | [113;128] | [81;96] | 0 | 0.00000 | 71 | 0.00114 | -0.00114 | 0.00000 | -0.00114 |
| [161;176] | [113;128] | [81;96] | 0 | 0.00000 | 5 | 0.00008 | -0.00008 | 0.00000 | -0.00008 |
| [97;112] | [97;112] | [97;112] | 8 | 0.00040 | 5 | 0.00008 | 0.00032 | 0.00032 | 0.00000 |
| [113;128] | [97;112] | [97;112] | 69 | 0.00342 | 196 | 0.00314 | 0.00029 | 0.00029 | 0.00000 |
| [129;144] | [97;112] | [97;112] | 5 | 0.00025 | 215 | 0.00344 | -0.00319 | 0.00000 | -0.00319 |
| [145;160] | [97;112] | [97;112] | 0 | 0.00000 | 40 | 0.00064 | -0.00064 | 0.00000 | -0.00064 |
| [161;176] | [97;112] | [97;112] | 0 | 0.00000 | 6 | 0.00010 | -0.00010 | 0.00000 | -0.00010 |
| [97;112] | [113;128] | [97;112] | 14 | 0.00069 | 13 | 0.00021 | 0.00049 | 0.00049 | 0.00000 |
| [113;128] | [113;128] | [97;112] | 624 | 0.03097 | 755 | 0.01208 | 0.01889 | 0.01889 | 0.00000 |
| [129;144] | [113;128] | [97;112] | 789 | 0.03916 | 1876 | 0.03002 | 0.00914 | 0.00914 | 0.00000 |
| [145;160] | [113;128] | [97;112] | 260 | 0.01290 | 1309 | 0.02094 | -0.00804 | 0.00000 | -0.00804 |
| [161;176] | [113;128] | [97;112] | 120 | 0.00596 | 575 | 0.00920 | -0.00324 | 0.00000 | -0.00324 |
| [177;192] | [113;128] | [97;112] | 4 | 0.00020 | 37 | 0.00059 | -0.00039 | 0.00000 | -0.00039 |
| [113;128] | [129;144] | [97;112] | 4 | 0.00020 | 1 | 0.00002 | 0.00018 | 0.00018 | 0.00000 |
| [129;144] | [129;144] | [97;112] | 22 | 0.00109 | 18 | 0.00029 | 0.00080 | 0.00080 | 0.00000 |
| [145;160] | [129;144] | [97;112] | 37 | 0.00184 | 75 | 0.00120 | 0.00064 | 0.00064 | 0.00000 |

Tabela 13. Comparação entre frequências absolutas e relativas para intervalos de cor relativamente à imagem original e imagem simulada (#2)

| Intervalos dos canais de cor | | | Frequências da imagem original em valores | | Frequências da imagem simulada em valores | | Diferenças | | |
|------------------------------|-----------|-----------|---|----------------|---|----------------|----------------|----------------|-----------------|
| R | G | B | absolutos | relativos | absolutos | relativos | totais | positivos | negativos |
| [161;176] | [129;144] | [97;112] | 68 | 0.00338 | 105 | 0.00168 | 0.00170 | 0.00170 | 0.00000 |
| [177;192] | [129;144] | [97;112] | 12 | 0.00060 | 15 | 0.00024 | 0.00036 | 0.00036 | 0.00000 |
| [113;128] | [97;112] | [113;128] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [145;160] | [97;112] | [113;128] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [97;112] | [113;128] | [113;128] | 1 | 0.00005 | 1 | 0.00002 | 0.00003 | 0.00003 | 0.00000 |
| [113;128] | [113;128] | [113;128] | 195 | 0.00968 | 198 | 0.00317 | 0.00651 | 0.00651 | 0.00000 |
| [129;144] | [113;128] | [113;128] | 650 | 0.03226 | 1031 | 0.01650 | 0.01577 | 0.01577 | 0.00000 |
| [145;160] | [113;128] | [113;128] | 606 | 0.03008 | 1788 | 0.02861 | 0.00147 | 0.00147 | 0.00000 |
| [161;176] | [113;128] | [113;128] | 304 | 0.01509 | 1477 | 0.02363 | -0.00854 | 0.00000 | -0.00854 |
| [177;192] | [113;128] | [113;128] | 9 | 0.00045 | 168 | 0.00269 | -0.00224 | 0.00000 | -0.00224 |
| [113;128] | [129;144] | [113;128] | 29 | 0.00144 | 33 | 0.00053 | 0.00091 | 0.00091 | 0.00000 |
| [129;144] | [129;144] | [113;128] | 429 | 0.02129 | 483 | 0.00773 | 0.01356 | 0.01356 | 0.00000 |
| [145;160] | [129;144] | [113;128] | 1809 | 0.08979 | 3157 | 0.05051 | 0.03927 | 0.03927 | 0.00000 |
| [161;176] | [129;144] | [113;128] | 4966 | 0.24648 | 13093 | 0.20949 | 0.03699 | 0.03699 | 0.00000 |
| [177;192] | [129;144] | [113;128] | 1447 | 0.07182 | 9261 | 0.14818 | -0.07636 | 0.00000 | -0.07636 |
| [193;208] | [129;144] | [113;128] | 0 | 0.00000 | 13 | 0.00021 | -0.00021 | 0.00000 | -0.00021 |
| [145;160] | [145;160] | [113;128] | 1 | 0.00005 | 0 | 0.00000 | 0.00005 | 0.00005 | 0.00000 |
| [161;176] | [145;160] | [113;128] | 28 | 0.00139 | 48 | 0.00077 | 0.00062 | 0.00062 | 0.00000 |
| [177;192] | [145;160] | [113;128] | 36 | 0.00179 | 163 | 0.00261 | -0.00082 | 0.00000 | -0.00082 |
| [193;208] | [145;160] | [113;128] | 0 | 0.00000 | 3 | 0.00005 | -0.00005 | 0.00000 | -0.00005 |
| [113;128] | [113;128] | [129;144] | 1 | 0.00005 | 0 | 0.00000 | 0.00005 | 0.00005 | 0.00000 |
| [129;144] | [113;128] | [129;144] | 0 | 0.00000 | 3 | 0.00005 | -0.00005 | 0.00000 | -0.00005 |
| [145;160] | [113;128] | [129;144] | 2 | 0.00010 | 1 | 0.00002 | 0.00008 | 0.00008 | 0.00000 |
| [161;176] | [113;128] | [129;144] | 1 | 0.00005 | 7 | 0.00011 | -0.00006 | 0.00000 | -0.00006 |
| [177;192] | [113;128] | [129;144] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [129;144] | [129;144] | [129;144] | 20 | 0.00099 | 38 | 0.00061 | 0.00038 | 0.00038 | 0.00000 |
| [145;160] | [129;144] | [129;144] | 433 | 0.02149 | 485 | 0.00776 | 0.01373 | 0.01373 | 0.00000 |
| [161;176] | [129;144] | [129;144] | 2572 | 0.12766 | 7234 | 0.11574 | 0.01191 | 0.01191 | 0.00000 |
| [177;192] | [129;144] | [129;144] | 1982 | 0.09837 | 11515 | 0.18424 | -0.08587 | 0.00000 | -0.08587 |
| [193;208] | [129;144] | [129;144] | 0 | 0.00000 | 56 | 0.00090 | -0.00090 | 0.00000 | -0.00090 |
| [145;160] | [145;160] | [129;144] | 21 | 0.00104 | 12 | 0.00019 | 0.00085 | 0.00085 | 0.00000 |
| [161;176] | [145;160] | [129;144] | 668 | 0.03315 | 925 | 0.01480 | 0.01835 | 0.01835 | 0.00000 |
| [177;192] | [145;160] | [129;144] | 1525 | 0.07569 | 3814 | 0.06102 | 0.01467 | 0.01467 | 0.00000 |
| [193;208] | [145;160] | [129;144] | 2 | 0.00010 | 90 | 0.00144 | -0.00134 | 0.00000 | -0.00134 |
| [193;208] | [161;176] | [129;144] | 1 | 0.00005 | 0 | 0.00000 | 0.00005 | 0.00005 | 0.00000 |
| [145;160] | [145;160] | [145;160] | 1 | 0.00005 | 0 | 0.00000 | 0.00005 | 0.00005 | 0.00000 |
| [161;176] | [145;160] | [145;160] | 16 | 0.00079 | 0 | 0.00000 | 0.00079 | 0.00079 | 0.00000 |
| [177;192] | [145;160] | [145;160] | 169 | 0.00839 | 9 | 0.00014 | 0.00824 | 0.00824 | 0.00000 |
| [193;208] | [145;160] | [145;160] | 13 | 0.00065 | 0 | 0.00000 | 0.00065 | 0.00065 | 0.00000 |
| [177;192] | [161;176] | [145;160] | 2 | 0.00010 | 0 | 0.00000 | 0.00010 | 0.00010 | 0.00000 |
| [193;208] | [161;176] | [145;160] | 5 | 0.00025 | 0 | 0.00000 | 0.00025 | 0.00025 | 0.00000 |
| Totais | | | 20148 | 1.00000 | 62500 | 1.00000 | 0.00000 | 0.22062 | -0.22062 |

Tabela 14. Comparação entre frequências absolutas e relativas para intervalos de cor relativamente à imagem original e imagem simulada (#2) (continuação)

| Intervalos dos canais de cor | | | Frequências da imagem original em valores | | Frequências da imagem simulada em valores | | Diferenças | | |
|------------------------------|-----------|----------|---|-----------|---|-----------|------------|-----------|-----------|
| R | G | B | absolutos | relativos | absolutos | relativos | totais | positivos | negativos |
| [81;96] | [81;96] | [65;80] | 6 | 0.00030 | 5 | 0.00008 | 0.00022 | 0.00022 | 0.00000 |
| [97;112] | [81;96] | [65;80] | 6 | 0.00030 | 85 | 0.00136 | -0.00106 | 0.00000 | -0.00106 |
| [113;128] | [81;96] | [65;80] | 0 | 0.00000 | 222 | 0.00355 | -0.00355 | 0.00000 | -0.00355 |
| [129;144] | [81;96] | [65;80] | 0 | 0.00000 | 56 | 0.00090 | -0.00090 | 0.00000 | -0.00090 |
| [145;160] | [81;96] | [65;80] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [97;112] | [97;112] | [65;80] | 5 | 0.00025 | 34 | 0.00054 | -0.00030 | 0.00000 | -0.00030 |
| [113;128] | [97;112] | [65;80] | 0 | 0.00000 | 326 | 0.00522 | -0.00522 | 0.00000 | -0.00522 |
| [129;144] | [97;112] | [65;80] | 0 | 0.00000 | 173 | 0.00277 | -0.00277 | 0.00000 | -0.00277 |
| [145;160] | [97;112] | [65;80] | 0 | 0.00000 | 19 | 0.00030 | -0.00030 | 0.00000 | -0.00030 |
| [161;176] | [97;112] | [65;80] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [113;128] | [113;128] | [65;80] | 0 | 0.00000 | 4 | 0.00006 | -0.00006 | 0.00000 | -0.00006 |
| [129;144] | [113;128] | [65;80] | 0 | 0.00000 | 4 | 0.00006 | -0.00006 | 0.00000 | -0.00006 |
| [161;176] | [113;128] | [65;80] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [97;112] | [81;96] | [81;96] | 2 | 0.00010 | 0 | 0.00000 | 0.00010 | 0.00010 | 0.00000 |
| [113;128] | [81;96] | [81;96] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [81;96] | [97;112] | [81;96] | 1 | 0.00005 | 0 | 0.00000 | 0.00005 | 0.00005 | 0.00000 |
| [97;112] | [97;112] | [81;96] | 53 | 0.00263 | 16 | 0.00026 | 0.00237 | 0.00237 | 0.00000 |
| [113;128] | [97;112] | [81;96] | 53 | 0.00263 | 380 | 0.00608 | -0.00345 | 0.00000 | -0.00345 |
| [129;144] | [97;112] | [81;96] | 2 | 0.00010 | 319 | 0.00510 | -0.00500 | 0.00000 | -0.00500 |
| [145;160] | [97;112] | [81;96] | 0 | 0.00000 | 30 | 0.00048 | -0.00048 | 0.00000 | -0.00048 |
| [161;176] | [97;112] | [81;96] | 0 | 0.00000 | 4 | 0.00006 | -0.00006 | 0.00000 | -0.00006 |
| [97;112] | [113;128] | [81;96] | 1 | 0.00005 | 2 | 0.00003 | 0.00002 | 0.00002 | 0.00000 |
| [113;128] | [113;128] | [81;96] | 32 | 0.00159 | 113 | 0.00181 | -0.00022 | 0.00000 | -0.00022 |
| [129;144] | [113;128] | [81;96] | 7 | 0.00035 | 183 | 0.00293 | -0.00258 | 0.00000 | -0.00258 |
| [145;160] | [113;128] | [81;96] | 0 | 0.00000 | 61 | 0.00098 | -0.00098 | 0.00000 | -0.00098 |
| [161;176] | [113;128] | [81;96] | 0 | 0.00000 | 11 | 0.00018 | -0.00018 | 0.00000 | -0.00018 |
| [97;112] | [97;112] | [97;112] | 8 | 0.00040 | 4 | 0.00006 | 0.00033 | 0.00033 | 0.00000 |
| [113;128] | [97;112] | [97;112] | 69 | 0.00342 | 135 | 0.00216 | 0.00126 | 0.00126 | 0.00000 |
| [129;144] | [97;112] | [97;112] | 5 | 0.00025 | 180 | 0.00288 | -0.00263 | 0.00000 | -0.00263 |
| [145;160] | [97;112] | [97;112] | 0 | 0.00000 | 23 | 0.00037 | -0.00037 | 0.00000 | -0.00037 |
| [161;176] | [97;112] | [97;112] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [97;112] | [113;128] | [97;112] | 14 | 0.00069 | 7 | 0.00011 | 0.00058 | 0.00058 | 0.00000 |
| [113;128] | [113;128] | [97;112] | 624 | 0.03097 | 641 | 0.01026 | 0.02071 | 0.02071 | 0.00000 |
| [129;144] | [113;128] | [97;112] | 789 | 0.03916 | 1843 | 0.02949 | 0.00967 | 0.00967 | 0.00000 |
| [145;160] | [113;128] | [97;112] | 260 | 0.01290 | 1404 | 0.02246 | -0.00956 | 0.00000 | -0.00956 |
| [161;176] | [113;128] | [97;112] | 120 | 0.00596 | 556 | 0.00890 | -0.00294 | 0.00000 | -0.00294 |
| [177;192] | [113;128] | [97;112] | 4 | 0.00020 | 19 | 0.00030 | -0.00011 | 0.00000 | -0.00011 |
| [113;128] | [129;144] | [97;112] | 4 | 0.00020 | 0 | 0.00000 | 0.00020 | 0.00020 | 0.00000 |
| [129;144] | [129;144] | [97;112] | 22 | 0.00109 | 17 | 0.00027 | 0.00082 | 0.00082 | 0.00000 |
| [145;160] | [129;144] | [97;112] | 37 | 0.00184 | 74 | 0.00118 | 0.00065 | 0.00065 | 0.00000 |

Tabela 15. Comparação entre frequências absolutas e relativas para intervalos de cor relativamente à imagem original e imagem simulada (#3)

| Intervalos dos canais de cor | | | Frequências da imagem original em valores | | Frequências da imagem simulada em valores | | Diferenças | | |
|------------------------------|-----------|-----------|---|----------------|---|----------------|----------------|----------------|-----------------|
| R | G | B | absolutos | relativos | absolutos | relativos | totais | positivos | negativos |
| [161;176] | [129;144] | [97;112] | 68 | 0.00338 | 98 | 0.00157 | 0.00181 | 0.00181 | 0.00000 |
| [177;192] | [129;144] | [97;112] | 12 | 0.00060 | 12 | 0.00019 | 0.00040 | 0.00040 | 0.00000 |
| [97;112] | [113;128] | [113;128] | 1 | 0.00005 | 1 | 0.00002 | 0.00003 | 0.00003 | 0.00000 |
| [113;128] | [113;128] | [113;128] | 195 | 0.00968 | 182 | 0.00291 | 0.00677 | 0.00677 | 0.00000 |
| [129;144] | [113;128] | [113;128] | 650 | 0.03226 | 1076 | 0.01722 | 0.01505 | 0.01505 | 0.00000 |
| [145;160] | [113;128] | [113;128] | 606 | 0.03008 | 1872 | 0.02995 | 0.00013 | 0.00013 | 0.00000 |
| [161;176] | [113;128] | [113;128] | 304 | 0.01509 | 1375 | 0.02200 | -0.00691 | 0.00000 | -0.00691 |
| [177;192] | [113;128] | [113;128] | 9 | 0.00045 | 124 | 0.00198 | -0.00154 | 0.00000 | -0.00154 |
| [113;128] | [129;144] | [113;128] | 29 | 0.00144 | 23 | 0.00037 | 0.00107 | 0.00107 | 0.00000 |
| [129;144] | [129;144] | [113;128] | 429 | 0.02129 | 530 | 0.00848 | 0.01281 | 0.01281 | 0.00000 |
| [145;160] | [129;144] | [113;128] | 1809 | 0.08979 | 3933 | 0.06293 | 0.02686 | 0.02686 | 0.00000 |
| [161;176] | [129;144] | [113;128] | 4966 | 0.24648 | 15268 | 0.24429 | 0.00219 | 0.00219 | 0.00000 |
| [177;192] | [129;144] | [113;128] | 1447 | 0.07182 | 7168 | 0.11469 | -0.04287 | 0.00000 | -0.04287 |
| [193;208] | [129;144] | [113;128] | 0 | 0.00000 | 2 | 0.00003 | -0.00003 | 0.00000 | -0.00003 |
| [145;160] | [145;160] | [113;128] | 1 | 0.00005 | 1 | 0.00002 | 0.00003 | 0.00003 | 0.00000 |
| [161;176] | [145;160] | [113;128] | 28 | 0.00139 | 59 | 0.00094 | 0.00045 | 0.00045 | 0.00000 |
| [177;192] | [145;160] | [113;128] | 36 | 0.00179 | 100 | 0.00160 | 0.00019 | 0.00019 | 0.00000 |
| [113;128] | [113;128] | [129;144] | 1 | 0.00005 | 1 | 0.00002 | 0.00003 | 0.00003 | 0.00000 |
| [145;160] | [113;128] | [129;144] | 2 | 0.00010 | 1 | 0.00002 | 0.00008 | 0.00008 | 0.00000 |
| [161;176] | [113;128] | [129;144] | 1 | 0.00005 | 4 | 0.00006 | -0.00001 | 0.00000 | -0.00001 |
| [177;192] | [113;128] | [129;144] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [113;128] | [129;144] | [129;144] | 0 | 0.00000 | 2 | 0.00003 | -0.00003 | 0.00000 | -0.00003 |
| [129;144] | [129;144] | [129;144] | 20 | 0.00099 | 42 | 0.00067 | 0.00032 | 0.00032 | 0.00000 |
| [145;160] | [129;144] | [129;144] | 433 | 0.02149 | 723 | 0.01157 | 0.00992 | 0.00992 | 0.00000 |
| [161;176] | [129;144] | [129;144] | 2572 | 0.12766 | 9488 | 0.15181 | -0.02415 | 0.00000 | -0.02415 |
| [177;192] | [129;144] | [129;144] | 1982 | 0.09837 | 9053 | 0.14485 | -0.04648 | 0.00000 | -0.04648 |
| [193;208] | [129;144] | [129;144] | 0 | 0.00000 | 4 | 0.00006 | -0.00006 | 0.00000 | -0.00006 |
| [145;160] | [145;160] | [129;144] | 21 | 0.00104 | 17 | 0.00027 | 0.00077 | 0.00077 | 0.00000 |
| [161;176] | [145;160] | [129;144] | 668 | 0.03315 | 1426 | 0.02282 | 0.01034 | 0.01034 | 0.00000 |
| [177;192] | [145;160] | [129;144] | 1525 | 0.07569 | 2922 | 0.04675 | 0.02894 | 0.02894 | 0.00000 |
| [193;208] | [145;160] | [129;144] | 2 | 0.00010 | 19 | 0.00030 | -0.00020 | 0.00000 | -0.00020 |
| [193;208] | [161;176] | [129;144] | 1 | 0.00005 | 0 | 0.00000 | 0.00005 | 0.00005 | 0.00000 |
| [113;128] | [97;112] | [145;160] | 0 | 0.00000 | 1 | 0.00002 | -0.00002 | 0.00000 | -0.00002 |
| [145;160] | [145;160] | [145;160] | 1 | 0.00005 | 0 | 0.00000 | 0.00005 | 0.00005 | 0.00000 |
| [161;176] | [145;160] | [145;160] | 16 | 0.00079 | 5 | 0.00008 | 0.00071 | 0.00071 | 0.00000 |
| [177;192] | [145;160] | [145;160] | 169 | 0.00839 | 12 | 0.00019 | 0.00820 | 0.00820 | 0.00000 |
| [193;208] | [145;160] | [145;160] | 13 | 0.00065 | 0 | 0.00000 | 0.00065 | 0.00065 | 0.00000 |
| [177;192] | [161;176] | [145;160] | 2 | 0.00010 | 0 | 0.00000 | 0.00010 | 0.00010 | 0.00000 |
| [193;208] | [161;176] | [145;160] | 5 | 0.00025 | 0 | 0.00000 | 0.00025 | 0.00025 | 0.00000 |
| [161;176] | [129;144] | [97;112] | 68 | 0.00338 | 98 | 0.00157 | 0.00181 | 0.00181 | 0.00000 |
| [177;192] | [129;144] | [97;112] | 12 | 0.00060 | 12 | 0.00019 | 0.00040 | 0.00040 | 0.00000 |
| [97;112] | [113;128] | [113;128] | 1 | 0.00005 | 1 | 0.00002 | 0.00003 | 0.00003 | 0.00000 |
| Totais | | | 20148 | 1.00000 | 62500 | 1.00000 | 0.00000 | 0.16518 | -0.16518 |

Tabela 16. Comparação entre frequências absolutas e relativas para intervalos de cor relativamente à imagem original e imagem simulada (#3) (continuação)

Na tabela 17, tal como foi efectuado para o caso anterior, apresenta-se um quadro resumo da percentagem de *pixels* que alteraram o seu intervalo de cor bem como a percentagem de *pixels* existentes em intervalos de cor que desaparecem nas imagens simuladas podendo constatar-se as percentagens de alterações de intervalo de cor bem como de *pixels* que desaparecem não apresentam valores significativos.

| Imagem simulada | #1 | #2 | #3 |
|---|-----------|-----------|-----------|
| % de <i>pixels</i> em novos intervalos de cor | 1,0464 | 1,6144 | 1,5168 |
| % de <i>pixels</i> em intervalos de cor que desaparecem | 0,0645 | 0,2084 | 0,1439 |

Tabela 17. Frequência de *pixels* (em %) convertidos em novos intervalos de cor e intervalos de cor que desaparecem nas imagens simuladas

Finalmente, na tabela 18 apresenta-se um quadro resumo dos modelos teóricos e respectivos parâmetros obtidos para a imagem original e para as três imagens simuladas.

| Canais | Imagens | Modelo | C₀ | C₁ | a |
|---------------|------------------------|-----------------|----------------------|----------------------|----------------|
| Red | Imagem original | Esférico | 96,4026 | 120,108 | 112,42 |
| | Imagem simulada 1 | Esférico | 118,672 | 91,5940 | 67,0233 |
| | Imagem simulada 2 | Esférico | 113,487 | 79,4755 | 128,0157 |
| | Imagem simulada 3 | Esférico | 108,0855 | 93,6961 | 114,7593 |
| Green | Imagem original | Esférico | 33,6754 | 33,6373 | 109,761 |
| | Imagem simulada 1 | Esférico | 37,2855 | 28,5324 | 66,5544 |
| | Imagem simulada 2 | Esférico | 43,0184 | 25,6381 | 127,934 |
| | Imagem simulada 3 | Esférico | 39,5503 | 36,0120 | 119,905 |
| Blue | Imagem original | Esférico | 41,3911 | 35,7869 | 107,941 |
| | Imagem simulada 1 | Esférico | 52,2240 | 36,9592 | 44,4895 |
| | Imagem simulada 2 | Esférico | 57,8088 | 31,0677 | 114,634 |
| | Imagem simulada 3 | Esférico | 55,2440 | 46,5218 | 116,227 |

Tabela 18. Comparação dos modelos teóricos da imagem original e das imagens simuladas

5.3.3. Análise dos resultados

Neste exemplo pretende-se mostrar a utilização das técnicas de simulação condicional na reprodução de padrões de cor em chapas de mármore.

Para isso calcula-se, para cada canal, por ajustamento dos variogramas direccionais, o modelo teórico no qual se baseia a simulação condicional de imagens de chapas de mármore.

Como a correlação entre os três canais de cor é sempre bastante elevada utilizou-se, nos dois exemplos, um modelo de coregionalização para a construção das imagens. Assim, para cada exemplo foram simuladas três imagens para o canal por Simulação Sequencial Gaussiana e três imagens para cada um dos outros canais por Co-Simulação Sequencial Gaussiana.

Finalmente, a análise das tabelas comparativas:

- das frequências de cor obtidas para intervalos de cor na imagem original e em cada uma das simuladas;
- da frequência de *pixels* nas imagens simuladas que desaparecem ou que existem em novos intervalos de cor;
- dos modelos obtidos nas imagens simuladas;

permitem constatar-se que as frequências de diferenças de cor bem como a frequência de *pixels* que desaparecem ou aparecem em novos intervalos de cor nas imagens simuladas são muito pouco significativas. No que se refere à concordância entre o modelo dos variogramas dos dados iniciais e aqueles que se calculam, a posteriori, sobre os dados simulados, verifica-se, em termos gerais, que os variogramas dos valores simulados aderem razoavelmente bem ao modelo de partida.

6. Conclusões

Na ideia subjacente ao sistema desenvolvido, este deveria fundamentalmente possibilitar:

- o processamento de variografia e estimação/simulação sobre imagens matriciais implementando operacionalidades relativas às suas características, como por exemplo, possibilitar a entrada de uma imagem matricial e, após o processamento, reconstruir as saídas se necessário sob forma de imagem matricial;
- a modelação automática 2D utilizando algoritmos genéticos e contemplando dois testes para validação de resultados: um teste englobado no algoritmo genético implementado (teste usual de diferença de quadrados entre os valores obtidos nos variogramas experimentais e os valores que se obtêm nos modelos propostos) e um teste de validação cruzada realizado à saída da modelação automático.

Assim, no que se refere ao processamento directo de variografia e estimação/simulação não foram encontradas quaisquer referências na literatura e, no que se refere à modelação automática utilizando algoritmos genéticos, apenas Bonduà ([BBM 2002]) refere um método de ajuste automático de variogramas experimentais baseado em algoritmos genéticos, sem contudo descrever o método que ainda se encontra em fase de implementação.

Para o desenvolvimento deste sistema foram muito importantes o estudo de casos, uma vez que foi através da própria problemática inerente à experimentação com imagem, que se foram realizando novas abordagens para encontrar uma metodologia eficiente que possibilitasse a modelação automática.

Quanto aos casos propostos - ampliação por mudança de escala de uma imagem, caracterização global e local dos valores da precipitação para toda a região do Algarve e a simulação de padrões de tonalidade e heterogeneidade da cor em chapas de rochas ornamentais - deve referir-se que são típicos de domínios tão diversos quanto o domínio do *design* gráfico, do Ambiente e da Arquitectura e Construção Civil.

Finalmente, é importante realçar que estes três exemplos, seleccionados para ilustrar as metodologias consideradas nesta dissertação, têm como denominador comum a intervenção de imagens sob diferentes perspectivas, com intervenção *principal* ou *secundária*.

Numa primeira abordagem pensou-se construir o modelo global à custa dos modelos direccionais obtidos, ou seja, o tipo do modelo global seria o tipo do modelo que ocorresse com maior frequência nos variogramas direccionais, sendo a amplitude dos modelos direccionais utilizada para calcular as anisotropias e o efeito de pepita e patamar gerados através de um algoritmo genético para o tipo de modelo considerado.

Esta abordagem que, pareceria *a priori* a mais intuitiva (mais perto da que é usualmente seguida pelos especialistas) revelou-se ineficiente na adequação dos tipos de modelos direccionais ao tipo de modelo omnidireccional pois:

- a sujeição do tipo de modelo global ao tipo de modelo direccional mais frequente pode conduzir a situações de erro encadeado porque dois tipos de modelos poderiam, com determinados valores nos seus parâmetros e sem a informação contextual relativa às hipóteses equiprováveis para modelos nas outras direcções, parecer igualmente prováveis;
- os resultados da validação cruzada não seleccionaram o modelo de potência em casos onde, segundo opinião pericial, este seria o mais adequado.

Nesta abordagem foram tentados vários ajustes no erro calculado no algoritmo genético citando-se, de seguida, os que se consideraram mais importantes:

- a não ponderação no algoritmo genético, numa primeira fase, dos pontos dos variogramas experimentais existentes acima da variância que, no caso dos modelos de potência é relevante para o cálculo do erro de ajuste. No caso dos outros modelos, apenas foram considerados, os pontos cujo valor fosse menor ou igual ao da variância);
- a escolha do melhor modelo através do teste de validação cruzada efectuado a dois tipos de modelos obtidos automaticamente em duas etapas: a primeira, referente à escolha de um modelo de tipo não potência onde os valores do efeito de pepita e do

patamar são condicionados pela variância; a segunda, específica para o modelo de potência, assegurando-se o menor erro no ajuste automático dos parâmetros.

Esta metodologia foi abandonada na prossecução dos trabalhos mas, por reflectir o comportamento usual dos especialistas, considerou-se pertinente a apresentação de alguns dos seus resultados em anexo.

Finalmente considerou-se o modelo global como um modelo com visão global dos direccionais, isto é, primeiro calculou-se um modelo global com todos os modelos direccionais em presença usando, numa segunda etapa, a geração de modelos direccionais apenas para a detecção da anisotropia (os modelos direccionais serão do tipo do modelo global variando apenas a amplitude).

Com esta nova abordagem, reforçou-se a convicção que os parâmetros do hipotético modelo do tipo potência deveriam ser gerados automaticamente com um algoritmo genético condicionado a esse modelo uma vez que não se poderia admitir que todos os modelos estavam todos sujeitos às mesmas restrições.

Por conseguinte, utilizaram-se dois algoritmos genéticos em paralelo para o cálculo dos dois tipos de modelos em presença (potência e não potência) e usou-se o menor erro da validação cruzada para seleccionar o que melhor ajusta os variogramas experimentais.

Com o mecanismo de modelação construído de acordo com esta metodologia obtiveram-se, como foi descrito no capítulo 3, resultados bastante semelhantes aos que se obtêm na modelação manual que permitem concluir que se deveria investir no aperfeiçoamento deste tipo de modelação pois a sua automação possibilitaria a aplicação da geoestatística a estudos em domínios onde os especialistas pretendam resultados que lhes permitam avançar na sua área de conhecimentos, sem recorrer a quaisquer artificios cuja validade só pode ser garantida através de uma longa prática de utilização das ferramentas geoestatísticas.

Quanto à eficiência em termos computacionais (tempos de processador) do tratamento geoestatístico de imagens, acredita-se que, o cálculo em paralelo, de variogramas para diferentes canais, implicando apenas uma maior utilização de espaço de memória (hoje tão facilmente disponibilizado) contribui para uma maior eficiência na produção de resultados,

pois as imagens contêm conjuntos extraordinariamente grandes de dados e o cálculo dos variogramas é o módulo mais pesado do sistema.

Ressalve-se ainda, no que respeita à eficiência computacional, que se optou por realizar testes de validação cruzada com conjuntos aleatórios de 30% dos *pixels* retirados das imagens originais, com o objectivo de diminuir o tempo para a realização dos referidos testes.

Quanto à melhoria futura do sistema desenvolvido deve referir-se a necessidade de:

- estudar efeitos de ponderação no conjunto dos pontos considerados para o ajuste de modelos do tipo não potência no sentido de otimizar o cálculo do erro possibilitando, assim, a obtenção de valores mais adequados aos patamares
- incorporar no *package* módulos de estimação e simulação não dependentes da dimensão dos dados (condicionante já referida para os módulos do GSLIB);
- otimizar o recurso a janelas permitindo a realização modular de tarefas em alturas diferentes de utilização do software, ou seja, permitindo a gravação dos resultados obtidos em cada tarefa (cálculo de variogramas, modelação, etc) e a sua posterior utilização para uma nova tarefa;
- testar, através da sua divulgação como *software* de uso livre, a facilidade de manuseamento do *package* e as situações de tratamento de imagem que poderiam ser complementarmente contempladas.

Referências

- [AS 04] ASHLOCK, D., *Evolutionary Computation for Modeling and Optimization*, Lectures Notes in <http://www.math.iastate.edu/danwell/ma378/ln378.html>, 2004
- [AJ 94] ALMEIDA, A. S.; JOURNEL, A. G., Joint simulation of multiple variables with a Markov-type coregionalization model, *Mathematical Geology*, 26(5), 1994
- [BBM 02] BONDUÀ, S.; BRUNO, R.; MUGE, F., Geostatistical Simulation of Ornamental Stone Images: Result Analysis by Mathematical Morphology, *IAMG*, 2002
- [DAV 04] DAVID, M., *Geostatistical Ore Reserve Estimation*, Elsevier, 1977
- [DJ 98] DEUTSCH, C. S.; JOURNEL, A. G., *GSLIB Geostatistical Software Library and User's Guide*, Second Edition, Oxford University Press, 1998
- [FO 90] FOLEY, J. D.; DAM, A. Van; FEINER, S. K.; HUGHES, J. F., *Computer Graphics Principles and Practice*, Addison-Wesley Publishing Company, 1990
- [FOW 66] FOGEL, L.; OWENS, A. J.; WALSH, M. J., *Artificial Intelligence through Simulated Evolution*, John Wiley and Sons, 1966
- [GO 97] GOOVAERTS, P., *Natural Resources Evaluation*, Oxford University Press, 1997
- [GO 99] GOOVAERTS, P., "Performance comparison of geostatistical algorithms for incorporating elevation into the mapping of precipitation" in *Proceedings of the International Conference on Geocomputaion*, 1999

- [GHJ 93] GOMEZ-HERNANDEZ, J. J.; JOURNAL, A.G, Joint sequential simulation of multiGaussian fields, *Geostatistics Troia '92*, A. Soares (ed.), Kluwer, 1993
- [HJ 75] HOLLAND, J. H., *Adaptative in Natural and Artificial Systems*, University of Michigan Press, 1975
- [HG 03] HOFFMAN, G., *Interpolations for Image Warping*, University of Applied Sciences in Emden, <http://www.fho-emden.de/~hoffmann/index.html>, 2003
- [IS 89] ISAAKS, Edward H.; SRIVASTAVA, R. M., *Applied Geostatistics*, 2th. Oxford University Press, 1989
- [JM 00] JOHNSON, M., *Graphic File Formats*, Columbia University – School of Arts, 2000
- [LA 03] LUÍS, A. A. G, *Utilização de Informação Auxiliar na Estimacão e Simulacão de Variáveis Regionalizadas*, PHD Thesis, IST-UTL, 2003
- [JH 78] JOURNAL, A. G.; HUIJBREGTS, Ch. J., *Mining Geostatistics*, Academic Press, 1978
- [LE 04] LEAL, M. M., *Análise de Dados Multivariados*, FC-UL
- [LO 01] LOUREIRO, N. S.; COUTINHO, M. A. (2001) - A new procedure to estimate the RUSLE EI30 index based on monthly rainfall data and applied to the Algarve reguion, Portugal, *Journal of Hydrology*, 2001
- [LF 00] LOBO, F. M. P. G., *The Parameter-Less Genetic Algorithm: Rational and Automated Parameter Selection for Simplified Genetic Algorithm Operation*, PHD Thesis, FCT-UNL, 2000
- [MA 65] MATHERON, G., *Les variable Regionalisées et leur estimation*, Ed. Masson, 1965
- [MA 78] MATHERON, G., *Estimating and Choosing: an Essay of Probability in Practice*, Springer-Verlag, 1978

- [MSB 91] MÜHLENBEIN, H.; SCHOMISH, M.; BORN, J., The Parallel Genetic Algorithm as Function Optimizer, *Parallel Computer*, 17, 1991
- [NR 01] NOGUEIRA, M. M.; ROCHA, C. S., *Edição Electrónica Panorâmica das Artes Gráficas III*, Plátano Edições Técnicas, 2001
- [PA 96] PANNATIER, Yvan, *Varwin: Software for Spatial Analysis*, Springer-Verlag, 1996
- [PE 79] PEREIRA, H. G., *Introdução às Variáveis Regionalizadas*, Revista Técnica, IST, Jan/Fev/Mar, 1979
- [PS 02] PEREIRA, H.G.; SOUSA, A. J., *Análise de Dados para o Tratamento de Quadros Multidimensionais*, CVRM-IST, 1998-2002
- [PTVF 02] PRESS, W. H.; TEUKOLSKY, S. A.; VETTERLING, W.T.; FLANNERY, B.P., *Numerical Recipes in C*, Cambridge University Press, 2002
- [SO 00] SOARES, A., *Geoestatística para as Ciências da Terra e do Ambiente*, IST Press, 2000
- [SM 87] SOUSA, A. J.; MUGE, F., *Elementos de Geoestatística*, Laboratório de Mineralogia e Planeamento Mineiro - IST, 1987
- [SOU 00] SOUSA, A. J., *Análise Geoestatística de Dados*, CVRM/Centro de GeoSistemas, IST, 2000
- [SJBFG 93] SPEARS, W. M.; JONG, K. A. De; BAECK, T.; FOGEL, D. B.; GARIS, H., "An Overview of Evolutionary Computation" in *Proceedings of the European Conference on Machine Learning*, 1993
- [SUN 99] Sun Microsystems, *Programming in Java Advanced Imaging*, Release 1.0.1, Sun Microsystems, Inc., 1999
- [TA 00] TAVARES, J. A. R., *Geração de Configurações de Sistemas Industriais com o Recurso à Tecnologia das Restrições e Computação Evolucionária*, PhD Thesis, Universidade do Minho, 2000

[YGV 98] YOUNG, Ian T.; GERBRANDS, Jan J.; VLIET, L. J., *Fundamentals of Image Processing Version 2.2*, Delf University of Technology, Netherlands, 1998

[HP AIGEOS] <http://www.ai-geostats.org/>, 30/8/2004

[HP CXIMG] http://www.xdp.it/cximage_reference.htm, 15/8/2004

[HP GEOMS] http://alfa.ist.utl.pt/~ncmrp/p_geoms.htm, 30/8/2004

[HP KUSIAC] <http://www.icaen.uiowa.edu/%7Eie238/lecture.html>, 30/8/2004

[HP MICRO]

<http://www.microscopy.fsu.edu/primer/java/digitalimaging/processing/panscrollzoom/>, 1/07/2005

[HP NOO] <http://www.noobeed.com/>, 31/07/2005

[HP OCX] http://www.codeproject.com/miscctrl/ntgraph_activex.asp, 15/8/2004

[HP PAG] <http://www.portaldasartesgraficas.com/resolucao.htm>, 20/4/2004 (2.2)

[HP PNG] <http://www.libpng.org/pub/png/png.html>, 20/4/2004

[HP RAND] <http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/emt.html>, 15/8/2004

[HP SURF] <http://www.goldensoftware.com/products/surfer/surfer.shtml>, 30/8/2004

[HP GSOF] <http://www.ibrae.ac.ru/~mkanev/eng/gsoffice>, 30/08/2004

<http://www.ibrae.ac.ru/~mkanev/eng/Varrose.html>, 30/08/2004

[HP WGSLIB] <http://www.staios.com/WinGslib/>, 30/8/2004

Apêndice I

Tabela comparativa de métodos de estimação

Neste apêndice considerou-se relevante apresentar, pela sua importância, os resultados que foram obtidos na interpolação de uma linha de uma imagem obtida através do estimador de vizinhança mais próxima, bilinear, bicúbica e por estimação geoestatística.

Assim, na tabela seguinte, pode observar-se para cada posição (x,y) o valor do *pixel* da imagem original (**Original**), os valores do *pixel* obtidos pelos 3 métodos de interpolação citados bem como o valor obtido por estimação geoestatística (**krigagem**) e os erros quadrados cometidos para cada tipo de interpolação.

| X | Y | Original | Nearest | Bilinear | Bicubic | Krigagem | NearErr | BilinErr | BicErr | KrigErr |
|----|----|----------|---------|----------|---------|----------|---------|----------|--------|---------|
| 0 | 75 | 62 | 0 | 0 | 0 | 99 | 3844 | 3844 | 3844 | 1352 |
| 1 | 75 | 132 | 153 | 159 | 153 | 111 | 441 | 729 | 441 | 460 |
| 2 | 75 | 131 | 161 | 168 | 161 | 125 | 900 | 1369 | 900 | 39 |
| 3 | 75 | 136 | 173 | 170 | 175 | 143 | 1369 | 1156 | 1521 | 51 |
| 4 | 75 | 142 | 163 | 158 | 154 | 151 | 441 | 256 | 144 | 73 |
| 5 | 75 | 129 | 156 | 158 | 157 | 151 | 729 | 841 | 784 | 467 |
| 6 | 75 | 137 | 166 | 158 | 155 | 150 | 841 | 441 | 324 | 162 |
| 7 | 75 | 141 | 155 | 162 | 163 | 147 | 196 | 441 | 484 | 34 |
| 8 | 75 | 149 | 168 | 170 | 174 | 145 | 361 | 441 | 625 | 14 |
| 9 | 75 | 148 | 170 | 180 | 184 | 145 | 484 | 1024 | 1296 | 9 |
| 10 | 75 | 154 | 224 | 178 | 183 | 147 | 4900 | 576 | 841 | 42 |
| 11 | 75 | 158 | 165 | 165 | 162 | 149 | 49 | 49 | 16 | 78 |
| 12 | 75 | 121 | 161 | 164 | 163 | 150 | 1600 | 1849 | 1764 | 854 |
| 13 | 75 | 152 | 163 | 166 | 166 | 147 | 121 | 196 | 196 | 25 |
| 14 | 75 | 149 | 175 | 168 | 165 | 142 | 676 | 361 | 256 | 44 |

| | | | | | | | | | | |
|----|----|-----|-----|-----|-----|-----|-------|-------|-------|-----|
| 15 | 75 | 136 | 172 | 169 | 170 | 138 | 1296 | 1089 | 1156 | 3 |
| 16 | 75 | 138 | 191 | 156 | 160 | 134 | 2809 | 324 | 484 | 14 |
| 17 | 75 | 137 | 136 | 137 | 129 | 131 | 1 | 0 | 64 | 36 |
| 18 | 75 | 128 | 142 | 143 | 138 | 128 | 196 | 225 | 100 | 0 |
| 19 | 75 | 128 | 149 | 158 | 161 | 125 | 441 | 900 | 1089 | 8 |
| 20 | 75 | 124 | 165 | 167 | 169 | 119 | 1681 | 1849 | 2025 | 22 |
| 21 | 75 | 118 | 164 | 167 | 168 | 113 | 2116 | 2401 | 2500 | 22 |
| 22 | 75 | 105 | 168 | 163 | 162 | 106 | 3969 | 3364 | 3249 | 1 |
| 23 | 75 | 99 | 165 | 160 | 160 | 94 | 4356 | 3721 | 3721 | 25 |
| 24 | 75 | 87 | 167 | 158 | 160 | 80 | 6400 | 5041 | 5329 | 51 |
| 25 | 75 | 60 | 162 | 154 | 156 | 67 | 10404 | 8836 | 9216 | 50 |
| 26 | 75 | 60 | 152 | 148 | 145 | 62 | 8464 | 7744 | 7225 | 2 |
| 27 | 75 | 48 | 142 | 151 | 149 | 67 | 8836 | 10609 | 10201 | 375 |
| 28 | 75 | 110 | 160 | 162 | 162 | 81 | 2500 | 2704 | 2704 | 852 |
| 29 | 75 | 91 | 159 | 168 | 168 | 91 | 4624 | 5929 | 5929 | 0 |
| 30 | 75 | 96 | 159 | 169 | 171 | 92 | 3969 | 5329 | 5625 | 20 |
| 31 | 75 | 78 | 158 | 168 | 168 | 82 | 6400 | 8100 | 8100 | 14 |
| 32 | 75 | 97 | 192 | 157 | 145 | 71 | 9025 | 3600 | 2304 | 653 |
| 33 | 75 | 63 | 156 | 156 | 148 | 63 | 8649 | 8649 | 7225 | 0 |
| 34 | 75 | 51 | 187 | 161 | 158 | 60 | 18496 | 12100 | 11449 | 90 |
| 35 | 75 | 53 | 192 | 152 | 142 | 61 | 19321 | 9801 | 7921 | 68 |
| 36 | 75 | 67 | 173 | 147 | 135 | 63 | 11236 | 6400 | 4624 | 16 |
| 37 | 75 | 83 | 162 | 151 | 139 | 62 | 6241 | 4624 | 3136 | 434 |
| 38 | 75 | 56 | 179 | 158 | 153 | 60 | 15129 | 10404 | 9409 | 16 |
| 39 | 75 | 44 | 182 | 156 | 146 | 56 | 19044 | 12544 | 10404 | 150 |
| 40 | 75 | 62 | 163 | 157 | 147 | 52 | 10201 | 9025 | 7225 | 92 |
| 41 | 75 | 58 | 179 | 161 | 155 | 48 | 14641 | 10609 | 9409 | 102 |
| 42 | 75 | 45 | 175 | 156 | 145 | 43 | 16900 | 12321 | 10000 | 5 |

| | | | | | | | | | | |
|----|----|-----|-----|-----|-----|-----|-------|-------|-------|------|
| 43 | 75 | 31 | 168 | 153 | 139 | 42 | 18769 | 14884 | 11664 | 123 |
| 44 | 75 | 30 | 161 | 158 | 149 | 56 | 17161 | 16384 | 14161 | 702 |
| 45 | 75 | 42 | 191 | 162 | 157 | 82 | 22201 | 14400 | 13225 | 1624 |
| 46 | 75 | 172 | 168 | 157 | 147 | 113 | 16 | 225 | 625 | 3425 |
| 47 | 75 | 222 | 160 | 151 | 139 | 129 | 3844 | 5041 | 6889 | 8625 |
| 48 | 75 | 231 | 159 | 148 | 135 | 133 | 5184 | 6889 | 9216 | 9509 |
| 49 | 75 | 76 | 163 | 152 | 142 | 130 | 7569 | 5776 | 4356 | 2873 |
| 50 | 75 | 91 | 159 | 158 | 149 | 139 | 4624 | 4489 | 3364 | 2318 |
| 51 | 75 | 91 | 163 | 158 | 147 | 159 | 5184 | 4489 | 3136 | 4618 |
| 52 | 75 | 239 | 167 | 156 | 144 | 189 | 5184 | 6889 | 9025 | 2478 |
| 53 | 75 | 246 | 161 | 155 | 144 | 209 | 7225 | 8281 | 10404 | 1367 |
| 54 | 75 | 250 | 160 | 155 | 142 | 212 | 8100 | 9025 | 11664 | 1467 |
| 55 | 75 | 187 | 154 | 159 | 149 | 188 | 1089 | 784 | 1444 | 1 |
| 56 | 75 | 165 | 152 | 176 | 179 | 154 | 169 | 121 | 196 | 124 |
| 57 | 75 | 105 | 184 | 184 | 188 | 129 | 6241 | 6241 | 6889 | 569 |
| 58 | 75 | 88 | 176 | 179 | 177 | 120 | 7744 | 8281 | 7921 | 1053 |
| 59 | 75 | 147 | 174 | 182 | 184 | 124 | 729 | 1225 | 1369 | 512 |
| 60 | 75 | 148 | 184 | 182 | 183 | 139 | 1296 | 1156 | 1225 | 80 |
| 61 | 75 | 150 | 180 | 179 | 177 | 149 | 900 | 841 | 729 | 1 |
| 62 | 75 | 156 | 180 | 180 | 179 | 149 | 576 | 576 | 529 | 44 |
| 63 | 75 | 153 | 180 | 182 | 182 | 148 | 729 | 841 | 841 | 25 |
| 64 | 75 | 153 | 182 | 184 | 185 | 149 | 841 | 961 | 1024 | 13 |
| 65 | 75 | 131 | 183 | 183 | 183 | 160 | 2704 | 2704 | 2704 | 825 |
| 66 | 75 | 177 | 181 | 182 | 181 | 169 | 16 | 25 | 16 | 69 |
| 67 | 75 | 201 | 177 | 183 | 186 | 176 | 576 | 324 | 225 | 612 |
| 68 | 75 | 169 | 188 | 185 | 192 | 182 | 361 | 256 | 529 | 174 |
| 69 | 75 | 178 | 194 | 167 | 162 | 179 | 256 | 121 | 256 | 2 |
| 70 | 75 | 181 | 132 | 152 | 138 | 180 | 2401 | 841 | 1849 | 2 |

| | | | | | | | | | | |
|----|----|-----|-----|-----|-----|-----|------|------|------|------|
| 71 | 75 | 182 | 132 | 151 | 139 | 176 | 2500 | 961 | 1849 | 37 |
| 72 | 75 | 188 | 128 | 150 | 139 | 174 | 3600 | 1444 | 2401 | 190 |
| 73 | 75 | 148 | 130 | 151 | 142 | 172 | 324 | 9 | 36 | 572 |
| 74 | 75 | 202 | 131 | 150 | 142 | 168 | 5041 | 2704 | 3600 | 1149 |
| 75 | 75 | 151 | 134 | 150 | 141 | 167 | 289 | 1 | 100 | 270 |
| 76 | 75 | 184 | 135 | 148 | 138 | 160 | 2401 | 1296 | 2116 | 565 |
| 77 | 75 | 140 | 136 | 146 | 136 | 157 | 16 | 36 | 16 | 288 |
| 78 | 75 | 146 | 135 | 147 | 138 | 160 | 121 | 1 | 64 | 195 |
| 79 | 75 | 179 | 135 | 147 | 140 | 165 | 1936 | 1024 | 1521 | 193 |
| 80 | 75 | 178 | 136 | 146 | 140 | 175 | 1764 | 1024 | 1444 | 9 |
| 81 | 75 | 180 | 135 | 144 | 139 | 180 | 2025 | 1296 | 1681 | 0 |
| 82 | 75 | 182 | 134 | 139 | 132 | 182 | 2304 | 1849 | 2500 | 0 |
| 83 | 75 | 189 | 128 | 139 | 131 | 184 | 3721 | 2500 | 3364 | 26 |
| 84 | 75 | 197 | 129 | 142 | 139 | 186 | 4624 | 3025 | 3364 | 128 |
| 85 | 75 | 184 | 136 | 142 | 136 | 189 | 2304 | 1764 | 2304 | 23 |
| 86 | 75 | 185 | 131 | 140 | 128 | 192 | 2916 | 2025 | 3249 | 45 |
| 87 | 75 | 200 | 136 | 137 | 126 | 188 | 4096 | 3969 | 5476 | 150 |
| 88 | 75 | 199 | 132 | 134 | 126 | 180 | 4489 | 4225 | 5329 | 379 |
| 89 | 75 | 145 | 131 | 131 | 123 | 160 | 196 | 196 | 484 | 217 |
| 90 | 75 | 171 | 134 | 127 | 119 | 144 | 1369 | 1936 | 2704 | 738 |
| 91 | 75 | 72 | 134 | 128 | 124 | 147 | 3844 | 3136 | 2704 | 5698 |
| 92 | 75 | 195 | 130 | 127 | 120 | 151 | 4225 | 4624 | 5625 | 1946 |
| 93 | 75 | 174 | 130 | 127 | 120 | 162 | 1936 | 2209 | 2916 | 152 |
| 94 | 75 | 149 | 134 | 128 | 123 | 168 | 225 | 441 | 676 | 362 |
| 95 | 75 | 159 | 128 | 126 | 121 | 161 | 961 | 1089 | 1444 | 5 |
| 96 | 75 | 160 | 128 | 127 | 123 | 161 | 1024 | 1089 | 1369 | 0 |
| 97 | 75 | 167 | 128 | 125 | 120 | 163 | 1521 | 1764 | 2209 | 19 |
| 98 | 75 | 176 | 129 | 122 | 117 | 161 | 2209 | 2916 | 3481 | 238 |

| | | | | | | | | | | |
|-----|----|-----|-----|-----|-----|-----|-------|-------|-------|------|
| 99 | 75 | 164 | 129 | 122 | 119 | 162 | 1225 | 1764 | 2025 | 5 |
| 100 | 75 | 135 | 127 | 122 | 118 | 166 | 64 | 169 | 289 | 958 |
| 101 | 75 | 198 | 132 | 122 | 116 | 166 | 4356 | 5776 | 6724 | 1020 |
| 102 | 75 | 174 | 127 | 131 | 131 | 171 | 2209 | 1849 | 1849 | 7 |
| 103 | 75 | 161 | 128 | 133 | 133 | 174 | 1089 | 784 | 784 | 160 |
| 104 | 75 | 174 | 122 | 129 | 125 | 171 | 2704 | 2025 | 2401 | 7 |
| 105 | 75 | 175 | 124 | 129 | 128 | 171 | 2601 | 2116 | 2209 | 14 |
| 106 | 75 | 169 | 130 | 132 | 132 | 171 | 1521 | 1369 | 1369 | 3 |
| 107 | 75 | 164 | 128 | 134 | 133 | 167 | 1296 | 900 | 961 | 10 |
| 108 | 75 | 173 | 127 | 139 | 139 | 162 | 2116 | 1156 | 1156 | 111 |
| 109 | 75 | 156 | 145 | 142 | 145 | 156 | 121 | 196 | 121 | 0 |
| 110 | 75 | 156 | 131 | 140 | 139 | 147 | 625 | 256 | 289 | 75 |
| 111 | 75 | 129 | 132 | 143 | 145 | 138 | 9 | 196 | 256 | 82 |
| 112 | 75 | 150 | 144 | 147 | 151 | 128 | 36 | 9 | 1 | 496 |
| 113 | 75 | 116 | 136 | 149 | 151 | 112 | 400 | 1089 | 1225 | 14 |
| 114 | 75 | 111 | 137 | 151 | 154 | 95 | 676 | 1600 | 1849 | 252 |
| 115 | 75 | 69 | 139 | 153 | 156 | 76 | 4900 | 7056 | 7569 | 52 |
| 116 | 75 | 57 | 142 | 155 | 159 | 56 | 7225 | 9604 | 10404 | 2 |
| 117 | 75 | 47 | 142 | 155 | 160 | 46 | 9025 | 11664 | 12769 | 2 |
| 118 | 75 | 26 | 140 | 151 | 154 | 54 | 12996 | 15625 | 16384 | 780 |
| 119 | 75 | 11 | 140 | 150 | 154 | 66 | 16641 | 19321 | 20449 | 2975 |
| 120 | 75 | 6 | 145 | 152 | 158 | 78 | 19321 | 21316 | 23104 | 5154 |
| 121 | 75 | 42 | 144 | 145 | 146 | 84 | 10404 | 10609 | 10816 | 1730 |
| 122 | 75 | 18 | 134 | 141 | 143 | 75 | 13456 | 15129 | 15625 | 3240 |
| 123 | 75 | 31 | 146 | 143 | 142 | 66 | 13225 | 12544 | 12321 | 1197 |
| 124 | 75 | 62 | 143 | 149 | 151 | 62 | 6561 | 7569 | 7921 | 0 |
| 125 | 75 | 55 | 148 | 157 | 166 | 74 | 8649 | 10404 | 12321 | 352 |
| 126 | 75 | 61 | 148 | 152 | 157 | 102 | 7569 | 8281 | 9216 | 1669 |

| | | | | | | | | | | |
|------------------------------|----|-----|-----|-----|-----|-----|-------|-------|-------|------|
| 127 | 75 | 189 | 147 | 147 | 150 | 132 | 1764 | 1764 | 1521 | 3263 |
| 128 | 75 | 173 | 149 | 145 | 150 | 163 | 576 | 784 | 529 | 105 |
| 129 | 75 | 153 | 149 | 146 | 151 | 178 | 16 | 49 | 4 | 600 |
| 130 | 75 | 179 | 150 | 144 | 148 | 179 | 841 | 1225 | 961 | 0 |
| 131 | 75 | 175 | 156 | 142 | 147 | 179 | 361 | 1089 | 784 | 20 |
| 132 | 75 | 182 | 157 | 141 | 148 | 179 | 625 | 1681 | 1156 | 10 |
| 133 | 75 | 177 | 155 | 141 | 147 | 187 | 484 | 1296 | 900 | 102 |
| 134 | 75 | 175 | 154 | 144 | 151 | 204 | 441 | 961 | 576 | 854 |
| 135 | 75 | 255 | 152 | 143 | 149 | 221 | 10609 | 12544 | 11236 | 1173 |
| 136 | 75 | 251 | 155 | 142 | 147 | 236 | 9216 | 11881 | 10816 | 212 |
| 137 | 75 | 242 | 157 | 141 | 145 | 241 | 7225 | 10201 | 9409 | 2 |
| 138 | 75 | 235 | 157 | 139 | 143 | 236 | 6084 | 9216 | 8464 | 1 |
| 139 | 75 | 231 | 155 | 142 | 148 | 233 | 5776 | 7921 | 6889 | 4 |
| 140 | 75 | 236 | 158 | 144 | 149 | 233 | 6084 | 8464 | 7569 | 9 |
| 141 | 75 | 234 | 149 | 145 | 150 | 235 | 7225 | 7921 | 7056 | 0 |
| 142 | 75 | 245 | 153 | 148 | 155 | 237 | 8464 | 9409 | 8100 | 63 |
| 143 | 75 | 239 | 147 | 146 | 151 | 239 | 8464 | 8649 | 7744 | 0 |
| 144 | 75 | 243 | 147 | 144 | 149 | 240 | 9216 | 9801 | 8836 | 7 |
| 145 | 75 | 240 | 141 | 140 | 144 | 240 | 9801 | 10000 | 9216 | 0 |
| 146 | 75 | 248 | 125 | 130 | 128 | 239 | 15129 | 13924 | 14400 | 76 |
| 147 | 75 | 244 | 118 | 127 | 126 | 238 | 15876 | 13689 | 13924 | 33 |
| 148 | 75 | 237 | 134 | 127 | 128 | 238 | 10609 | 12100 | 11881 | 1 |
| 149 | 75 | 238 | 123 | 126 | 125 | 238 | 13225 | 12544 | 12769 | 0 |
| Erro Quadrático Médio | | | | | | | 4847 | 4600 | 4613 | 604 |

Apêndice II

Resultados obtidos com uma metodologia *bottom-up*

Neste apêndice considerou-se relevante apresentar, pela sua importância, os resultados obtidos pela abordagem *bottom-up*, isto é, partindo da geração dos modelos direccionais para o modelo global e utilizando todos os pontos do variograma (acima ou abaixo da variância) para o cálculo do erro.

Assim, de seguida, apresentam-se alguns resultados relativos à modelação automática dos variogramas referentes às seguintes imagens: foto em tons de cinzento constante na página 133, foto a cores constante na página 136, foto de satélite constante na página 142 e amostra 1 de mármore constante na figura 160 devendo notar-se que, em algumas situações, a soma do efeito de pepita com o patamar é superior à variância pois, ao momento, não estava a ser validado.

Foto em tons de cinzento

Modelo teórico omnidireccional (anisotropia a 90°) e modelos direccionais referentes ao canal R:

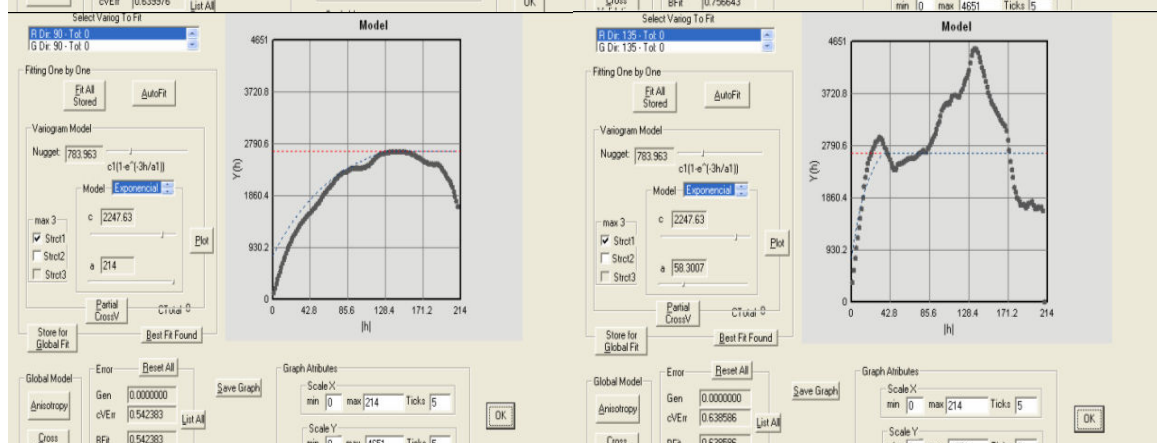
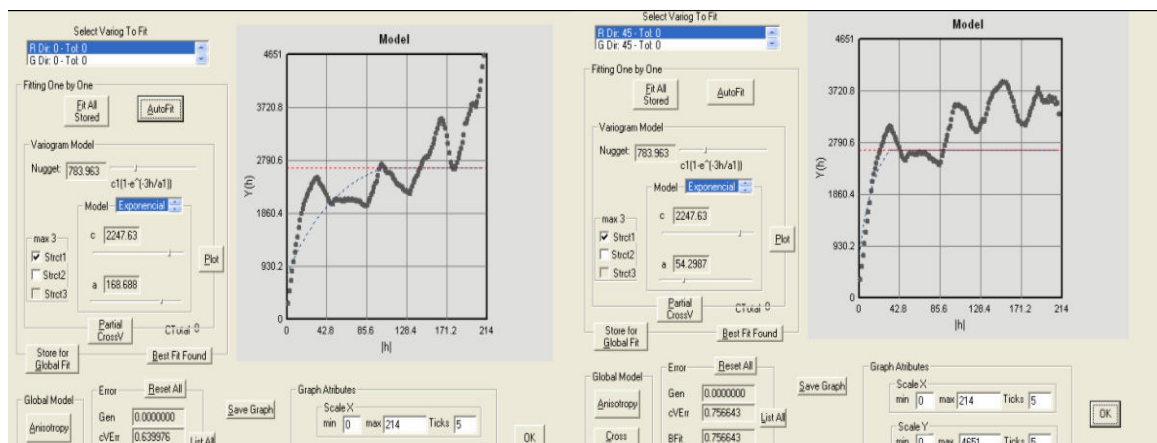
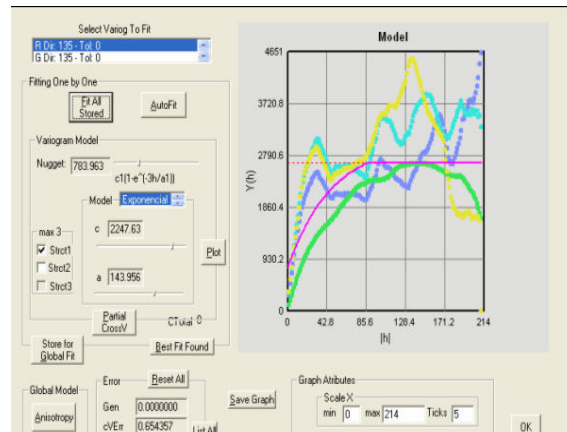
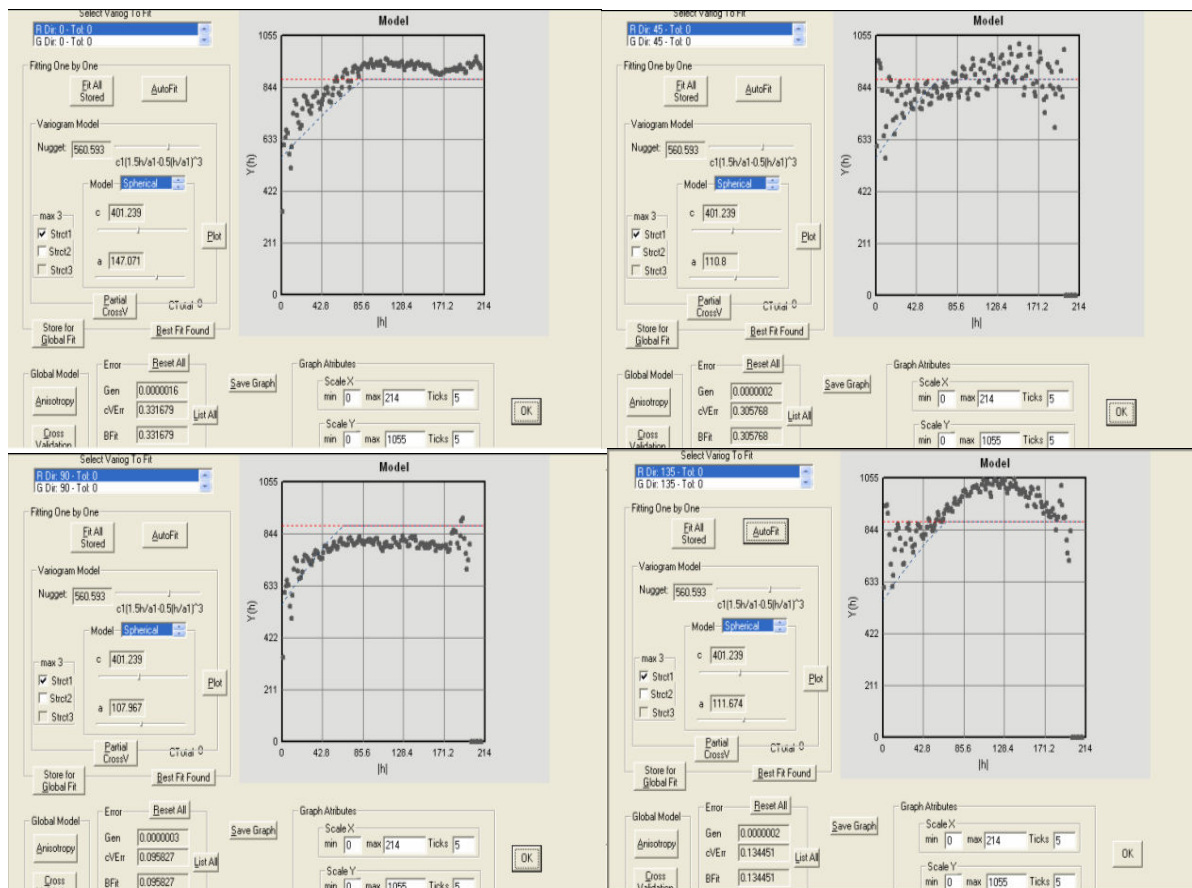
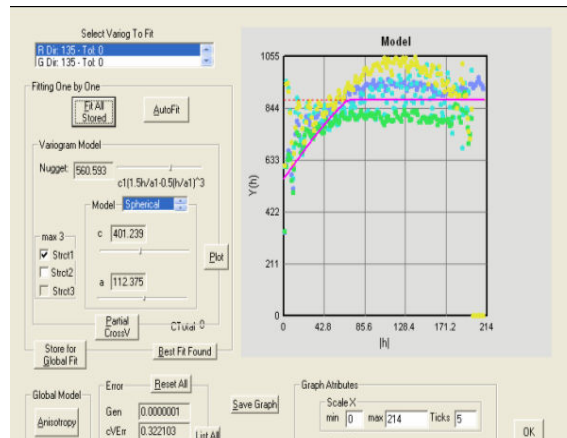
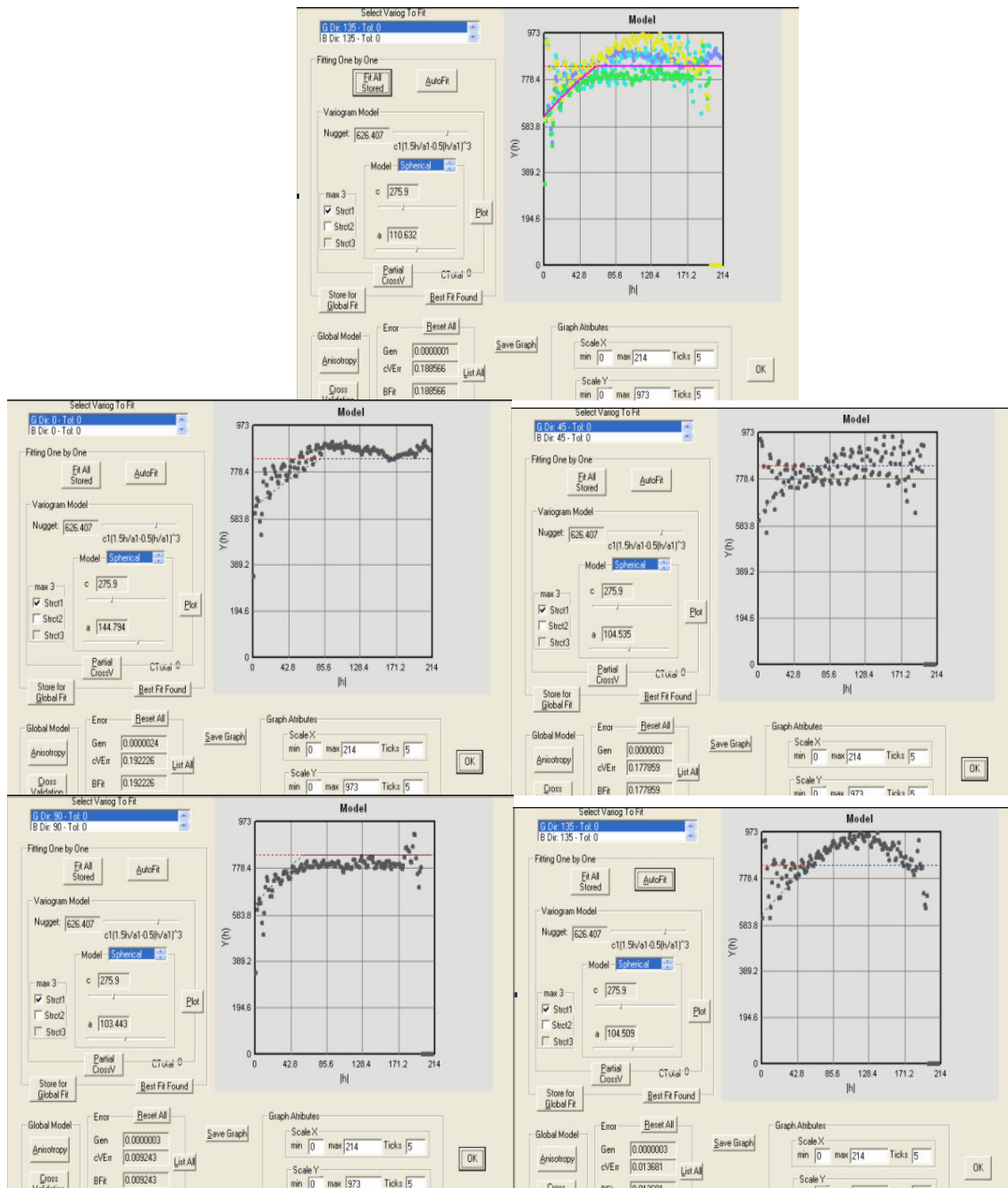


Foto da calçada portuguesa

Modelo teórico omnidireccional isótopo e modelos direccionais referentes ao canal R:



Modelo teórico omnidireccional isótropo e modelos direccionais referentes ao canal G:



Modelo teórico omnidireccional isótropo e modelos direccionais referentes ao canal B:

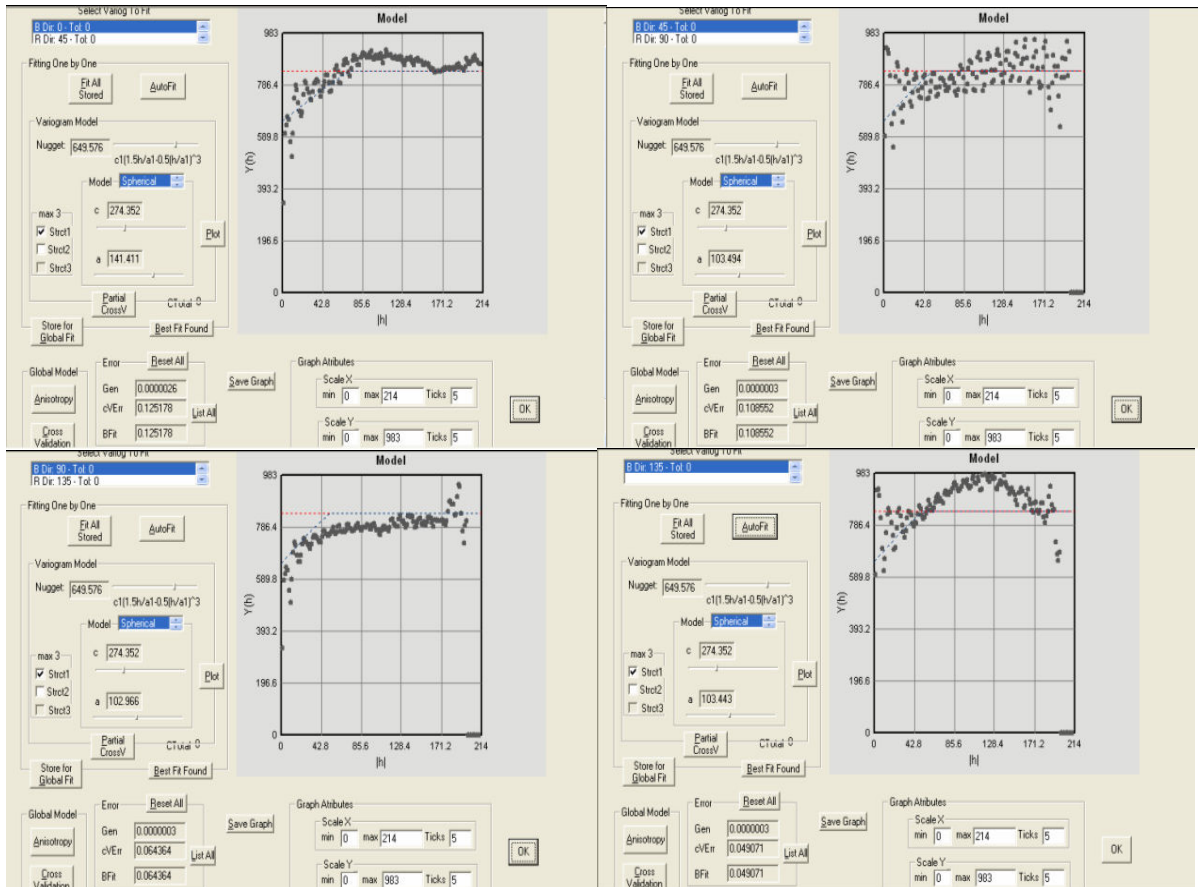
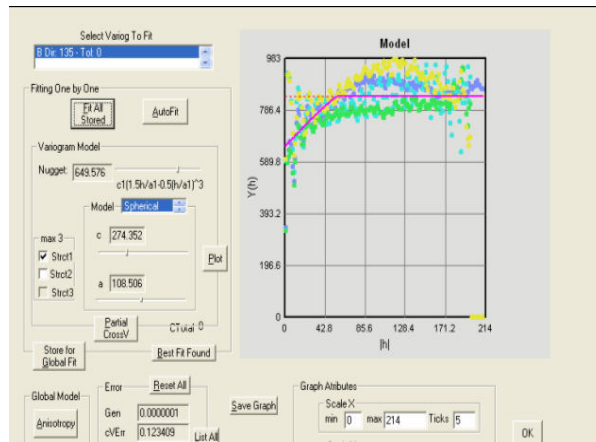
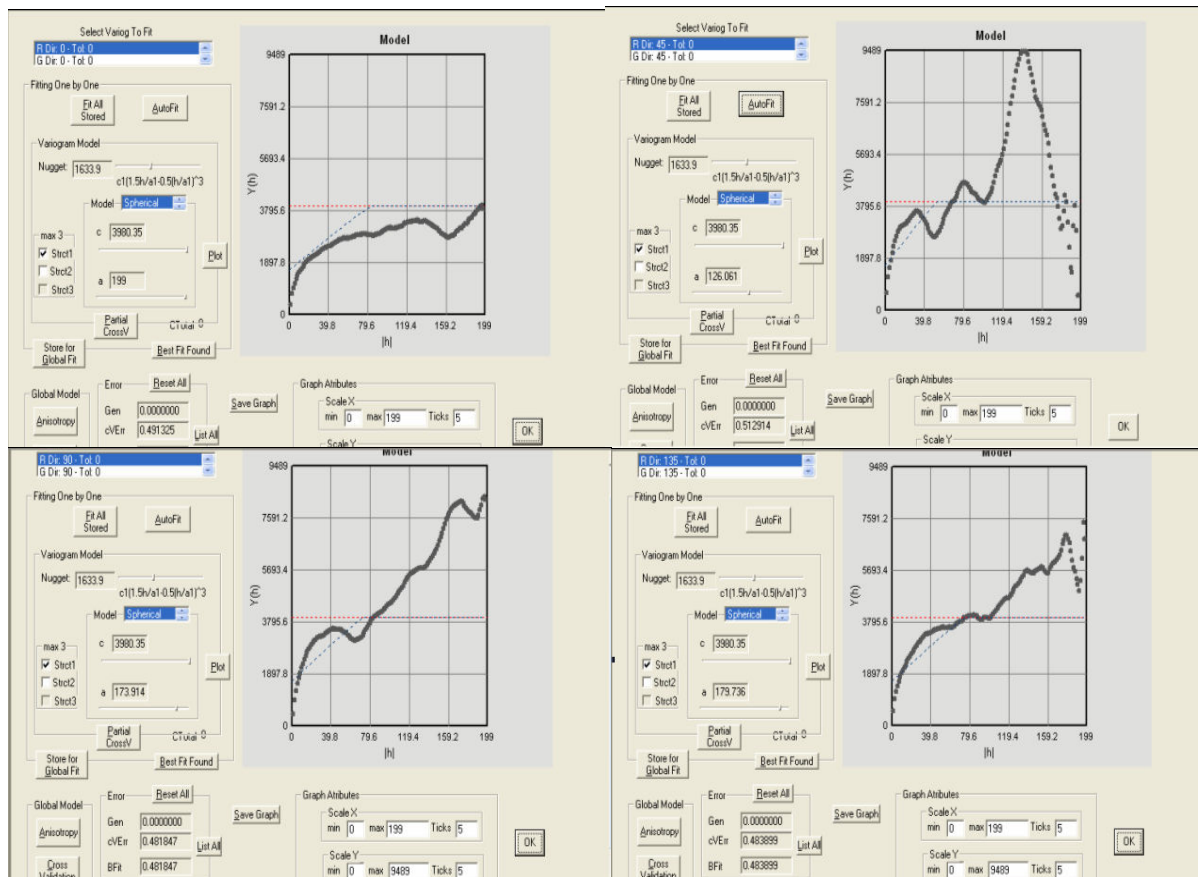
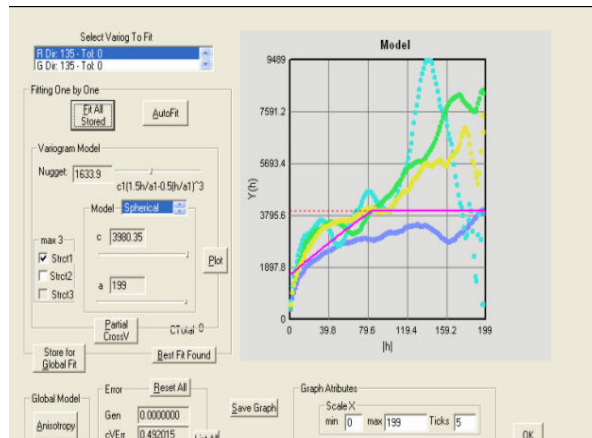
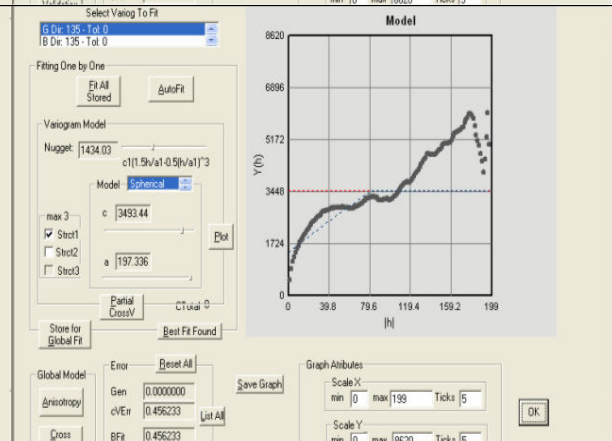
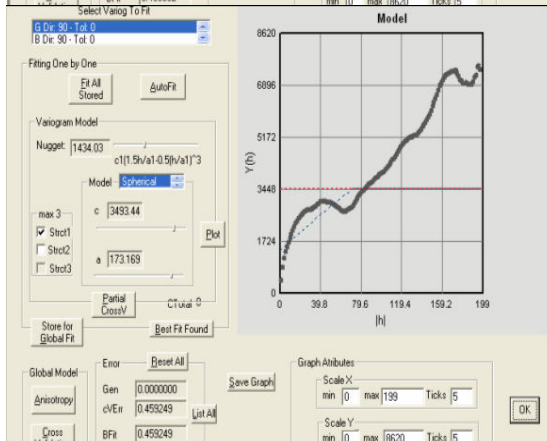
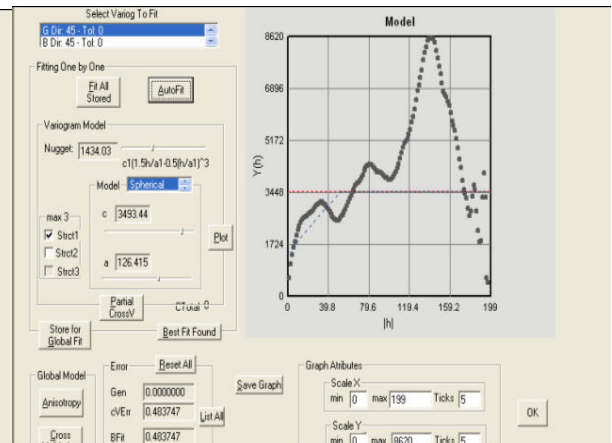
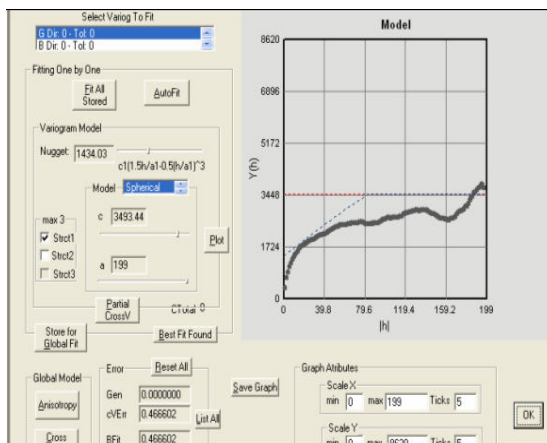
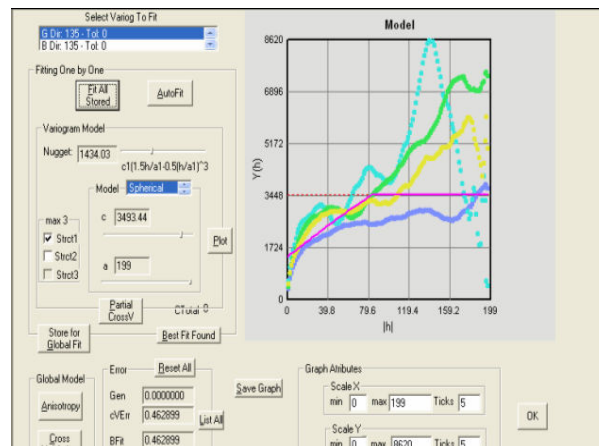


Imagem de satélite

Modelo teórico omnidireccional (anisotropia a 0°) e modelos direccionais referentes ao canal R:

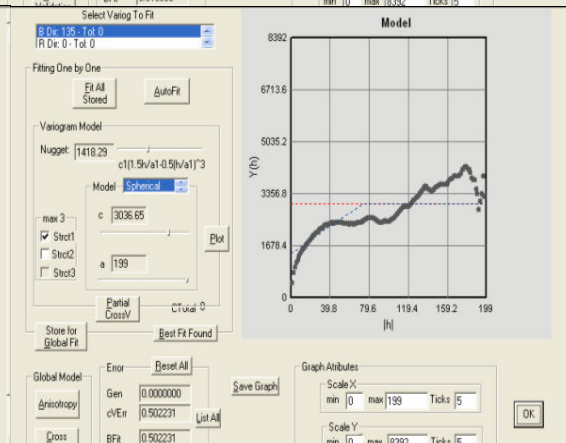
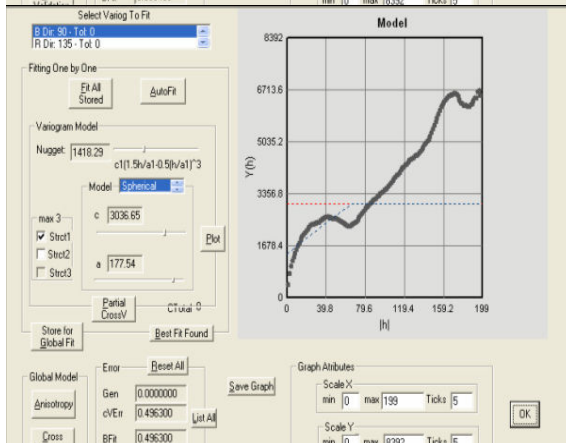
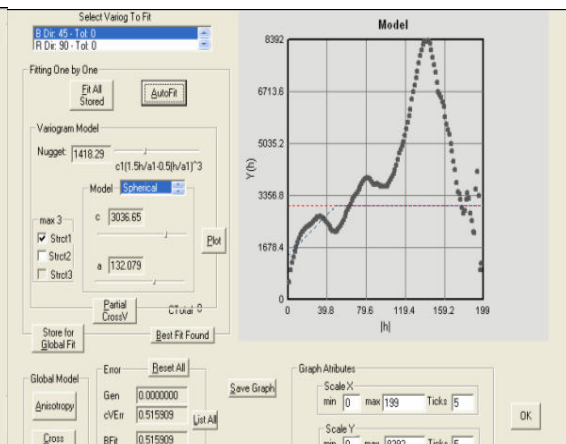
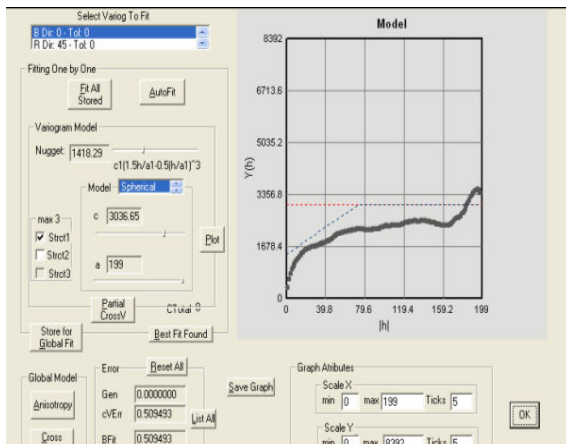
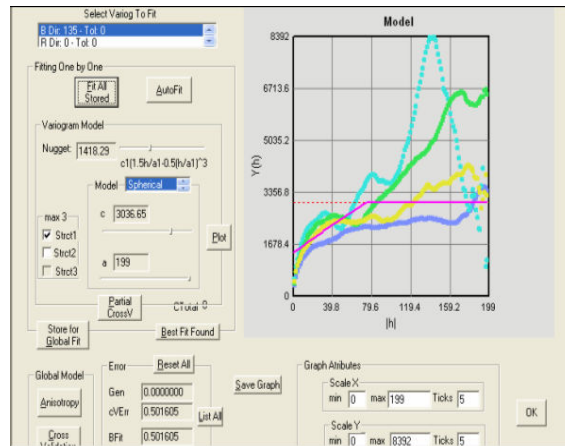


Modelo teórico omnidireccional anisotropia a 0° e modelos direccionais referentes ao canal G:



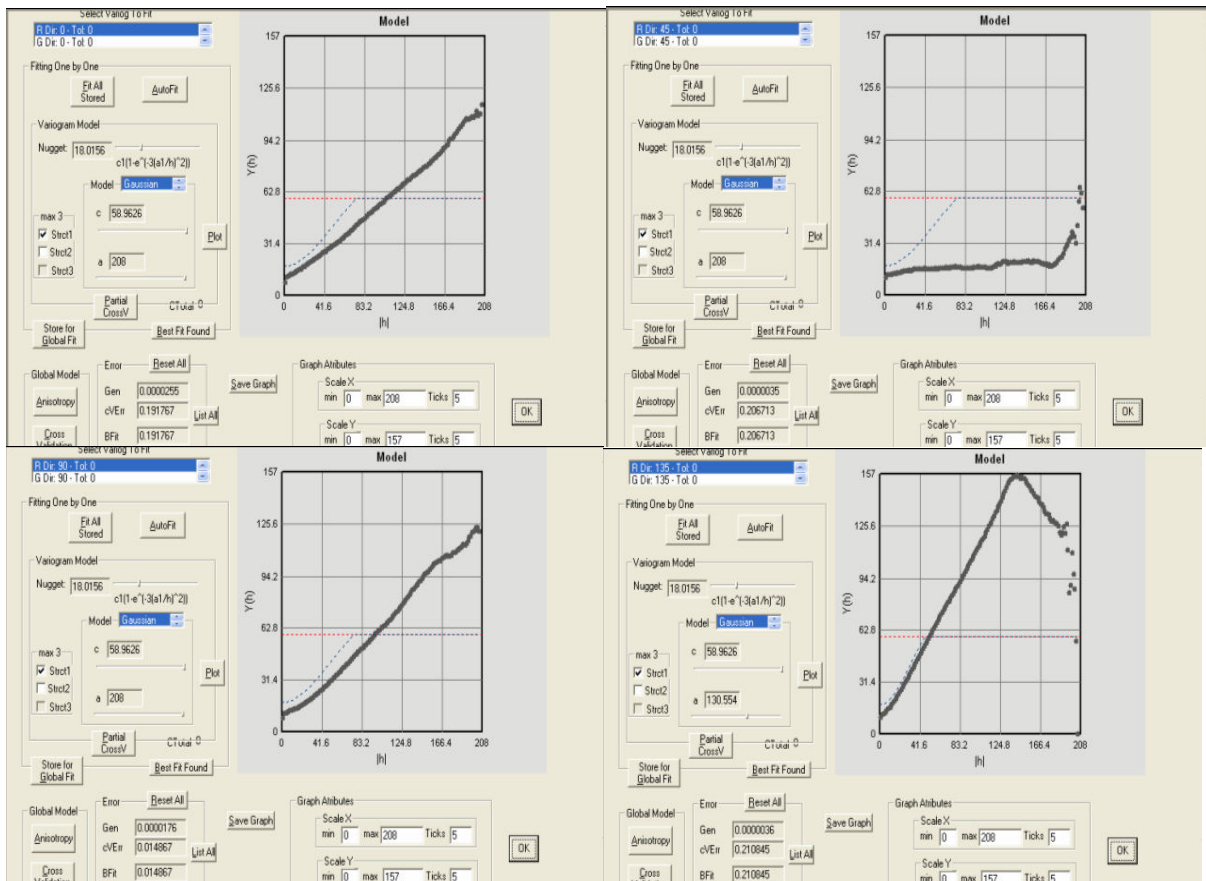
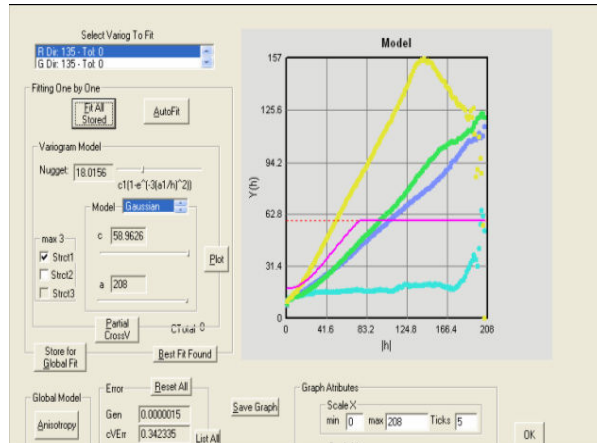
Modelo teórico omnidireccional (anisotropia a 0°) e modelos direccionais referentes ao canal

B:

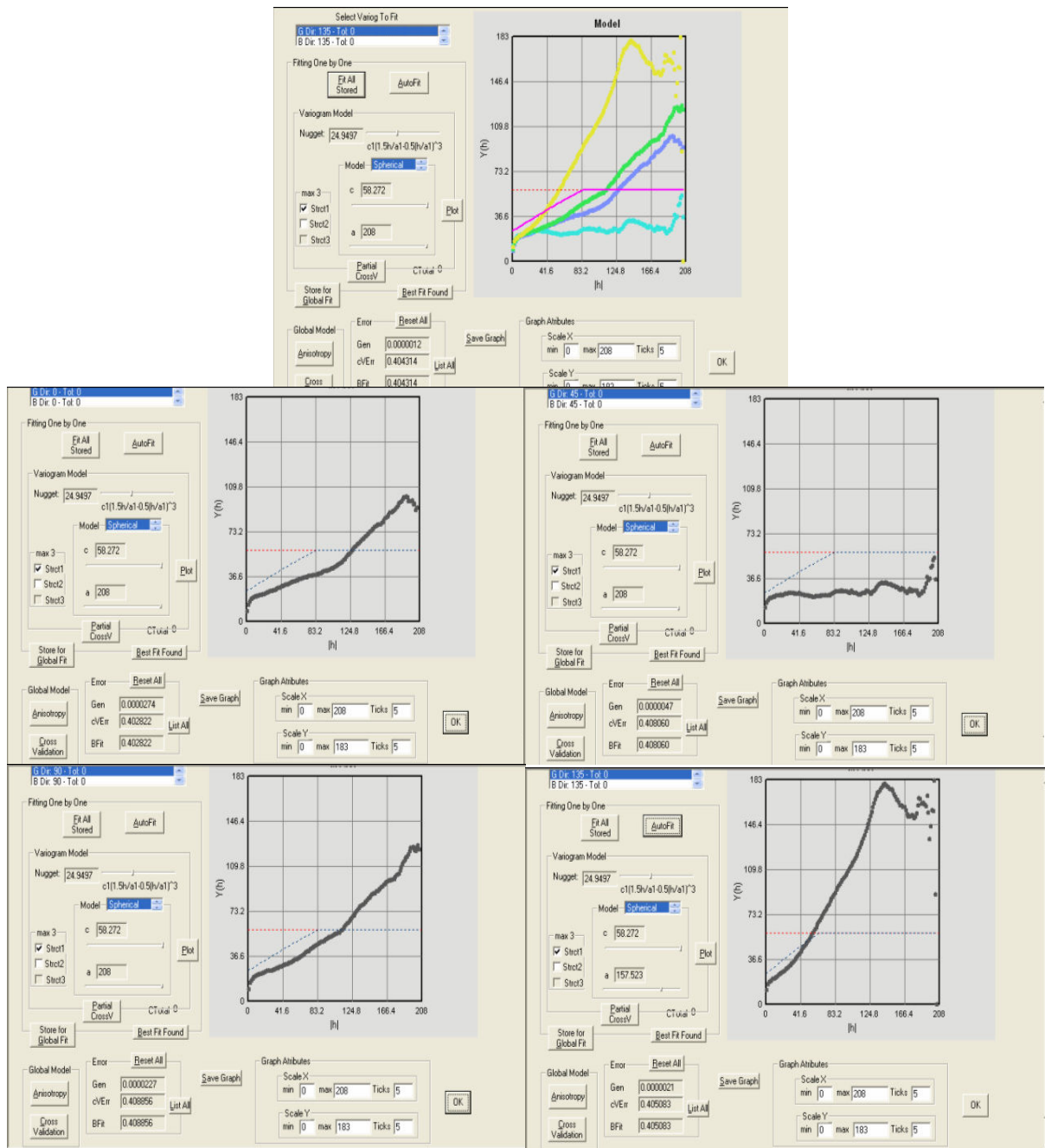


Amostra 1 de mármore

Modelo teórico omnidireccional (anisotropia a 0°) e modelos direccionais referentes ao canal R:



Modelo teórico omnidireccional isótopo e modelos direccionais referentes ao canal G:



Modelo teórico omnidireccional isótropo e modelos direccionais referentes ao canal B:

