

UNIVERSIDADE ESTADUAL DE CAMPINAS  
FACULDADE DE ENGENHARIA AGRÍCOLA

**DESENVOLVIMENTO DE UM SISTEMA  
COMPUTACIONAL PARA ANÁLISE GEO-ESTATÍSTICA**

---

**LAURIMAR GONÇALVES VENDRUSCULO**

CAMPINAS  
DEZEMBRO DE 2001

UNICAMP  
BIBLIOTECA CENTRAL

UNIDADE:	80
Nº CHAMADA:	V553a
V:	
TOMBO BO:	48795
PROG.:	16.837102
C:	<input type="checkbox"/>
PREÇO:	R\$11,00
DATA:	07/05/02
Nº CPD:	

CM001671BB-B

319 ID 239683

FICHA CATALOGRÁFICA ELABORADA PELA  
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

V553d Vendrusculo, Laurimar Gonçalves  
Desenvolvimento de um sistema computacional para  
análise geo-estatística / Laurimar Gonçalves  
Vendrusculo.--Campinas, SP: [s.n.], 2001.

Orientadores: Paulo Sérgio G. Magalhães e Sidney  
Rosa Vieira.

Dissertação (mestrado) - Universidade Estadual de  
Campinas, Faculdade de Engenharia Agrícola.

1. Solos-amostragem. 2. Análise espacial  
(Estatística). 3. Delphi (Programa de computador). I.  
Magalhães, Paulo Sérgio Graziano. II. Vieira, Sidney  
Rosa. III. Universidade Estadual de Campinas. Faculdade  
de Engenharia Agrícola. IV. Título.

UNIVERSIDADE ESTADUAL DE CAMPINAS  
FACULDADE DE ENGENHARIA AGRÍCOLA

**DESENVOLVIMENTO DE UM SISTEMA  
COMPUTACIONAL PARA ANÁLISE GEO-ESTATÍSTICA**

---

Dissertação submetida à banca examinadora  
para obtenção do título de Mestre em Engenharia  
Agrícola na área de concentração de  
Planejamento e Desenvolvimento Rural  
Sustentável

**LAURIMAR GONÇALVES VENDRUSCULO**

ORIENTADOR: Prof. Dr. PAULO S. GRAZIANO MAGALHÃES

CO-ORIENTADOR: Prof. Dr. SIDNEY ROSA VIEIRA

CAMPINAS  
DEZEMBRO DE 2001

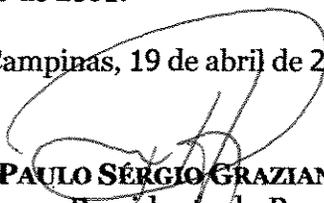
6966100

---

## **PARECER**

Este exemplar corresponde a redação final da Dissertação de Mestrado defendida por e aprovada **LAURIMAR GONÇALVES VENDRUSCULO** pela Comissão Julgadora em 20 de dezembro de 2001.

Campinas, 19 de abril de 2002.



**Prof. Dr. PAULO SÉRGIO GRAZIANO MAGALHÃES**  
Presidente da Banca

---

Para o meu esposo Edson, minha mãe Maria Auxiliadora e irmãos: Marymar, Anderson,  
Everton, Peterson, Geicimar, Gleicimar e Debora com muito carinho.

## AGRADECIMENTOS

Ao meu esposo Edson, pelo carinho, apoio e compreensão durante todo o tempo de desenvolvimento deste trabalho.

Aos meus pais, em especial a D. Auxiliadora, pelo constante incentivo durante todos os momentos da minha vida.

Aos meus irmãos que, mesmo distante, apoiaram e torceram pela realização deste trabalho.

Ao Prof. Dr. Paulo Sérgio Graziano Magalhães, pela orientação, confiança e estímulo durante o desenvolvimento deste trabalho.

Ao Prof. Dr. Sidney Rosa Vieira, do Instituto Agronômico de Campinas, que prontamente cedeu suas rotinas computacionais e co-orientou este trabalho, estimulando continuamente a busca pelo conhecimento geo-estatístico.

Ao Dr. José Ruy Porto, da Embrapa Informática Agropecuária, pelo apoio como Conselheiro Acadêmico e pelas consultorias de cunho estatístico.

Aos colegas da Embrapa Informática Agropecuária pelas ações de apoio a esta pesquisa.

Aos colegas da Faculdade de Engenharia Agrícola e a coordenação da Pós-Graduação durante o ano de 2001 em especial pelo empenho na difusão desta pesquisa.

À Empresa Brasileira de Pesquisa Agropecuária – Embrapa e Fapesp, pelo auxílio financeiro.

## LISTA DE FIGURAS

Figura 3.1 - Amostra aleatória de 32 pontos localizados dentro da área quadrada. (Fonte: WOLLENHAUPT et al., 1997).....	06
Figura 3.2 - Estágios para desenvolvimento de amostragem aninhada. (Fonte: WOLLENHAUPT et al., 1997).....	07
Figura 3.3 - Amostragem sistemática estratificada em grade quadrada. (Fonte: WOLLENHAUPT et al., 1997).....	08
Figura 3.4 - Amostragem estratificada sistemática não alinhada. (Fonte: WOLLENHAUPT et al., 1997).....	09
Figura 3.5 - Amostragem em duas dimensões (Fonte: CAMARGO, 1997).....	14
Figura 3.6 - Variograma típico e seus parâmetros. (Fonte: Modificado de VIEIRA et al., 1983).....	15
Figura 3.7 - Modelos teóricos de variograma. (Fonte: CAMARGO,1997) .....	17
Figura 3.8 – Fases do Desenvolvimento de um software (Fonte: Modificado de PRESSMAN, 1982).....	22
Figura 4.1 - Distribuição espacial dos pontos amostrados nos dois campos.....	27
Figura 4.2 – Estrutura do arquivo de dados Waynicks.dat .....	28
Figura 4.3 - Diagrama de funcionalidades do sistema.....	30
Figura 4.4 – Esquema genérico de interação entre módulos dos sistemas e arquivos gerados pelas rotinas Fortran.....	32
Figura 4.5 – Interface para geração de estatísticas clássicas e janela de resultados.....	33
Figura 4.6 – Interface de entrada dos parâmetros para cálculo do semivariograma médio.....	34
Figura 4.7 – Interface do Semivariograma médio experimental .....	36
Figura 4.8 – Interface de entrada dos parâmetros para cálculo do semivariograma médio, quando existirem dados perdidos .....	37

Figura 4.9 – Interface para ajuste de modelos matemáticos teóricos.....	38
Figura 4.10 – Porção da interface utilizada na retirada de tendência dos dados .....	41
Figura 4.11 – Gráfico gerado pela função Retirada de tendência, com o ajuste da curva de resíduos de Carbono em Davis (semivariograma experimental - ●, curva de resíduos - ■, superfície de tendência ▲) .....	42
Figura 4.12 – Interface para verificar se o número de vizinhos escolhidos para o processo de krigagem é adequado ou se os parâmetros escolhidos para o modelo de ajuste são corretos pela técnica de <i>Jack-Knifing</i> .....	44
Figura 4.13 – Especificação de vizinhança através do critério de número constante de vizinhos. A interface ilustra a pesquisa por 8 vizinhos dentro do raio da distância máxima.....	45
Figura 4.14 – Especificação de vizinhança através do critério de distância constante de 10 metros.....	45
Figura 4.15 – Malha de amostragem genérica (a) e malha de referência usada na krigagem (b), destacando os pontos medidos (pontos cheios) e aqueles a serem estimados (pontos vazados).....	47
Figura 4.16 – Interface de entrada para o processo de krigagem.....	48
Figura 4.17 – Estrutura de um arquivo relativo aos pontos do polígono que envolve a área de amostragem.....	48
Figura 4.18 – Interface para o módulo de otimização de malha da amostragem.....	51
Figura 4.19 – Parâmetros de uma malha fictícia usada no módulo de projeto de malha.....	51
Figura 4.20 – Interface para o cálculo dos semivariogramas direcionais .....	53
Figura 4.21 – Interface resultante dos semivariogramas direcionais.....	54
Figura 4.22 – Interface do módulo para geração do semivariograma cruzado.....	55
Figura 5.1 - Semivariograma escalonado para distância máxima igual a 73 (a) e 83 (b) metros e Número de passos = 4. ....	59
Figura 5.2 - Semivariograma escalonado para distância máxima igual a 73 (a) e 83 (b) metros e Número de passos = 1. ....	60

Figura 5.3 - Semivariograma experimental para C/N em Davis e Oakley. ....	60
Figura 5.4 - Semivariograma escalonado da variável C/N de Davis ajustada pelos modelos esférico, exponencial, gaussiano, <i>Hole effect</i> e pentaesférico. ....	61
Figura 5.5 – Ajuste fino do modelo exponencial ( $C_0 = 0,32$ ; $C=0,7$ e $a=25$ ) para a relação C/N em Davis .....	62
Figura 5.6 – Ajuste de Carbono por meio da curva de resíduos, utilizando superfície de tendência parabólica.....	63
Figura 5.7 – Semivariogramas de Carbono (a) e Nitrogênio (b e c) aplicando-se retirada de tendência por meio de superfície parabólica.....	64
Figura 5.8 – Resultados do <i>Jack-knifing</i> para várias configurações do semivariograma teórico.....	65
Figura 5.9 – Visão parcial dos pontos referentes a malha de referência .....	66
Figura 5.10 – Interface do processo de Krigagem para carbono em Davis.....	67
<u>Figura 5.11 – Arquivo resultante do processo de krigagem para a variável Carbono em Davis.....</u>	<u>68</u>
Figura 5.12 - Operações para integração do arquivo interpolado com o sistema Surfer, visando geração do mapa de classes para Carbono em Davis.....	69
Figura 5.13 – Mapa simbólico dos valores interpolados de Carbono (Davis) .....	69
Figura 5.14 - Mapa simbólico dos valores interpolados de Carbono (Oakley) .....	70
Figura 5.15 - Mapa simbólico dos valores interpolados de Nitrogênio (Davis) .....	70
Figura 5.16 - Mapa simbólico dos valores interpolados de Nitrogênio (Oakley).....	71
Figura 5.17 – Comportamento da variância da estimativa para várias configurações de malha.....	73
Figura 5.18 – Gráfico da variância da estimativa para diversos projetos de amostragem utilizando a curva de resíduos de Carbono em Davis.....	74
Figura 5.19 - Semivariogramas direcionais da variável Carbono em Davis para diferentes valores de ângulo de tolerância, NSTEP e NLAG.....	76
Figura 5.20 – Parâmetros de entrada para análise do semivariograma cruzado entre Carbono e Nitrogênio em Davis.....	77
Figura 5.21 – Semivariograma cruzado de Carbono e Nitrogênio em Davis .....	78
Figura 5.22 – Semivariograma cruzado de Carbono e Nitrogênio em Oakley .....	78

## LISTA DE TABELAS

Tabela 4.1 – Parte inicial dos valores interpolados e medidos de Carbono (Davis) e sua respectiva variância .....	49
Tabela 5.1 – Estatísticas clássicas para o conjunto de dados Waynick e Sharp (1919).....	57
Tabela 5.2- Matriz de Correlação .....	58
Tabela 5.3– Configurações pré-definidas para malhas de amostragem.....	72

---

## RESUMO

A adoção de técnicas e métodos de gerenciamento, mais eficazes, em propriedades rurais visando o aumento da produção agropecuária, em bases sustentáveis, deve merecer mais atenção por parte dos agentes do sistema produtivo. Destaca-se neste trabalho a amostragem de solos com incorporação de métodos sistemáticos e preditivos. É sabido que as propriedades do solo diferenciam-se por seus padrões de variabilidade espacial e temporal. Ao contrário da estatística clássica, a geo-estatística considera a distância para melhor explicar a variância entre as amostras e predizer valores em locais não amostrados. A geo-estatística e outras técnicas integram a abordagem denominada de Agricultura de Precisão, a qual preconiza o tratamento diferenciado para cada área de acordo com as necessidades e potencial produtivo. Dentre seus objetivos a Agricultura de Precisão procura: reduzir desperdícios, aumentar produtividade e garantir a qualidade do meio ambiente. O objetivo deste estudo compreende a concepção e desenvolvimento de um sistema para análise de variabilidade espacial com uso da técnica de geo-estatística. A ferramenta é constituída de módulos que permitem: a modelagem de semivariogramas médios, direcionais e cruzados, o ajuste de modelos matemáticos, o tratamento da estacionariedade dos dados e teste de autovalidação. Permite ainda a interpolação de pontos não amostrados, disponibilizando informações para a geração de mapas de atributos de solo. O sistema utilizou duas ferramentas básicas, a saber: biblioteca de rotinas geo-estatísticas escritas anteriormente em Fortran 77, e ambiente de programação Delphi 5. O sistema desenvolvido foi validado por meio de um conjunto de dados relatado em literatura. Foram estudados os comportamentos espaciais das variáveis Carbono, Nitrogênio e da relação C/N, num esquema de amostragem denso. Os resultados obtidos se mostraram semelhantes aqueles encontrados em literatura específica e a integração das rotinas geo-estatísticas agilizou o processo de análise em relação ao uso modular das mesmas em ambiente DOS.

---

**Palavras-chave:** amostragem de solo, variabilidade espacial, krigagem, tecnologias de informação, linguagem de programação Delphi

# Development of computational system to geostatistical analysis

## ABSTRACT

The growing demand for agricultural land is obliging the productive agents to adopt modern methods and techniques. The soil sampling makes up as an important technique since it can incorporate systematic and predictive tools. It is known that the soil properties have difference when analyzing patterns of spatial and temporal variability. Nevertheless, different of the classical statistics, geostatistics takes into account the distance between samples pairs to better explain the variance and to estimate values in areas not sampled. Precision Farming comprises geostatistics and other techniques, which deal with a specific management system to each field. The purpose of the precision farming is to increase the production while guaranteeing the environmental quality and reducing the waste. The goal of this study is to design and develop a computer system able to analyze soil spatial variability through geostatistical methods. It has modules for: modeling of average, directional and cross semivariogram; fitting mathematical models, data stationarity treating and autovalidation testing. Moreover, attributes soil maps could be generate by kriging module. Two basics tools have been used to develop the system: Geoestatistical FORTRAN routines and application development environment for Windows Borland Delphi 5. Dataset available in the literature were used to validate the system performance. Spatial structure of Carbon, Nitrogen and C/N were studied in dense soil sampling. The results were close with those obtained in specific literature and the modules integration provided by the system allowed to accelerate the data analysis.

---

**Key Words:** soil sampling, spatial variability, kriging, information technology, Delphi programming language.

# SUMÁRIO

	Página
LISTA DE FIGURAS.....	v
LISTA DE TABELAS.....	vii
RESUMO.....	ix
ABSTRACT.....	x
1. INTRODUÇÃO E JUSTIFICATIVA .....	1
2. OBJETIVOS .....	3
3. REVISÃO DE LITERATURA .....	4
3.1 Amostragem de solos com ênfase na variabilidade de suas propriedades.....	4
3.2 - Projetos para amostragem de solo .....	5
3.2.1 - Amostragem Proposital.....	6
3.2.2 - Amostragem aleatória simples.....	6
3.2.3 - Aleatória estratificada .....	6
3.2.4 - Cluster .....	7
3.2.5 - Aninhada ou múltiplos estágios.....	7
3.2.6 - Sistemática.....	8
3.2.7 - Sistemática Estratificada não alinhada.....	8
3.2.8 - Informação Complementar.....	9
3.3 - Métodos de Interpolação.....	10
3.4 - Hipóteses fundamentais da geo-estatística .....	11
3.4.1 - Fundamentos .....	11
3.4.2. - Hipótese de estacionaridade de ordem 2 .....	12
3.4.3 - Hipótese intrínseca .....	13
3.4.4 - Semivariograma .....	14
3.4.4.1 - Parâmetros de um semivariograma ideal .....	15
3.4.5 - Modelos Teóricos clássicos .....	16
3.4.5.1- Modelo Esférico .....	17
3.4.5.2 - Modelo Exponencial.....	18
3.4.5.3 - Modelo Gaussiano.....	18

3.4.5.4 - Modelo <i>Hole Effect</i> .....	18
3.4.5.5 - Modelo Pentaesférico .....	19
3.4.6 - Krigagem.....	19
3.5 - Estudos de variabilidade de solos utilizando a técnica da geo-estatística.....	20
3.6 - Desenvolvimento de sistemas para tratamento geo-estatístico .....	21
4. MATERIAL E MÉTODOS .....	26
4.1 - Descrição da área estudada .....	26
4.2 - Desenvolvimento Computacional .....	28
4.3 - Descrição dos módulos do sistema .....	32
4.3.1 - Módulo de análise exploratória .....	32
4.3.2 - Módulo para modelagem do semivariograma experimental.....	33
4.3.3 - Módulo de ajuste do semivariograma aos modelos matemáticos.....	37
4.3.4 - Módulo para retirada de tendência .....	40
4.3.5 - Módulo de Autovalidação – Jack-Knifing .....	42
4.3.6 - Módulo de Krigagem.....	46
4.3.7 - Módulo para projeto da malha de amostragem.....	50
4.3.8 - Módulo para estudo de semivariogramas direcionais .....	52
4.3.9 - Módulo para estudo de semivariograma cruzado.....	54
5. RESULTADOS E DISCUSSÕES.....	57
5.1 - Análise exploratória de um conjunto dos dado .....	57
5.2 - Modelagem do semivariograma médio.....	58
5.3 - Ajuste do semivariograma aos modelos matemáticos .....	61
5.4 - Tratamento da estacionaridade dos dados.....	62
5.5 - Aplicação do módulo de Autovalidação.....	64
5.6 - Krigagem .....	66
5.7 - Projeto de configuração de amostragem .....	71
5.8 - Semivariogramas direcionais.....	74
5.9 - Semivariograma cruzado .....	77
6. CONSIDERAÇÕES FINAIS.....	80
7. REFERÊNCIAS BIBLIOGRÁFICAS.....	84

# CAPÍTULO 1

## INTRODUÇÃO E JUSTIFICATIVA

A atividade agropecuária tem sido direcionada para a sustentabilidade, como resultado das pressões impostas pela sociedade e pelo governo. Esta abordagem implica em empregar os recursos humanos e naturais de modo a suprir as necessidades correntes, sem comprometer a capacidade de sobrevivência das gerações futuras, por meio do equilíbrio de variáveis econômicas, ambientais e sociais (PINAZZA e ALIMANDRO, 2000).

Nas últimas três décadas o Brasil tem adotado soluções tecnológicas para a agricultura embasada pelo fenômeno "Revolução Verde", o qual prevê o emprego de sementes melhoradas, fertilizantes químicos, defensivos e mecanização dentre outros. Esse processo de transferência de tecnologia foi adotado dos países industrializados do ocidente e contempla uma contribuição significativa para o desenvolvimento econômico e tecnológico desses países.

Entretanto, no Brasil, segundo FAGNANI (1997) a incorporação destes pacotes foi feita sem a adequação, principalmente, dos fatores climáticos, ambientais e sociais locais. Ressalta-se que à área produtiva total, nesse modelo tecnológico, associa-se um valor médio das propriedades de solo ou parâmetros de produtividade.

Contrapondo a questão do tratamento discretizado preconizada pela "Revolução Verde", surge nas décadas de 70 e 80 a chamada agricultura de precisão, a qual reúne um conjunto de técnicas e equipamentos que tem por objetivo tratar diferenciadamente as unidades/zonas que compõem uma área agrícola produtiva. A agricultura de precisão considera a variabilidade espacial e temporal das culturas e das propriedades de solo, buscando racionalizar a aplicação de insumos e pesticidas. Considerada como uma nova abordagem de *gerenciamento agrícola* integrada, ROBERT (1999), dentro do contexto americano, enumera alguns potenciais ganhos advindos de sua adoção: produtividade, sustentabilidade, qualidade da produção, segurança alimentar, proteção ambiental, qualidade de vida e desenvolvimento econômico rural.

É sabido que propriedades do solo diferenciam-se por suas grandezas e pelos padrões de variabilidade espacial e temporal. Segundo DAMPNEY e MOORE (1998) a variabilidade abrange importantes fatores, tais como: nutrientes do solo, padrões de crescimento de culturas

e incidência de pragas. Porém, quando há dependência espacial entre as propriedades, detectado pelo processo de amostragem, a estatística clássica, por meio de parâmetros como média e desvio padrão, não deve ser usada como única ferramenta de análise do fenômeno.

Desde 1951, quando o matemático e engenheiro de minas D. G. KRIGE (1951) conseguiu explicar a variância de concentrações de ouro, considerando a distância entre observações, o estudo e aplicações da chamada geo-estatística continuam eficientes quando se deseja obter mais informações sobre a continuidade do solo.

O processo de amostragem de solos, no contexto da agricultura de precisão, encontra-se em estágio altamente tecnificado. Estão em uso amostradores de solos automatizados (TEIXEIRA, 2000), que além da coleta realizam análise, por meio de sensores, de propriedades físico-químicas (ADAMCHUK et al., 1999), em coordenadas geográficas definidas por equipamento GPS (Global Positioning System). Esta infra-estrutura permite o registro de variáveis em tempo real, agilizando o processo de coleta e reduzindo custos em relação aos métodos tradicionais de amostragem. As propriedades, objeto de análises laboratoriais, são incorporadas posteriormente aos programas computacionais com finalidade de estudo e simulações das condições do campo.

A otimização de esquemas de amostragem de solos, utilizando a técnica de geo-estatística, melhora o conhecimento da variabilidade espacial dos elementos de solo e sugere alternativas para distanciamento entre pontos baseado no conceito da variância da estimativa. Por meio do seu estimador (Krigagem) pontos que não foram amostrados podem ter seus valores estimados. A krigagem é um método que se diferencia por apresentar variância mínima e ser não tendencioso.

Estão disponíveis sistemas de informação, agregados ou não a instrumentos e sensores eletrônicos, com finalidade de facilitar a análise espacial das propriedades do solo. Estas tecnologias agilizaram tratamentos diferenciados em áreas de estudo, durante operações de, por exemplo, aplicação de nutrientes a taxa variável e elaboração de mapas de produtividade.

É oportuno o desenvolvimento de novas ferramentas computacionais que permitam o estudo da variabilidade por meio da geo-estatística, e que suportem com eficácia projetos para amostragem de solos.

## CAPÍTULO 2

### OBJETIVOS

O objetivo geral deste trabalho compreende a concepção, desenvolvimento e teste de um sistema computacional para estudo da variabilidade de atributos de solo por meio da técnica da geo-estatística.

De forma específica pretende-se:

- avaliar um conjunto de dados citado em literatura, estudando seu comportamento espacial e aferindo os resultados obtidos;
- desenvolver um ambiente integrado baseado no sistema Windows que agilize o processo de análise geo-estatística em relação a procedimentos modulares que utilizam o sistema MS-DOS;
- Verificar a distância máxima entre pontos amostrados que permitam (ou que não comprometa) uma adequada interpretação da variabilidade espacial do solo na região de estudo.

# CAPÍTULO 3

## REVISÃO DE LITERATURA

Este capítulo reforça a importância da variabilidade dos solos no contexto dos processos de amostragem. Alguns projetos de amostragem são descritos, bem como introduzem-se os conceitos básicos da geo-estatística. Cita-se finalmente alguns estudos de variabilidade dos solos utilizando técnicas de geo-estatística

### 3.1 Amostragem de solos com ênfase na variabilidade de suas propriedades

A demanda crescente por terras objetivando produção agropecuária aliada a condições de baixa sustentabilidade dos solos obriga os agentes do sistema produtivo a adoção de técnicas e métodos mais eficazes. Destacamos no escopo deste trabalho a técnica de amostragem de solos com incorporação de métodos sistemáticos e preditivos.

A amostragem tradicional de solos, de acordo com AMARAL (1997), é a mais crítica no programa de recomendações de adubação e calagem, pelas seguintes razões:

- O solo é um corpo heterogêneo em suas propriedades químicas;
- A heterogeneidade química do solo é acentuada pelas práticas de manejo, adubação, calagem e cultivo;
- Insuficiência de conhecimento dos princípios de amostragem de solos por parte daqueles que a realizam;
- Insuficiência de informações complementares para a interpretação de análise: adubação, calagem, rendimento de cultivos anteriores, topografia etc.

Os métodos mais comuns usados em amostragem de solo caracterizam-se pelas seguintes etapas:

- 1 - selecionar uma área com características homogêneas, considerando aspectos como: tipo de solo, topografia, vegetação e histórico da utilização do solo;
- 2 - caminhar e amostrar aleatoriamente seis a sete amostras dos primeiros vinte centímetros de solo;

3 - misturá-las para formar uma amostra composta para análises de rotina das propriedades químicas,

4 - fazer as recomendações de fertilizantes, com base nas análises, como se toda a área fosse homogênea.

Os processos de amostragem de solo estão sofrendo modificações para considerar a variabilidade espacial de suas propriedades, melhorando a abordagem de generalização de valores médios para uma determinada área. Cientistas e agricultores há muito tempo reconhecem as variações do solo e as características específicas de seu manejo bem com o potencial de sua produtividade.

WOLLENHAUPT et al. (1997) classificou a discussão da variabilidade das propriedades de solos em três passos independentes, compreendidos por:

- 1) Amostragem de solos;
- 2) Interpolação;
- 3) Mapeamento.

---

Padrões e intensidade da amostragem podem afetar o resultado do procedimento utilizado para estimar valores de propriedades de solo em pontos não amostrados quando se usa o processo de interpolação.

### **3.2 - Projetos para amostragem de solo**

O solo, para fins agrícolas, compreende a camada superficial que varia de zero a aproximadamente 0,60 m e apresenta uma grande variabilidade espacial em todas as direções. Determinar as propriedades do solo em um nível de confiança adequado torna-se, portanto, uma tarefa onerosa e de difícil execução. A alternativa à este problema é tomar amostras que representem tais propriedades por meio da divisão da área de interesse em elementos discretos ou individuais.

O projeto de amostragem influencia a acurácia dos resultados obtidos. WOLLENHAUPT et al. (1997) descrevem algumas destas estratégias.

### 3.2.1 - Amostragem Proposital

É aquela em que o investigador decide quais são os pontos (locais) que farão parte da população estudada. A inconveniência deste método é a incorporação de vício, característica inseparável de todo comportamento humano. Este método caracteriza-se pela generalização dos resultados obtidos para toda uma área.

### 3.2.2 - Amostragem aleatória simples

Como amostragem aleatória entende-se a escolha de indivíduos de uma população de tal modo que todos os membros da população tenham igual chance (probabilidade) de serem escolhidos.

Na modalidade simples desta técnica divide-se o campo em pequenas áreas de amostras (pontos) segundo uma seleção aleatória de um subconjunto de amostras. A Figura 3.1 ilustra a referida amostragem. Frequentemente um programa computacional que possibilite a geração de números aleatórios é usado para determinar os pontos que serão amostrados.

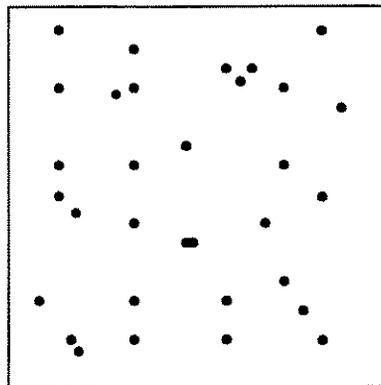


Figura 3.1 - Amostra aleatória de 32 pontos localizados dentro da área quadrada.

(Fonte: WOLLENHAUPT et al., 1997).

### 3.2.3 - Aleatória estratificada

Nesta técnica os solos são agrupados em classes conhecidas também por extratos. A área é dividida em células de igual tamanho (normalmente quadrados), posteriormente são

escolhidos aleatoriamente pontos dentro de cada célula.

### 3.2.4 - Cluster

Agrupamentos de unidades individuais são selecionados aleatoriamente. Selecionando-se em seguida uma unidade do agrupamento. Utiliza-se esta técnica em indivíduos que ocorram naturalmente agrupados como população de ervas daninhas ou infestações de doenças.

### 3.2.5 - Aninhada ou múltiplos estágios

Apresenta-se como uma extensão da técnica de *cluster* onde os agrupamentos são subdivididos. Em seguida as subdivisões são aleatoriamente selecionadas e novamente subdivididas. Este processo só é interrompido quando a menor unidade é identificada. A Figura 3.2 ilustra o processo.

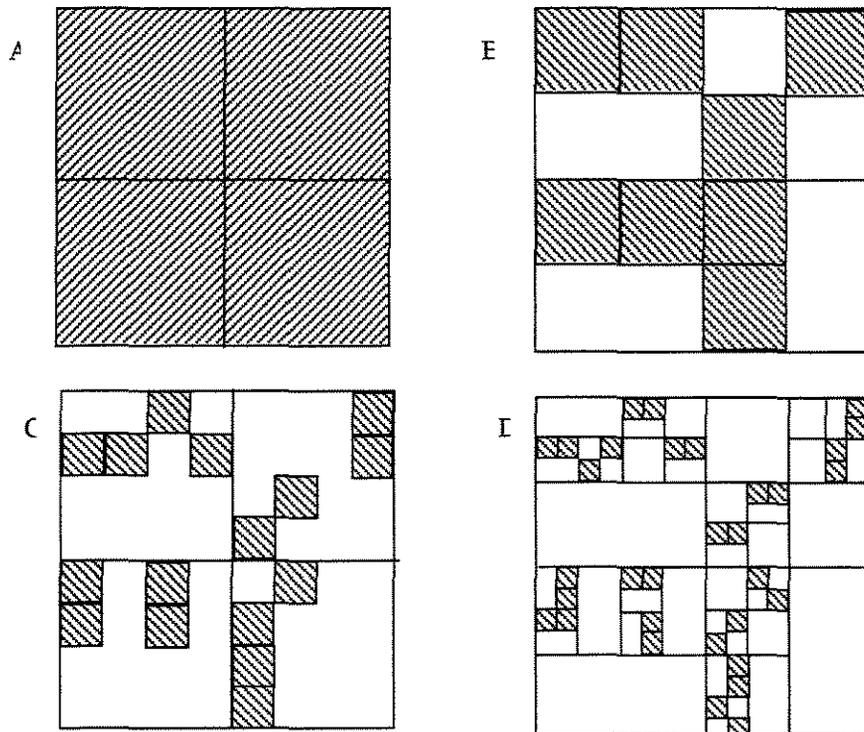


Figura 3.2 - Estágios para desenvolvimento de amostragem aninhada.

(Fonte: WOLLENHAUPT et al., 1997).

A principal desvantagem apresentada pela técnica relaciona-se com a cobertura irregular da área. Todavia, se o número de amostras selecionadas for suficiente, a população poderá ser estudada em diferentes escalas e níveis.

### 3.2.6 - Sistemática

Amostra em intervalos regulares de uma grade ou malha. A simplicidade é sua maior vantagem. Atualmente o GPS agiliza o trabalho do investigador no traçado de linhas ou localização de pontos.

Se existe um padrão ou periodicidade na população, os dados podem ter forte inclinação a incorporarem vício. Se o período de variação e direção forem conhecidos pode-se criar padrões de amostragem usando figuras geometricamente conhecidas como quadrados (Figura 3.3), triângulos, retângulos, hexágonos, etc.

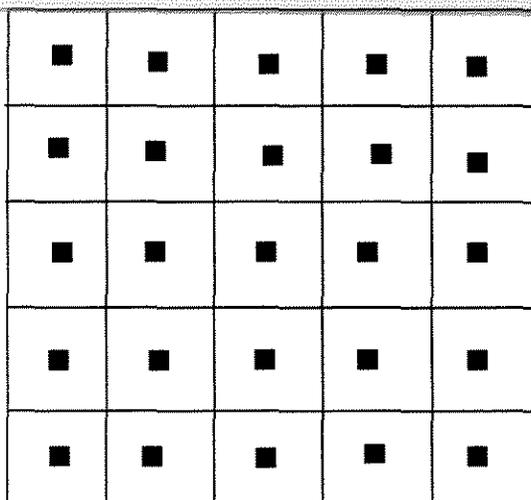


Figura 3.3 - Amostragem sistemática estratificada em grade quadrada. (Fonte: WOLLENHAUPT et al., 1997).

### 3.2.7 - Sistemática Estratificada não alinhada

Inicialmente o campo é intensivamente subdividido ou estratificado. A partir de um primeiro ponto na célula os outros são colocados de maneira a formar um conjunto de pontos espaçados igualmente, mas não alinhados em relação às linhas e colunas, como mostra a

Figura 3.4. Este esquema se mostrou superior a amostragem aleatória estratificada e é recomendada em atividades de manejo localizado.

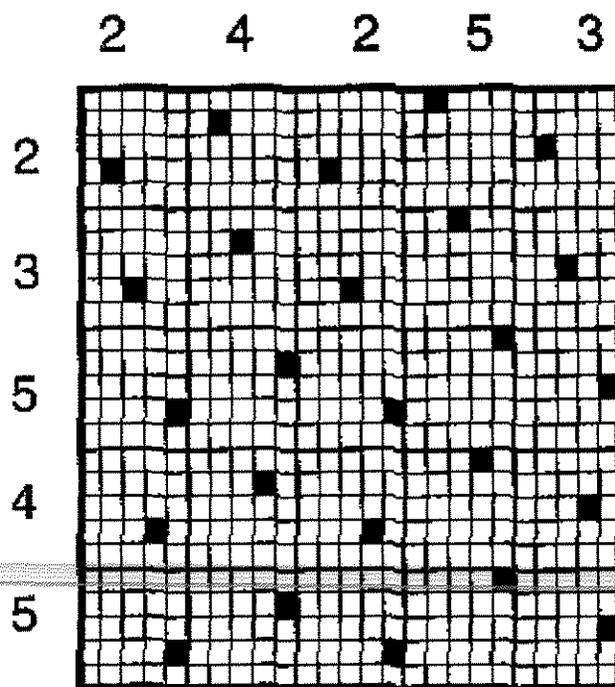


Figura 3.4 - Amostragem estratificada sistemática não alinhada. (Fonte: WOLLENHAUPT et al., 1997).

### 3.2.8 - Informação Complementar

De uma maneira geral, poucas são as propriedades ou áreas que regularmente armazenam informações sobre: operações de manejo, resultados de amostragens de solo e variações climáticas, as quais constituam um histórico de dados atualizado e recuperável para a tomada de decisão. Esta informação pode estar na forma de mapas de solos, fotografias aéreas, mapas de produtividade, mapas topográficos, históricos de manejo e/ou climáticos e experiências pessoais do local em estudo.

Estratégias de amostragem envolvem sobretudo o número de amostras a coletar, qual (quais) esquema(s) de amostragem utilizar e principalmente onde amostrar.

### 3.3 - Métodos de Interpolação

Interpolação é o processo utilizado para estimação de valores de alguma propriedade de interesse em pontos ou áreas onde não se processou a amostragem. De fato, os valores nos pontos amostrados próximos daqueles a serem estimados contribuirão sobremaneira para a composição da sua grandeza.

O mapeamento exige a interpolação sistemática de muitos valores sobre uma região de interesse. Várias são as técnicas de interpolação usadas. Dentre elas destaca-se as descritas por WOLLENHAUPT et al. (1997):

- métodos locais de tendência central - usa os pontos da vizinhança para interpolar o nó na grade;

- método do inverso da distância - média ponderada dos pontos vizinhos, o peso dado a cada observação é função da distância entre o ponto na grade e dos dados de localização;

- ~~Superfícies de tendência - Modelos são usados para descrever a variação espacial de~~  
uma propriedade usando regressão, equações polinomiais e mínimos quadrados.

O interpolador com embasamento geo-estatístico, denominado krigagem, parte do pressuposto de que as amostras mais próximas têm, frequentemente, maior influência do que aquelas mais afastadas ou pertencentes a agrupamentos (clusters). Esta abordagem explicita a dependência espacial entre valores de amostras vizinhas e é detalhada na seção 3.4.

Existem outros interpoladores exatos ou não que podem ser implementados com processamento computacional razoavelmente rápido, que também cumprem a finalidade de estimar valores em pontos não amostrados. Porém não há nenhum outro não tendencioso (sem viés) e com variância mínima. Estes interpoladores estão disponíveis, em maior frequência, nos sistemas de informação geográfica (SIG) para confecção de mapas de interesse.

Se alguma prevenção ou premeditação é colocada na seleção de uma estratégia de amostragem, os dados provavelmente conterão informações suficientes sobre a variação espacial da propriedade de interesse. A natureza desta variação e sua escala tornam-se fatores decisivos para a seleção do método de interpolação. Os métodos de interpolação usados na geo-estatística fornecem variância da estimativa em locais não amostrados. Em geral, com o aumento da variância dada pela krigagem incremental também, a incerteza de um valor estimado (WOLLENHAUPT et al., 1997).

## 3.4 - Hipóteses fundamentais da geo-estatística

### 3.4.1 Fundamentos

Adota-se neste trabalho as definições e convenções utilizadas no âmbito da estatística e na literatura corrente. Como a geo-estatística tem suas bases nos conceitos de funções e variáveis aleatórias, estes serão denotados por caracteres maiúsculos (Exemplo:  $Z(x_i)$  e  $Z(x_0)$ ), valores observados em minúsculos (Exemplo: valor da variável aleatória  $Z$  medido na posição  $x_k$  é  $z(x_k)$ ) e vetores em negrito ( $\mathbf{K}, \lambda$ ).

São duas as hipóteses que fundamentam a krigagem: a) Hipótese de estacionaridade de ordem 2 e intrínseca. VIEIRA (2000) explana de maneira clara e completa estes conceitos.

Inicialmente, considera-se um campo de área  $S$  que tenha um conjunto de valores amostrados  $\{z(x_i), i=1, n\}$ , onde a  $x_i$  associa-se uma posição no espaço e tempo e  $(x_i, y_i)$  representam pares de coordenadas. Considera-se cada valor amostrado da variável de interesse  $z(x_k)$  como uma realização de uma certa variável aleatória,  $Z(x_k)$ , para uma dada posição fixa  $x_k$ . A variável regionalizada  $Z(x_k)$ , para qualquer  $x_i$  dentro da área  $S$ , por sua vez, pode ser considerada uma realização do conjunto de variáveis aleatórias  $\{Z(x_i), \text{para qualquer } x_i \text{ dentro de } S\}$ .

Então, para se estimar valores para os locais não amostrados, aplica-se a restrição que permite que a variável regionalizada seja, necessariamente, estacionária estatisticamente. Formalmente, uma variável regionalizada é estacionária se os momentos estatísticos da variável aleatória  $Z(x_i+h)$  forem os mesmos para qualquer vetor  $h$  (módulo e direção). De acordo com o número  $k$  de momentos estatísticos que são constantes, a variável é chamada de estacionária de ordem  $k$ . Estacionariedade de ordem 2 é tudo que é requerido em geo-estatística (OLEA, 1975).

Supondo-se que a função aleatória  $Z(x_i)$  tenha valores esperados  $E\{Z(x_i)\} = m(x_i)$  e  $E\{Z(x_i+h)\} = m(x_i+h)$  e variâncias  $\text{VAR}\{Z(x_i)\}$  e  $\text{VAR}\{Z(x_i+h)\}$ , respectivamente, para os locais  $x_i$  e  $x_i+h$ , e qualquer vetor  $h$ , então, a covariância  $C(x_i, x_i+h)$  entre  $Z(x_i)$  e  $Z(x_i+h)$  é definida por VIEIRA (2000) como:

$$C(x_i, x_i + h) = E\{Z(x_i) Z(x_i + h)\} - m(x_i) m(x_i + h) \quad (1)$$

e o variograma  $2\gamma(x_i, x_i+h)$  é definido por:

$$2\gamma(x_i, x_i + h) = E \{Z(x_i) - Z(x_i + h)\}^2 \quad (2)$$

Na equação 1, a estacionaridade da covariância implica na estacionaridade da variância de  $Z(x_i)$ , dada por:

$$Var \{Z(x_i)\} = E \{Z(x_i) Z(x_i + 0) - m(x_i) m(x_i + 0)\} \quad (3)$$

$$Var \{Z(x_i)\} = E \{Z^2(x_i) - m^2(x_i)\} = C(x_i, x_i)$$

e a variância de  $Z(x_i+h)$  :

$$Var \{Z(x_i + h)\} = E \{Z^2(x_i + h) - m^2(x_i + h)\} = C(x_i + h, x_i + h) \quad (4)$$

Assim, existem três hipóteses de estacionaridade de uma função aleatória  $Z(x_i)$ , e pelo menos uma delas deve ser satisfeita antes de se fazer qualquer aplicação de geo-estatística.

### 3.4.2. - Hipótese de estacionaridade de ordem 2

Uma função aleatória  $Z(x_i)$  é estacionária de ordem 2 se:

$$E \{Z(x_i)\} = m \quad (5)$$

a) O valor esperado  $E\{Z(x_i)\}$  exista e não dependa da posição  $x$ , ou seja para qualquer  $x_i$  dentro da área  $S$ .

b) Para cada par de variáveis aleatórias,  $\{Z(x_i), Z(x_i+h)\}$ , a função covariância,  $C(h)$ , exista e seja função de  $h$ :

$$C(h) = E\{Z(x_i) Z(x_i+h)\} - m^2 \quad (6)$$

para qualquer  $x_i$  dentro da área S.

A hipótese de estacionaridade de ordem 2 supõe a existência de uma covariância  $C(h)$  e, então, de uma variância finita.

Se a hipótese de estacionaridade de ordem 2 puder ser satisfeita, a covariância  $C(h)$  e o variograma  $2\gamma(h)$  são ferramentas de igual peso para caracterizar a dependência espacial, ou seja de caracterizar a auto-correlação dos pares  $Z(x_i)$  e  $Z(x_i+h)$  separados pelo vetor  $h$ . A existência de estacionaridade dá a oportunidade de repetir um experimento mesmo que as amostras devam ser coletadas em pontos diferentes, porque todas são consideradas pertencentes a populações com os mesmos momentos estatísticos (VIEIRA, 2000).

### 3.4.3 - Hipótese intrínseca

Segundo VIEIRA (2000) essa hipótese requer apenas a existência e estacionaridade do variograma, sem nenhuma restrição quanto à existência de variância finita. Uma função aleatória é intrínseca quando, além de satisfazer a condição expressa na equação 5, a estacionariedade do primeiro momento estatístico, também o incremento  $\{Z(x_i) - Z(x_i+h)\}$  tem variância finita, e não depende de  $x_i$  para qualquer vetor  $h$ . Matematicamente, isso pode ser escrito:

$$Var ([Z(x_i) - Z(x_i + h)]) = E [Z(x_i) - Z(x_i + h)]^2 = 2\gamma(h) \quad (7)$$

para qualquer  $x_i$  dentro da área S.

A função  $\gamma(h)$  é o semivariograma. A razão para o prefixo "semi" é que a equação 7 pode ser escrita na forma:

$$\gamma(h) = 1 / 2 E[Z(x_i) - Z(x_i + h)]^2 \quad (8)$$

O fator 2 foi introduzido na definição do variograma, sendo mais usado o valor de  $\gamma(h)$  e não  $2\gamma(h)$ . Por ser menos restritiva a hipótese intrínseca é a mais usada em geo-estatística.

### 3.4.4 - Semivariograma

Constitui-se numa ferramenta importante para representação quantitativa da variabilidade de determinado fenômeno existente no espaço e tempo. Segundo VIEIRA (2000), esta ferramenta é mais adequada para medir a dependência espacial quando amostras forem coletadas em duas dimensões no campo e é necessário interpolar locais não amostrados com intuito de gerar mapas de isolinhas.

Ao variograma associa-se as noções de continuidade e de relacionamento devido a posição no espaço, segundo CLARK (1979).

Considera-se, então, duas variáveis regionais, X e Y, onde  $X = Z(x_i)$  e  $Y = Z(x_i + h)$  são, por exemplo, a quantidade de Carbono no solo, associadas a pares de coordenadas  $(x_1, y_1)$  e  $(x_2, y_2)$  respectivamente e separadas pelo vetor distância h (módulo e direção), conforme mostra a Figura 3.5.

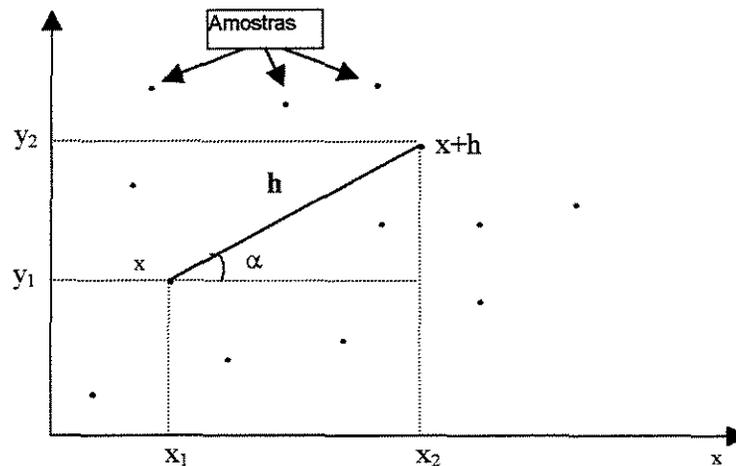


Figura 3.5 - Amostragem em duas dimensões (Fonte: CAMARGO, 1997).

O semivariograma pode ser definido como :

$$\gamma(h) = 1/2 E\{ Z(x_i) - Z(x_i + h) \}^2 \quad (9)$$

ou seja é a metade da esperança matemática do quadrado da diferença entre os valores dos pontos em um campo, distanciados do vetor distância  $h$ . A equação para estimação pode ser dada por:

$$\gamma^*(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(X_i) - Z(x_i + h)]^2 \quad (10)$$

onde  $N(h)$  é o número de pares de valores medidos  $Z(x_i)$ ,  $Z(x_i+h)$ , separados por um vetor e  $Z(x_i)$ ,  $Z(x_i+h)$ , são valores da  $i$ -ésima observação da variável regionalizada amostrados nos pontos  $x_i$  e  $x_i+h$ .

Pela observação do gráfico do semivariograma podemos identificar sua isotropia. Esta característica é identificada quando o semivariograma tem praticamente a mesma forma para qualquer direção do vetor  $h$ . Seu sentido contrário é chamado anisotropia onde por meio de raios de busca encontraram-se variações graduais em uma direção e irregulares e rápidas em outras. Na natureza a maioria dos fenômenos são anisotrópicos.

#### 3.4.4.1 - Parâmetros de um semivariograma ideal

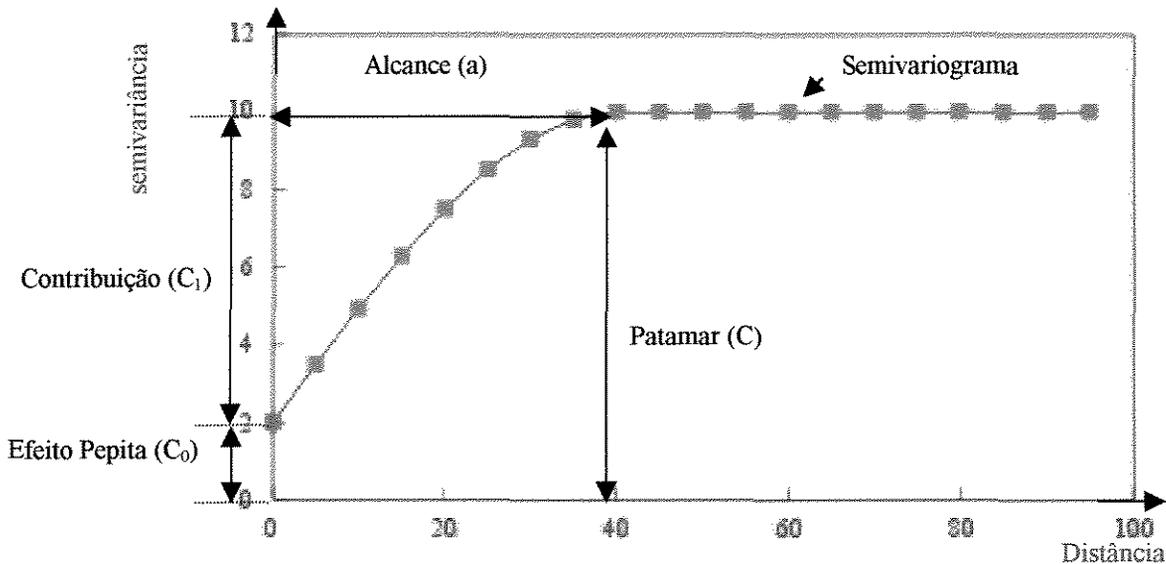


Figura 3.6 - Variograma típico e seus parâmetros. (Fonte: Modificado de VIEIRA et al., 1983).

A Figura 3.6 representa o gráfico de um semivariograma ideal, o qual são destacados os seguintes parâmetros:

- **Efeito Pepita** ( $C_0$ ) ou *Nugget Effect* - Quando o semivariograma tende a zero ( $\gamma(h) = 0$ ) o valor  $C_0$  é observado, revelando a descontinuidade do fenômeno para valores menores que a menor distância entre as amostras. ISAAKS e SRIVASTAVA (1989) atribuem as condições adversas na época de medição ou a variabilidade de menor escala não percebida pelo processo de amostragem.

- **Alcance** ( $a$ ) ou *Range*: distância na qual as amostras se encontram correlacionadas espacialmente.

- **Patamar** ( $C$ ) ou *Sill*: corresponde ao alcance ( $a$ ) no gráfico do semivariograma.

~~Considera-se que a partir deste ponto não exista mais dependência espacial entre as amostras. É aproximadamente igual a variância dos dados.~~

Intuitivamente espera-se que as diferenças  $\{Z(x_i) - Z(x_i+h)\}$  decresçam assim que  $h$ , a distância que os separa, decresça. É esperado que medições localizadas próximas sejam mais parecidas entre si do que aquelas separadas por grandes distâncias. Dessa maneira, é de se esperar que  $\gamma(h)$  aumente com a distância  $h$ .

Um passo importante na análise do semivariograma é a averiguação do comportamento do mesmo em várias direções, com objetivo de determinar a máxima e mínima continuidade da estrutura de variabilidade. CAMARGO (1997) diz que a direção máxima está relacionada com o semivariograma de maior alcance. De maneira similar, o semivariograma de menor alcance indica direção de mínima continuidade.

### 3.4.5 - Modelos Teóricos clássicos

Uma vez que o gráfico do semivariograma  $\gamma(h)$  versus  $h$  é calculado e construído deve-se ajustar os pontos discretos em uma função contínua que melhor represente o comportamento do fenômeno. Serão apresentados neste trabalho os mais usuais: esférico

(Sph), exponencial (Exp) e gaussiano (Gau), representados na Figura 3.7. Estes modelos são básicos e denominados por ISAAKS e SRIVASTAVA (1989) como isotrópicos e com patamar. Os modelos com patamar são identificados como transitivos e alguns destes atingem o patamar ( C ) assintoticamente. Existem também os modelos sem patamar, os quais nunca atingem o patamar, continuam aumentando indefinidamente com a distância. Adicionamos neste estudo os modelos Pentaesférico, com o comportamento semelhante ao modelo esférico e o *Hole Effect*, caracterizado por oscilações no seu patamar.

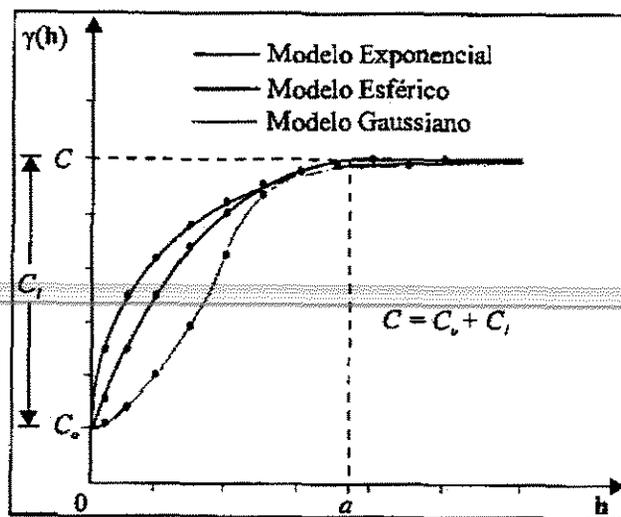


Figura 3.7 - Modelos teóricos de variograma. (Fonte: CAMARGO,1997).

### 3.4.5.1- Modelo Esférico

A equação deste modelo é representada por:

$$\begin{aligned}
 \text{Sph}(h) &= C_0 + C_1[1,5(h/a) - 0,5(h/a)^3] && \text{para } h > 0 \\
 \text{Sph}(h) &= C_0 + C_1 && \text{para } h > a \\
 \text{Sph}(h) &= 0 && \text{para } h = 0
 \end{aligned}
 \tag{11}$$

A vantagem de um modelo que atinge o seu patamar é que C pode ser considerado a variância amostral da medida de interesse.

### 3.4.5.2 - Modelo Exponencial

A equação deste modelo é dada por:

$$\begin{aligned} \text{Exp}(h) &= C_0 + C_1[1 - \exp(-h/a)] && \text{para } h > 0 \\ \text{Exp}(h) &= 0 && \text{para } h = 0 \end{aligned} \quad (12)$$

Este modelo atinge o patamar assintoticamente. O alcance prático é definido como a distância na qual o valor do modelo é 95% do patamar (ISAAKS e SRIVASTAVA, 1989).

### 3.4.5.3 - Modelo Gaussiano

O modelo Gaussiano é um modelo de transição usado com frequência quando o fenômeno a se estudar for extremamente contínuo (ISAAKS e SRIVASTAVA, 1989). A formulação deste modelo é dada por:

$$\begin{aligned} \text{Gau}(h) &= C_0 + C_1[1 - \exp(-(h/a)^2)] && \text{para } h > 0 \\ \text{Gau}(h) &= 0 && \text{para } h = 0 \end{aligned} \quad (13)$$

Semelhante aos modelos Gaussiano e Exponencial, alcança o patamar assintoticamente. Sua diferenciação em relação aos outros modelos é o comportamento parabólico próximo a origem.

### 3.4.5.4 - Modelo *Hole Effect*

A fórmula para modelo *Hole Effect* dada por JIAN et al. (1996) é:

$$\text{HE}(h) = C_0 + C_1[1 - (\sin(4.4934 h/a) / (4.4934h/a))] \quad (14)$$

O patamar desta curva caracteriza-se por várias oscilações na região do patamar.

### 3.4.5.5 - Modelo Pentaesférico

A fórmula para modelo pentaesférico dada por JIAN et al. (1996) é:

$$\text{PEN}(h) = C_0 + C_1[15/8 \cdot h/a - 5/4(h/a)^2 + 3/8(h/a)^5] \quad (15)$$

O formato da curva resultante deste modelo se assemelha ao modelo exponencial.

É esperado que fenômenos mais complexos não se expliquem somente com os modelos básicos, mas utilizando os chamados modelos aninhados ou que possuem estrutura entrelaçada. Estes modelos possuem mais de um patamar e alcance sendo necessário ajustar mais de um modelo ou um modelo para cada estrutura de variabilidade. MCBRATNEY et al. (1982) utilizaram modelos aninhados para estudar a variação de cobre e cobalto no solo.

---

### 3.4.6 - Krigagem

MATHERON (1963) homenageia o engenheiro de minas sul-africano Daniel G. Krige com o método de interpolação denominado Krigagem, o qual permite-se interpolar valores em qualquer posição no campo de estudo, sem tendência e com variância mínima.

O método da Krigagem assemelha-se ao método de interpolação por média móvel ponderada, diferindo dos métodos de interpolação linear simples e inverso do quadrado das distâncias por determinar os pesos segundo análise espacial, originada do semivariograma experimental. O estimador da Krigagem fornece estimativas não tendenciosas, ou seja, em média, a diferença entre valores estimados e verdadeiros para um mesmo ponto deve ser nula. Outra característica importante é variância mínima, significando que, embora possam existir diferenças ponto por ponto entre o valor estimado e o medido, essas diferenças devem ser mínimas.

Como estimador linear, a krigagem pode ser representada pelas combinações lineares ponderadas de dados medidos, ou uma média móvel que considera a estrutura de variabilidade da variável medida, expressa pelo semivariograma e pela localização dos valores conhecidos. Neste conceito, os pontos mais próximos da posição a ser interpolada contribuem com maior peso que aqueles mais distantes.

O estimador da krigagem apresenta a seguinte formulação:

$$Z^*(x_0) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (16)$$

Onde  $Z(x_i)$  é uma realização da função aleatória de estacionaridade de ordem 2,  $N$  é o número de valores medidos ou pontos vizinhos usados na predição,  $z(x_i)$ , envolvidos na estimativa, e  $\lambda_i$  são os pesos associados a cada valor medido,  $z(x_i)$ .

As formas mais usadas da Krigagem a Simples e Ordinária são detalhadas em JOURNAL(1989).

### 3.5 – Estudos de variabilidade de solos utilizando a técnica da geo-estatística

---

Considerando a variabilidade dos solos, estudos mais acurados podem ser realizados com o objetivo de: detectar áreas com diferentes potenciais produtivos, estudar poluição em águas subterrâneas e prevenir desperdícios no uso de fertilizantes e corretivos. A seguir, são apresentados, trabalhos que utilizam a técnica da geo-estatística no estudo de variabilidade de solos.

FLATMAN e YFANTIS (1984) preconizam a divisão do projeto de amostragem utilizando geo-estatística em duas etapas. No primeiro estágio, foi usado malha radial para coletar dados que permitiram, pela análise do semivariograma, definir os intervalos de influência e a orientação da estrutura de correlação (*survey*). No segundo estágio a amostragem é projetada/planejada, visando cobrir exaustivamente a área de interesse (*census*), no caso do estudo, contaminada por chumbo. Com relação a direção da amostragem CHUNG et al., (1995) recomendam que, para fins de fertilidade, as amostras sejam tomadas no sentido perpendicular às atividades de manejo agrícola.

ENTZ e CHANG (1991) avaliaram o impacto de três esquemas de amostragem. Utilizaram um bloco de solo de 1 m<sup>2</sup>, estudando-o por meio de variograma direcionais e krigagem. O estudo visava o conhecimento da variabilidade do atributo densidade. Foram utilizadas configurações em malha sistemática, aleatória e a combinação das duas. Os resultados mostraram que, para propostas práticas, os variogramas de ambos os esquemas de

amostragem apresentaram resultados semelhantes para a variabilidade espacial da densidade. Os autores recomendam o esquema aleatório estratificado em malha como o mais adequado para produzir variogramas em espaçamentos (*lags*) curtos. Indicam ainda que o mesmo mantém uniformemente amostras em toda a área, requerendo ainda menor esforço para a localização dos pontos de amostragem em relação a outros esquemas.

Visando contribuir para a quantificação da natureza espacial do solo MCBRATNEY e PRINGLE (1997) compilaram vários trabalhos acerca de atributos de solos estudados por meio da geo-estatística, reescalando seus respectivos semivariogramas. As propriedades estudadas foram: pH, argila (dag/kg), carbono (%), potássio (mg/kg), fósforo (mg/kg). A análise dos semivariogramas parece indicar que nutrientes do solo, com o potássio e fósforo são melhor descritos quando ajustados pelo modelo esférico. As outras propriedades foram descritas pelo modelo exponencial. O estudo contribui para a generalização dos modelos matemáticos utilizados no semivariograma experimental, porém cada caso de ajuste deve ser analisado individualmente, pois os atributos de solo sofrem influência de fatores adversos.

Uma estratégia para o projeto de uma rede de amostragem para estimação de variograma é proposta por RUSSO (1984). O método requer uma rede de amostragem inicial que pode ser sistemática ou aleatória. Para um dado tamanho de amostras  $N$  e usando um número constante de pares de ponto para cada classe de lags, o método utiliza como critério para seleção da localização dos pontos de amostragem a uniformidade dos valores de separação (ou do lag) do vetor  $h$ , dentro de uma dada classe de lag. O método fornece um conjunto de fatores de escala, que em última instância é usado para calcular a nova localização dos pontos de amostragem por processo iterativo.

Os projetos de amostragem em sua maioria baseiam-se em esquemas pré-determinados de grades regulares, nas suas várias configurações. As diferenças geológicas e geomórficas são aspectos poucos considerados para a formação de um mapa de delineamento ou para mapas convencionais de levantamentos como alerta ODEH et al. (1990).

### **3.6. Desenvolvimento de sistemas para tratamento geo-estatístico**

Várias são as metodologias utilizadas para desenvolvimento de um sistema computacional. A escolha apropriada depende de fatores como: complexidade do projeto,

tamanho e tipo do sistema, além do tamanho da equipe. Porém em todas as metodologias a fase que absorve maior percentual de custo e tempo no ciclo de vida de um novo software é o desenvolvimento. Uma vez que os requisitos do aplicativo tenham sido especificados, a fase de desenvolvimento inicia-se, obedecendo a quatro etapas distintas, conforme a Figura 3.8.

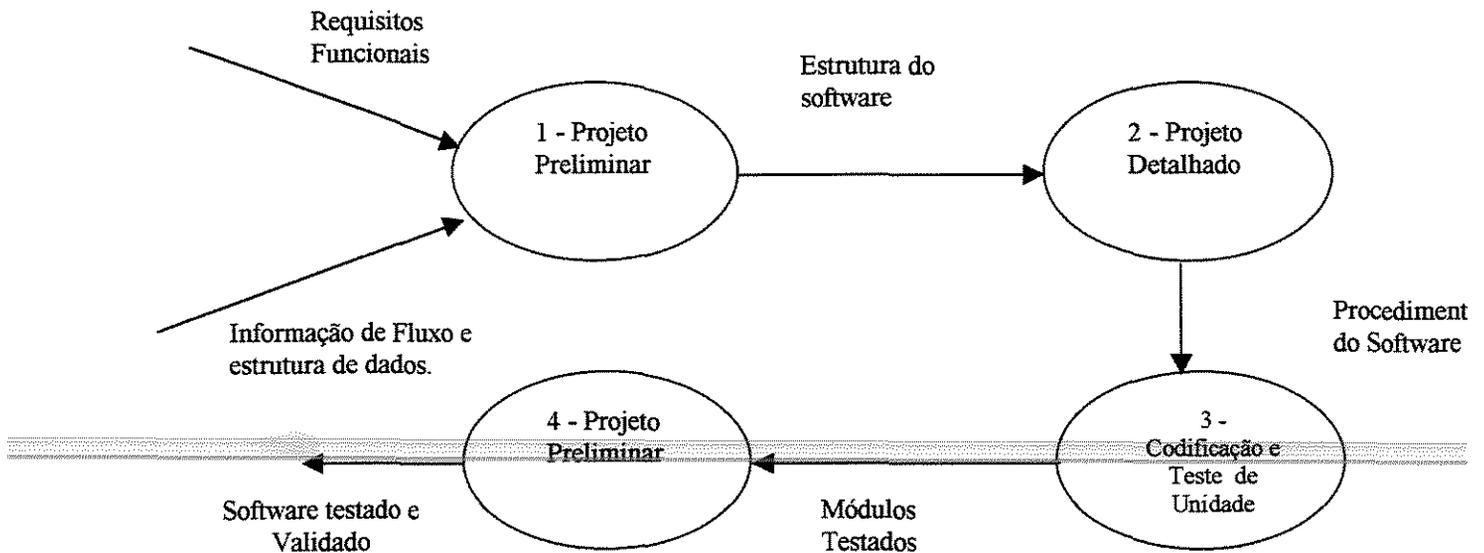


Figura 3.8 – Fases do Desenvolvimento de um software (Fonte: Modificado de PRESSMAN, 1982).

Considerando as rápidas e profundas transformações das necessidades do usuário de programas aplicativos, tem-se direcionado esforços na obtenção ou melhoria de técnicas que agilizem a geração de software dentro de padrões de qualidade estabelecidos.

Uma destas técnicas, chamada Reuso de software, tem como objetivo o uso de software existente em um novo contexto dentro de sistemas existentes ou em novas aplicações de forma a aumentar a produtividade da etapa de desenvolvimento e melhorar a qualidade do produto. São passíveis de Reuso de software todos os elementos do sistema, citamos aqueles mais usados: requisitos, projeto, código, documentação, procedimentos e manuais.

A atividade de reuso é dividida em algumas etapas (KANG et al., 1992):

- 1) Desenvolvimento um plano de reuso ou estratégia que disponibilize soluções para o problema;

- 2) Identificação da solução para o problema, seguindo o plano de reuso ou estratégia;
- 3) Aquisição ou modificação de componentes pré-definidos;
- 4) Integração dos componentes dentro do produto de software;
- 5) Avaliação do produto.

Seguindo esta ótica, vários sistemas computacionais foram construídos com o objetivo de implementar os conceitos da geo-estatística de maneira isolada ou integrada à outras ferramentas, utilizando-se de rotinas pré-existentes. Observa-se a tendência de oferta destas funcionalidades em Sistemas de Informação Geo-referenciadas (SIG).

Uma dos conjuntos mais completos de rotinas voltadas à geo-estatística denomina-se *Geostatistical Software LIBrary* (GSLIB) e foi implementada por DEUTSCH e JOURNAL (1998), na Universidade de Stanford. Constitui-se de uma biblioteca de programas escritos em linguagem de programação Fortran 77 padrão ANSI (*American National Standard Institute*). O padrão ANSI permite que o GSLIB seja executado em várias plataformas computacionais

A biblioteca GSLIB é um software não comercial, na qual onde não existem garantias e/ou suporte técnico formal. Porém, a biblioteca tem sido atualizada e incrementada de maneira sistemática, permitindo a geração de outros programas geo-estatísticos de domínio público. Um exemplo de aplicação é descrito por CAMARGO (1997). A versão para o sistema Windows, denominada WinGslib está disponível no site da [gslib.com](http://gslib.com) (GSLIB, 2001).

GSTAT é outro programa de computador de domínio público que permite a modelagem de variograma, interpolações por meio de krigagem e simulações condicionais ou não em dados com uma ou mais variáveis, através de ambiente de uma, duas ou três dimensões (PEBESMA, 1998).

Utilizando também a linguagem Fortran, VIEIRA et al. (1983) desenvolveram e validaram programas computacionais para cálculo de funções matemáticas associadas a geo-estatística. Os principais módulos são relacionados a seguir:

- AVAR e AVARIO - Calculam os semivariogramas médios, na ausência ou na presença de pontos perdidos respectivamente;
- DIVAR - Implementa o semivariograma direcional, para 4 principais direções;
- XGAMA - Calcula o semivariograma cruzado de duas variáveis distintas;

- TREND - Uma vez que se detecte a presença de tendência nos pontos do semivariogramas, esta é retirada por meio desse programa;
- JACK - Implementa o teste de *Jack-knifing*. Este teste é uma das principais ferramentas utilizadas, além do exame do semivariograma, para o conhecimento indireto da estacionaridade dos dados;
- KRIGE e COKRI - Usados em operações de krigagem e Co-krigagem.

As rotinas desenvolvidas por VIEIRA et al. (1983) são executadas por meio de chamadas diretas ao ambiente DOS. Com a entrada de parâmetros específicos são gerados arquivos de saídas em formato ASCII. Os programas são manuseados de maneira modular.

Como todo o processo de análise geo-estatística envolve incerteza e subjetividade, faz-se necessário avaliar o desempenho da estimativa em pontos onde os valores são conhecidos. Esta avaliação do erro da estimativa pode ser realizada pelo procedimento conhecido como *Jack-Knifing* ou autovalidação.

Existem outras formas de validação da estimativa, como o estudo dos resíduos (diferença entre os valores medidos e estimados), porém o *Jack-Knifing* é um processo mais completo, pouco implementado pelas bibliotecas geo-estatísticas mas disponível em VIEIRA et al.(1983).

Observa-se a tendência dos programas computacionais, com abordagem científica, em incorporarem recursos das chamadas linguagens visuais como Delphi (INPRISE CORPORATION, 1999) e C++ Builder. Tais linguagens permitem a construção de ambientes integrados para resolução de problemas.

CORNELL e STRAIN (1996) citam algumas das vantagens do ambiente Delphi, voltado para criação de interfaces gráficas de usuário (GUI – Graphic User Interface), em relação as linguagens clássicas orientadas a objeto. São elas:

- Rapidez no desenvolvimento das interfaces pela facilidade de uso e disponibilidade dos componentes visuais (botões de comando, caixa de diálogo e edição, etc) ou não (controle de mensagens de erro, acesso a banco de dados, etc);
- possibilidade de construir componentes reutilizáveis segundo o paradigma da programação orientada a objetos.

O capítulo discorreu sobre os principais projetos de amostragem, utilizados no estudo dos atributos de campo. Projetos que envolvem maior complexidade fazem uso atualmente de equipamentos precisos de localização como o GPS. Porém a etapa de projeto de amostragens carece de maior apoio computacional.

A geo-estatística é base deste estudo e tem seus conceitos descritos e comentados neste capítulo. O uso de seu interpolador (Krigagem) vem aumentando continuamente conforme demonstram os estudos descritos em encontros e publicações técnico-científicas nas mais diversas áreas e mais precisamente na Agricultura de precisão.

O desenvolvimento de sistemas que implementam a técnica da geo-estatística se modernizou, incluindo técnicas de qualidade que garantam a acurácia dos resultados e a facilidade de uso por parte dos usuários.

O próximo capítulo apresentará as interfaces correspondentes ao sistema desenvolvido para estudo da variabilidade de atributos de solos por meio de análise geo-estatística.

## CAPÍTULO 4

### MATERIAL E MÉTODOS

Este capítulo tem como objetivo apresentar os módulos do sistema para análise geoestatística. Inicia-se com uma breve descrição dos dados de Waynick e Sharp citados por Vieira (2000), o qual subsidiou testes no sistema. Posteriormente, são descritos funcionalmente os módulos relativos às operações de modelagem do semivariograma, autovalidação e krigagem, seguidos da apresentação de suas respectivas interfaces.

#### **4.1 - Descrição da área estudada**

Os dados utilizados originaram-se dos estudos de WAYNICK e SHARP (1919) e foram posteriormente utilizados por Vieira (2000) quando da revisão de conceitos geoestatísticos.

As observações referem-se a dois campos na Califórnia, EUA. Foram tomadas 100 amostras em cada um dos dois campos, um deles na chamada Fazenda Universitária de Davis (solo franco argiloso) e outro na cidade de Oakley (solo arenoso). Os campos possuíam declividade praticamente nula.

O estudo de caso apresentado neste trabalho estuda a análise de Carbono, Nitrogênio e da relação Carbono/Nitrogênio numa malha retangular com diferentes densidades nos cantos e na parte central da malha. Segundo os autores dos dados, a forma de amostragem visava verificar o efeito da distância entre amostras.

Os campos eram aproximadamente uniformes e encontravam-se sem vegetação na época da amostragem. As amostras foram coletadas com um trado de 0,0762 metros de diâmetro e encaminhadas ao laboratório para análise de nitrogênio e carbono totais.

A distribuição espacial dos pontos amostrados é ilustrada pela Figura 4.1. A malha de amostragem possuía faces de 63,693 metros por 81,891 metros.

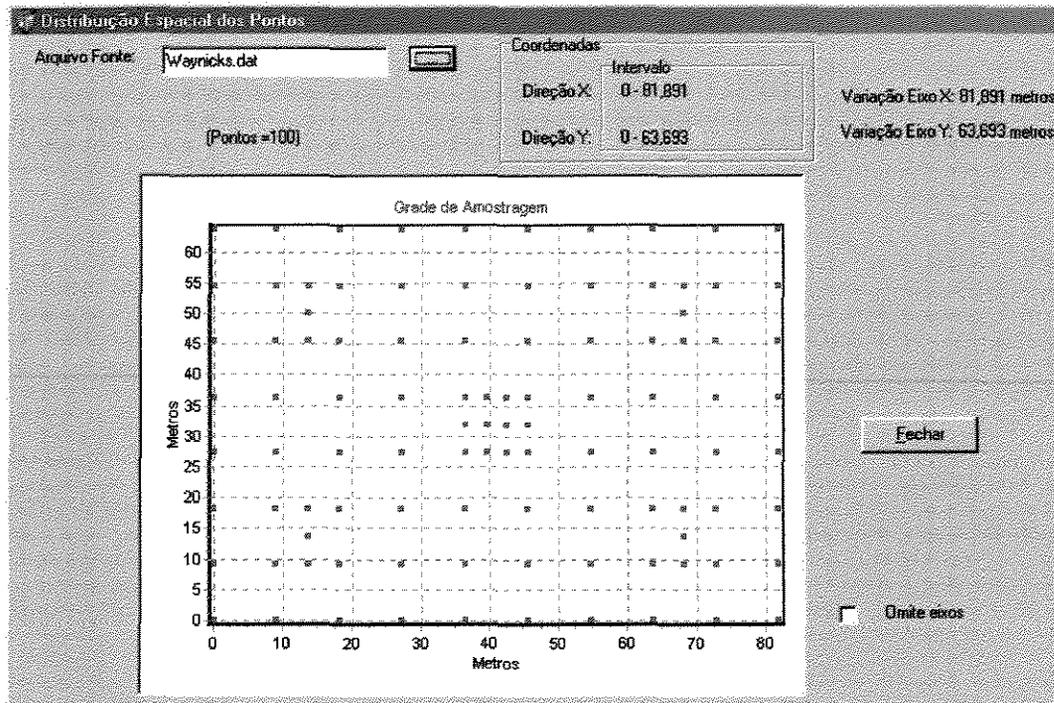


Figura 4.1 - Distribuição espacial dos pontos amostrados nos dois campos.

Os dados de WAYNICK e SHARP (1919) foram amostrados originalmente em pés. Neste trabalho estas medidas foram convertidas para metros. A distância entre as amostras variava entre 3,033 (10 pés); 4,5494 (15 pés) e 9,099 (30 pés) metros.

As unidades resultantes de Carbono e Nitrogênio totais são dadas em grama/kg e a relação (C/N) em porcentagem.

A Figura 4.2 representa a estrutura do arquivo de dados **Waynicks.dat** estudado bem como uma porção inicial de seus valores.

```

Waynick&Sharp(1919) Sidney R. Vieira California 1919
100,6
CDaviNDaviCOakyNOakyC-NDaC-NOk
g/kg g/kg g/kg % %
0 0 11.67 1.04 6.24 0.42 11.221 14.857
0 9.099 10.48 0.86 4.17 0.32 12.186 13.031
0 18.198 9.58 0.8 3.65 0.27 11.975 13.519
0 27.297 10.69 0.92 1.82 0.22 11.62 8.273
0 36.396 10.71 0.99 3.55 0.29 10.818 12.241
0 45.495 11.32 0.98 3.83 0.26 11.551 14.731
0 54.594 11.24 1.07 2.79 0.27 10.505 10.333
0 63.693 11.95 1.17 4.04 0.31 10.214 13.032
9.099 0 10.59 1.05 5.65 0.42 10.086 13.452
9.099 9.099 10.84 0.93 4.2 0.32 11.656 13.125
9.099 18.198 10.77 0.86 3.3 0.24 12.523 13.75
9.099 27.297 9.97 0.91 3.61 0.24 10.956 15.042
9.099 36.396 8.96 0.82 3.49 0.31 10.927 11.258
9.099 45.495 11.98 1.05 3.14 0.32 11.41 9.813
9.099 54.594 12.18 1.14 4.75 0.33 10.684 14.394
9.099 63.693 11.57 1.08 3.3 0.42 10.713 7.857
18.198 0 9.7 0.94 4.99 0.42 10.319 11.881

```

Figura 4.2 – Estrutura do arquivo de dados Waynicks.dat.

A primeira linha deste arquivo corresponde a autoria do projeto, local e data de amostragem. A segunda elenca o total de registros e o número de variáveis em estudo, separados por vírgula. Neste estudo, são seis (6) as variáveis e suas iniciais estão na seguinte ordem: Carbono e Nitrogênio em Davis (CDavi, NDavi), Carbono e Nitrogênio em Oakley (COaky, NOaky), Relação C/N em Davis (C-NDa) e Relação C/N em Oakley (C-NOk).

A linha seguinte enumera o nome das variáveis, citadas no parágrafo anterior e quarta linha, as unidades de cada variável. Tanto o nome quanto a unidade das variáveis ocupam individualmente até 5 caracteres.

Após este cabeçalho são associadas as coordenadas geográficas dos pontos de amostragem e os respectivos valores de cada variável. No caso em estudo, as coordenadas são cartesianas, dadas em metros, com ponto de origem em 0,0.

## 4.2 - Desenvolvimento Computacional

A metodologia empregada no desenvolvimento deste sistema baseia-se nos princípios

e paradigmas da Engenharia de Software e nas etapas do Ciclo de Vida Tradicional de desenvolvimento de sistemas (PRESSMAN, 1982) associada a técnica de Reuso de software. O sistema foi desenvolvido para o ambiente Windows e utiliza, principalmente, duas ferramentas computacionais, a saber:

- Rotinas geo-estatísticas previamente desenvolvidas por VIEIRA et al. (1983) – Conjunto de programas executáveis escritos em linguagem Fortran 77 e executadas modularmente no ambiente DOS. A disponibilidade do código fonte e a rotina de autovalidação foram fatores importantes na escolha desta biblioteca como parte deste estudo;
- Ambiente e linguagem de programação Borland Delphi. A versão 5 deste sistema foi utilizada.

---

A estrutura do sistema, incluindo seus principais módulos, é fornecida pela Figura

4.3. O sistema proposto tem como objetivo a análise em duas dimensões (eixos X e Y) de dados espacialmente distribuídos.

Uma forte motivação para o trabalho foi o manuseio restrito que o ambiente DOS impõe aos usuários deste tipo de aplicação. Um exemplo desta restrição é a não consistência dos parâmetros de entrada para determinada etapa. Uma vez que valores incorretos ou errôneos são usados, o ambiente DOS ou as rotinas apresentam mensagens de erro pouco explicativas, obrigando o usuário a um cuidado redobrado na elaboração destes.

Outro fator de motivação foi falta de integração entre as rotinas, advinda da própria natureza modular das mesmas. Um exemplo, para esta situação, ocorre na etapa de tratamento da estacionaridade dos dados do semivariograma experimental, onde dois módulos específicos devem ser aplicados aos dados. No ambiente DOS o usuário teria de ativar o módulo de retirada de tendência e posteriormente, executar o módulo para cálculo do semivariograma médio para o novo arquivo de valores de resíduos. No sistema implementado estas operações são executadas automaticamente.

A abordagem aplicada a este sistema resultou na reutilização integral da biblioteca de rotinas geo-estatística na sua forma original por meio de seus arquivos executáveis. Esta abordagem deveu-se ao cronograma restrito, o qual inviabilizaria a tradução dos

procedimentos geo-estatístico em Fortran para a linguagem de programação Delphi com a devida acurácia.

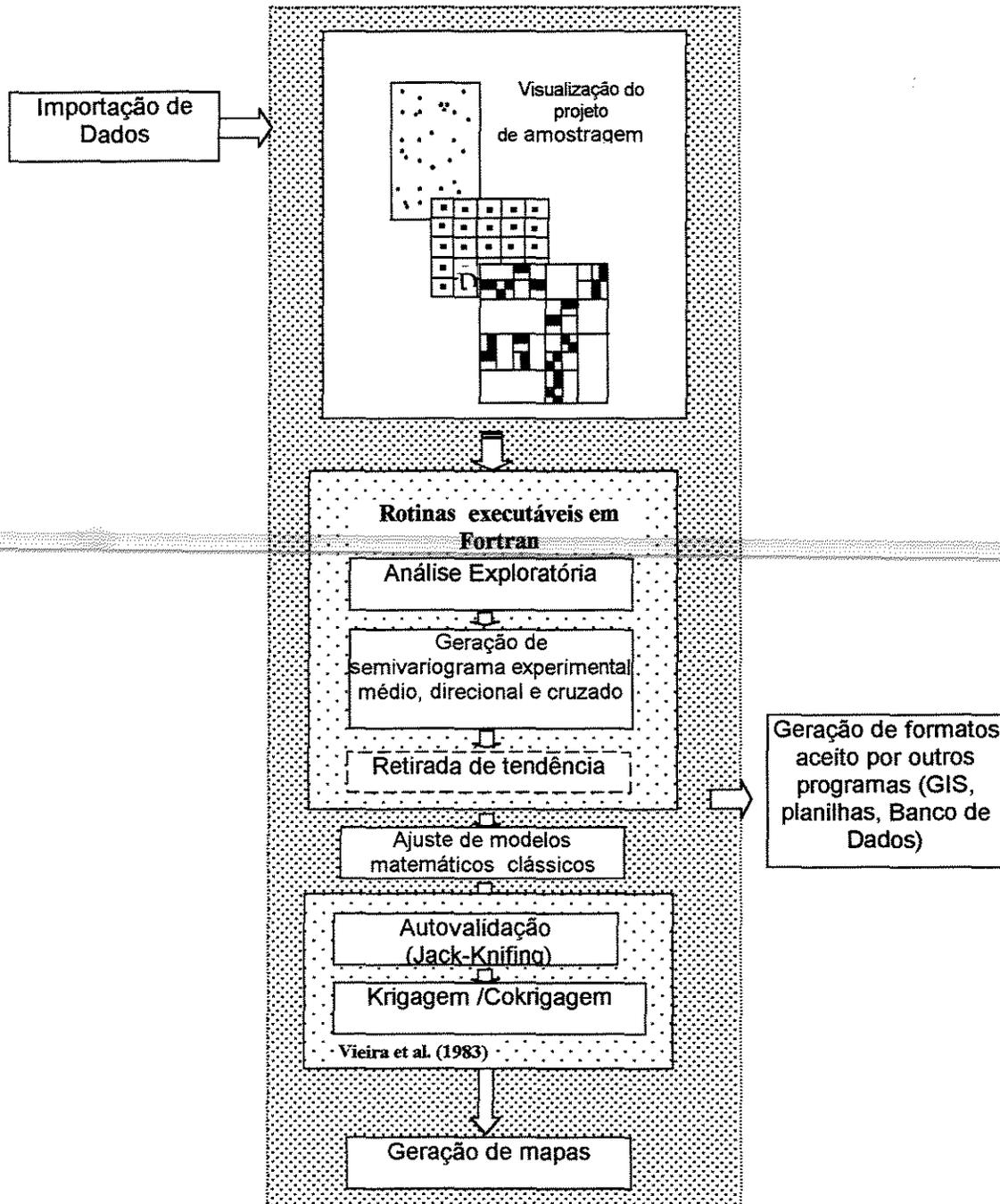


Figura 4.3 - Diagrama de funcionalidades do sistema.

Optou-se pela implementação de interfaces que permitissem a entrada de dados de maneira amigável e realizassem o respectivo processo de consistência dos dados.

O sistema gera posteriormente o arquivo de parâmetros e de maneira transparente executa a rotina geo-estatística apropriada. Uma vez que o nome do arquivo resultante é conhecido, este é lido e tratado de maneira adequada pelas interfaces gráficas do sistema. Um exemplo da interação entre módulos implementados pelo sistema e os arquivos textos gerados pela execução das rotinas Fortran é mostrado na Figura 4.4.

Outro fator importante para o auxílio do tratamento dos dados é a formatação fixa dos arquivos de saída. Tal fato permite a elaboração de rotinas automáticas para leitura e tratamento dos arquivos.

O procedimento análogo no ambiente DOS, seria execução de vários aplicativos pelo usuário, além das próprias rotinas, tais como: editores de texto e planilhas para a obtenção dos mesmos resultados.

O sistema foi desenvolvido sob a plataforma operacional do Microsoft Windows 98, não foram efetuados testes para outros sistemas operacionais. As mensagens de erros retornadas pelo sistema são divididas em duas classes:

- proveniente do sistema em si, resultado de consistência de dados ou impossibilidade de execução de determinado módulo;
- mensagens dos sistemas operacionais Windows ou DOS.

Os detalhes dos módulos implementados pelo sistema são descritos no próximo item.

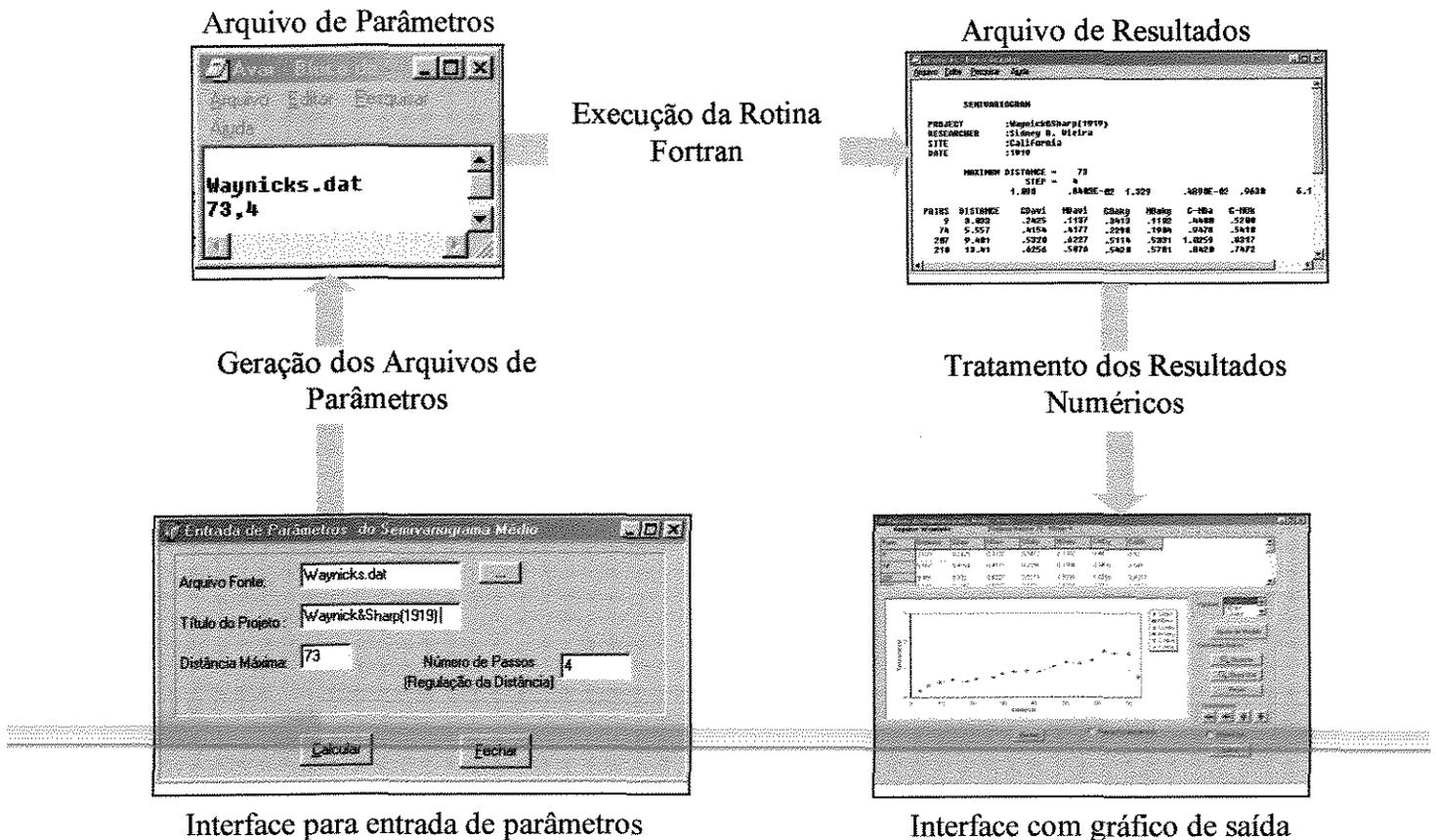


Figura 4.4 – Esquema genérico de interação entre módulos dos sistemas e arquivos textos gerados pelas rotinas Fortran.

### 4.3 – DESCRIÇÃO DOS MÓDULOS DO SISTEMA

#### 4.3.1 - Módulo de análise exploratória

O início do todo estudo geo-estatístico consiste em analisar os parâmetros apresentados pela estatística descritiva, a qual supõe a análise de médias, valores máximos e mínimos, matrizes de correlação e covariância e indicadores de distribuição normal de frequência. Este passo permite um melhor conhecimento do conjunto de dados. Observando a matriz de correlação, por exemplo, pode-se indicar aquelas variáveis que são fortes candidatas a análise conjunta, por meio semivariograma cruzado.

A geração do arquivo de análises clássicas inicia-se quando a opção *Estatística descritiva* incluída no quadro *Geo-estatística* do menu principal é ativada (Figura 4.5). Nesta versão do sistema são realizadas somente estatísticas univariadas.

O módulo permite a padronização dos valores de dados perdidos. Nesta situação os pontos onde por algum motivo não se conseguiu gerar ou coletar os valores específicos, associa-se um valor único (Exemplo: -99,99) e fictício de forma que o sistema reconheça este valor não o incluindo no cálculo das estatísticas.

A opção para cálculo da matriz de correlação e covariância entre as variáveis também está disponível nesta interface.

É permitida a seleção de variáveis para fins de transformação logarítmica. Esta operação permite uma visão mais próxima da distribuição normal dos dados.

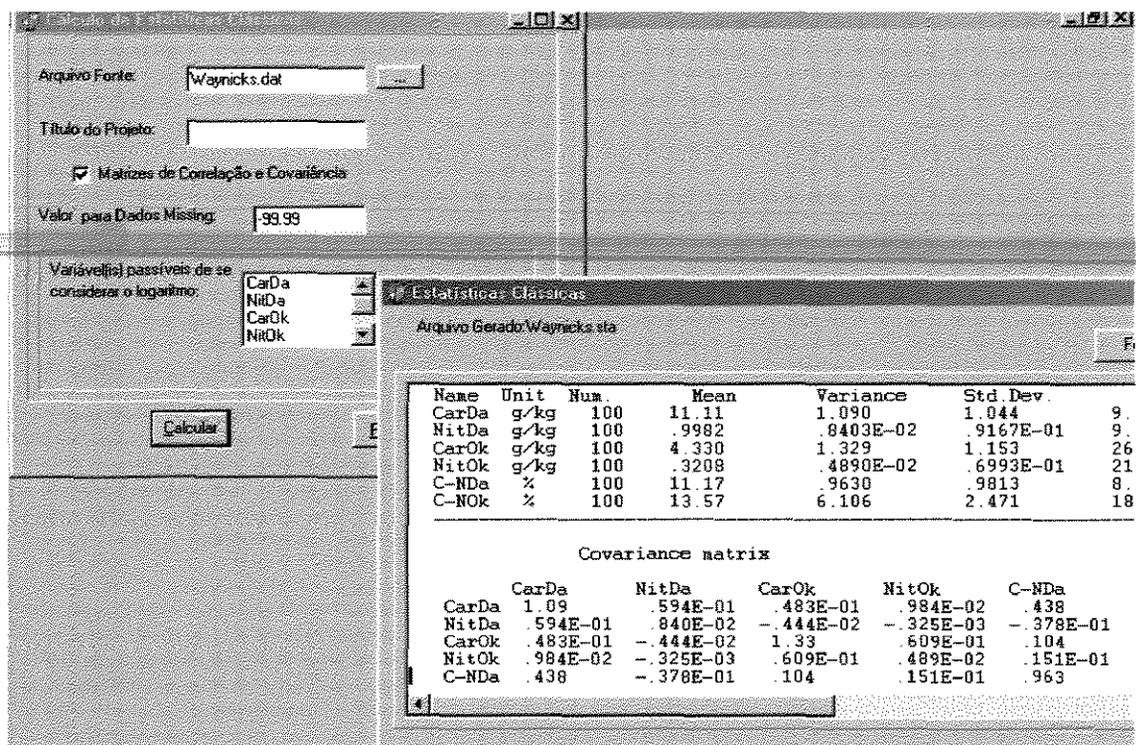


Figura 4.5 – Interface para geração de estatísticas clássicas e janela de resultados.

### 4.3.2 - Módulo para modelagem do semivariograma experimental

Na geo-estatística, uma das etapas mais críticas para uma boa representação da estrutura espacial é a modelagem do semivariograma. Esta se inicia com a geração de um modelo experimental, fruto das interações entre os pares de pontos amostrados. Descreve-se inicialmente o processo de um semivariograma médio, definido como aquele que considera a

interação em todas as direções. O semivariograma que estuda direções específicas é descrito na seção 4.3.8.

As etapas para análise do semivariograma construídas neste sistema seguem a mesma filosofia daquelas relatadas por (CAMARGO, 1997). São elas:

1. Geração do semivariograma experimental;
2. Ajuste do semivariograma a uma família de modelos matemáticos;
3. Existindo não estacionaridade dos dados, aplicação de superfície matemática para retirada de tendência;
4. Validação dos modelos pela técnica de *Jack-Knifing*.

A geração do semivariograma médio experimental inicia-se quando a opção *Cálculo do Semivariograma* incluída no quadro *Geo-estatística* do menu principal é ativada. Esta opção abre a interface mostrada na Figura 4.6, a qual solicita os seguintes parâmetros: Nome do arquivo de dados, título do projeto, distância máxima e número de passos.

As considerações no âmbito deste sistema serão pertinentes a malhas com configuração quadrada ou retangular.

A distância máxima utilizada neste estudo foi identificada inicialmente como a faixa que abrange 70 a 80% da diagonal da área estudada. A *largura dos passos* estabelece o chamado “amaciamento” da curva, onde a escolha de largura alta resulta num menor número de classes de distâncias de pontos, gerando uma curva de semivariograma sem grandes oscilações e de mais fácil interpretação.

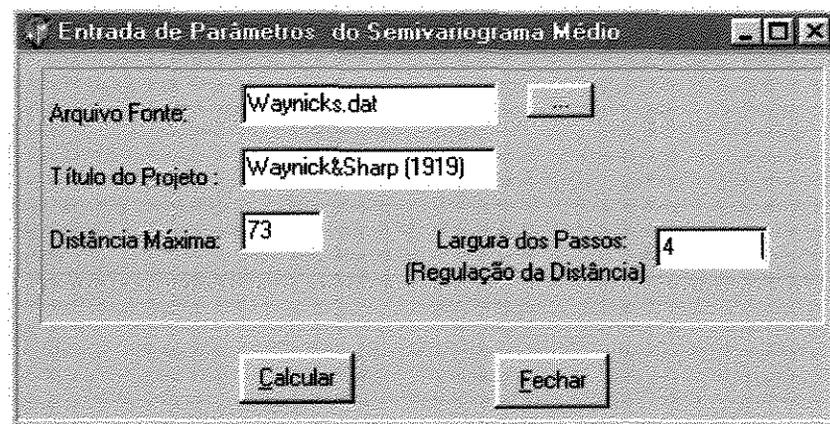


Figura 4.6 – Interface de entrada dos parâmetros para cálculo do semivariograma médio.

Com os parâmetros preenchidos, o cálculo do semivariograma experimental é realizado clicando o botão *Calcular*.

São três as mensagens de erro geradas por esta interface. Se os campos *Distância Máxima* e *Largura dos Passos* são deixados em branco, as respectivas mensagens são criadas: Distância máxima inválida e Nstep Inválido. No caso do campo *Arquivo Fonte*, o número de caracteres do nome do arquivo de dados deve ser de oito letras seguidas do caracter ponto mais três letras da extensão. Se o Título do projeto diferir desta formação a mensagem a seguir é apresentada: “Deve conter 8 dígitos para o arquivo seguido da extensão .dat”

Internamente, ao se acionar o botão *Calcular*, o sistema executará o módulo geoestatístico *Avar.exe* para produzir os resultados numéricos, segundo os parâmetros desta interface. O arquivo de saída é do tipo texto e segue uma regra específica para formação de seu nome. A identificação do arquivo de saída adota a seguinte formação: Nome inicial do arquivo de dados (8 caracteres) + extensão do arquivo do tipo *.GAM* (Ex. Waynicks.gam).  
Este arquivo geralmente localiza-se no mesmo diretório do arquivo de dados.

O resultado é a interface dada pela Figura 4.7.

Os valores numéricos dos pares, classes de distâncias, e a semivariância de cada variável são ilustrados na planilha situada na parte superior da interface da Figura 4.7. A Lista de variáveis é preenchida automaticamente, na parte superior direita da interface. Ao se clicar sobre o nome de alguma variável a respectiva curva do semivariograma é atualizada no gráfico da interface. Para proceder ao desenho múltiplo de semivariograma deve-se clicar nos demais nomes de variáveis com a tecla **Ctrl** pressionada.

Este procedimento desenha inicialmente o semivariograma escalonado. Este tipo de semivariograma é obtido dividindo-se cada valor de semivariância pela respectiva variância do atributo, resultando em grandezas adimensionais com valores máximos próximos à unidade. A opção *Retirar Escalonamento* redesenha o gráfico multiplicando cada ponto pelo respectivo valor da variância.

A vantagem da análise dos semivariogramas escalonados consiste em comparar o comportamento espacial de vários atributos numa mesma escala. Opções de *Zoom In*, *Zoom Out* e deslocamento do gráfico melhoram a análise e identificação dos parâmetros essenciais ao semivariograma (efeito pepita, alcance e patamar) dentre outras facilidades.

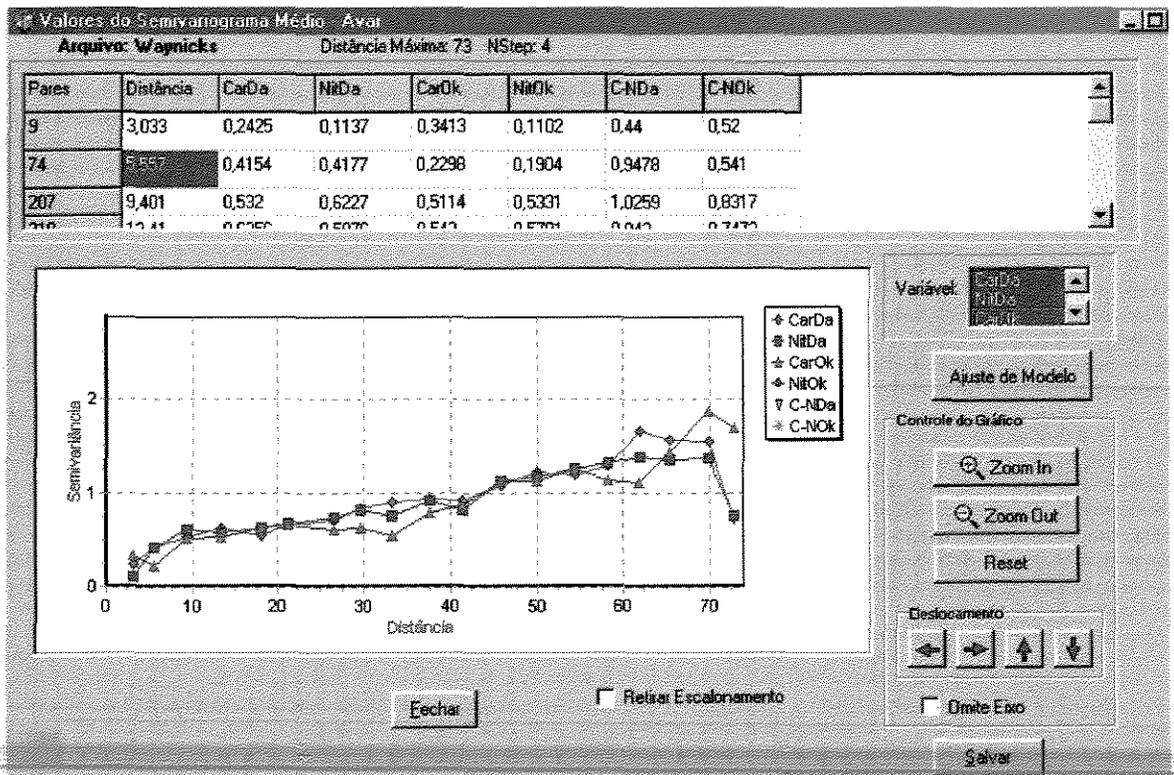


Figura 4.7 – Interface do Semivariograma médio experimental.

Quando o conjunto de dados possuir dados perdidos, ou seja, existe a localização do ponto mas o valor do atributo não é conhecido e expresso por um número fictício (ex: -99.99, -1.00, etc), a biblioteca geo-estatística disponibiliza o módulo Avario.exe. A interface com os parâmetros de entrada é mostrada pela Figura 4.8. O programa Avario permite o escalonamento do semivariograma. Uma vez selecionada esta opção, as semivariâncias resultantes serão divididas pela correspondente variância.

Pode-se excluir variáveis para cálculo do semivariograma, neste módulo, selecionando a(s) variável(is) no ListBox *Variáveis não consideradas no cálculo*. A opção *Considera Logaritmo no Cálculo*, aplica a função logarítmica aos resultados da coluna desejada.

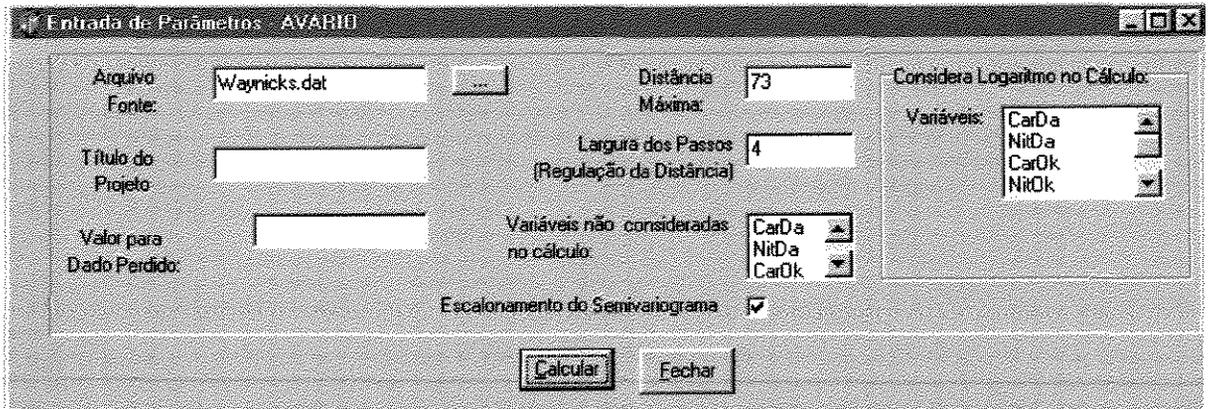


Figura 4.8 – Interface de entrada dos parâmetros para cálculo do semivariograma médio, quando existirem dados perdidos.

### 4.3.3 - Módulo de ajuste do semivariograma aos modelos matemáticos

Este módulo permite o ajuste do modelo teórico ao semivariograma experimental de maneira visual, onde todos os parâmetros são definidos por inspeção.

Neste tipo de ajuste, o conhecimento ou afinidade com o fenômeno estudado bem como informações auxiliares influenciarão na definição dos parâmetros do modelo.

A interface mostrada pela Figura 4.9 é ativada quando se pressiona o botão *Ajuste de Modelo*, abaixo da lista de variáveis, presente na Figura 4.7.

São 5 (cinco) os modelos matemáticos implementados nesta interface. São eles: Esférico, exponencial, gaussiano, *Hole effect* e pentaesférico.

Na atual versão do sistema, utiliza-se uma estrutura apenas para representação do modelo matemático do semivariograma, ou seja, estruturas entrelaçadas não são possíveis de serem descritas.

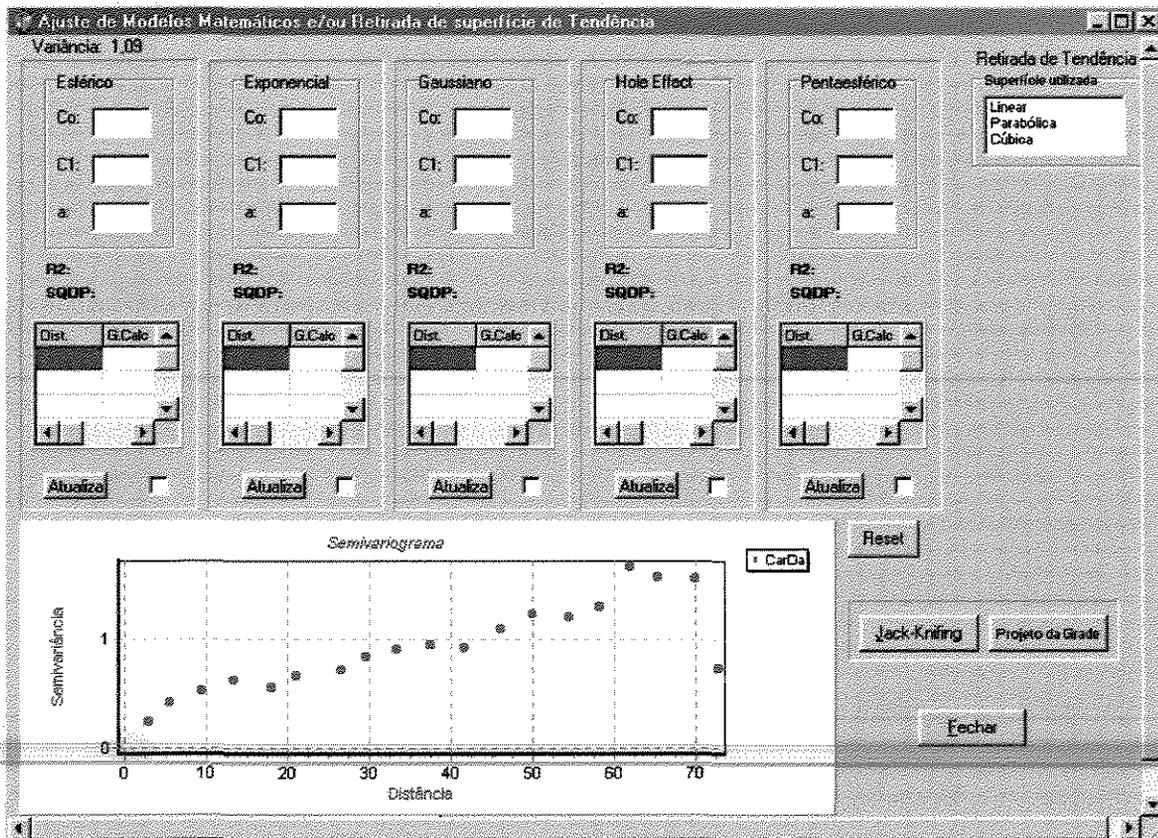


Figura 4.9 – Interface para ajuste de modelos matemáticos teóricos.

Tendo como base o semivariograma escalonado do atributo, a escolha dos parâmetros é realizada. Ao se pressionar o botão *Atualiza*, o modelo é sobreposto ao desenho original do semivariograma. Permite-se que todos modelos sejam sobrepostos ao original, possibilitando opções de escolha com vistas ao melhor. As planilhas abaixo dos parâmetros de cada modelo contêm os valores das classes de distância (Dist.) e o valor teórico de semivariância de cada modelo (G. Calc.). O quadrado ao lado do botão *Atualiza* ativa ou desativa no gráfico o desenho do modelo, de forma a prover melhor comparação visual entre os modelos.

Para auxílio na escolha do melhor modelo, dois importantes parâmetros são calculados. São eles: Coeficiente de Determinação ( $R^2$ ) e Soma dos Quadrados dos Desvios Ponderados (SQDP).

Pode-se calcular o coeficiente de determinação pela seguinte fórmula:

$$R^2 = \left( 1 - \frac{\sum_{i=1}^{C(h)} (\gamma(h)_{\text{exp}} - \gamma(h)_{\text{mod}})^2}{\sum_{i=1}^{C(h)} (\gamma(h)_{\text{exp}} - \gamma(h)_{\text{orig médio}})^2} \right) \quad (17)$$

$$\text{onde } \gamma(h)_{\text{exp médio}} = \frac{\sum_{i=1}^{C(h)} \gamma(h)_{\text{exp}}}{N}$$

O cálculo da Soma dos Quadrados dos Desvios Ponderados é obtida pela seguinte fórmula:

$$SQDP = \sum_{i=1}^{C(h)} \left( \frac{P_i}{\sum_{i=1}^{C(h)} P_i} * (\gamma(h)_{\text{mod}_i} - \gamma(h)_{\text{exp}_i})^2 \right) \quad (18)$$

onde:

$P_i$  correspondem aos números de pares para cada classe de distância;

$\gamma(h)_{\text{mod}_i}$  é a semivariância do modelo matemático;

$\gamma(h)_{\text{exp}_i}$  corresponde a semivariância do semivariograma experimental;

$C(h)$  representa o número de classes de distâncias;

$\gamma(h)_{\text{orig médio}}$  é a semivariância média do semivariograma experimental.

Ressaltamos que o cálculo de  $R^2$  e SQDP é realizado sobre todos os pontos do semivariograma experimental e do modelo. Na comparação entre os modelos procura-se valores de  $R^2$  próximos à unidade e baixos valores de SQDP.

#### 4.3.4 – Módulo para retirada de tendência

Em situações onde o semivariograma cresce indefinidamente para todos os valores de  $h$  calculados, não estabilizando em um patamar nítido, pode-se pensar na presença de um fenômeno denominado capacidade infinita de dispersão, ou seja não possui variância finita. Esta situação pode advertir que o tamanho da área amostrada não possui grandeza o suficiente para expressar toda a variância, podendo existir tendência significativa nos dados em determinada direção. Vieira (2000) cita duas alternativas para tratamento deste problema:

- a) Remove-se a tendência e trata-se a curva de resíduos (diferença entre a curva de tendência e o semivariograma experimental);
- b) Estuda-se a hipótese da tendência dos dados originais.

---

Neste estudo o sistema implementado contempla a abordagem expressa no item a

---

O sistema implementou as três superfícies de tendência polinomiais disponíveis na biblioteca geo-estatística, conforme as equações 19, 20 e 21, segundo o método dos quadrados mínimos:

Superfície linear:

$$Z_{supL} = A_0 + A_1.x + A_2.y \quad (19)$$

Superfície parabólica:

$$Z_{supP} = A_0 + A_1.x + A_2.y + A_3.x^2 + A_4.x.y + A_5.y^2 \quad (20)$$

Superfície cúbica:

$$Z_{supC} = A_0 + A_1.x + A_2.y + A_3.x^2 + A_4.x.y + A_5.y^2 + A_6.x^2.y + A_7.x.y^2 + A_8.x^3 + A_9.y^3 \quad (21)$$

Onde:

$A_n$  corresponde aos parâmetros das superfícies;

X ao eixo das distâncias;

Y ao eixo das semivariâncias.

A opção de retirada de tendência encontra-se no canto direito da Figura 4.9. A função é ativada ao se selecionar umas das superfícies individualmente. Informações como o melhor ajuste, coeficiente de correlação e os respectivos parâmetros da superfície são calculados e atualizados na interface da Figura 4.10.

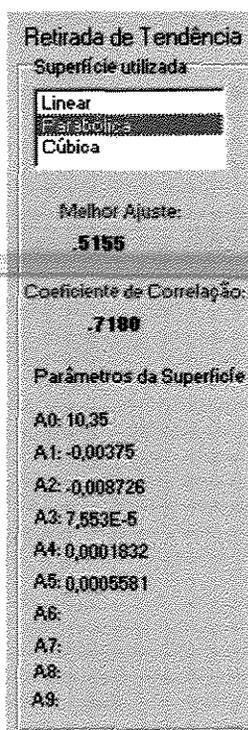


Figura 4.10 – Porção da interface utilizada na retirada de tendência dos dados.

Uma vez selecionada a superfície de tendência, o gráfico do semivariograma é atualizado e passa a exibir as curvas de tendência e de resíduos, além do semivariograma experimental. Por inspeção visual, observa-se o comportamento da curva de resíduos, aquela superfície de tendência de maior simplicidade que propiciar o alcance do patamar pela curva de resíduos será a escolhida. Dentre as opções implementadas, a superfície mais simples é a superfície linear e a mais complexa a cúbica. Posteriormente, a análise do coeficiente de

correlação e da variável “Melhor ajuste”, com valores ideais próximos unidade, poderão auxiliar num critério de desempate entre resultados próximos.

Escolhida a superfície de tendência o passo seguinte é ajustar os parâmetros do semivariograma (efeito pepita, alcance e contribuição) à curva de resíduos, conforme mostra o exemplo da Figura 4.11.

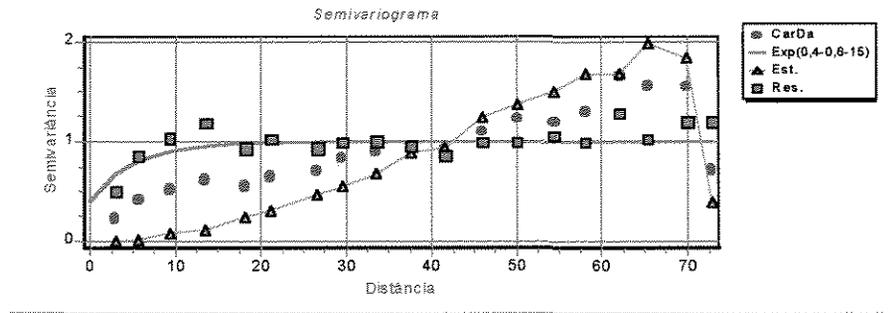


Figura 4.11 – Gráfico gerado pela função Retirada de tendência, com o ajuste da curva de resíduos de Carbono em Davis ( semivariograma experimental - ●, curva de resíduos - ■, superfície de tendência Δ ).

Para desativar a opção de retirada de tendência clica-se duplamente naquela superfície selecionada.

#### 4.3.5 - Módulo de Autovalidação – *Jack-Knifing*

A despeito dos parâmetros auxiliares ( $R^2$ , SQDP e superfície de tendência), há ainda um certo grau de incerteza na escolha dos parâmetros ideais para o semivariograma. Essa incerteza é representada pelo erro da estimativa, o qual pode ser avaliado pelo procedimento de autovalidação denominado *Jack-Knifing*.

A técnica de *Jack-Knifing* permite verificar se o número de vizinhos escolhidos para o processo de krigagem é adequado, ou se os parâmetros escolhidos para o modelo de ajuste são corretos. Esta técnica consiste em re-estimar todos os valores medidos usando todos os pontos vizinhos a este ponto com exceção dele mesmo. A razão de se desconsiderar o próprio ponto deve-se a característica de exatidão do interpolador. Uma vez que o ponto é

desconsiderado, este é estimado em função dos pontos que estão próximos a este, e não se fornece peso unitário ao ponto medido.

Como produto do processo de autovalidação cita-se o gráfico do valor medido  $Z(x_i)$  versus estimado  $Z^*(x_i)$  pela krigagem ou co-krigagem como base para a análise de alguns parâmetros. A este gráfico aplica-se regressão linear entre os pontos obtendo a equação da reta:

$$Z^*(x_i) = A + B \cdot Z(x_i) \quad (22)$$

Onde **A** é a intercessão com o eixo Y, **B** o coeficiente angular da reta e  $r^2$  o coeficiente de correlação ente  $Z^*(x_i)$  e  $Z(x_i)$ . Se o valor estimado fosse o mesmo do medido, o valor de **A** tenderia a 0 (zero) e o coeficiente **B** e  $r^2$  à unidade (um). Verifica-se então que o aumento do valor de **A** em relação ao ponto de origem (0,0) para valores positivos causa a superestimação de valores pequenos de  $Z(x_i)$  e a subestimação de valores grandes. Se o crescimento de **A** é para valores negativos, a situação contrária acontece.

Outra possibilidade é a análise do erro reduzido (ER), que introduz o conceito da variância da estimativa  $\sigma^2_k(x_i)$ :

$$ER [Z (X_i)] = [ ( Z^*(x_i) - Z(x_i) ) / \sigma_k Z^*(x_i) ] \quad (23)$$

O sistema apresenta estes parâmetros na forma de gráficos. O eixo X de todos os gráficos corresponde ao tamanho da vizinhança. A interface desta operação é mostrada pela Figura 4.12.

A técnica de *Jack-knifing* leva em conta os vizinhos para a estimativa do ponto medido. Os critérios implementados pelo sistema para cálculo da melhor vizinhança abrangem o método do o número constante de vizinhos e a vizinhança escolhida segundo uma distância constante. Estes dois critérios podem ser escolhidos na parte superior da interface apresentada na Figura 4.12.

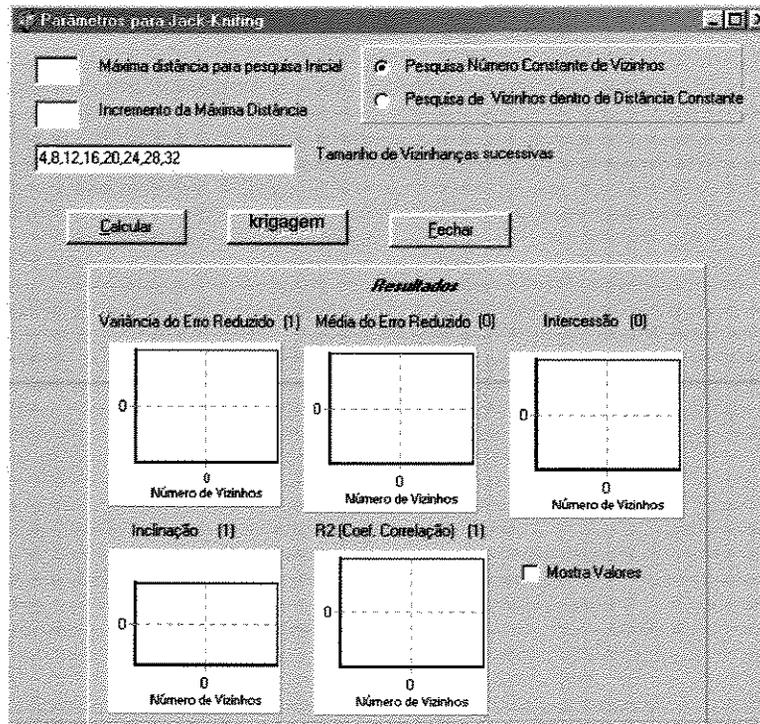


Figura 4.12 – Interface para verificar se o número de vizinhos escolhidos para o processo de krigagem é adequado ou se os parâmetros escolhidos para o modelo de ajuste são corretos pela técnica de *Jack-Knifing*.

O primeiro critério mantém constante o número de vizinhos em qualquer posição no campo. Neste método, determina-se um raio de distância máxima (DMAX) para a localização dos respectivos vizinhos. Se o número encontrado for menor que aquele especificado, o raio é incrementado e o processo de busca reiniciado. Outra situação se apresenta quando uma quantidade maior de pontos em relação àquele especificado é encontrada, estes então, são ordenados em função da distância e somente os mais próximos são considerados. A Figura 4.13 ilustra o critério.

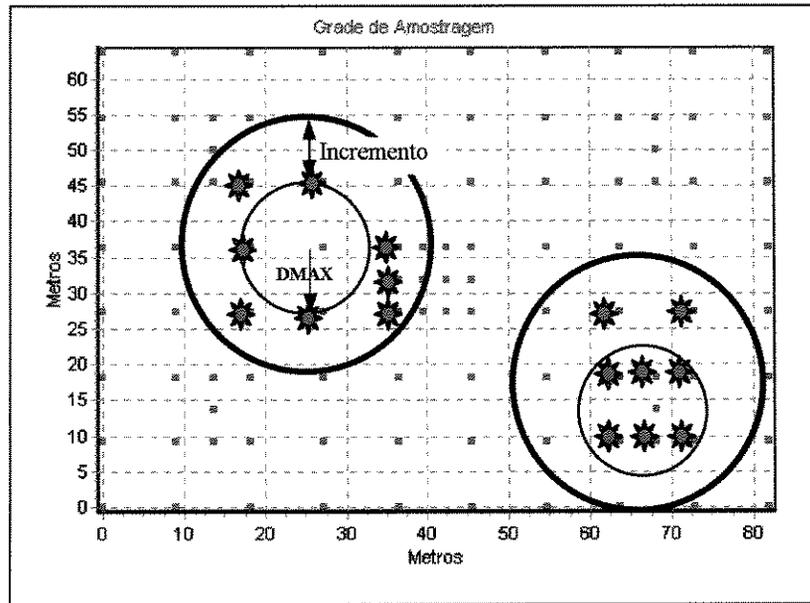


Figura 4.13 – Especificação de vizinhança pelo critério de número constante de vizinhos. A interface ilustra a pesquisa por 8 vizinhos dentro do raio da distância máxima.

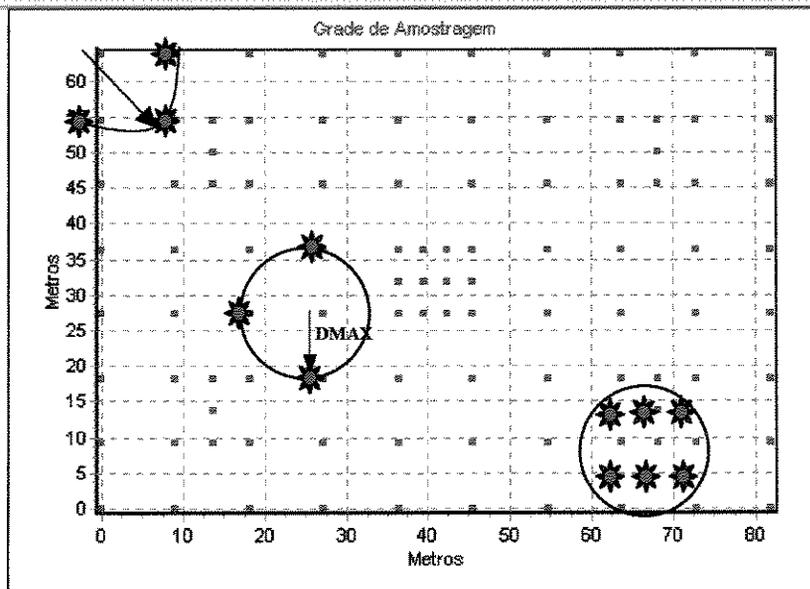


Figura 4.14 – Especificação de vizinhança através do critério de distância constante de 10 metros.

No critério da distância constante utiliza-se também uma distância máxima constante, para a seleção dos vizinhos. A observação feita para este tipo de busca é para aqueles pontos localizados nos cantos do campo, que a priori, teriam seu raio de ação limitado a  $\frac{1}{4}$  do círculo, conforme ilustra a Figura 4.14.

A interface representada na Figura 4.12 enumera uma série inicial de possíveis números de vizinhança (4, 8, 12, 16, 20, 24, 28, 32). Com o conhecimento da distribuição espacial dos pontos, pode-se alterar esta formação para as necessidades próprias de cada configuração de amostragem. A razão para a escolha de múltiplos de 4 para o número de vizinhança deve-se a tentativa de se selecionar os pontos vizinhos que se localizem igualmente em cada quadrante da área que circunda os pontos amostrados.

Ao se pressionar o botão *Calcular*, são atualizados os gráficos dos parâmetros citados anteriormente.

Intuitivamente, posicionou-se os gráficos em uma ordem de prioridade a ser verificada pelo especialista.

A escolha do número de vizinhos mais adequado deve ser feita segundo a análise dos valores ótimos de cada parâmetro.

---

#### **4.3.6 - Módulo de Krigagem**

Uma vez que os parâmetros do semivariograma teórico foram ajustados segundo suas condições específicas isto permitiu determinar a melhor vizinhança por meio da técnica de *Jack-Knifing*. O próximo passo da análise geo-estatística é a ação de estimar os pontos que não foram amostrados. Este constitui o passo da Krigagem simples, a qual pode ser ativada por meio do botão *Krigagem* na interface da Figura 4.12.

Este botão encontra-se desativado até o final dos cálculos relativo aos parâmetros de *Jack-Knifing*. Os parâmetros solicitados pela Krigagem encontram-se ilustrados na Figura 4.16.

Segundo a análise dos parâmetros da operação de autovalidação, a vizinhança que mais reduz o vício do estimador é o primeiro parâmetro solicitado.

O segundo conjunto de parâmetros solicitados pela Figura 4.16 (Nome do arquivo com pontos envoltórios, Incremento em Metros do eixo X e Y) é relativo a uma malha de referência usada para a estimação dos pontos não amostrados. Esta é uma malha superdensada situada nas mesmas coordenadas da malha de amostragem. Posteriormente, os valores estimados em cada nó da malha de referência são usados para construção de mapas

simbólicos, isolinhas e tridimensionais. Um exemplo de malha de referência é dado pela Figura 4.15.

É necessário fornecer ao sistema o nome do arquivo com os pontos do polígono que envolve a malha de amostragem, bem como os tamanhos dos incrementos no eixo X e Y que constituem esta malha.

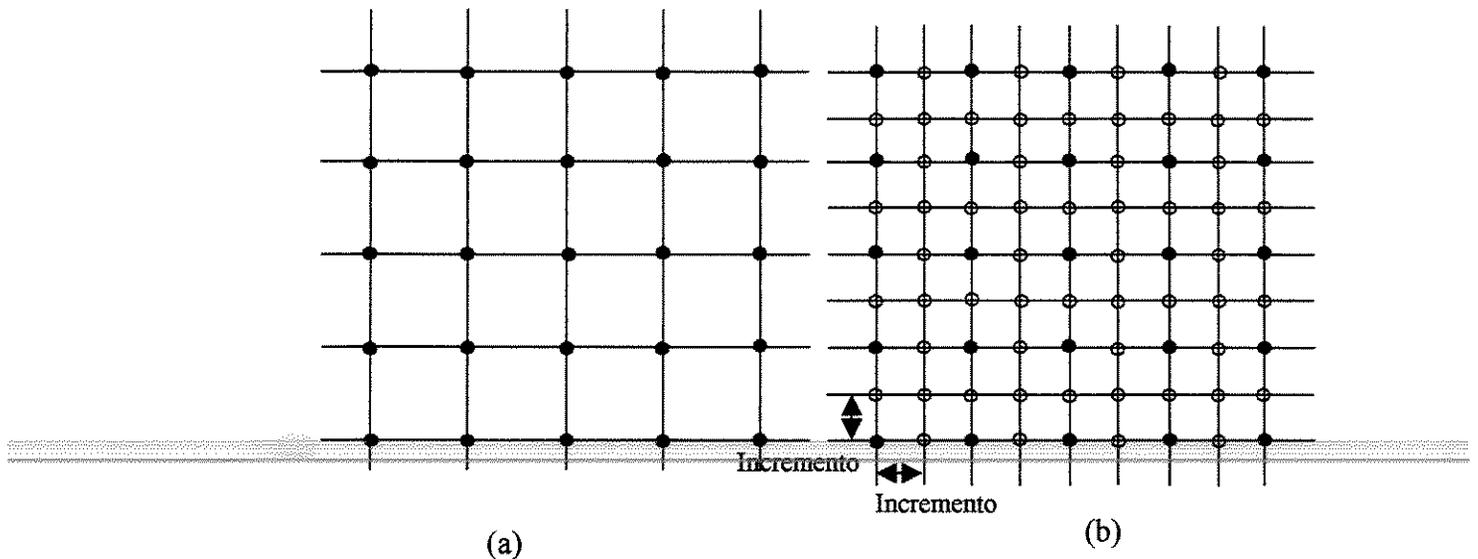


Figura 4.15 – Malha de amostragem genérica (a) e malha de referência usada na krigagem (b), destacando os pontos medidos (pontos cheios) e aqueles a serem estimados (pontos vazados).

A malha de referência, no sistema implementado pode alcançar até 20.000 pontos em combinações de linhas e colunas diversas, conforme a necessidade do estudo.

Na parte inferior da Figura 4.16 são apresentados os dados relativos aos parâmetros da superfície de tendência, se foi esta a opção escolhida para o tratamento da estacionaridade dos dados. Os valores não estão disponíveis para serem manipulados, apenas visualizados constituindo parâmetros para o cálculo do processo de krigagem.

Um exemplo da estrutura do arquivo dos pontos envoltórios é mostrado na Figura 4.17.

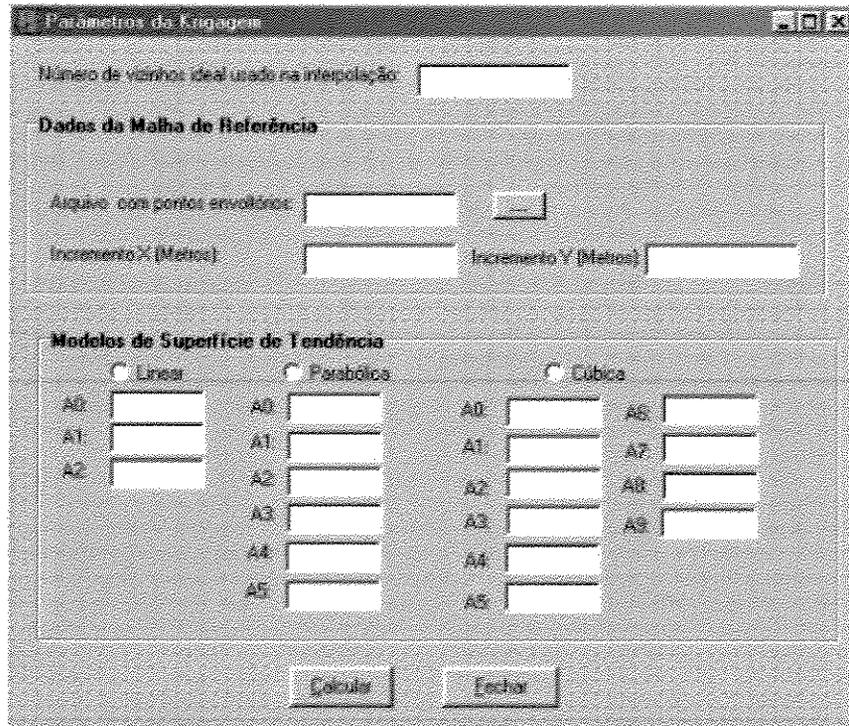


Figura 4.16 – Interface de entrada para o processo de krigagem.

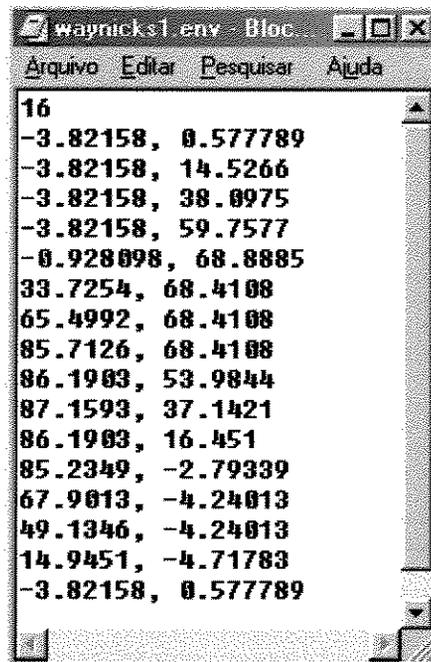


Figura 4.17 – Estrutura de um arquivo relativo aos pontos do polígono que envolve a área de amostragem.

A primeira linha do arquivo descrito na Figura 4.17 corresponde ao número de coordenadas do polígono. As linhas seguintes enumeram as coordenadas geográficas propriamente ditas. As coordenadas devem seguir o sentido horário do relógio e a primeira coordenada deve se repetir no final do arquivo para assegurar o fechamento do polígono. O sistema implementado aceita formatos de coordenadas cartesianas e UTM (Projeção Transversa de Mercator), as quais utilizam escala em metros.

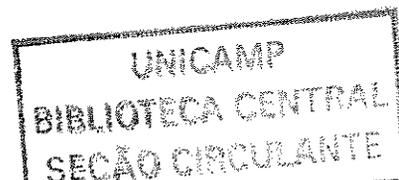
Este arquivo pode ser obtido com auxílio de software de mapeamento, como o *Surfer*<sup>1</sup>, por meio de opção de digitalização.

O resultado da operação de krigagem é a geração de um arquivo constituído dos pontos integrantes da malha de referência, dos valores estimados do atributo estudado e das variâncias da estimativa, conforme ilustra a Tabela 4.1.

Tabela 4.1 – Parte inicial dos valores interpolados e medidos de Carbono (Davis) e sua respectiva variância

Eixo X	Eixo Y	Valor interpolado	Variância da estimativa
0,00	0,00	11,67	0,0000
0,00	1,00	11,11	0,6776
0,00	2,00	10,97	0,7818
0,00	3,00	10,86	0,8439
0,00	4,00	10,77	0,8732
0,00	5,00	10,72	0,8727
0,00	6,00	10,66	0,8455
0,00	7,00	10,60	0,7863
0,00	8,00	10,54	0,6865
0,00	9,10	10,48	0,0000

<sup>1</sup> (c) Surfer é um produto comercial da Golden Software Incorporated



#### 4.3.7 - Módulo para projeto da malha de amostragem

Este módulo tem como objetiva contribuir para a otimização da configuração de amostragem, pelo conhecimento da variância da estimativa resultante da operação de krigagem. Lembrando que a características básicas da krigagem são permitir a estimativa de valores sem tendência para locais não amostrados e adicionalmente obter informações sobre a confiança dessas estimativas, ditas ótimas.

Sabe-se que o cálculo da variância da estimativa, não depende diretamente dos valores medidos, e o peso associado a cada ponto reflete o como o atributo varia no espaço. Conclui-se que uma vez que se conheça o semivariograma de um determinado atributo, pode-se desenhar então esquemas de amostragens para variâncias de estimativas pré-especificadas. Pode-se aplicar tal situação em atributos que foram amostrados para fins de reconhecimento e otimizar o processo de amostragem definitiva segundo condições pré-definidas.

Clicando-se botão *Projeto da Grade* presente na Figura 4.9 o módulo para o projeto da malha da amostragem é ativado e resulta no surgimento da interface mostrada pela Figura 4.18. A opção é válida apenas para esquemas de amostragem sistemáticas estratificadas em grade.

O módulo captura os parâmetros que foram usados no modelo do semivariograma teórico do atributo. Solicita também valores correspondentes a coordenadas de malhas fictícias para o cálculo das respectivas variâncias da estimativa. Alguns parâmetros de malhas estão pré-definidos na interface, podendo adquirir outras configurações conforme necessidades do estudo.

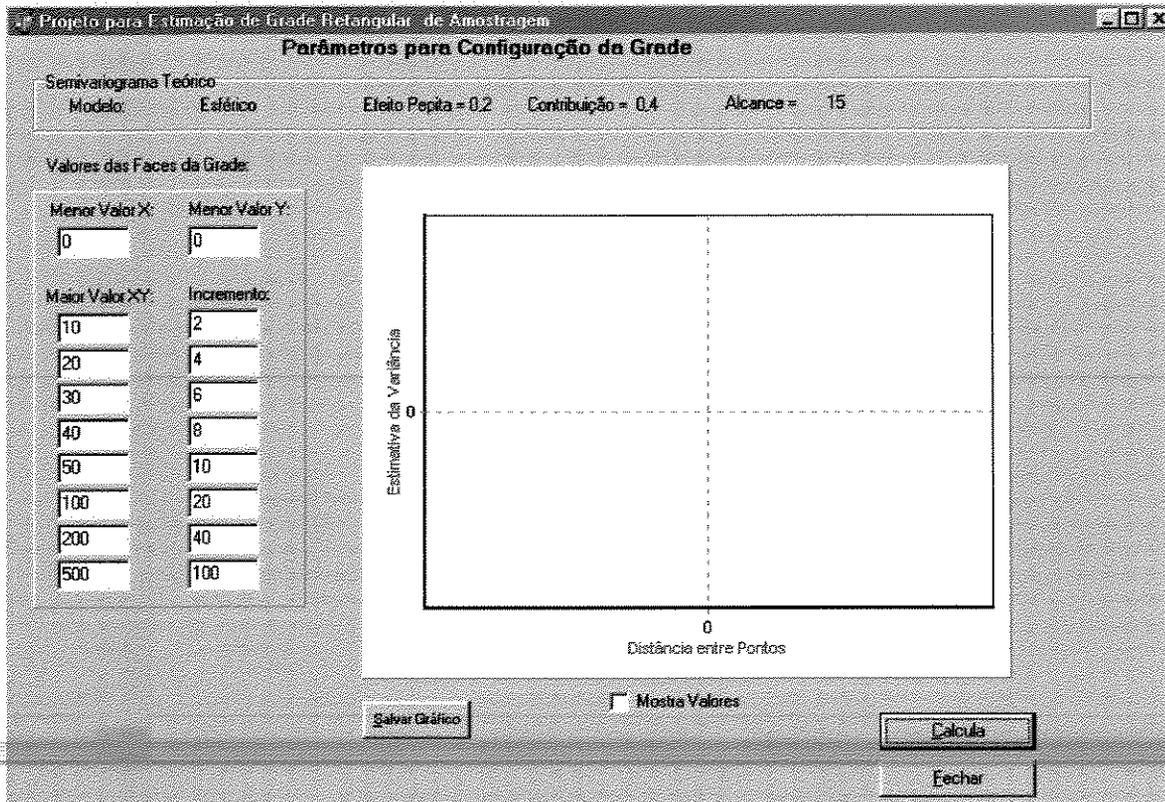


Figura 4.18 – Interface para o módulo de otimização de malha da amostragem.

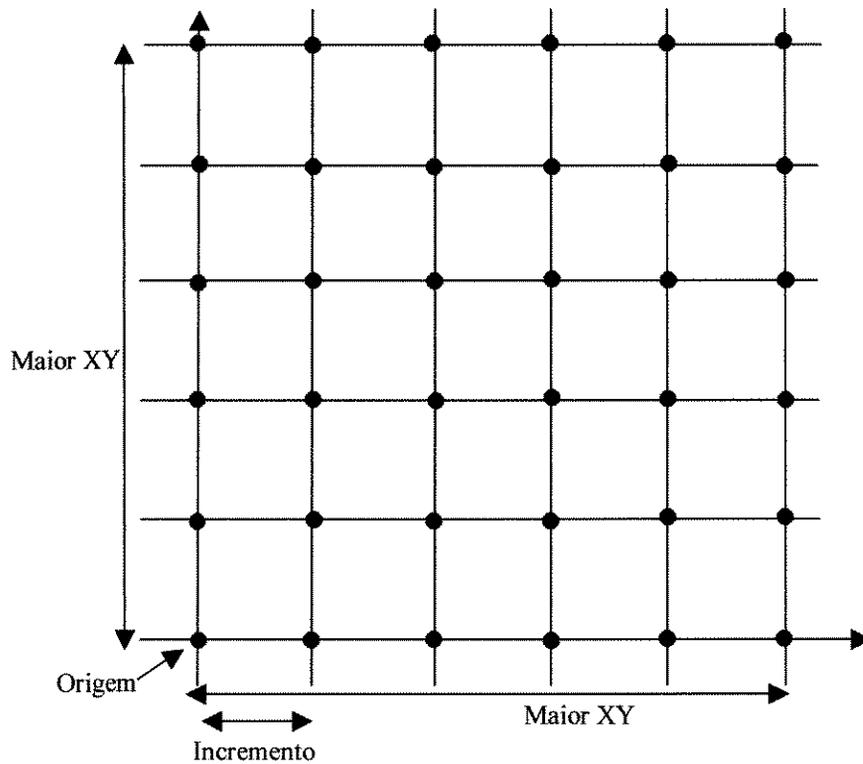


Figura 4.19 – Parâmetros de uma malha fictícia usada no módulo de projeto de malha.

Os parâmetros “Menor valor X” e “Menor valor Y” correspondem aos pontos de origem dos eixos X e Y. Todas as malhas a serem criadas terão sempre a mesma origem.

O “Maior valor XY” corresponde a face dos eixos X e Y da malha e o incremento as subdivisões a que esta malha está dividida, conforme ilustra a Figura 4.19.

Uma vez definida os parâmetros das várias malhas de amostragem o resultado pode ser gerado pressionando o botão *Calcular*, o qual atualizará o gráfico da interface. Os eixos deste gráfico são a separação entre as distâncias versus a variância da estimativa.

A análise do gráfico é análoga ao semivariograma médio experimental.

#### **4.3.8 - Módulo para estudo de semivariogramas direcionais**

O módulo para modelagem do semivariograma médio apresentado na seção 4.3.2, considera a combinação de pares de pontos em todas as direções do campo. Se houve necessidade de um estudo mais aprofundado em outras direções é utilizado o módulo para geração dos semivariograma em direções específicas.

O sistema implementado gera basicamente 4 semivariogramas considerando as principais direções. São elas:

- 0° - direção do eixo X;
- 90° - direção do eixo Y;
- 45 ° e – 45° nas direções das diagonais.

O módulo pode ser ativado segundo a Figura 4.20 e contribui para o estudo da anisotropia, a qual investiga as mudanças da variável em diferentes direções.

De maneira análoga à interface do semivariograma médio, a interface da Figura 4.20 solicita o nome do arquivo e o título do projeto. Uma vez selecionado o arquivo de dado, a lista de variáveis é preenchida automaticamente. Pode-se marcar a transformação dos valores da variável para a escala logarítmica, por meio da opção *Considera logaritmo*.

Valores pré-definidos de 5° são sugeridos para o ângulo de tolerância em cada direção, podendo ser modificado conforme a necessidade do estudo. Este ângulo influencia o agrupamento das classes de distâncias.

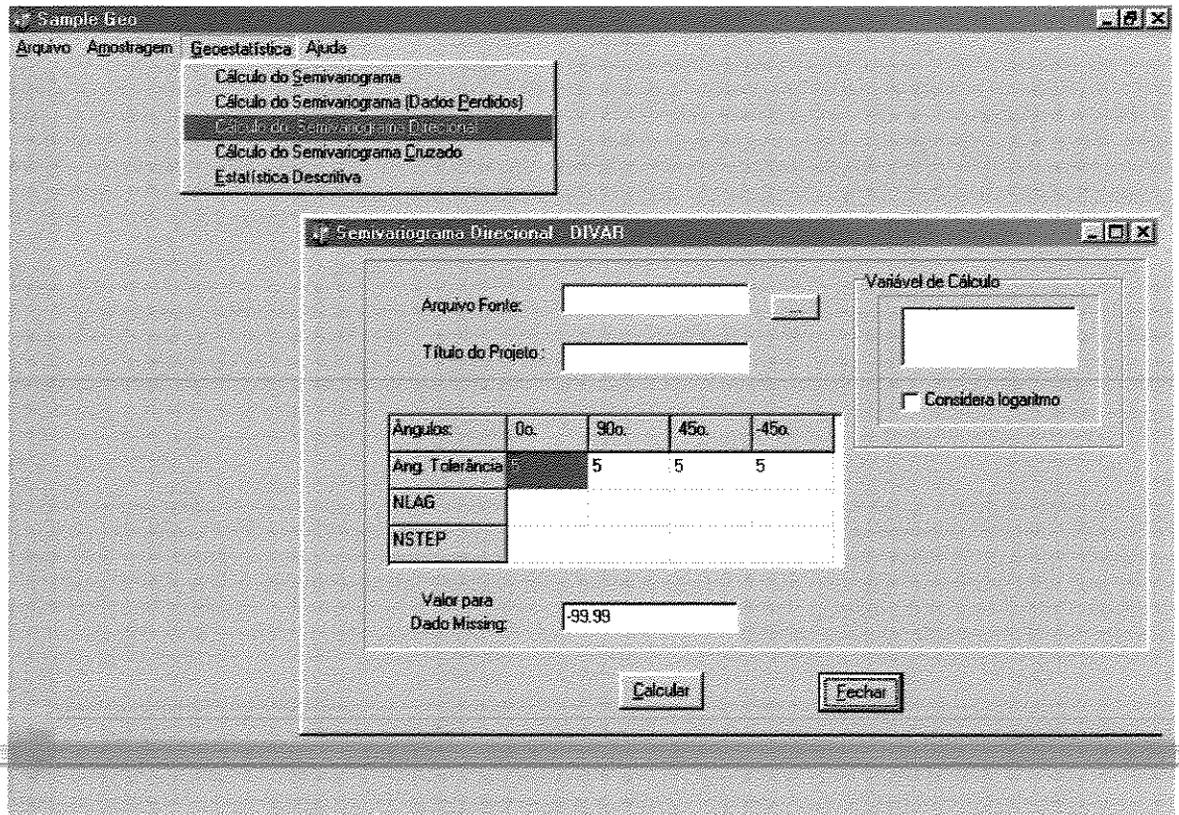


Figura 4.20 – Interface para o cálculo dos semivariogramas direcionais.

O parâmetro NLAG é o número de semivariância a ser calculado em cada direção e o NSTEP corresponde ao número de passos presente na interface para modelagem do semivariograma médio.

O gráfico e valores resultantes são disponíveis por meio do botão *Calcular*. Esta operação gera a interface mostrada na Figura 4.21.

Os semivariogramas direcionais podem ser desenhados individualmente ou em grupo. A medida que seu desenho é anexado no gráfico da interface a planilha com os valores de pares, classes de distâncias e respectivas variâncias é ativada nesta tela.

Segundo Camargo (1997) a máxima continuidade tem relação com o semivariograma de maior alcance. O pensamento equivalente é realizado para a mínima continuidade, a qual tem seu valor associado a um semivariograma de menor alcance.

A possibilidade do desenho simultâneo dos semivariogramas por este módulo revela as direções de maior e menor continuidade.

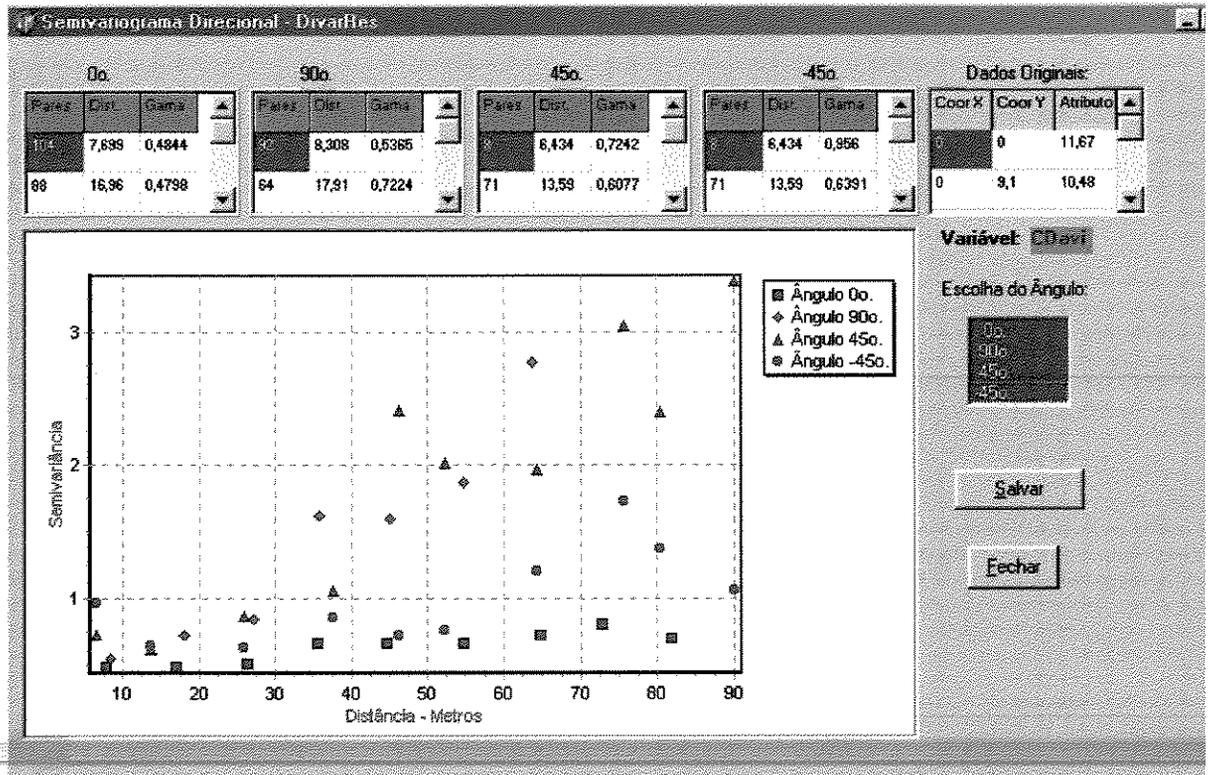


Figura 4.21 – Interface resultante dos semivariogramas direcionais.

#### 4.3.9 - Módulo para estudo de semivariograma cruzado

É comum em várias áreas das ciências naturais a existência de relacionamento entre duas ou mais variáveis. Em situações onde esta correlação exista, a estimativa de uma delas pode ser feita usando-se informações de ambas por meio do semivariograma cruzado e do método de interpolação apropriado denominado Cokrigagem (Vieira, 2000).

Neste estudo é disponibilizado o módulo para análise de duas variáveis. A condição para esta operação é que haja coincidência das coordenadas geográficas dos pontos de amostragem. Não importa se determinada variável apresentou mais pontos amostrados que outra, mas que existam pontos coincidentes em que as duas variáveis foram medidas.

Como o cálculo da semivariância no semivariograma cruzado envolve o produto da diferença entre duas variáveis distintas, o significado dos parâmetros do semivariograma teórico ajustado a curva é diferente daquele realizado em uma variável. O alcance neste caso representa a distância máxima da dependência espacial entre as variáveis. O patamar, se

semivariograma cruzado negativo, se a covariância for negativa, significando que para o incremento de uma variável há decréscimo de outra.

Os parâmetros de entrada solicitados para este módulo são ilustrado pela Figura 4.22.

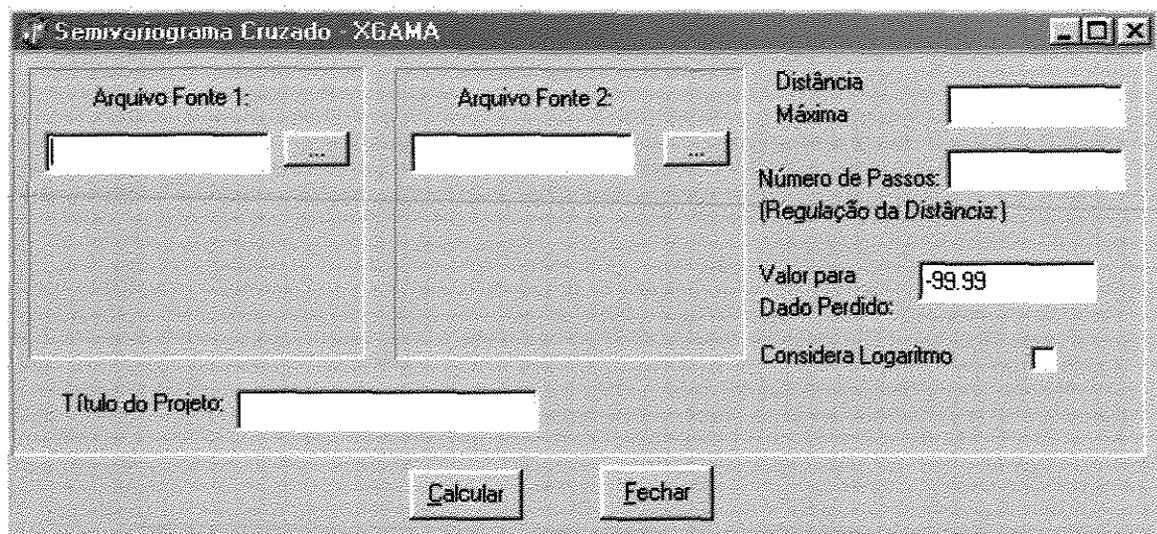


Figura 4.22 – Interface do módulo para geração do semivariograma cruzado.

Pode-se comparar arquivo de dados distintos ou um mesmo arquivo, como é feito neste estudo, escolhendo-se em seguida as variáveis de interesse.

O módulo permite a escolha de apenas uma variável por arquivo. Pode-se especificar um valor para dados perdidos se for o caso.

A interface resultante possui sumariamente as mesmas informações disponibilizadas pelo módulo do semivariograma experimental

No próximo capítulo o conjunto de dados de Waynick e Sharp (1919) é investigado visando avaliar o sistema por meio de suas interfaces e processos internos bem como aferir os resultados numéricos apresentados confrontando-os com aqueles encontrados por Vieira (2000).

resultados numéricos apresentados confrontando-os com aqueles encontrados por Vieira (2000).

## CAPÍTULO 5

### RESULTADOS E DISCUSSÕES

Este capítulo tem como objetivo avaliar o sistema desenvolvido por meio de um estudo de caso. Todos os módulos desenvolvidos foram testados utilizando o conjunto de dados elaborado por Waynick e Sharp em 1919 e reutilizado por Vieira (2000). Neste capítulo é estudado o comportamento da estrutura espacial das variáveis Carbono, Nitrogênio e a relação C/N.

Os resultados obtidos pelo sistema não serão objeto de análise agrônômica aprofundada, pois já representam um estudo de caso documentado na literatura. Seu uso principal foi aferir os resultados encontrados pelo sistema com aqueles encontrados por Vieira (2000).

Outra finalidade desta etapa do trabalho foi avaliar a seqüência e o relacionamento entre as rotinas Fortran dentro do ambiente integrado.

#### 5.1 - Análise exploratória de um conjunto dos dados

As estatísticas clássicas do conjunto de dados estudados são mostradas na Tabela 5.1.

**Tabela 5.1 – Estatísticas clássicas para o conjunto de dados Waynick e Sharp (1919)**

	Nú- mero	Média	Variância	C.V.	Mínimo	Máximo	Coef. Simetria	Curtose
<b>C-Davis (g/kg)</b>	100	11,110	1,0900	9,394	8,960	13,83	0,464	-0,154
<b>N-Davis (g/kg)</b>	100	0,998	0,008403	9,183	0,770	1,18	-0,254	-0,0183
<b>C-Oakley (g/kg)</b>	100	4,330	1,3290	26,620	1,820	9,50	1,468	3,881
<b>N-Oakley (g/kg)</b>	100	0,3208	0,0048	21,800	0,210	0,60	1,582	4,233
<b>C/N-Davis (%)</b>	100	11,170	0,9630	8,784	9,415	17,26	2,731	14,320
<b>C/N- Oakley (%)</b>	100	13,570	6,1060	18,210	7,857	23,23	0,926	3,055

**Tabela 5.2- Matriz de Correlação**

	C-Davis	N-Davis	C-Oakley	N-Oakley	C/N-Davis	C/N-Oakley
C-Davis	1,0000	0,6268	0,0406	0,1362	0,4322	- 0,1130
N-Davis	0,6268	1,0000	- 0,0425	- 0,0511	- 0,4243	- 0,0101
C-Oakley	0,0406	- 0,0425	1,0000	0,7637	0,0925	0,5205
N-Oakley	0,1362	- 0,0511	0,7637	1,0000	0,2224	- 0,1337
C/N-Davis	0,4322	- 0,4243	0,0925	0,2224	1,0000	- 0,1333
C/N-Oakley	- 0,1130	- 0,0101	0,5205	- 0,1337	- 0,1333	1,0000

Observa-se na Tabela 5.1 que as maiores médias e os valores máximos e mínimos de Carbono e Nitrogênio são encontrados no campo Davis. A variável C/N em Oakley é a que melhor se aproxima de uma distribuição normal, apesar do valor elevado de simetria.

Com exceção da variável C/N- Oakley, todas as outras apresentam baixos valores de variância, indicando que a maioria dos valores está próxima a média.

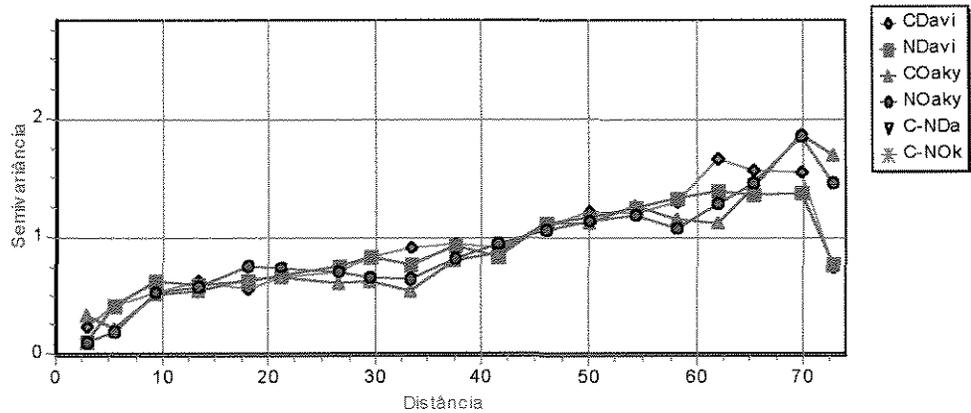
Os valores elevados do coeficiente de curtose representam uma curva de frequência mais fechada que a normal, confirmando uma concentração de valores próximos ao valor médio.

Na Tabela 5.2, o índice de 0,7637, entre as medidas de Carbono e Nitrogênio no campo de Oakley e o índice 0,6268 entre Carbono e Nitrogênio em Davis estabelecem correlações positivas médias entre as variáveis. As demais variáveis têm correlação baixa ou não se correlacionam.

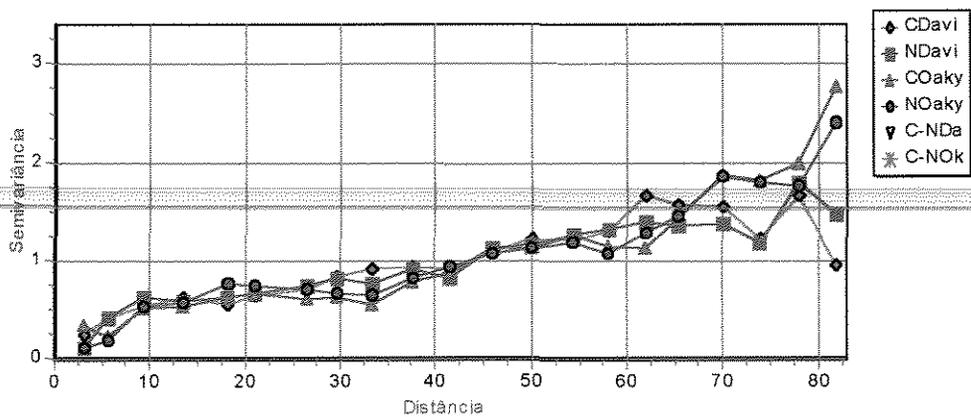
## 5.2 - Modelagem do semivariograma médio

Os parâmetros iniciais para o cálculo do semivariograma médio no sistema, são distância máxima e número de passos.

Para a área da Figura 4.1 do capítulo de Material e Métodos, a diagonal calculada tem valor de 103,74 m e o intervalo da distância máxima situa-se entre 72,62 e 82,99 m. As Figuras 5.1a e 5.1b demonstram que a escolha arredondada destes valores, 73 e 83 metros respectivamente, não trouxe diferença significativa no desenho final do semivariograma. As Figuras 5.1a e 5.1b foram geradas por meio do botão *Salvar* disponível na interface da Figura 4.7 do capítulo anterior.



a



b

Figura 5.1 - Semivariograma escalonado para distância máxima igual a 73 (a) e 83 (b) metros e Número de passos = 4.

A análise do semivariograma médio escalonado da Figura 5.1 revela que a variabilidade de Carbono e Nitrogênio nos dois campos pode ser dita isotrópica, caracterizada por um valor aproximadamente igual em todas as direções.

Analisando a Figura 5.1, apenas por inspeção visual, observa-se que os valores de semivariância para o atributo Nitrogênio nos campos de Davis e Oakley são os mais altos em relação a variável Carbono nos dois campos.

Observa-se também que a uma distância superior a 15 metros, todos os semivariogramas médios não conseguem atingir um patamar definido indicando não estacionaridade e sugerindo o tratamento de retirada de tendência nas operações seguintes.

Com objetivo de comparar os de semivariogramas obtidos na Figura 5.1 apresenta-se a Figura 5.2 que utilizou o número passos igual a 1 (um). Observa-se que em função da maior quantidade de classes de distância entre os pares a curva resultante apresenta maior oscilações de valores.

Por estabelecerem uma boa representação da estrutura espacial dos fenômenos estudados e facilidade na análise visual os valores empregados na Figura 5.1a serão aqueles utilizados nas análises que se seguirem.

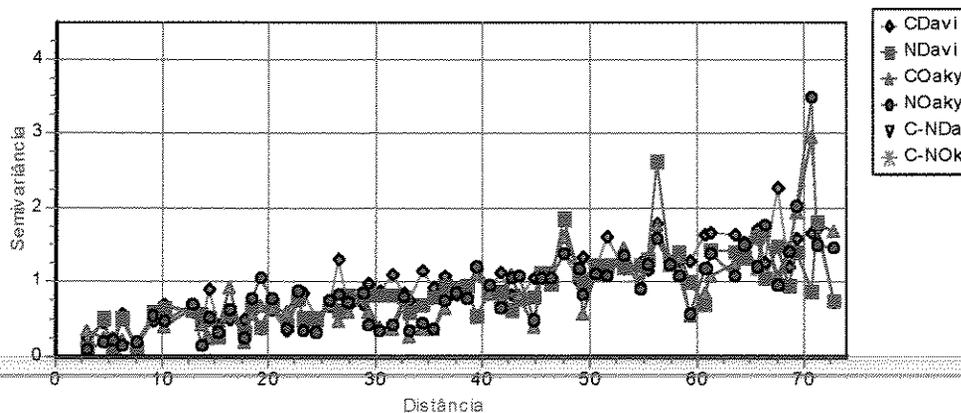


Figura 5.2 - Semivariograma escalonado para distância máxima igual a 73 (a) e 83 (b) metros e Número de passos = 1.

As relações C/N em Davis e Okley são ilustradas pela Figura 5.3. Observa-se que os semivariogramas de C/N para os dois campos apresentam alto efeito pepita em relação ao patamar, denotando pequena dependência espacial.

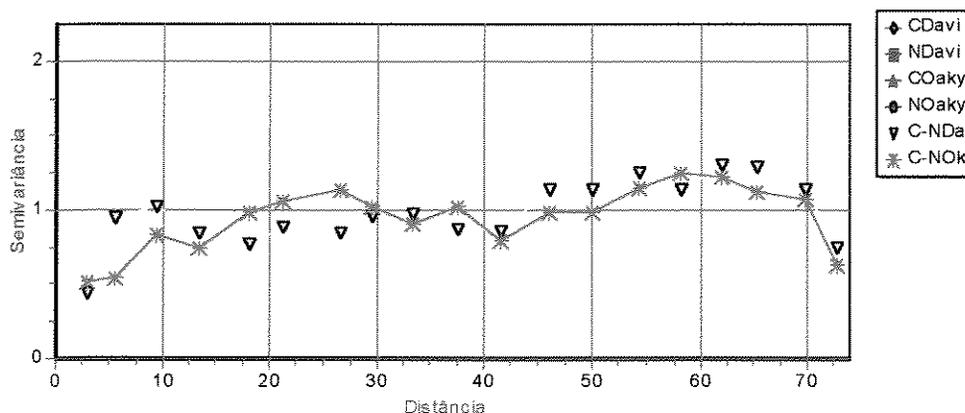


Figura 5.3 - Semivariograma experimental para C/N em Davis e Oakley.

### 5.3 - Ajuste do semivariograma aos modelos matemáticos

Este procedimento realiza o ajuste dos modelos matemáticos à curva do semivariograma gerado pelos valores experimentais.

Escolheu-se a variável C/N em Davis, para ilustrar o módulo de ajuste de modelos matemáticos, pelo fato desta apresentar um patamar bem definido.

Por meio de inspeção visual os valores para os parâmetros dos modelos matemáticos foram adicionados à interface da Figura 5.4. Apesar do baixo coeficiente de determinação o modelo exponencial, foi aquele que em relação aos demais modelos melhor obteve valores de  $R^2$  e SQDP.

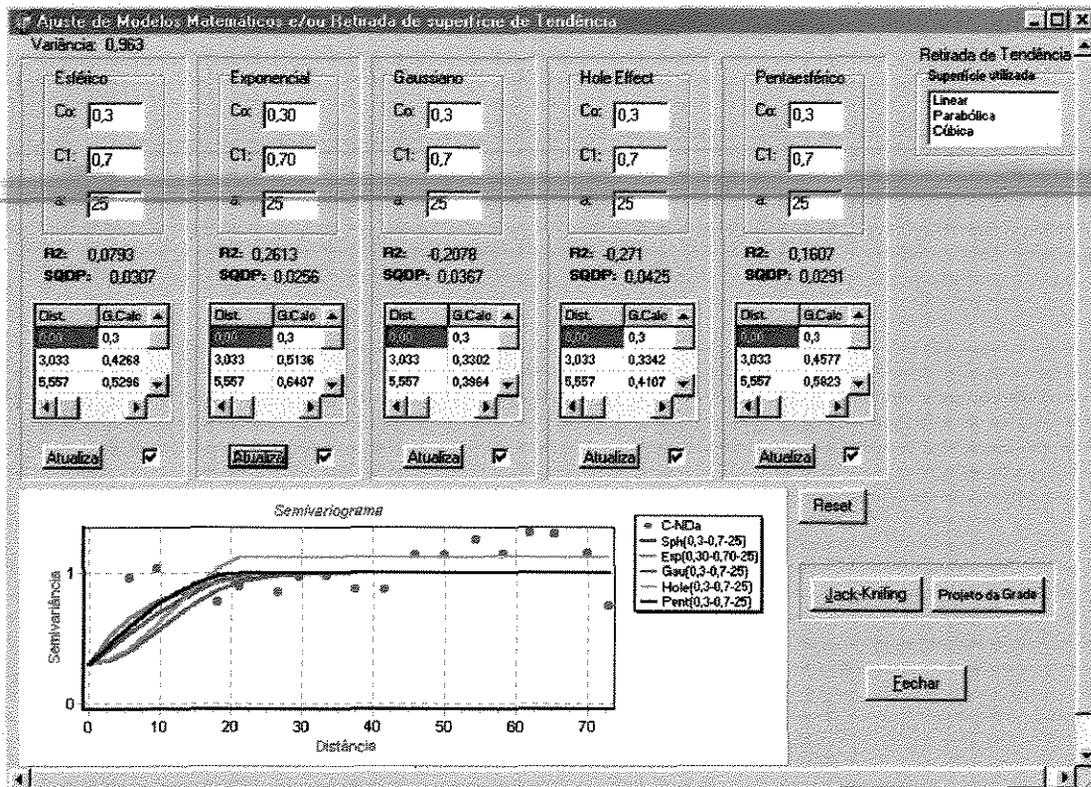


Figura 5.4 - Semivariograma escalonado da variável C/N de Davis ajustada pelos modelos: esférico, exponencial, gaussiano, *Hole effect* e pentaesférico.

Prosseguiu-se com um ajuste mais detalhado dos parâmetros de Co, C e alcance para o modelo exponencial, visando melhorar os valores de  $R^2$  e SQDP. Várias tentativas resultaram na modificação apenas do efeito pepita conforme mostra a Figura 5.5, melhorando o valor de  $R^2$  para 0,3001 e SQDP igual a 0,0242.

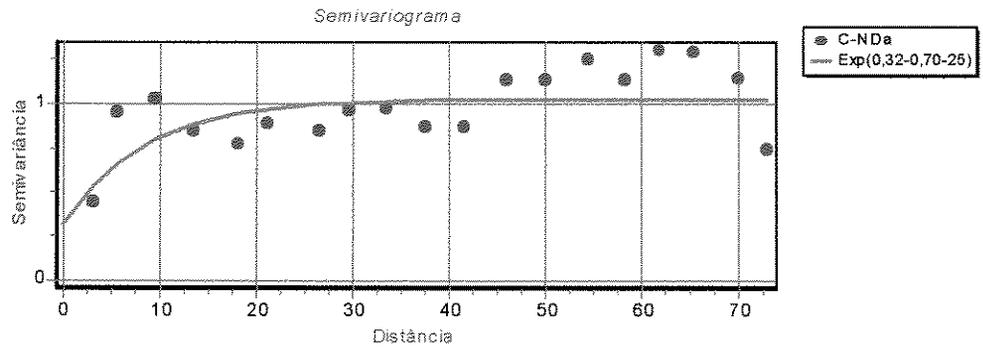


Figura 5.5 – Ajuste fino do modelo exponencial ( $C_0 = 0,32$ ;  $C=0,7$  e  $a=25$ ) para a relação C/N em Davis.

#### 5.4 – Tratamento da estacionaridade dos dados

Para as variáveis Carbono e Nitrogênio nota-se porém uma certa tendência dos dados. Estes não possuem estacionaridade, aumentando continuamente. Uma solução para esta situação constitui-se em aplicar uma superfície teórica de tendência, de forma a amenizar este efeito. De forma prática, o novo ajuste será feito na curva de resíduos (diferença entre os valores experimentais e teóricos).

Retirou-se a tendência dos dados por meio da superfície parabólica na variável Carbono (Davis). Notou-se que a curva de resíduos aproximava do formato de um semivariograma ideal, melhor representando o fenômeno estudado. O novo ajuste foi feito usando o modelo exponencial. Os detalhes das curvas de resíduos com a aplicação de superfície de tendência parabólica são mostrados pela Figura 5.6.

O patamar é estabelecido aplicando-se a superfície parabólica nas variáveis de Carbono e Nitrogênio dos dois campos, conforme ilustra a Figura 5.7.

**Retirada de Tendência**  
 Superfície utilizada

Linear  
**Parabólica**  
 Cúbica

Melhor Ajuste  
**.5155**

Coefficiente de Correlação:  
**.7190**

Parâmetros da Superfície

A0: 10,35  
 A1: -0,00375  
 A2: -0,008726  
 A3: 7,553E-5  
 A4: 0,0001932  
 A5: 0,0005581  
 A6:  
 A7:  
 A8:  
 A9:

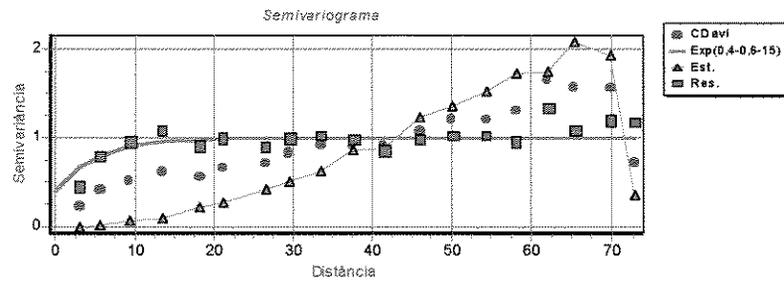
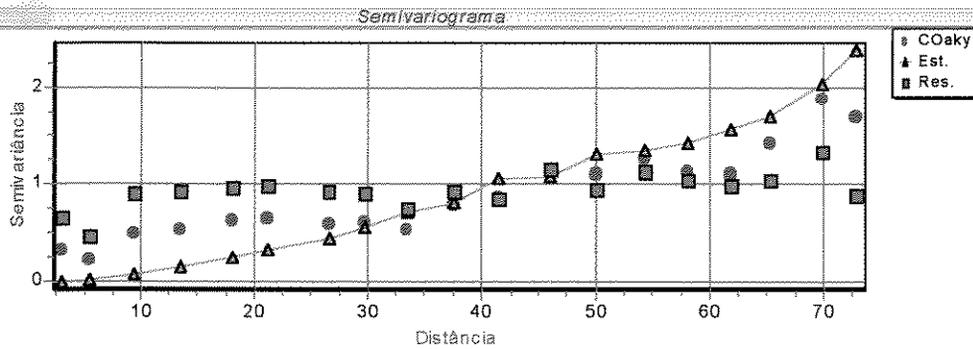
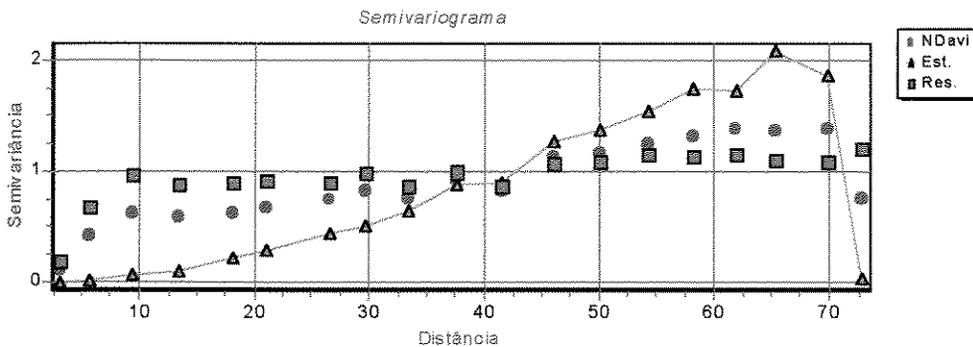


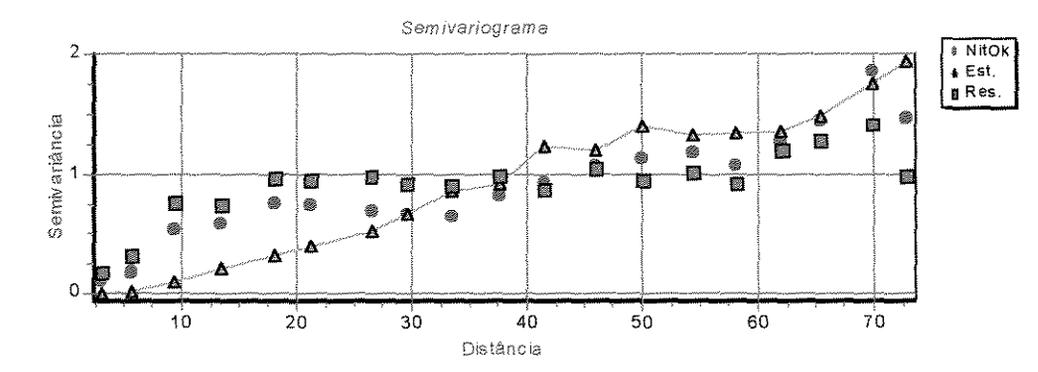
Figura 5.6 – Ajuste de Carbono por meio da curva de resíduos, utilizando superfície de tendência parabólica.



a – Carbono – Oakley



b - Nitrogênio – Davis



c- Nitrogênio – Oakley

Figura 5.7 – Semivariogramas de Carbono (a) e Nitrogênio (b e c) aplicando-se retirada de tendência por meio de superfície parabólica.

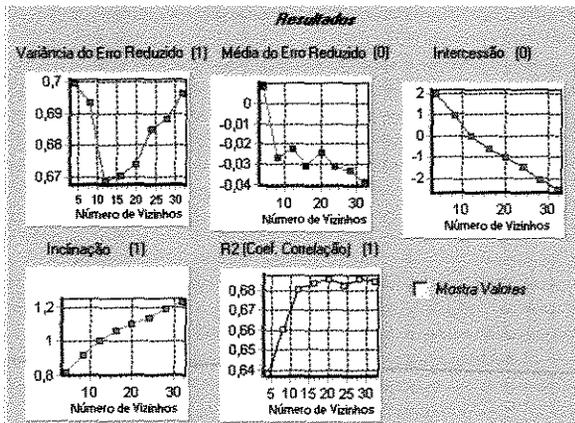
## 5.5 - Aplicação do módulo de Autovalidação

O módulo de autovalidação estudou o comportamento da variável Carbono em Davis.

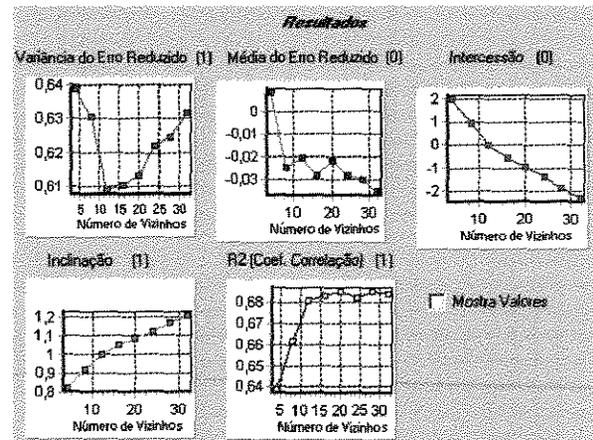
Foi utilizado o critério do número fixo de vizinhos, por ser o mais popular, optando-se por um raio de busca (distância máxima) com valor de 10 metros. Este valor foi escolhido por assegurar que dentro desta distância seriam encontrados pelo menos 4 vizinhos (com exceção das áreas de borda), visto que grande parte das amostras está distanciada de 9,099 m. O incremento da distância, parâmetro utilizado pelo sistema no caso de não se alcançar o número de vizinhos pré-determinado, também foi de 10 metros pelo mesmo motivo exposto anteriormente.

A série de 4, 8, 12, 16, 20, 24, 28 e 32 vizinhos foi mantida.

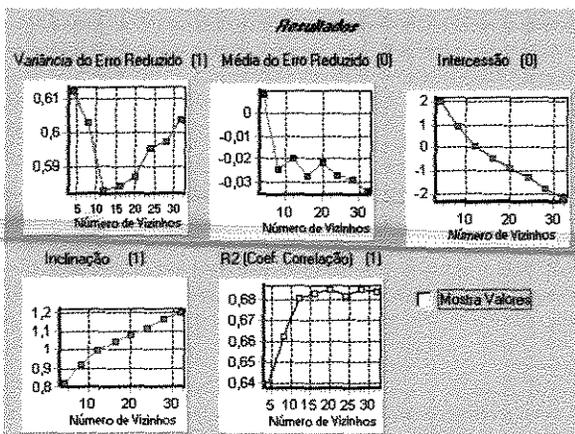
Visando verificar se os parâmetros do semivariograma teórico ( $C_0 = 0,4$ ;  $C = 0,6$  e  $a = 15$ ) obtido no módulo de retirada de tendência, são os mais adequados, realizou-se testes com 3 conjuntos de parâmetros com valores próximos aqueles citados. Estas tentativas estão registradas na Figura 5.8.



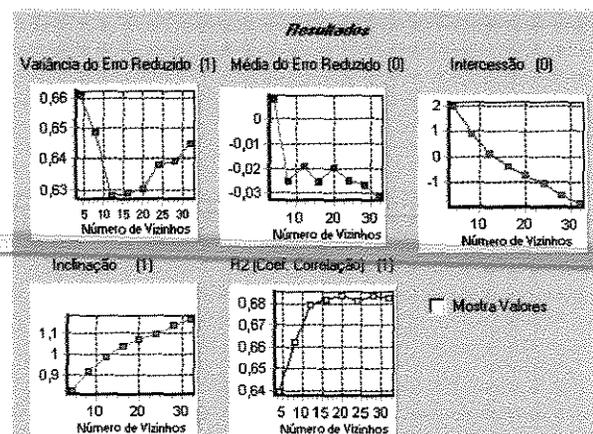
a -  $Co = 0,4$  ;  $C = 0,5$  ;  $a = 15$



b -  $Co = 0,4$  ;  $C = 0,6$  ;  $a = 15$



c -  $Co = 0,4$  ;  $C = 0,65$  ;  $a = 15$



d -  $Co = 0,4$  ;  $C = 0,6$  ;  $a = 17$

Figura 5.8 – Resultados do *Jack-knifing* para várias configurações do semivariograma teórico.

Observou-se por inspeção que a variação dos valores no semivariograma teórico não trouxe mudança significativa no comportamento das curvas na Figura 5.8, bem como em suas grandezas.

Os parâmetros da Figura 5.8b, da variável foram mantidos para continuidade da análise geo-estatística da variável Carbono em Davis.

Pela análise da Figura 5.8b observou-se que 12 pontos representam a melhor vizinhança. Este valor alcança os melhores índices de interseção e inclinação e a segunda melhor posição dentre os valores da média dos erros reduzidos. Este processo de escolha da vizinhança concentra uma certa carga de subjetividade, pois encontramos o número de 20 pontos alcançando bons resultados dos critérios estabelecidos na interface da Figura 5.8.

## 5.6 - Krigagem

Esta seção dá continuidade à análise geo-estatística para a variável Carbono em Davis, a qual sofreu tratamento de retirada de tendência por meio de superfície parabólica e definiu-se o número de 12 pontos como a melhor vizinhança.

O arquivo de pontos envoltórios usados foi aquele mostrado pela Figura 4.16 do capítulo de Material e Métodos.

Em função do módulo para geração da malha de referência (**grid.exe**) aceitar somente valores inteiros, procedeu-se a um passo intermediário no preparo deste arquivo. Gerou-se inicialmente uma malha de referência com 5312 pontos com incremento de 1 em 1 metro nas direções X e Y. No arquivo com os pontos da malha de referência substituiu-se manualmente os 100 pontos correspondentes ao arquivo original (waynicks.dat), conforme ilustra a Figura 5.9.

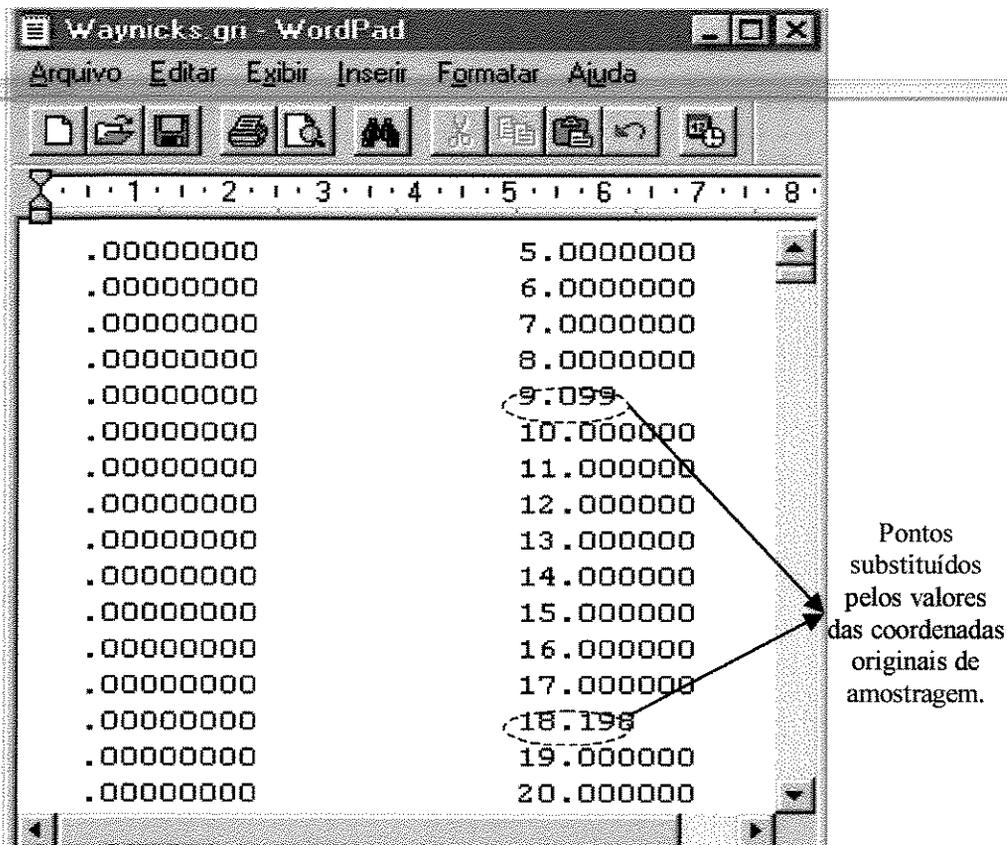


Figura 5.9 – Visão parcial dos pontos referentes a malha de referência.

Uma alternativa para evitar o manuseio do arquivo da malha de referência, seria manter as coordenadas originais dadas em pés, como usado por Vieira (2000). Preferiu-se

trabalhar com as distâncias em metros para manter coerência com os resultados apresentados neste capítulo, e por ser uma grandeza adotada no sistema internacional de medidas.

A Figura 5.10 representa a interface do processo de krigagem para a variável Carbono – Davis. Os parâmetros da superfície de tendência parabólica utilizada na operação de retirada de tendência são mostrados automaticamente pelo sistema.

Figura 5.10 – Interface do processo de Krigagem para Carbono em Davis.

O resultado deste procedimento foi um arquivo com os valores estimados da variável segundo os pontos da malha de referência. A parte inicial deste arquivo é mostrada na Tabela 4.1. Os valores gerados possuem alta variância da estimativa oscilando entre 0,6 e 0,8. Os menores valores da variância estão próximos aos pontos medidos e os maiores na metade da distância entre uma amostra e outra. Estes valores fundamentam a teoria de que o erro da interpolação é menor nas áreas próximas aos pontos medidos em campo. O formato do arquivo com os valores interpolados, medidos e suas respectivas variância da estimativa é apresentado pela Figura 5.11.

X	Y	KRIGED	ESTIMATION VARIANCE
.00	.00	11.67	.0000
.00	1.00	11.11	.6776
.00	2.00	10.97	.7818
.00	3.00	10.86	.8439
.00	4.00	10.77	.8732
.00	5.00	10.72	.8727
.00	6.00	10.66	.8455
.00	7.00	10.60	.7863
.00	8.00	10.54	.6865
.00	9.10	10.48	.0000
.00	10.00	10.46	.6606
.00	11.00	10.42	.7696
.00	12.00	10.37	.8356
.00	13.00	10.32	.8682
.00	14.00	10.24	.8723
.00	15.00	10.18	.8484
.00	16.00	10.09	.7932

Figura 5.11 – Arquivo resultante do processo de krigagem para a variável Carbono em Davis.

A etapa final do processo de interpolação é o mapeamento da variável estudada. Neste estudo, optou-se por implementar mapas simbólicos dos atributos por meio de classes de valores. Esta forma mais simplificada da representação espacial da variável permite uma análise clara da variabilidade do atributo no campo.

Os gráficos dos valores interpolados visualizados por meio de classes são apresentados pelas Figuras 5.13 a 5.16 foram realizados por meio do software de mapeamento *Surfer*<sup>1</sup> 7.0.

Para a integração com o sistema Surfer, o arquivo apresentado pela Figura 5.11 tem sua primeira linha retirada e é novamente salvo. Por meio das opções *Map* e *Classed Post*, ilustradas na Figura 5.12, o arquivo é selecionado e aberto. As escolhas dos valores limites para cada classe foram baseadas naquelas usadas por Vieira (2000).

<sup>1</sup> Surfer™ é um produto comercial da Golden Software Inc.

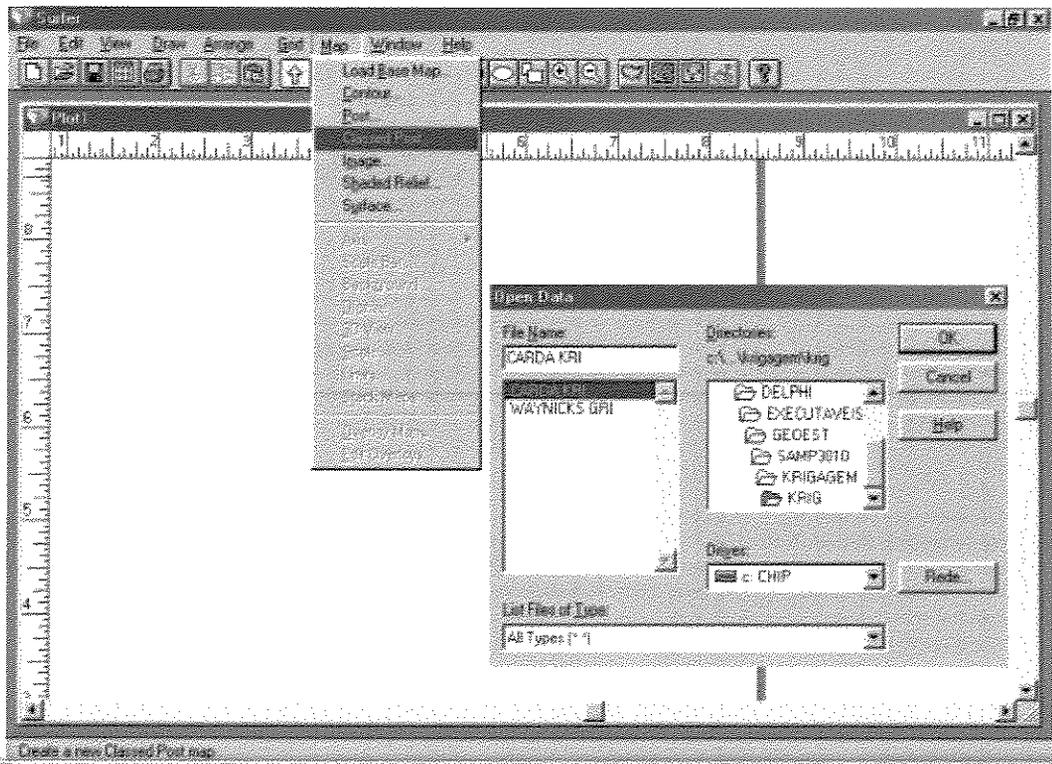


Figura 5.12 – Operações para integração do arquivo interpolado com o sistema Surfer, visando geração do mapa de classes para Carbono em Davis.

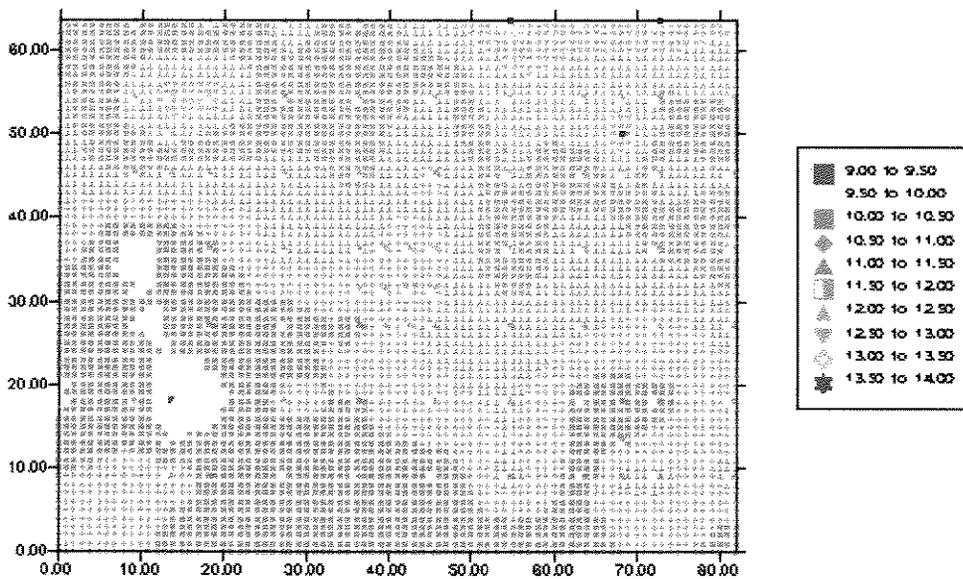


Figura 5.13 – Mapa simbólico dos valores interpolados de Carbono (Davis).

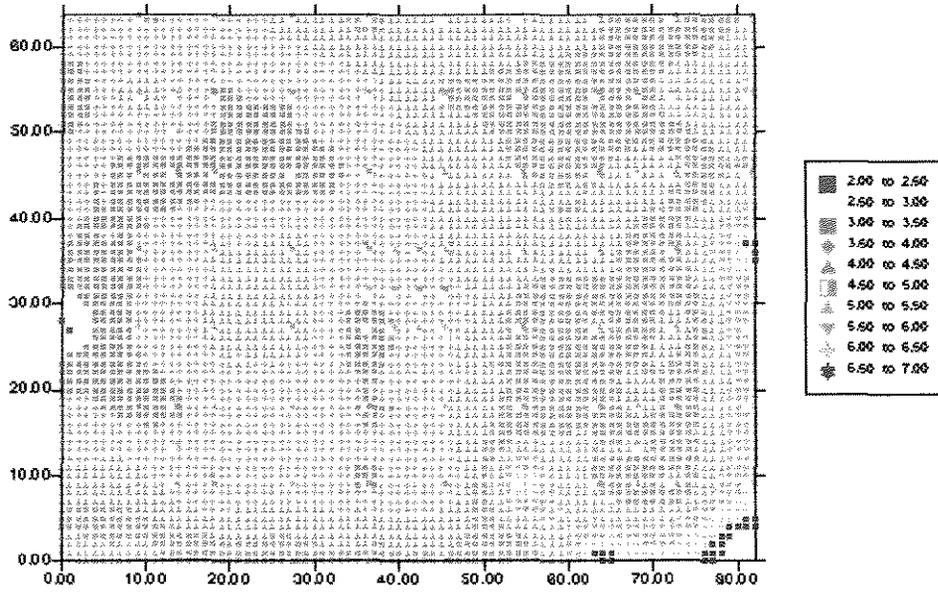


Figura 5.14 - Mapa simbólico dos valores interpolados de Carbono (Oakley).

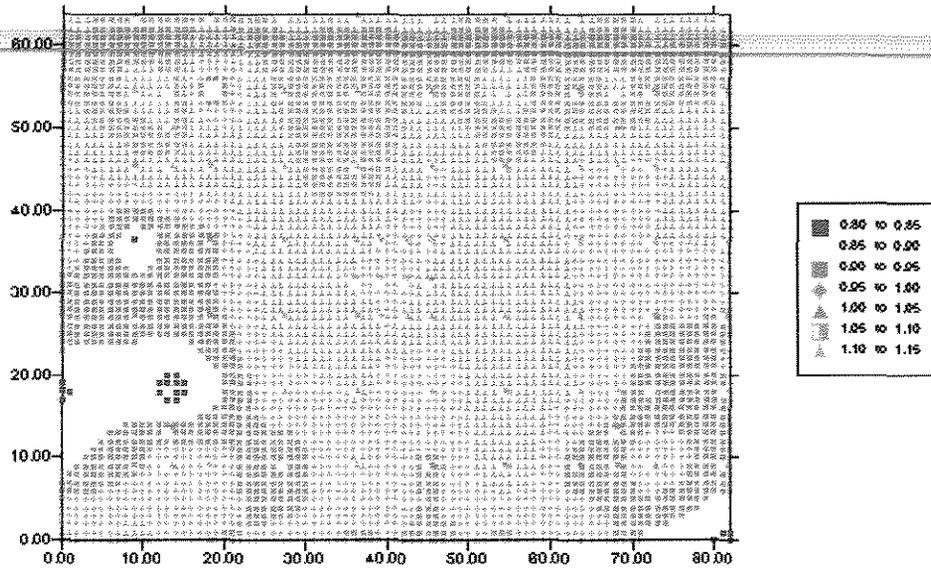


Figura 5.15 - Mapa simbólico dos valores interpolados de Nitrogênio (Davis).

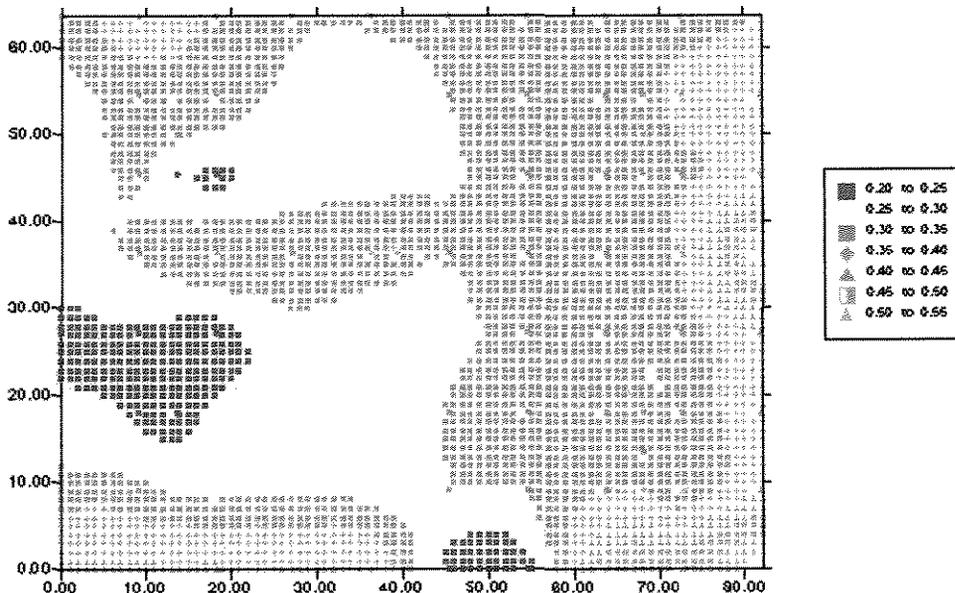


Figura 5.16 - Mapa simbólico dos valores interpolados de Nitrogênio (Oakley).

Apesar dos altos valores da variância da estimativa quando o mapa foi classificado sua forma assemelhou-se àquela encontrada por Vieira (2000). Percebe-se, na Figura 5.13, que o crescimento dos valores de Carbono se processa na diagonal de 45°, e valores entre 10 e 10,5 g kg<sup>-1</sup> ocorrem na maior parte da área amostrada.

O mapa interpolado de Carbono em Oakley (Figura 5.14) mostra na direção -45° a sua maior continuidade.

Outro fato notado por inspeção visual é a semelhança entre os mapas simbólicos de Carbono e Nitrogênio em Davis, mostrado pela Figura 5.15, onde locais com altas concentrações de Carbono correspondem a altos valores de Nitrogênio. Esta correlação foi detectada também na matriz de correlação da seção 5.1.

O mapa de Nitrogênio em Oakley, ilustrado pela Figura 5.16, apresenta em seu eixo Y (direção 90°) a menor continuidade, contrastante a alta continuidade apresentada pela mesma variável em Davis.

### 5.7 - Projeto de configuração de amostragem.

A variância da estimativa é utilizada neste estudo para a escolha do melhor distanciamento entre amostras segundo configurações de malhas pré-determinadas e parâmetros do semivariograma.

Estudou-se a variável Carbono em Davis com os mesmos parâmetros do modelo exponencial obtido no módulo de retirada de tendência da Figura 5.6 na seção 5.4.

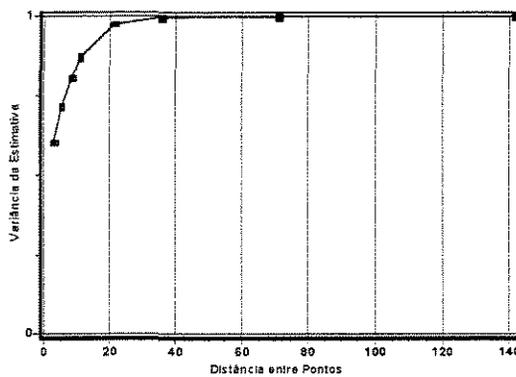
Foram realizadas algumas variações nos parâmetros das malhas visando estudar o comportamento das distâncias entre os pontos de amostragem. Considerou-se o ponto 0,0 como origem de todas as malhas. As variações das faces das áreas de amostragem e suas divisões (incrementos) são descritas na Tabela 5.3. Os parâmetros do semivariograma teórico foram mantidos para todas as configurações.

A interface exige que a relação entre a face e o incremento seja constante para todas as malhas. As configurações 1, 2, 3, 4 e 6 usam fator 5 e a configuração 5 fator 7.

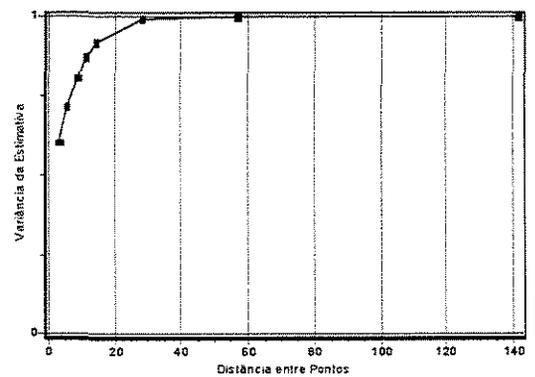
**Tabela 5.3– Configurações pré-definidas para malhas de amostragem**

Configuração 1		Configuração 2		Configuração 3		Configuração 4		Configuração 5		Configuração 6	
Face (m)	Increment. (m)										
10	2	10	2	15	3	5	1	70	10	10	2
20	4	20	4	25	5	10	2	140	20	20	4
30	6	30	6	35	7	15	3	210	30	30	6
40	8	40	8	45	9	20	4	280	40	40	8
75	15	50	10	65	13	25	5	350	50	50	10
125	25	100	20	125	25	30	6	420	60	100	20
250	50	200	40	225	45	35	7	490	70	200	40
500	100	500	100	350	70	60	12	560	80	500	100

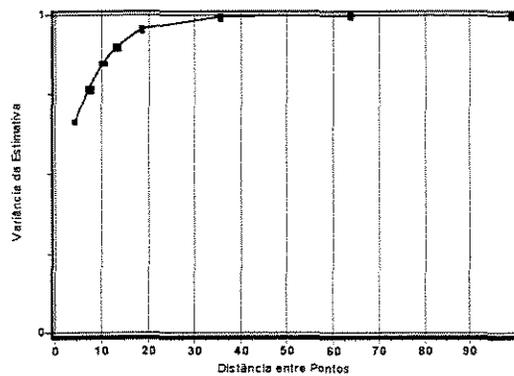
Os respectivos gráficos resultantes das várias configurações são mostrados pela Figura 5.17.



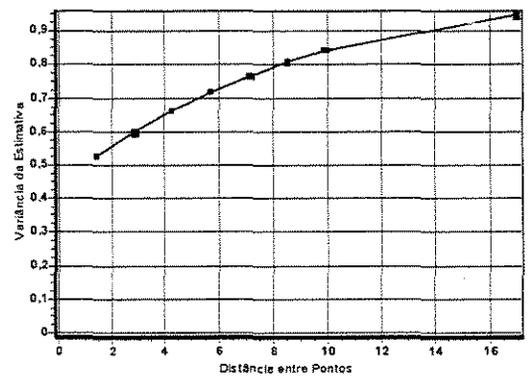
a - Configuração 1



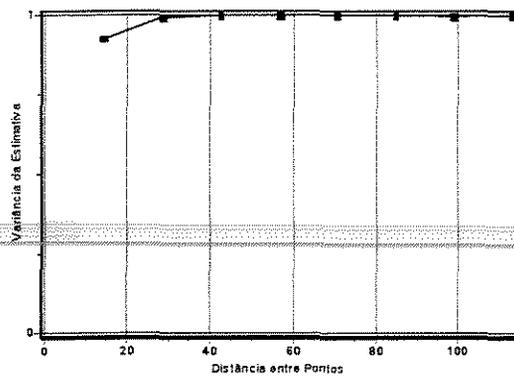
b - Configuração 2



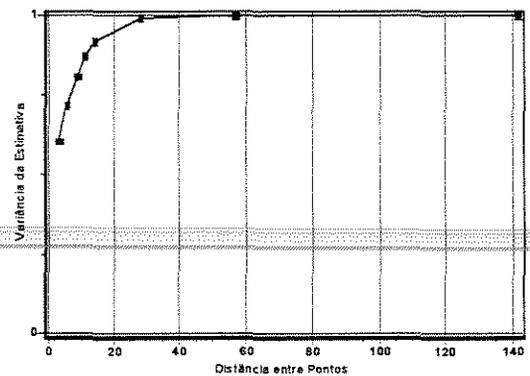
c - Configuração 3



d - Configuração 4



e - Configuração 5



f - Configuração 6

Figura 5.17 – Comportamento da variância da estimativa para várias configurações de malha.

Os gráficos relativos às Figuras 5.17a, 5.17b, 5.17c, 5.17e e 5.17f indicam um patamar com um alcance localizado entre 35 a 40 metros. Esta faixa sugere que um novo distanciamento entre os pontos de amostragem no campo, para efeito de estudo do carbono. Na configuração representada pela Figura 5.17 d, nota-se o não estabelecimento de um patamar definindo, sugerindo testes com maiores incrementos, como foi realizado nos gráficos restantes da Figura 5.17. Ressalta-se que esta sugestão de distanciamento é válida para uma configuração regular de amostragem, para qual este módulo foi projetado.

A Figura 5.18 mostra com maior detalhamento a configuração 6 no sistema, para a variável Carbono (Davis).

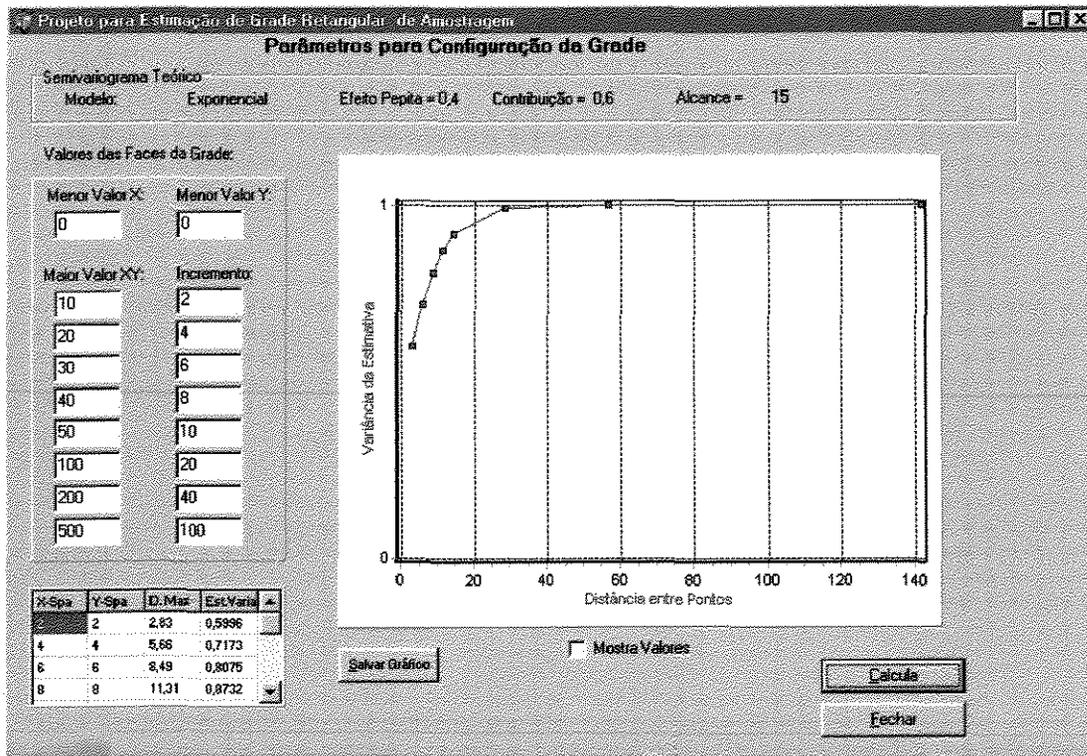
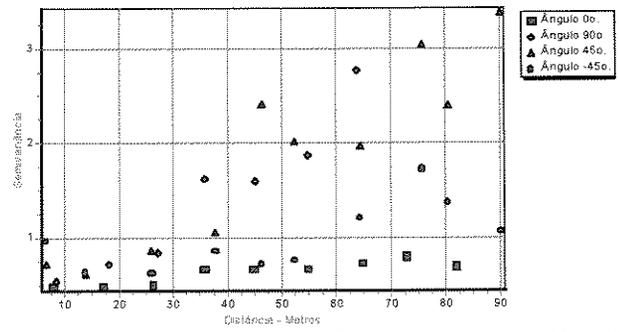


Figura 5.18 - Gráfico da variância da estimativa para diversos projetos de amostragem utilizando a curva de resíduos de Carbono em Davis.

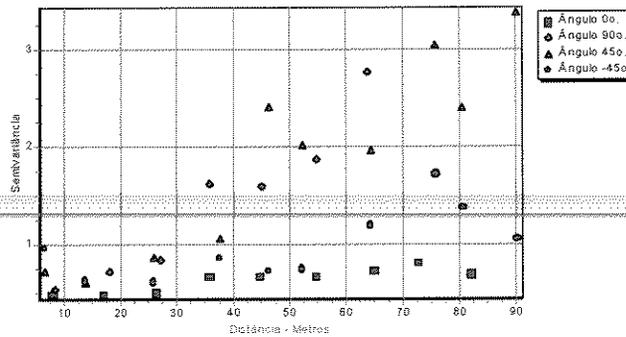
## 5.8 - Semivariogramas direcionais

Foi analisada neste estudo a anisotropia da variável Carbono em Davis ou o seu comportamento em várias direções, por meio de seus semivariogramas direcionais. É um processo caracterizado por várias interações, onde o objetivo é detectar as direções de maior e menor continuidade do fenômeno na área em estudo.

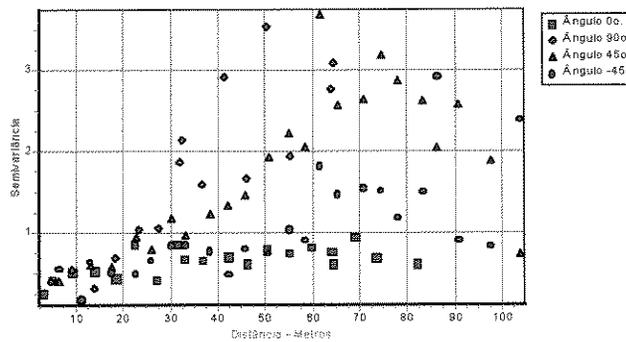
Realizaram-se testes no módulo específico com variações nos parâmetros de entrada. Os resultados são apresentados na Figura 5.19.



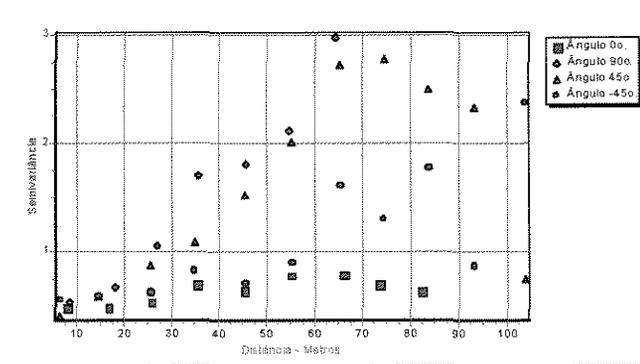
a - Ângulo de tolerância de 5°, NLAG=50 m,  
NSTEP=10



b- Ângulo de tolerância de 5°, NLAG=73 m,  
NSTEP=10



c - Ângulo de tolerância de 25°, NLAG=73 m,  
NSTEP=4



d - Ângulo de tolerância de 25°, NLAG=73 m,  
NSTEP=10

Figura 5.19 - Semivariogramas direcionais da variável Carbono em Davis para diferentes valores de ângulo de tolerância, NSTEP e NLAG.

Observou-se que o incremento do ângulo de tolerância e da variável NStep não deformava o comportamento dos semivariogramas direcionais, apenas tornava mais denso o número de pontos pesquisados.

É clara que a direção de menor continuidade é dada pelo ângulo 0°. Contudo a maior continuidade oscila entre os ângulos de 45° e 90°.

Analisando a Figura 5.19, observa-se que a direção de 0° alcança um patamar em torno de 15 metros mantendo-se constante com o crescimento da distância. As demais direções (90°, 45°, e - 45°) têm crescimento abrupto, não indicando a consolidação do patamar.

Em especial o semivariograma para direção de 90° e 45 ° não tem a definição de seus respectivos patamares, ao contrário da direção 0° que tem seu patamar bem definido. Por apresentar alta continuidade, a direção de 90° sugere que um maior detalhamento na amostragem seja realizado posteriormente, para registrar com maior acurácia a ocorrência das várias classes de valores nestas direções. A baixa variabilidade de direção 0°, lembrando que apresenta geometria de amostragem análoga em relação à de 90°, indica que a densidade e o projeto da amostragem expressaram bem os níveis de variação de Carbono no campo em Davis.

As etapas seguintes da análise geo-estatística como aquelas do semivariograma médio experimental deveriam ser efetuadas, visando o mapeamento completo desta variável. Porém nesta versão, estas opções não foram implementadas.

## 5.9 - Semivariograma cruzado

Como as amostragens nos campos de Davis e Oakley geraram resultados de concentrações de Carbono e Nitrogênio em todos os 100 pontos, foi possível estudar a correlação entre estas duas variáveis por meio de seus semivariogramas cruzados.

Os parâmetros de entrada utilizados na geração destes semivariogramas no campo em Davis são ilustrados pela Figura 5.20. Em função da não estacionaridade dos semivariogramas de carbono e nitrogênio em Davis e Oakley, realizou-se a retirada de tendência por meio de superfície parabólica. Nesta etapa, optou-se por utilizar a curva de resíduos gerada por cada variável.

A interface de usuário do software XGAMA para a geração de um semivariograma cruzado. O título da janela é "Semivariograma Cruzado - XGAMA". A interface é dividida em seções para configuração de arquivos, parâmetros de distância e opções de cálculo.

- Arquivo Fonte 1:** Contém o campo "CaDaRePa.dat" e uma lista de arquivos com "CaDa", "CaEst" e "CaRes" selecionados.
- Arquivo Fonte 2:** Contém o campo "NDaRePa.dat" e uma lista de arquivos com "NiDa", "NiEst" e "NiRes" selecionados.
- Distância Máxima:** Campo de entrada com o valor "73".
- Número de Passos:** Campo de entrada com o valor "4", com o subtítulo "(Regulação da Distância)".
- Valor para Dado Perdido:** Campo de entrada vazio.
- Considera Logaritmo:** Caixa de seleção desmarcada.
- Título do Projeto:** Campo de entrada vazio.
- Botões:** "Calcular" e "Fechar".

Figura 5.20 – Parâmetros de entrada para análise do semivariograma cruzado entre Carbono e Nitrogênio em Davis, utilizando a curva de resíduos.

Os semivariogramas cruzados das variáveis Carbono e Nitrogênio nos dois campos têm, predominantemente, patamar positivo, como revela as Figuras 5.21 e 5.22. Apesar de não terem tido um modelo teórico ajustado, sugerem por inspeção visual um alcance em torno de 18 a 22 metros. Esta faixa é superior aos 13 metros encontrado pelo modelo gaussiano em Vieira (2000).

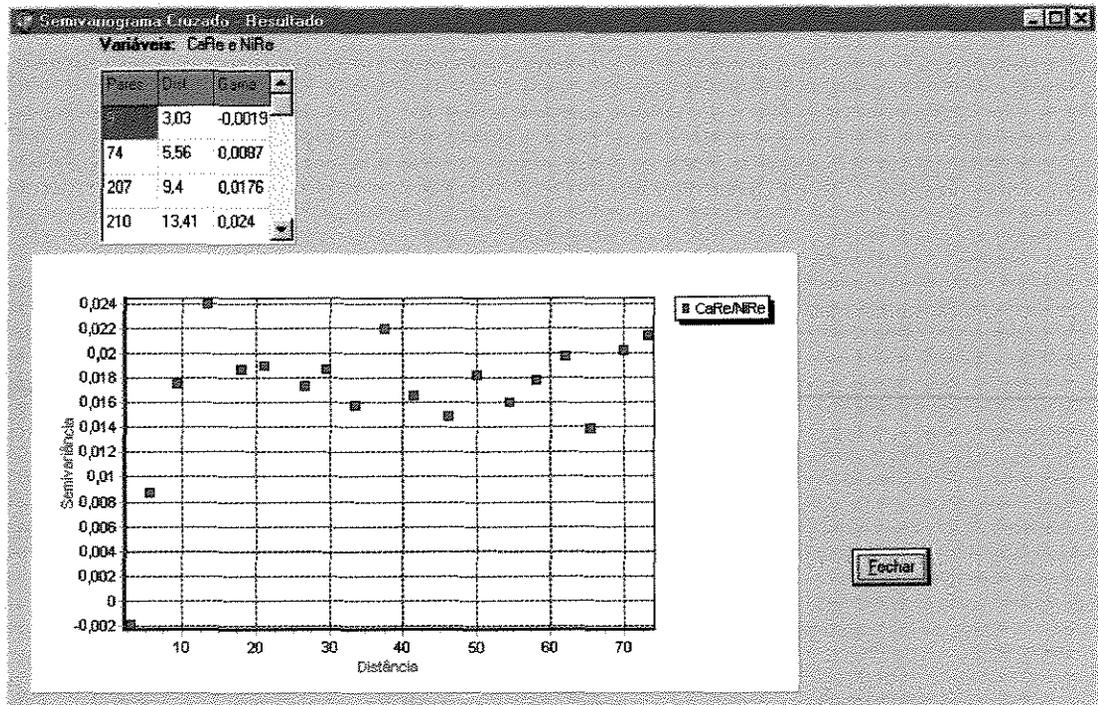


Figura 5.21 – Semivariograma cruzado da curva de resíduos de Carbono e Nitrogênio em Davis.

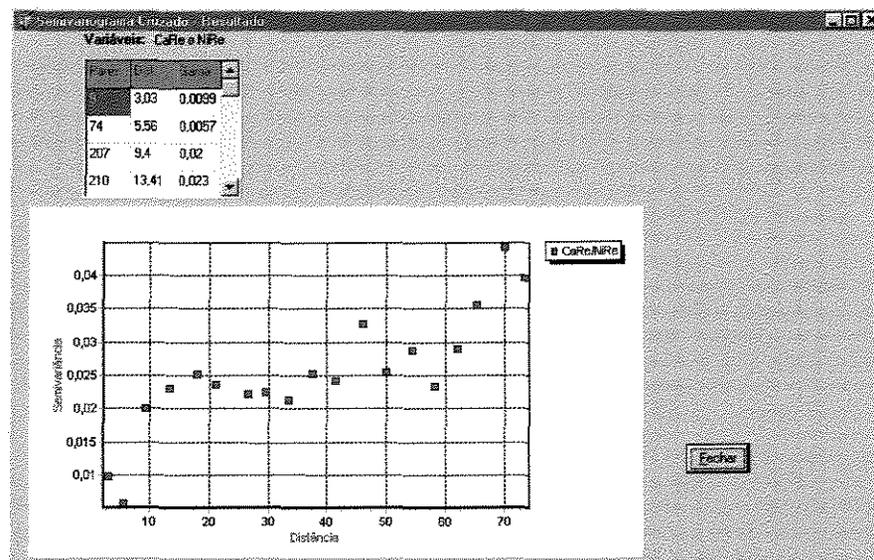


Figura 5.22 – Semivariograma cruzado da curva de resíduos de Carbono e Nitrogênio em Oakley.

A primeira porção até o alcance dos semivariogramas (Figuras 5.21 e 5.22), localizada em torno de 20 metros indica que as duas variáveis tem correlação espacial. Este

tipo de resultado viabiliza, por exemplo, uma amostragem em menor escala da variável Nitrogênio do que aquela usada para o Carbono.

A análise com a curva de resíduos das variáveis mostradas pela Figura 5.21 mostrou ser mais eficiente do que aquela com os dados originais. Ainda que esta curva apresente uma certa oscilação no início do patamar, o semivariograma cruzado entre carbono e nitrogênio em Davis pode ser ajustado.

Observa-se na curva apresentada pela Figura 5.22, que para distâncias superiores a 40 não há estabilidade do patamar. Este fato demonstra que mesmo para a curva de resíduos ou para os dados originais obteve-se um semivariograma complexo de ser modelado.

É contemplado em versões futuras o desenvolvimento dos mesmos módulos contidos no semivariograma médio experimental (ajuste de modelos matemáticos, tratamento de estacionaridade, *Jack-knifing*, etc), com a finalidade de finalizar o ciclo da análise geo-estatística. Novos métodos de krigagem são também previstos de serem incorporados ao sistema, o qual nesta versão incorpora somente a krigagem simples.

No próximo capítulo são apresentadas as considerações finais desta pesquisa.

## **CAPÍTULO 6**

### **Considerações Finais**

A operação de amostragem de solos convencional aborda a variabilidade de maneira discreta associando a grandes áreas características homogêneas. A geo-estatística vem contemplar esta abordagem, subdividindo estas zonas homogêneas em sub áreas de continuidade espacial. Este tipo de abordagem prepara o campo para o monitoramento e dentro de um escopo mais completo para o gerenciamento de modo localizado. O comportamento temporal acrescenta mais um vértice a esta abordagem.

Nesta pesquisa procurou-se disponibilizar uma ferramenta alternativa para a análise da variabilidade com uso da técnica da geo-estatística. Foi dada ênfase no processo de otimização de projetos de amostragem de solo.

Por meio da ferramenta desenvolvida tem-se um ambiente integrado para estudo das estruturas espaciais de até 16 variáveis. No estudo inicial do semivariograma médio mostrado pela Figura 4.7, a possibilidade de se desenhar conjuntamente todos os gráficos numa mesma escala constitui-se um importante recurso para análise do comportamento espacial das mesmas e características mais intrínsecas como a anisotropia.

As funcionalidades disponíveis na ferramenta contemplam quase a totalidade da biblioteca geo-estatística elaborada por Vieira et al. (1983). A única exceção foi o módulo de interpolação de Co-krigagem, definido como aquele que gera a interpolação entre duas variáveis.

A organização dos módulos do sistema foi sendo construída gradativamente por meio da interação com o pesquisador Sidney Rosa Vieira, do Instituto Agrônomo de Campinas, durante o ano de 2001. Esta estrutura reflete de certa maneira um processo cognitivo particular de se tratar fenômenos de variabilidade segundo a técnica da geo-estatística. Apesar de pessoal, o processo incorpora as principais etapas comumente usadas por outros softwares geo-estatísticos.

Procurou-se padronizar os termos e conceitos correntes da geo-estatística nas interfaces dos módulos visando produzir um ambiente sem significados ambíguos ou inconsistentes para o usuário.

O sistema foi implementado no software Delphi<sup>2</sup> 5. Este ambiente de programação integrado providencia módulos para a criação amigável de interfaces gráficas e utiliza como base a linguagem Pascal orientada a objetos (*ObjectPascal*). Testou-se o sistema nos sistemas operacionais Windows 95 e 2000 em um computador Pentium Pro com 128 Mb de memória RAM. Para manuseio do sistema exige-se o arquivo executável, gerado pelo ambiente Delphi e os arquivos da biblioteca geo-estatística de Vieira et al. (1983).

Deve-se ressaltar porém que a ferramenta é ainda um protótipo, que precisa ser melhor testada. Detectou-se que há necessidade de se traduzir as rotinas fontes em Fortran para a linguagem ObjectPascal. Tal necessidade almeja:

- Melhorar o desempenho do sistema, por meio de cálculos diretos. Rotinas como *Jack-Knifing* e krigagem demandam tempo considerável de processamento comprometendo a apresentação dos resultados pelo módulo Delphi específico.
- Incrementar o número de variáveis em estudo concomitante. Em algumas situações da Agricultura de Precisão precisa-se analisar conjuntamente um número superior a 16 variáveis.

No processo de avaliação da ferramenta foram ressaltados vários aspectos positivos em relação ao uso isolado das rotinas Fortran em ambiente MS-DOS. Percebeu-se que a rápida geração do gráfico do semivariograma médio, direcional ou cruzado permitia que um maior número de interações fossem realizadas, de modo a determinar os melhores parâmetros para a análise. A escolha mais acurada dos parâmetros do semivariograma teórico bem como a melhor vizinhança são fatores que influenciam diretamente a qualidade da interpolação bem como um projeto de amostras mais confiável.

Apesar de não terem sido realizados testes específicos, constatou-se que tempo para uma análise geo-estatística completa foi diminuído pela integração dos módulos.

Para avaliar a ferramenta, optou-se por realizar a análise geo-estatística em um conjunto de dados obtidos por Waynick e Sharp em 1919 e reutilizado por Vieira (2000), quando discorreu sobre os fundamentos da geo-estatística. O conjunto de dados estuda o comportamento do Carbono e Nitrogênio totais e sua relação (C/N) em dois campos dos Estados Unidos em um esquema de amostragem denso e variado.

---

<sup>2</sup> Delphi é um produto comercial da Borland.

A análise destes dados constituiu-se em um estudo inicial sem intenção de aprofundar os resultados no que concerne a questão agrônômica, porém resultados que ampliaram o conhecimento dos dados foram encontrados.

Com relação a análise do conjunto de dados percebeu-se a importância da análise exploratória ou descritiva como instrumento inicial de conhecimento das particularidades dos dados. Esta anunciou a correlação entre Carbono e Nitrogênio no campo de Oakley.

Os atributos Carbono e Nitrogênio nos dois campos não apresentaram estabilização de seus patamares, pela análise do semivariograma médio. A operação de retirada de tendência por meio de superfície parabólica mostrou-se eficiente na resolução deste problema.

A análise do semivariograma direcional indica uma alta continuidade na direção de 90°, confirmada pelo mapa simbólico resultante do processo de krigagem. Este tipo de resultado sustenta a decisão de, por exemplo, reamostrar o campo em maior intensidade na direção de 90 ° para alguma finalidade específica. Esta decisão pode ser complementada de forma a amostrar com menor intensidade o atributo Nitrogênio, como demonstra a análise de correlação dada pelo semivariograma cruzado.

De acordo com o WebSite do laboratório de solos da Embrapa Solos<sup>3</sup>, em novembro de 2001, uma análise de nitrogênio custava 5 reais em contrapartida aos 4 reais cobrados para análise de Carbono orgânico. Esta redução no número de amostras de Nitrogênio, que é na média uma análise mais cara, constitui-se um benefício financeiro para operações de amostragem com finalidades científica ou comercial.

O módulo para projeto de configuração de amostragem, descrito na seção 5.7 recomenda um distanciamento maior (em torno de 35 a 40 metros) para o Carbono em Davis, do que aquele usado na malha original. Num estudo que investiga também o comportamento no tempo destas variáveis, este resultado deveria ser considerado, de forma a otimizar e racionalizar o projeto de amostragem.

Os resultados obtidos pelo sistema, neste estudo, em relação aos encontrados por Vieira (2000) foram similares. Os semivariogramas, gerados pelo programa, de Carbono e Nitrogênio nos dois campos mostraram a tendência dos dados e os gráficos apresentaram a mesma quantidade de classes de distância. Porém os valores de semivariância, nestes gráficos, não coincidiram com aqueles apresentados em Vieira (2000). Como este autor não apresentou os parâmetros básicos de Distância máxima e Largura de passos utilizado

---

<sup>3</sup> Embrapa Solos – <http://www.cnps.embrapa.br>

nesta publicação, pode ser este sido este o fator da discrepância de valores. Este fato, porém, não invalidou o processo de análise geo-estatística. Outra semelhança encontrada foi o ajuste de modelo na curva de resíduos de tendência para carbono. Os parâmetros utilizados foram os mesmos de Vieira (2000). Por conseguinte, todo o processo de estatística espacial culminou em mapas de solos interpolados pelo sistema que seguem os mesmos padrões de variabilidade obtidos por Vieira.

Todos os resultados das análises geo-estatística de modo individual ou em grupo contribuíram para um maior entendimento da estrutura de variabilidade nos campos amostrados, bem com a interação entre estas variáveis. As considerações feitas nos parágrafos anteriores são advindas somente da análise dos resultados apresentados pelo sistema. Para que o sistema torne-se uma ferramenta útil e prática é interessante agregar mais informações ou dados, como: mapas de solos, imagens de sensoriamento remoto, histórico de manejo da área, etc. Um cenário o mais realista e abrangente possível é sempre desejável para embasar as tomadas de decisão relativas as operações em campo.

---

## CAPÍTULO 7

### Referências Bibliográficas

ADAMCHUK, V.I., MORGAN, M.T, ESS, D.R. An automated sampling system for measuring soil pH. **Transactions of the ASAE**, St Joseph, v. 42, n. 4, p. 885-891, 1999.

AMARAL, E. F. do. **Amostragem de solo para avaliação da fertilidade**. 4p. Instruções Técnicas, 6 - EMBRAPA-CPAF/AC, 1997.

CAMARGO, E. C. G. **Desenvolvimento, implementação e teste de procedimentos geoestatísticos (krigeagem) no sistema de processamento de informações georreferenciadas (SPRING)**. 115 p. Dissertação (Mestrado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 1997.

CHUNG, C.K., CHONG, S., VARSA, E.C. Sampling strategies for fertility on a stoy silt loam soil. **Communications in Soil Science and Plant Analysis**, Nova York, v. 26, n. 5-6, p. 741-763, 1995.

CLARK, I. **Practical geostatistics**. London: Applied Science Publishers, 1979. 130 p.

CORNELL, G.; STRAIN, T. **Delphi: segredos e soluções**. São Paulo: Makron Books do Brasil, 1996. 367 p.

DAMPNEY, P. M. R.; MOORE, M. Precision agriculture in England: current practice and research-based advice to farmer. In: Robert, P.C.; Rust, R. H. Larson, W. E. **Proceedings of the Fourth International Conference on Precision Agriculture**. Madison: ASA-CSSA-SSA, Part A, p. 661-673, 1998.

DEUTSCH, C. V., JOURNEL A.G. **GSLIB: geostatistical software library and user's guide**. New York: Oxford University Press, 1998. 369p.

ENTZ, T; CHANG, C. Evaluation of soil sampling schemes for geostatistical analyses: a case study for soil bulk density. **Canadian Journal Soil Science**, Ottawa, v. 71, p. 165-176, Maio, 1991.

FAGNANI, Maria Angela **A questão ecológica na formação do engenheiro agrícola**. 184p. Tese (Doutorado em Educação) - Faculdade de Educação. Universidade Estadual de Campinas, Campinas, 1997. Cap. 2: A modernização da agricultura. p. 35-36.

FLATMAN, G.T, YFANTIS, A. A. Geostatistical strategy for soil sampling: The survey and the census. **Environmental Monitoring and Assessment**, Dordrechtv, v. 4, n. 4, p. 335-349, 1984.

GSLIB.com (2001) Disponível: <http://www.gslib.com> Consultado em 13 de novembro de 2001.

INPRISE CORPORATION. **Borland Delphi 5 for Windows 98, Windows 95, & Windows NT: developer's guide**. Scotts Valley, 1999. Paginação irregular

ISAAKS, E. H., SRIVASTAVA, R.M. **An Introduction to applied geostatistics**, New York: Oxford University Press, 1989. 561p.

---

JIAN, X., OLEA R.A., YU, Y. Semivariogram modeling by weighted least squares. **Computers & Geosciences**, Oxford, v. 22, n. 4, p. 387-397, 1996.

JOURNEL, A.G. **Fundamentals of geostatistics in five lessons**. Short Course in Geology. Washington: American Geophysical Union, 1989, 40 p.

KANG, K.C., COHEN, S., HOLIBAUH, R., PERRY, J., PETERSON, A S. **A reuse-based software Development Methodology**, Special Report SMU/SEI-92-SR-4, Pittsburgh: Cargenie Mellon University, 1992. 36 p.

KRIGE, D.C., A statistical approach to some mine evaluation problems on the Witwatersrand. **Johannesburg Chemistry Metallurgy Mining Society South African**, v. 52, n. 6, p. 119-139, 1951.

MATHERON, G. Principles of geostatistics. **Economic Geology**, v. 58, p. 1246-1266, 1963.

McBRATNEY, A.B., PRINGLE, M. J. Spatial variability in soil - implications for precision agriculture. **Precision agriculture**, Hingham, v. 1, p. 3-30, 1997.

McBRATNEY, A.B.; WEBSTER, R.; McLAREN, R.G.; SPIERS, R.G.. Regional variation of extractable copper and cobalt in the topsoil of Southwest Scotland. **Agronomie**, Paris, v.2, n.10, p.969-982, 1982.

ODEH, I.O.A., McBRATNEY, A.B., CHITTLEBOROUGH, D.J. Design of optimal sample spacing for mapping soil using fuzzy-k-means and regionalized variable theory. **Geoderma**, Amsterdam, v. 47, p.93-122, 1990.

OLEA, R.A. **Optimum mapping techniques using regionalized variable theory**. 137p Series on Spatial Analysis, 2. University of Kansas/Kansas Geol. Survey, Lawrence 1975.

PEBESMA, E. J.; WESSELING, C. G. GSTAT: a program for geostatistical modelling, prediction and simulation. **Computer & Geosciences**, Oxford, v. 24, n. 1, p. 17-32, 1998.

PRESSMAN, R. S. **Software engineering: a practitioner's approach**. New York: McGraw-Hill, 1982. 352 p. (McGraw-Hill Series in Software Engineering and Technology).

---

PINAZZA, L.A., ALIMANDRO, R. Novo mundo rural. **AgroAnalysis**, Rio de Janeiro, v.20, n.4, p.12-24, abril, 2000.

ROBERT, P.C. Precision agriculture: research needs and status in the USA. In: STAFFORD, J.V. **Precision Agriculture'99**. Sheffield: Sheffield Academic Press, v. 1, p. 19-331, 1999.

RUSSO, D. Design of an optimal sampling network for estimation the variogram. **Soil Science Society American Journal**, Madison, v. 48, p. 708-716, 1984.

TEIXEIRA, M. A.; MAGALHÃES, P. S. G.; BRAUNBECK, O. A. Equipamento para extração de amostras indeformadas de solo. **Revista Brasileira de Ciência do Solo**, Viçosa, v. 24, n. 4, p. 693-699, 2000.

VIEIRA, S. R.; HATFIELD, T. L.; NIELSEN, D. R.; BIGGAR, J. W. Geostatistical theory and application to variability of some agronomical properties. **Hilgardia**, Berkeley, v. 51, n. 3, p. 1-75, 1983.

VIEIRA, S. R. Geoestatística em estudos de variabilidade espacial do solo. In: NOVAIS, R. F. de; ALVAREZ V., V. H; SCHAEFER, C. E. G. R. (Ed.). **Tópicos em ciência do solo**. Viçosa: Sociedade Brasileira de Ciência do Solo, v. 1, p. 1-54, 2000.

WAYNICK, D. D.; SHARP, L.T Variability in soils and its significance to past and future soil investigations. II. Variation in nitrogen and carbon in field soils and their relation to the accuracy of field trials. **Agriculture Science**, v. 4, p. 121-139, 1919.

WOLLENHAUPT, N. C.; MULLA, D. J.; CRAWFORD, C. A. G. Soil sampling and interpolation techniques for mapping spatial variability of soil properties. In: PIERCE, F. J. L.; SADLER, E. J. **The state of site-specific management for agriculture**. Madison: ASA/CSSA/SSA, p. 19-53, 1997.

---