



Governo do Estado de São Paulo  
Secretaria de Agricultura e Abastecimento  
Agência Paulista de Tecnologia dos Agronegócios  
**Instituto Agrônomo**

# **CANA-DE-AÇÚCAR**

Editores

Leila Luci DINARDO-MIRANDA  
Antonio Carlos Machado de VASCONCELOS  
Marcos Guimarães de Andrade LANDELL

Campinas (SP)  
2008

**Governador do Estado**

José Serra

**Secretário de Agricultura e Abastecimento**

João de Almeida Sampaio Filho

**Secretário-Adjunto**

Antônio Júlio Junqueira de Queiroz

**Chefe de Gabinete**

Antonio Vagner Pereira

**Coordenador da Agência Paulista de Tecnologia dos Agronegócios**

João Paulo Feijão Teixeira

**Diretor Técnico de Departamento do Instituto Agrônômico**

Orlando Melo de Castro

**Diretor do Centro de Cana**

Marcos Guimarães de Andrade Landell

**Comitê Editorial do Instituto Agrônômico**

Oliveiro Guerreiro Filho, Ricardo Marques Coelho e Cecília Alzira Ferreira Pinto Maglio

**Equipe de Edição**

**Revisão de vernáculo**

Bia Nogueira e Vitorio Barato Neto

**Revisão de referências bibliográficas**

Maria Bernardete Malerbo

**Secretaria para correções ortográficas**

Lívia Vasconcelos

**Diagramação**

João Paulo de Carvalho

**Capa**

Lau Batista – Fazedores de imagem

Ficha elaborada pelo Núcleo de Informação e Documentação do Instituto Agrônômico

C212 Cana-de-açúcar / (eds.) Leila Luci Dinardo-Miranda; Antonio Carlos Machado de Vasconcelos; Marcos

Guimarães de Andrade Landell

Campinas: Instituto Agrônômico, 2008.

882 p

ISBN: 978-85-85564-17-9

1. Cana-de-açúcar 2. Instituto Agrônômico 3. Cultura da cana-de-açúcar  
I. Dinardo-Miranda, Leila Luci II. Vasconcelos, Antonio Carlos Machado de  
III. Landell, Marcos Guimarães de Andrade IV. Título

CDD. 633.61

**Todos os direitos reservados**

A reprodução não autorizada desta publicação, no todo ou em parte, constitui violação do copyright (Lei nº 9.610).

Centro de Cana – Instituto Agrônômico

Rodovia SP333, km 321

Caixa Postal 206

14001-970 Ribeirão Preto (SP)

## CAPÍTULO 39

# APLICAÇÕES DE GEOESTATÍSTICA EM PESQUISAS COM CANA-DE-AÇÚCAR

Sidney Rosa Vieira, Mauro Alexandre Xavier e Célia Regina Grego

## 1. CONCEITOS GERAIS DE GEOESTATÍSTICA

### 1.1. O que é geoestatística

Todas as amostras retiradas de algum ponto no espaço ou no tempo devem ser consideradas como parte de uma função contínua e são pontos discretos desta função. Por isso, pode-se dizer que todas as amostras são, de algum modo, relacionadas com seus vizinhos, mesmo que não se conheça exata e matematicamente qual é a expressão para este relacionamento. Nessa condição, pode-se dizer que amostras separadas por pequenas distâncias são mais parecidas umas com as outras do que amostras separadas por grandes distâncias (VIEIRA, 2000). É óbvio que isto é totalmente dependente da intensidade de amostragem em relação à escala de trabalho. Geoestatística é a maneira mais correta e contém as ferramentas ideais para analisar dados com dependência espacial, ou seja, cujos vizinhos próximos são mais semelhantes entre si do que aqueles separados por distâncias maiores. A condição absoluta para o uso de geoestatística é que se tenha as coordenadas geográficas da posição de onde as amostras foram retiradas. Sem o conhecimento das coordenadas dos pontos onde foram efetuadas as medidas, sejam elas obtidas por GPS, por métodos topográficos, como distâncias a partir de uma origem arbitrária, ou qualquer outro método, não é possível fazer uso da geoestatística.

A geoestatística surgiu na África do Sul em trabalho com dados de concentração de ouro, no qual se concluiu que não era possível encontrar sentido nas variâncias, se não fosse levado em conta a distância entre as amostras. Baseando-se nestas observações, foi desenvolvida uma teoria chamada de teoria das variáveis regionalizadas, que contém os fundamentos da geoestatística. Uma variável regionalizada é definida como uma função espacial numérica, que varia de um local para outro com uma continuidade aparente e cuja variação não pode ser repre-

sentada por uma função matemática simples (VIEIRA *et al.*, 1983). Essa continuidade, ou dependência espacial, pode ser estimada através do semivariograma. A geoestatística tem um método de interpolação chamado krigagem, que usa a dependência espacial entre amostras vizinhas, expressa no semivariograma, para estimar valores em qualquer posição dentro do campo, sem tendência e com variância mínima (VIEIRA, 2000).

### 1.2. Variabilidade

O que é variabilidade? Esta pergunta pode ter várias respostas. O Dicionário *Michaelis* define variabilidade como qualidade ou caráter de variável. Pode-se dizer que a variabilidade de números significa que podem existir medidas diferentes de uma mesma grandeza e, nesse caso, a variabilidade pode ser expressa pelo coeficiente de variação. Já a genética define variabilidade como a característica das plantas em apresentar diferenças para um mesmo caráter. Por outro lado, a variabilidade espacial significa que medidas tomadas em locais diferentes podem apresentar valores diferentes. O fato incontestável é que, caso se façam várias medições de um mesmo parâmetro, certamente haverá variabilidade.

Experimentos com culturas em campo, em geral, constam de uma série de tratamentos para os quais se espera que a cultura irá reagir de alguma maneira, proporcional ao tratamento aplicado. Estes modelos baseiam-se em técnicas como casualização e repetição, e o melhor conhecimento de funções de distribuição e de amostragem ao acaso, desprezando-se assim as coordenadas geográficas do ponto amostrado. Este procedimento somado à distribuição normal de freqüências era, e ainda é, usado para assumir independência entre as amostras, e assim garantir a validade do uso da média e do desvio-padrão para representar o fenômeno. Estes conceitos da estatística clássica e seus fundamentos podem ser encontrados em livros textos de esta-

tística experimental (SNEDECOR e COCHRAN, 1967). Por outro lado, a distribuição normal não garante, de maneira alguma, a independência entre amostras, a qual pode ser verificada através da autocorrelação. A principal razão para isto é que o cálculo da frequência de distribuição não leva em conta a distância na qual as amostras foram coletadas no campo.

Para dados como estes, para os quais não se tem conhecimento das coordenadas dos locais de onde as amostras foram tomadas, as medidas de dispersão, como coeficiente de variação (CV), expressam simplesmente a variabilidade entre os números. Um exemplo de dados como estes pode ser a altura de um grupo de pessoas.

### 1.3. Variabilidade espacial

A variabilidade espacial de propriedades do solo vem sendo uma das grandes preocupações de pesquisadores há muito tempo. Praticamente desde que os pesquisadores iniciaram as amostragens ou medições no campo ficou evidente que alguma medida deveria ser tomada para que se pudesse confiar nos números obtidos. A variação de qualquer fenômeno, no espaço ou no tempo, seja ela causada por processos naturais ou por ações impostas pelo homem, sempre existiu e trouxe problemas para qualquer usuário do meio ambiente.

Por outro lado, tem havido um grande avanço nos métodos empregados tanto para medições no campo, como nos equipamentos e métodos de laboratório. Isto tem causado uma melhoria na sensibilidade dos medidores. Devido a isto, medições de muitas características que antes se mostravam semelhantes, hoje em dia revelam diferenças que podem alterar os resultados. Dito de maneira simplista, é como se há alguns anos as régua para medir distâncias tivessem somente graduação a cada dez centímetros, enquanto hoje apresentam graduação que permite medir milímetros.

Para analisar ou expressar a variabilidade espacial de alguma propriedade, é absolutamente necessário ter as coordenadas dos pontos de onde as amostras foram retiradas (VIEIRA, 2000).

## 2. AMOSTRAGEM

À medida que os conhecimentos científicos evoluíram e, conseqüentemente, também evoluíram os métodos de medição, os métodos de cálculo numérico, a velocidade, a eficiência e a pre-

cisão de equipamentos de computação, foi se tornando evidente a necessidade de se conhecer e se considerar a variação dos dados para quaisquer usos das medidas efetuadas. Ao mesmo tempo, técnicas desenvolvidas para análise de dados de ciências sociais e engenharia elétrica foram adaptadas para análise de dados em uma dimensão, no espaço ou no tempo, revelando aspectos da variação não disponíveis por outras técnicas. Dessa maneira, tornou-se conhecido que as medidas variam no espaço e no tempo, não necessariamente porque os instrumentos de medição têm erros, mas também porque as medidas variam naturalmente. De fato, existe, para qualquer medida efetuada, no espaço ou no tempo, uma continuidade perfeita, constando de uma variação gradual a qual pode ou não ser conhecida na escala de trabalho. Assim, não é nenhum exagero afirmar que, a única solução para não se encontrar variação em medições é tomar uma única medição, embora essa possa não ser uma escolha muito inteligente. Dessa maneira, toda amostragem deverá considerar que as amostras retiradas de algum ponto no espaço ou no tempo fazem parte de uma função contínua, e são pontos discretos desta função (VIEIRA *et al.*, 1983).

É praticamente impossível planejar uma amostragem com sucesso sem nenhum conhecimento prévio sobre o ambiente que se quer caracterizar. Em termos ideais, o planejamento da amostragem deve fazer uso de todas as facilidades e todo o conhecimento a respeito da área. Tudo depende da escala ou resolução desejada, uma vez que a variabilidade que se está investigando pode acontecer em milímetros ou quilômetros. Seu conhecimento adequado dependerá dos recursos disponíveis para caracterização do ambiente.

Em geral, a amostragem pode ser **adequada** se o espaçamento entre as amostras for menor do que as "manchas" que se quer caracterizar; **inteligente**, caso se faça uso da variabilidade de variáveis indicativas para alterar a densidade de amostragem. Assim, é desperdício de recursos amostrar com grande intensidade locais para os quais pode se prever (baseando-se em variáveis indicativas como topografia, cor do solo, pH, por exemplo) que irão ter baixa variabilidade. Por outro lado, locais onde se pode prever que terão grande variabilidade deverão ser amostrados com maior intensidade.

Existe também uma alternativa de se amostrar em duas transeções (linhas retas com espaçamento regular entre amostras) perpendiculares entre si e inferir o espaçamento ideal para as duas direções. O risco que se corre com esta

alternativa é de que nem todas as posições de transeções são iguais quanto à variabilidade.

## 2.1. Número de amostras

Vários fatores podem influir na seleção do número de amostras a ser adotado para qualquer projeto, entre os quais: recursos financeiros e humanos disponíveis, objetivo da amostragem, tempo disponível. Caso existam ilimitados recursos financeiros e humanos e tempo sem limite, situação esta que raramente ocorre na vida real, então a amostragem poderá fazer uso total destes.

A tradicional maneira de se calcular um número de amostras baseando-se na variância de números previamente obtidos (SNEDECOR e COCHRAN, 1967), simplesmente oferece a possibilidade de se saber o número de amostras para se estimar o valor médio dentro de níveis especificados, caso uma nova amostragem seja efetuada na mesma área e nas mesmas condições. É desnecessário dizer que esta é uma situação raramente ocorrente.

Por esta razão, um conhecimento mínimo sobre a variável que está sendo amostrada é imprescindível no momento da decisão do espaçamento entre amostras. A variabilidade da variável em estudo somente será suficientemente caracterizada, se o espaçamento entre amostras for fixado usando-se este conhecimento. Os recursos financeiros, humanos e de tempo necessários para efetuar a amostragem são outros fatores que podem influenciar no número final de amostras. Portanto, a definição da quantidade de amostras é tarefa difícil e um número não pode ser simplesmente fixado sem que todos os fatores citados sejam adequadamente avaliados.

## 2.2. Espaçamento entre amostras

Assumindo que a amostragem deverá ser suficiente para se analisar a variabilidade espacial, então o espaçamento entre amostras deverá ser menor do que as "manchas" que se quer caracterizar. O grande problema está em se definir qual é a dimensão aproximada destas "manchas". Uma alternativa é efetuar a amostragem de uma variável indicativa em transeções. Quando se quer avaliar algum caráter na cultura de cana, por exemplo, outra alternativa é plantar uma cultura indicativa como *Crotalaria juncea*, sem adubação e avaliar a variabilidade do caráter na *C. juncea* (SALVIANO *et al.*, 1998).

Há que se considerar também que variáveis, como teor de fósforo, infiltração de água no solo e resistência à penetração têm, normalmente, variabilidade muito maior do que densidade e pH do solo. A amostragem deverá ser projetada para caracterizar a variabilidade daquelas que mais variam.

Somando-se as idéias apresentadas com bom senso, acredita-se que seja possível estabelecer uma densidade de amostragem que poderá satisfazer às mais variadas necessidades. O fato mais evidente é que não existem fórmulas mágicas que possam ser aplicadas sem um conhecimento prévio do ambiente que se quer amostrar.

## 2.3. Esquemas de amostragem

Os esquemas mais comuns de amostragem são:

- a) Grade retangular ou quadrada marcada no campo com trena, teodolito ou estação total (figura 1).
- b) Grade retangular ou quadrada com "ilhas" com concentração de amostras mais próximas. Estas "ilhas" são usadas com a intenção de fornecer informações em pequenas distâncias sem aumentar exageradamente o número de amostras (figura 2).
- c) Grade trapezoidal, na qual as colunas pares têm um deslocamento de metade da distância entre linhas para a direita e metade para cima. Assim, a menor distância entre pontos será a diagonal (figura 3).
- d) Ao acaso é a amostragem espacialmente distribuída sem nenhuma ordem. Suas coordenadas são normalmente determinadas com aparelhos GPS (figura 4).

Para qualquer um destes esquemas, é importante que sejam fixadas referências, pelo menos nos quatro cantos da área, para próximas amostragens nos mesmos locais e que sejam anotadas as modificações efetuadas no campo em relação ao planejado.

## 3. SEMIVARIOGRAMA

### 3.1. Definição

O semivariograma é a grandeza mais apropriada para decidir se a dependência espacial existe ou não. Dependência espacial existe quando a amostragem é efetuada em espaçamentos menores do que as manchas na área. Portanto, a análise dos dados resultará em um

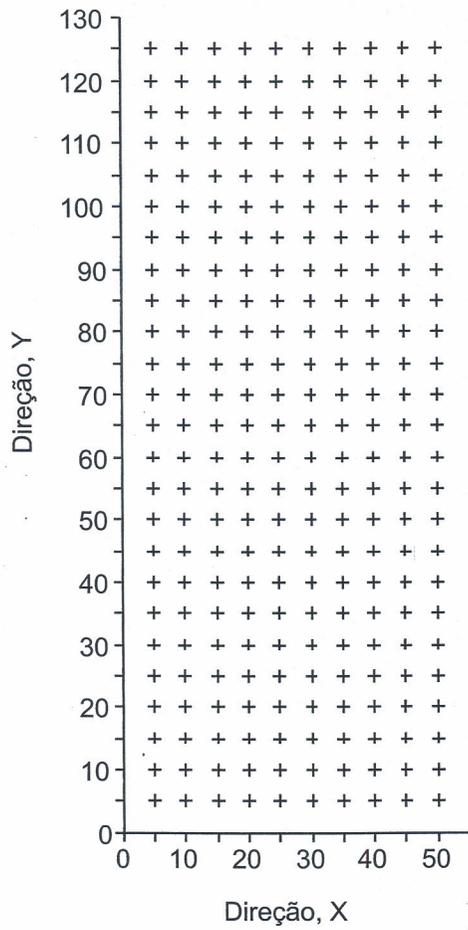


Figura 1. Amostragem em grade quadrada.

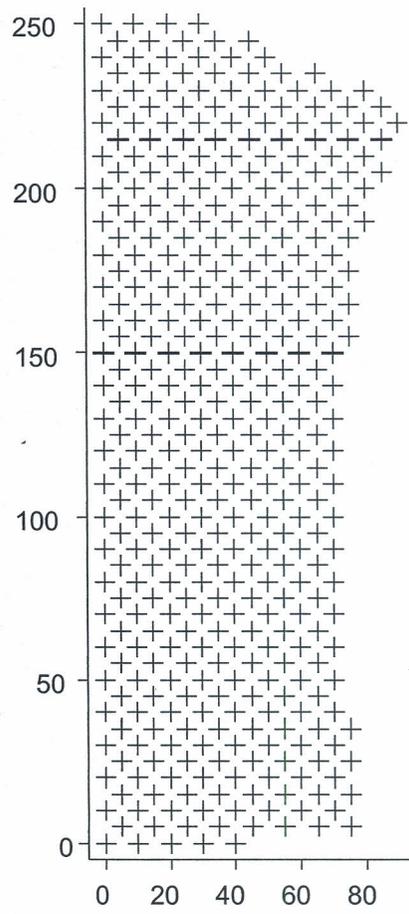


Figura 3. Amostragem em grade trapezoidal.

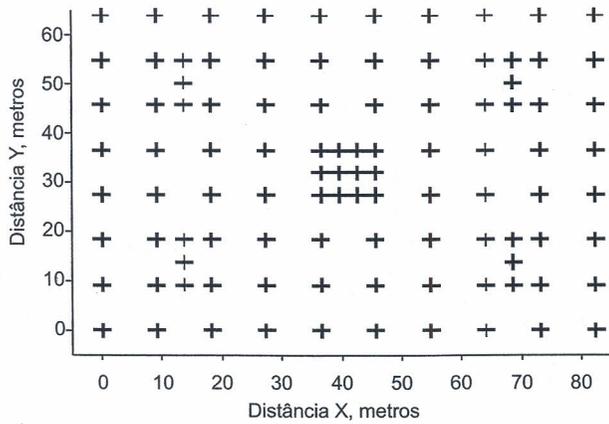


Figura 2. Amostragem em grade com "ilhas".

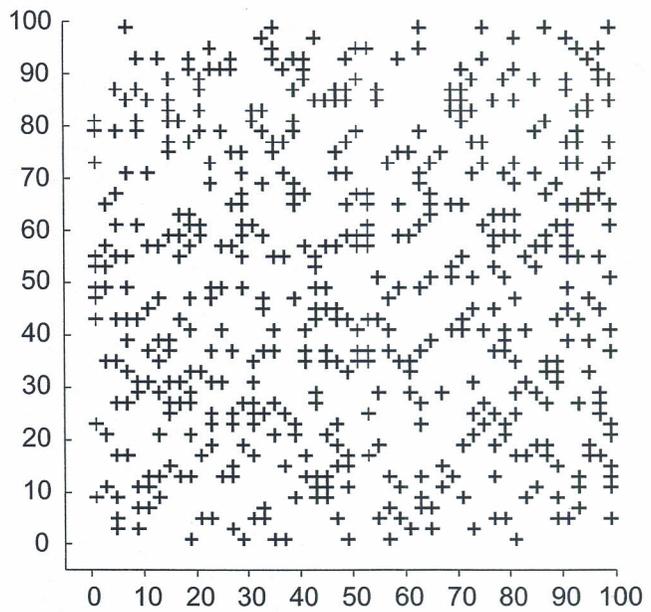


Figura 4. Amostragem ao acaso.

semivariograma que apresenta crescimento na semivariância até uma determinada distância, a partir da qual ele se estabiliza (VIEIRA, 2000). Desta maneira, a semivariância “depende” da distância. Caso o semivariograma não tenha nenhum crescimento com a distância e se constitua somente de flutuação em torno do valor da variância, então as amostras são independentes umas das outras, e têm distribuição totalmente ao acaso. Nesta condição, diz-se que a variável tem efeito pepita puro.

O semivariograma é, por definição:

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

Esta equação pode ser lida como a metade do valor esperado da diferença quadrada entre valores separados por um vetor  $h$ . O valor esperado,  $E$ , de uma variável é o valor que mais ocorre. Quando se assume que uma variável tem distribuição normal, o valor que mais ocorre é a média. O vetor  $h$ , como tal, tem módulo, direção e sentido, o que significa que podem haver diferenças na variabilidade para direções diferentes. Se isto ocorrer, a variável estudada é dita anisotrópica. Em condições isotrópicas (variabilidade igual em todas as direções),  $h$  passa a ser simplesmente a distância de separação entre amostras (VIEIRA *et al.*, 1983).

É razoável pensar que a equação do semivariograma realmente expressa o grau de semelhança entre vizinhos por ela ser uma “semivariância”, que pode variar em razão da distância entre amostras.

### 3.2. Estimativa

O semivariograma pode ser estimado por:

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

em que  $N(h)$  é o número de pares de valores medidos  $Z(x_i)$ ,  $Z(x_i+h)$ , separados por um vetor  $h$ . O gráfico de  $\gamma^*(h)$  versus os valores correspondentes de  $h$ , chamado semivariograma, é uma função do vetor  $h$  e, portanto, depende de ambos, magnitude e direção de  $h$ . Quando o gráfico do semivariograma é idêntico para qualquer direção de  $h$ , ele é chamado isotrópico e representa uma situação bem mais simples do que quando é anisotrópico. Neste último caso, o semivariograma

deve sofrer transformações antes de ser usado. É importante notar que a maioria das variáveis de ciência do solo poderá ter um comportamento anisotrópico, isto é, mudar de maneira diferente para direções diferentes. É óbvio que isso depende muito da propriedade em estudo, das dimensões do campo de estudo e do tipo de solo envolvido. Existem algumas maneiras de se transformar um semivariograma anisotrópico em isotrópico. Em geral, a precisão da interpolação ou o tipo de hipótese satisfeita não são afetados se, em vez de se preocupar com a escolha do método de transformação da anisotropia, apenas se limitar a faixa de distância na qual se utiliza o semivariograma. De qualquer maneira, é sempre aconselhável examinar semivariogramas para várias direções, antes de tomar decisões. As principais direções que devem ser examinadas são:  $0^\circ$ , na direção do eixo X;  $90^\circ$ , na direção do eixo Y e  $45^\circ$ , nas duas diagonais. O método *jack-knifing* é também de grande utilidade para se determinar a faixa de distância na qual o semivariograma pode ser, na prática, considerado isotrópico (VIEIRA, 2000).

De qualquer maneira, sob isotropia ou não, a equação (2) é a que é usada para o cálculo do semivariograma. Os programas de computador utilizados para calcular o semivariograma, simplesmente calculam aquela equação (VIEIRA *et al.*, 2002). Quando os dados forem coletados em transecto, o semivariograma é unidirecional e nada pode ser dito a respeito de anisotropia, mas por outro lado é uma preocupação a menos.

### 3.3. Parâmetros e componentes de um semivariograma

Por definição,  $\gamma(0) = 0$ , como pode ser visto pela equação (2), quando  $h = 0$ . Entretanto, na prática, à medida que  $h$  tende para 0 (zero),  $\gamma(h)$  se aproxima de um valor positivo chamado efeito pepita (*nugget effect*), que recebe o símbolo  $C_0$ .

Resultados com valores de efeito pepita maiores que zero foram encontrados para precipitação pluvial (VIEIRA *et al.*, 2002), pH (CAMPBELL, 1978) e distribuição de tamanho de partículas de solo (VIEIRA, 1997). O valor de  $C_0$  revela a descontinuidade do semivariograma para distâncias menores do que a menor distância entre as amostras (VIEIRA, 2000). Parte dessa descontinuidade pode ser também devida a erros de medição, mas é impossível quantificar qual contribui mais, se os erros de medição ou a variabilidade em uma escala menor do que aquela amostrada.

À medida que  $h$  aumenta,  $\gamma(h)$  também aumenta até um valor máximo no qual se estabiliza. Esse valor no qual  $\gamma(h)$  se estabiliza chama-se patamar (*sill*), e é aproximadamente igual à variância dos dados,  $\text{VAR}(z)$ . A distância na qual  $\gamma(h)$  atinge o patamar é chamada de alcance (*range*), recebe a denominação de  $a$ , e é a distância limite de dependência espacial. Medições localizadas a distâncias maiores que  $a$  têm distribuição espacial aleatória e por isto são independentes entre si. Para essas amostras, a estatística clássica pode ser aplicada sem restrições. Por outro lado, amostras separadas por distâncias menores que  $a$  são correlacionadas umas às outras, o que permite que se faça interpolações para espaçamentos menores do que os amostrados. Dessa maneira, o alcance  $a$  é a linha divisória para a aplicação de geoestatística ou estatística clássica, e, por isso, o cálculo do semivariograma deveria ser feito rotineiramente para dados de campo para garantir as hipóteses estatísticas sob as quais serão analisados. Dados que apresentarem semivariogramas semelhantes ao da figura 5, muito provavelmente poderão ser estacionários de ordem 2, porque têm um patamar claro e definido e, com toda certeza, estarão sob a hipótese intrínseca.

A figura 5 mostra um semivariograma típico com características bem próximas do ideal, as quais serão discutidas a seguir. Seu comportamento representa o que, intuitivamente, se deve esperar de dados de campo. É de se esperar 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 separa-

das por grandes distâncias. Dessa maneira, é de se esperar que  $\gamma(h)$  aumente com a distância  $h$ .

Se o semivariograma, ao invés de ser crescente e dependente de  $h$  como o mostrado na figura 5, for constante e igual ao patamar para qualquer valor de  $h$ , então tem-se um efeito pepita puro ou ausência total de dependência espacial. Isto significa que o alcance  $a$ , para os dados em questão, é menor do que o menor espaçamento entre amostras. Para esses dados, tem-se uma distribuição espacial completamente aleatória, e a única estatística aplicável é a estatística clássica.

É bastante comum um semivariograma que, partindo do valor do efeito pepita,  $C_0$ , cresce além do valor do patamar (*sill*), até uma determinada distância e depois cai e apresenta flutuações abaixo do valor do patamar. Pode até apresentar flutuações abaixo do valor do patamar para pequenas distâncias. Isso é indicação de periodicidade (repetição de valores semelhantes em determinadas distâncias) nos dados. A periodicidade nos dados requer um tratamento específico chamado densidade espectral.

Outro tipo de semivariograma que pode ocorrer é aquele que cresce, sem limites, para todos os valores de  $h$  calculados. Um exemplo é o semivariograma para pH do horizonte A, mostrado na figura 6, identificado no gráfico pelos quadrados azuis. Esse semivariograma indica a presença de fenômeno com capacidade infinita de dispersão, o qual não tem variância finita e para o qual a covariância não pode ser definida. Ele indica, também, que o tamanho do campo amostrado não foi suficiente para exibir toda a variância, e é provável que exista uma grande tendência nos

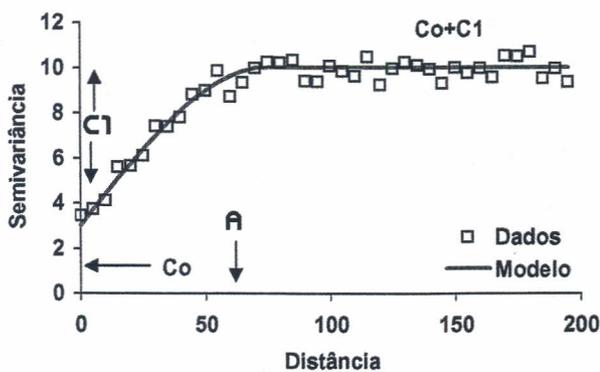


Figura 5. Semivariograma.

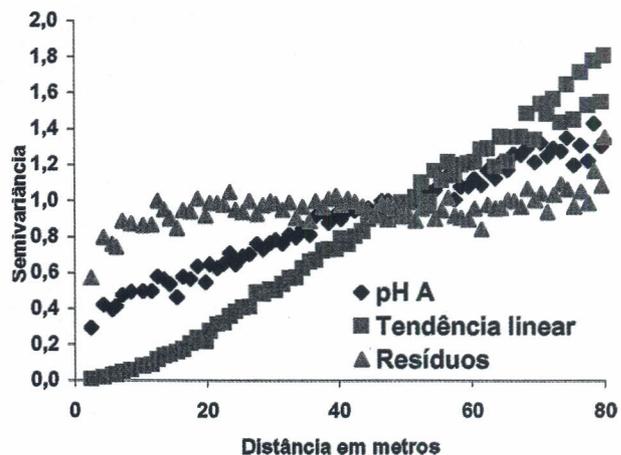


Figura 6. Semivariograma para pH do horizonte A, tendência linear e resíduos.

dados em determinada direção. Se isto puder ser constatado, tem-se duas alternativas distintas: a) remove-se a tendência e trabalha-se nos resíduos para examinar se se enquadram nas hipóteses de estacionariedade de ordem 2 ou intrínseca; b) trabalha-se com hipótese de tendência nos dados originais, segundo recomenda VIEIRA (2000). Por simplicidade, deve-se preferir a primeira alternativa. Um método bastante eficiente para retirada da tendência é pela superfície de tendência. Se, após retirar a tendência, não houver nenhuma dependência espacial expressa no semivariograma dos resíduos, isto significa que a superfície de tendência encontrada é a melhor representação espacial do fenômeno. Este procedimento foi adotado com os dados referentes ao semivariograma apresentado na figura 7, e o resultado foi um semivariograma com patamar definido mostrado para resíduos de tendência parabólica com triângulos verdes.

#### 3.4. Modelos para o semivariograma

O gráfico do semivariograma experimental,  $\gamma(h)$  versus  $h$ , calculado usando-se a equação (2), mostrará uma série de pontos discretos de  $\gamma(h)$ , correspondendo a cada valor de  $h$  e para o qual uma função contínua deve ser ajustada. Neste capítulo serão discutidos apenas os principais.

O ajuste de um modelo teórico ao semivariograma experimental é um dos aspectos mais importantes das aplicações da teoria das variáveis regionalizadas e pode ser uma das maio-

res fontes de ambigüidade e polêmica nessas aplicações. Todos os cálculos de geoestatística dependem do valor do modelo do semivariograma para cada distância especificada. Por isso, se o modelo ajustado estiver errado, todos os cálculos seguintes também estarão. Outro ponto importante é que o ajuste de curvas com calculadoras ou computador, empregando-se métodos automáticos, embora possa ser usado, não é necessário. Existem programas comerciais que fazem ajuste pelo método dos quadrados mínimos, considerando-se o número de pares como pesos nas ponderações. Da mesma maneira, esses também podem ser usados, embora não seja necessário. O método de tentativa e erro, aliado ao exame dos resultados do *jack-knifing* (validação dos modelos) são suficientes (VIEIRA *et al.*, 1983). Como regra, quanto mais simples puder ser o modelo ajustado, melhor, e não se deve dar importância excessiva a pequenas flutuações que podem ser artifícios referentes a um pequeno número de dados. É importante que o modelo ajustado represente a tendência de  $\gamma(h)$  em relação a  $h$ . Matematicamente e estatisticamente, é obrigatório que o modelo ajustado tenha positividade definida condicional, embora o significado desta exigência esteja além dos objetivos deste trabalho. Além disso, essa condição não é fácil de entender nem de testar. A grosso modo, o modelo que satisfaça a condição acima garantirá que  $\gamma(h) > 0$  e  $\gamma(-h) = \gamma(h)$ , qualquer que seja  $h$ . De qualquer modo, os modelos apresentados neste trabalho satisfazem a exigência de positividade definida condicional e são suficientes para praticamente qualquer situação.

Dependendo do comportamento de  $\gamma(h)$ , para altos valores de  $h$ , os modelos podem ser classificados em: modelos com patamar (*sill*) e modelos sem patamar. Os modelos sem patamar representam um processo para o qual o campo amostrado não foi suficientemente grande para mostrar toda a sua variabilidade e nesta condição não se enquadra na hipótese intrínseca que prevê a estacionariedade da média e da variância em razão de distâncias de separação entre amostras. Esses modelos correspondem a fenômenos que têm uma capacidade infinita de dispersão e, por isso, não têm variância finita e a covariância não pode ser definida. Devido a isto, os modelos sem patamar não serão tratados neste trabalho.

Nos modelos seguintes,  $C_0$  é o efeito pepita,  $C_0 + C_1$  é o patamar e  $a$  é o alcance do semivariograma.

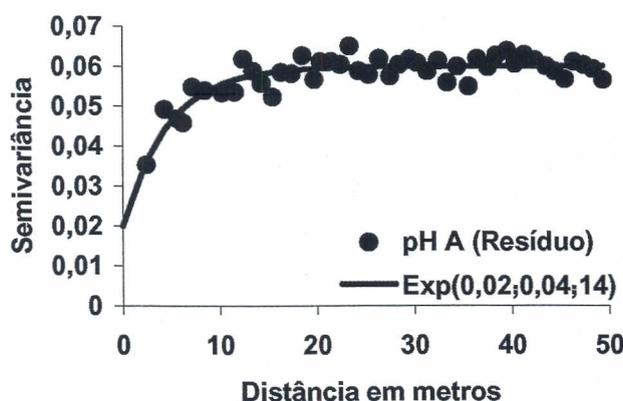


Figura 7. Semivariograma para resíduos de tendência de pH do horizonte A com modelo ajustado.

**a) Modelo linear:**

$$\gamma(h) = C_0 + \frac{C_1}{a} h \quad 0 < h < a \quad (3)$$

$$\gamma(h) = C_0 + C_1 \quad h > a$$

onde  $C_1/a$  é o coeficiente angular para  $0 < h < a$ . Nesse modelo, o patamar é determinado por inspeção; o coeficiente angular,  $C_1/a$ , é determinado pela inclinação da reta que passa pelos primeiros pontos de  $g(h)$ , dando-se maior peso àqueles que correspondem ao maior número de pares; o efeito pepita,  $C_0$ , é determinado pela interseção da reta no eixo  $\gamma(h)$ ; o alcance,  $a$ , é o valor de  $h$  correspondente ao cruzamento da reta inicial com o patamar; e  $C_1 = \text{patamar} - C_0$ .

**b) Modelo esférico:**

$$\gamma(h) = C_0 + C_1 \left[ \frac{3}{2} \left( \frac{h}{a} \right) - \frac{1}{2} \left( \frac{h}{a} \right)^3 \right] \quad 0 < h < a \quad (4)$$

$$\gamma(h) = C_0 + C_1 \quad h > a$$

O modelo esférico é obtido selecionando-se os valores do efeito pepita,  $C_0$ , e do patamar,  $C_0 + C_1$ , depois passando-se uma reta que intercepte o eixo  $y$  em  $C_0$  e seja tangente aos primeiros pontos próximos de  $h = 0$ . Essa reta cruzará o patamar à distância,  $a' = 2/3 a$ . Assim, o alcance,  $a$ , será  $a = 3a'/2$ . O modelo esférico é linear até aproximadamente  $1/3 a$ .

**c) Modelo exponencial:**

$$\gamma(h) = C_0 + C_1 \left[ 1 - \exp\left(-3 \frac{h}{a}\right) \right] \quad 0 < h < d \quad (5)$$

onde  $d$  é a máxima distância na qual o semivariograma é definido. Uma diferença fundamental entre o modelo exponencial e o esférico é que o exponencial atinge o patamar apenas assintoticamente, enquanto o modelo esférico o atinge no valor do alcance. O parâmetro  $a$  é determinado visualmente como a distância após a qual o semivariograma se estabiliza. Os parâmetros  $C_0$  e  $C_1$  para os modelos exponencial

e gaussiano (explicado a seguir) são determinados da mesma maneira que para o esférico.

**d) Modelo gaussiano:**

$$\gamma(h) = C_0 + C_1 \left[ 1 - \exp\left(-3 \left( \frac{h}{a} \right)^2 \right) \right] \quad 0 < h < d \quad (6)$$

Alguns fenômenos podem ter semivariogramas que mostram estrutura entrelaçada, ou seja, mais de um patamar e mais de um alcance. Isso acontece quando se tem diferentes escalas de variabilidade nos dados.

A figura 8 mostra os principais modelos usados. O modelo *hole effect* representa um fenômeno que tem periodicidade.

#### 4. UM EXEMPLO DE APLICAÇÃO DE CONCEITOS DE GEOESTATÍSTICAS: MAPEAMENTO DA VARIABILIDADE DE SOLO E DA PRODUTIVIDADE DE VARIEDADES DE CANA EM CAMPO

Ensaio de competição de variedades pressupõe que o efeito de variabilidade do solo não será suficiente para alterar as respostas das variedades. No entanto, a uniformidade de propriedades do solo raramente existe e não pode ser simplesmente pressuposta sem uma adequada averiguação.

Uma das maneiras úteis de averiguar a variabilidade antes do plantio de parcelas em ensaios de competição de variedades de cana-de-açúcar é avaliar certos parâmetros em *C. juncea*, plantada na área no verão anterior, sem adubação. Desta maneira, a variabilidade em parâmetros da crotalária poderá ser usada para determinar regiões no campo onde o solo é mais uniforme do que em outras.

Como exemplo, é apresentado a seguir um conjunto de dados coletados no Centro de Cana em Ribeirão Preto, em março de 2005.

O local escolhido foi uma faixa entre dois terraços, com aproximadamente 50 m de largura e 350 m de comprimento, cuja coleta de amostras foi efetuada a cada 10 m, segundo mostra a figura 9.

Assim, na área, inicialmente cultivada com crotalária, foram marcados 205 pontos para amostragem, nos quais se coletaram informações sobre altura, massa verde e massa seca de plantas. Após a colheita da crotalária, plantou-se cana-de-açúcar no terço final do campo experimental (a partir da coordenada  $Y = 200$  da figura 9) e a amostragem dos caracteres de plantas de cana

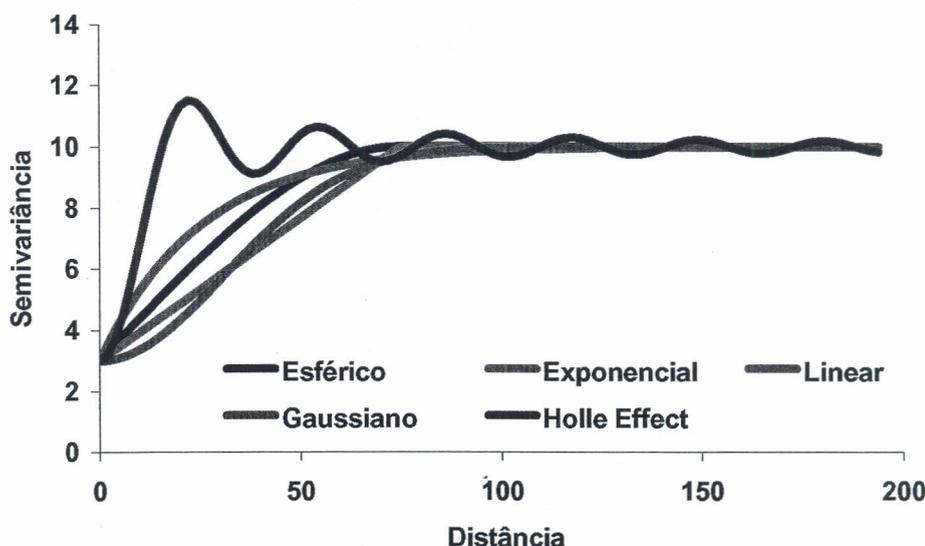


Figura 8. Modelos de ajuste ao semivariograma.

(perfilhos por metro e produtividade) foi feita em 97 pontos.

As perguntas que se queria responder então eram:

- a) Seria possível colocar um experimento de competição de variedades neste campo e supor que as respostas eram devidas ao potencial genético da planta?
- b) Seria possível usar mapas de produção de massa de crotalária para selecionar locais onde se pudesse alocar parcelas?

A geoestatística pode ajudar a responder a estas perguntas, como mostrado a seguir.

A Tabela 1 expõe os resultados de estatística descritiva para altura, massa verde e massa seca de crotalária. Nota-se que, embora o coeficiente de variação (CV) da variável altura seja baixo (8,3%), há uma variação extremamente grande entre máximo e mínimo. Isto indica que pode haver uma grande variabilidade de solo. Isto é melhor ilustrado pelos dados de massa verde, que apresentaram um CV duas vezes maior (16,2%) e uma variação de quase quatro vezes entre o valor máximo e mínimo. As duas variáveis têm distribuição normal de freqüências, como indicam os coeficientes de simetria e de curtose aproximadamente iguais a zero (0). O CV para dados de massa seca é ainda maior do que para os dados de massa verde, o que também indica que deve haver uma variabilidade grande nestes dados.

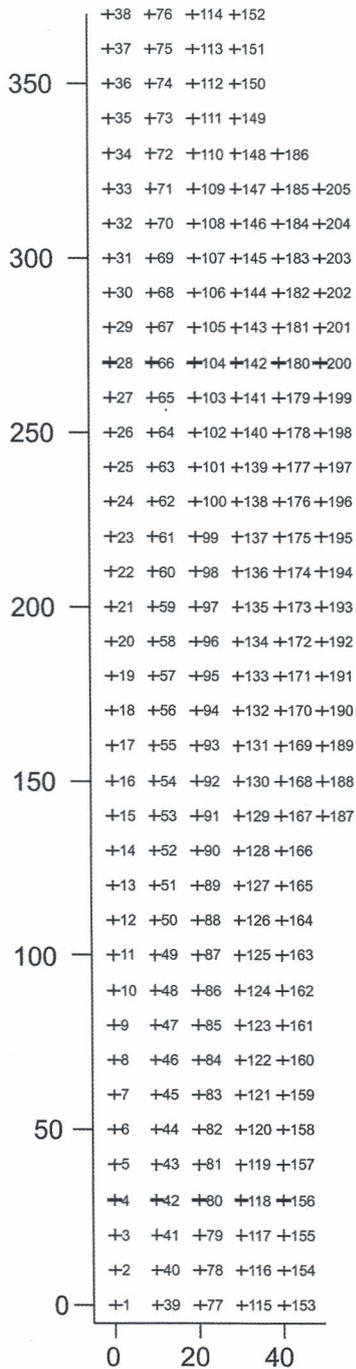
A figura 10 mostra os semivariogramas para os dados de altura, massa verde e massa seca, cujos parâmetros estão na Tabela 2. Como indi-

cam os valores baixos para alcance (a) e para grau de dependência (GD), existe uma variabilidade extremamente grande nesta área. O grau de dependência (GD) foi calculado, segundo recomenda ZIMBACK (2001). Os alcances indicam o tamanho das manchas existentes para os caracteres avaliados. Somente pelos parâmetros da estatística descritiva e por estes dois semivariogramas já é possível dizer que é praticamente impossível ter locais homogêneos para disposição de parcelas nesta área. A resposta para a pergunta a (Seria possível colocar um experimento de competição de variedades neste campo e supor que as respostas eram devidas ao potencial genético da planta ?) é, então, negativa.

O grande número de valores extremos em amostras vizinhas, provocando o baixo grau de dependência (GD) e altos valores de efeito pepita (Co), também pode ser visto na figura 11. Da mesma forma, a alta variabilidade deste campo é também revelada pelas muitas manchas de valores diferentes no mapa de massa verde (figura 12), obtido por krigagem. Nesse mapa, parece que existe alguma continuidade maior no sentido da diagonal.

As mesmas análises foram efetuadas com os dados obtidos na cana-de-açúcar, plantada após colheita da crotalária. A análise estatística descritiva (Tabela 3) dos dados de número de perfilhos por metro e produtividade de colmos mostrou que os coeficientes de variação eram baixos, segundo classificação de WARRICK e NIELSEN (1980). Os dados de produtividade apre-

sentaram maior coeficiente de variação, o que pode indicar maior variabilidade deste do que para número de perfilhos por metro. Houve distribuição normal de freqüência, como indicam os valores de simetria e curtose (próximos de 0) e o histograma para os resultados de produtividade (figura 13).



A Tabela 4 e a figura 14 mostram os parâmetros de ajuste e os semivariogramas para perfilhos e produtividade. Houve dependência espacial para as duas variáveis com alcance de dependência (a) de 45 metros. O grau de dependência apresentou-se baixo, devido ao alto valor do efeito pepita ( $C_0$ ).

Os mapas resultantes da interpolação por krigagem tanto para perfilho quanto para produtividade, apresentam manchas com maiores valores concentradas no lado esquerdo da área (figura 15). Este resultado pode ser decorrente de alguma variação espacial referente ao solo. Para que se possa averiguar a causa desta variabilidade-

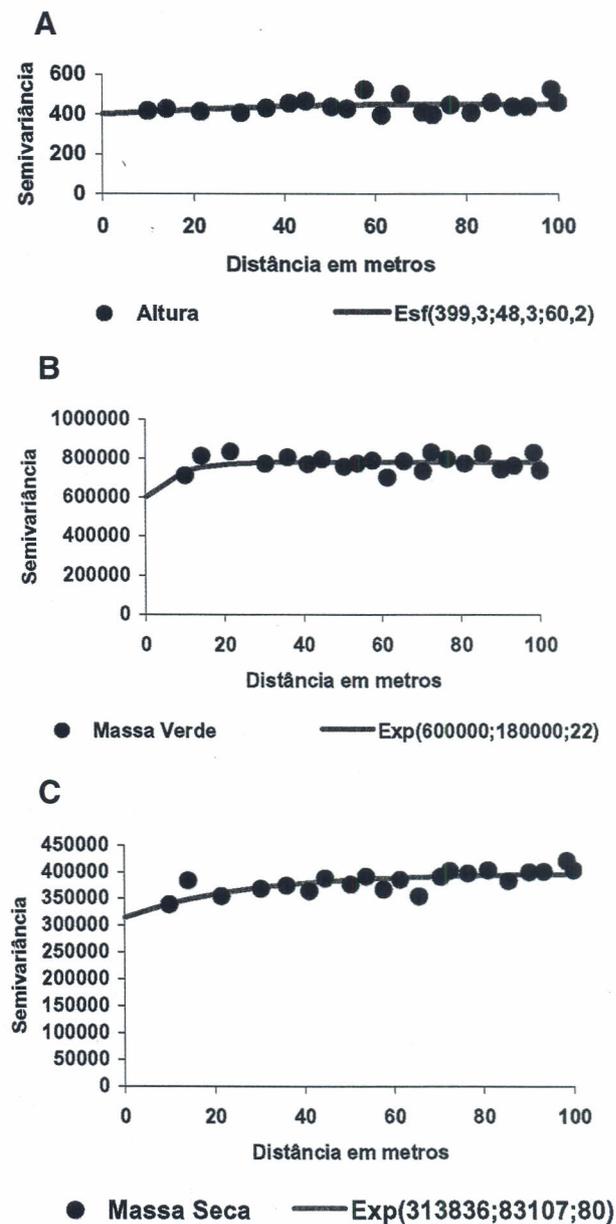


Figura 9. Esquema de amostragem para dados de crotalária.

Figura 10. Semivariogramas para altura, massa verde e massa seca de crotalária.

Tabela 1. Resultados estatísticos de parâmetros avaliados na crotalária.

Variável	Unidade	Número	Média	Desvio padrão	CV	Valor mínimo	Valor máximo	Simetria	Curtose
Altura	cm	205	260,9	21,8	8,3	190,00	315,00	-0,006	0,421
Massa verde	g	205	5.514,00	893,1	16,2	2.460,00	8.020,00	-0,161	0,412
Massa seca	g	205	3.478,00	699,9	20,1	1.815,00	5.550,00	0,198	0,061

Tabela 2. Parâmetros de ajuste de modelos de semivariogramas para altura, massa verde e massa seca de crotalária.

Variável	Modelo	C0	C1	a	r <sup>2</sup>	sqdp	GD
Altura	Esférico	399,26	48,2	60,23	0,0885	968,09	10,80
Massa verde	Exponencial	600.000	180.000	22,00	0,0230	173.062.022,9	23,08
Massa seca	Exponencial	313.836	83.107	80,00	0,446	309.881.736	20,94

de espacial ocorrida na cana, outras investigações seriam necessárias nesta área de estudo.

Considerando-se o mapa de massa seca da crotalária (figura 12) e o de produtividade de cana (figura 15), notam-se muitas semelhanças entre eles. Manchas de menores valores de massa seca de crotalária correspondem às de menores produtividades de cana, o que leva a uma resposta positiva à pergunta *b* (Seria possível usar mapas de produção de massa de crotalária para selecionar locais onde se pudesse alocar parcelas?).

Outra útil aplicação de geoestatística é a confecção de mapas de indicadores de concentração de sacarose, para planejamento de colheita. Tais mapas podem ser facilmente obtidos desde que os valores medidos se enquadrem nos modelos apresentados, ou seja, as amostragens devem ser suficientemente próximas para conseguir caracterizar as possíveis manchas de variabilidade. Se houve fotografias aéreas com as quais se possa identificar regiões com diferentes níveis de verde, é possível usar estas fotos para dirigir as amostragens, concentrando-se nas regiões onde existir maior variabilidade e diminuindo-se a densidade nos locais mais uniformes. As coordenadas podem ser medidas com GPS.

Existem muitas outras aplicações de geoestatística em estudos de variabilidade espacial de característica agrônomicas, por exemplo, em entomologia (DINARDO-MIRANDA *et al.*, 2007; FARIAS *et al.*, 2001; FARIAS *et al.*, 2004; GREGO *et al.*, 2006a), propriedades químicas do solo (CAMPBELL *et al.*, 1978; VIEIRA, 1997) e propriedades físicas (GREGO *et al.*, 2006b, VIEIRA, 1997), que podem fornecer uma vasta diversidade de análises de dados.

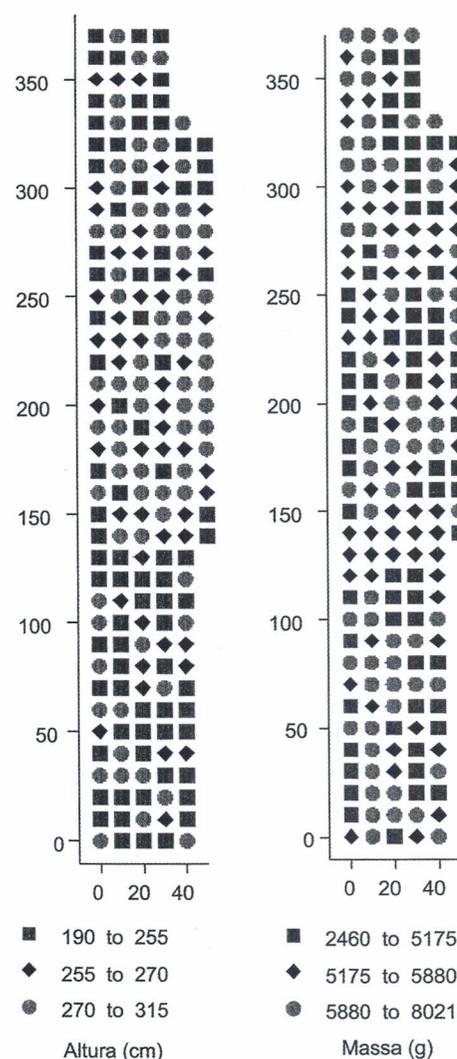


Figura 11. Mapas de postagem classificados para altura e massa verde de crotalária.

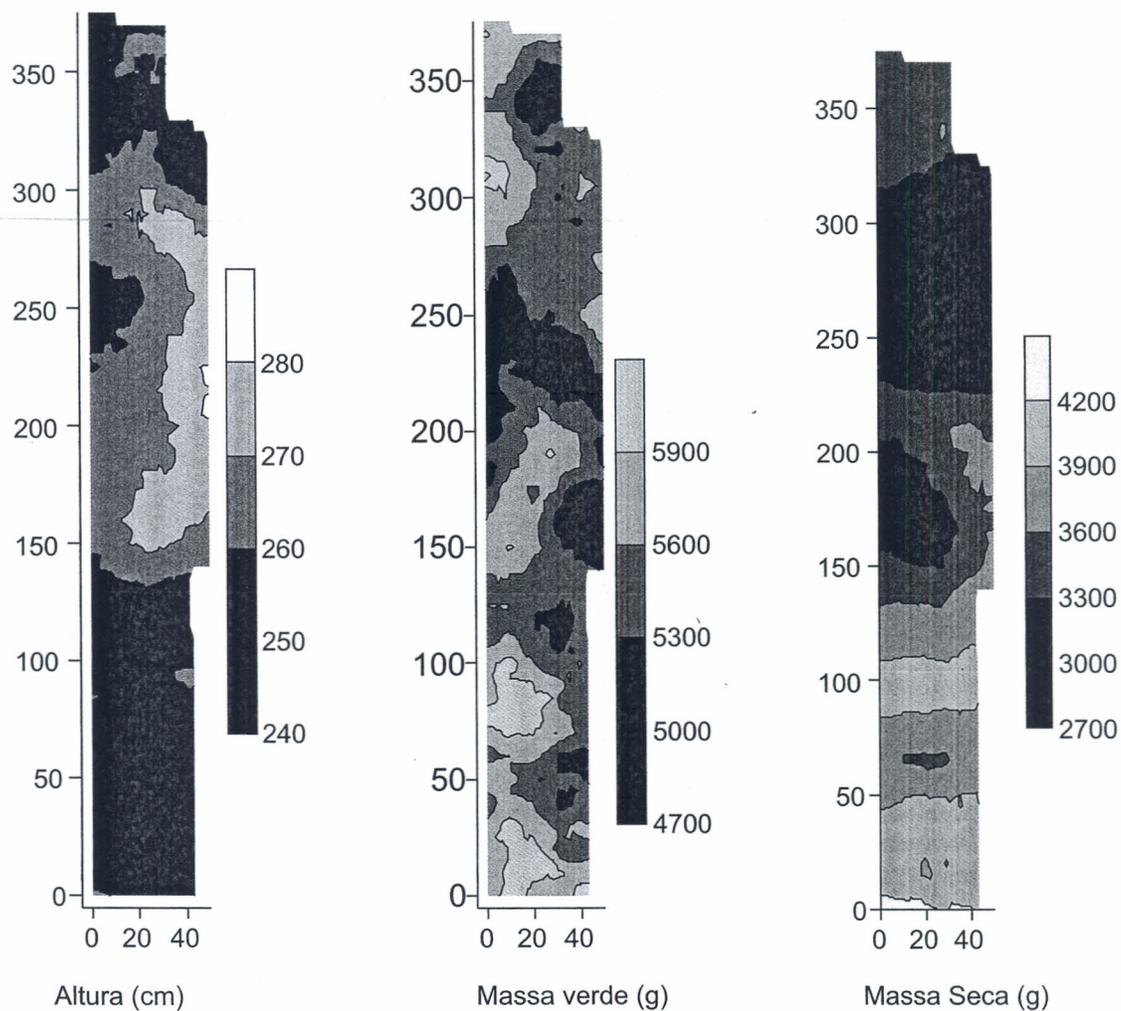


Figura 12. Mapas obtidos por krigagem para dados de crotalária.

Tabela 3. Análise estatística para parâmetros avaliados na cana-de-açúcar variedade IACSP 93-3046.

Variável	Unidade	Número	Média	Desvio padrão	CV	Valor mínimo	Valor máximo	Simetria	Curtose
Perfilhos	Nº m <sup>-1</sup>	97	13,54	1,11	8,20	11,00	17,11	0,494	0,876
Produtividade	t ha <sup>-1</sup>	97	97,71	14,23	14,57	71,98	136,50	0,384	-0,616

Tabela 4. Parâmetros de semivariograma para dados biométricos da variedade IACSP 93-3046 de cana-de-açúcar.

Variável	Modelo	C <sub>0</sub>	C <sub>1</sub>	a	Média dos erros reduzidos	Variância dos erros reduzidos	GD
Perfilho	Esférico	0,95	0,22	45	0,0091	1,021	18,8
TCH e	Esférico	170,00	30,00	45	-0,0096	1,005	15,0

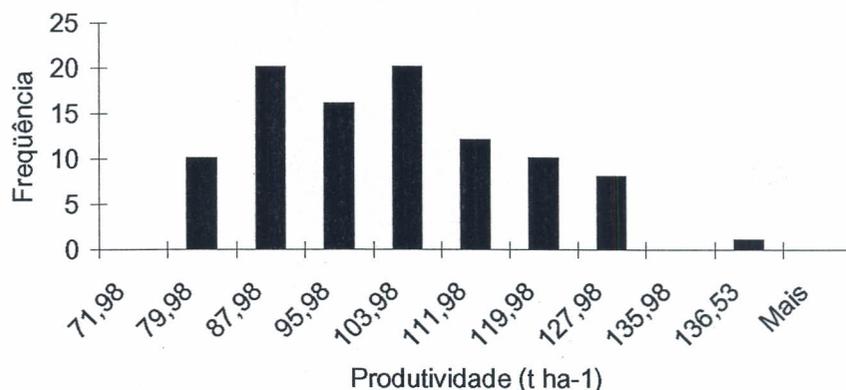


Figura 13. Histograma da produtividade da variedade IACSP 93-3046.

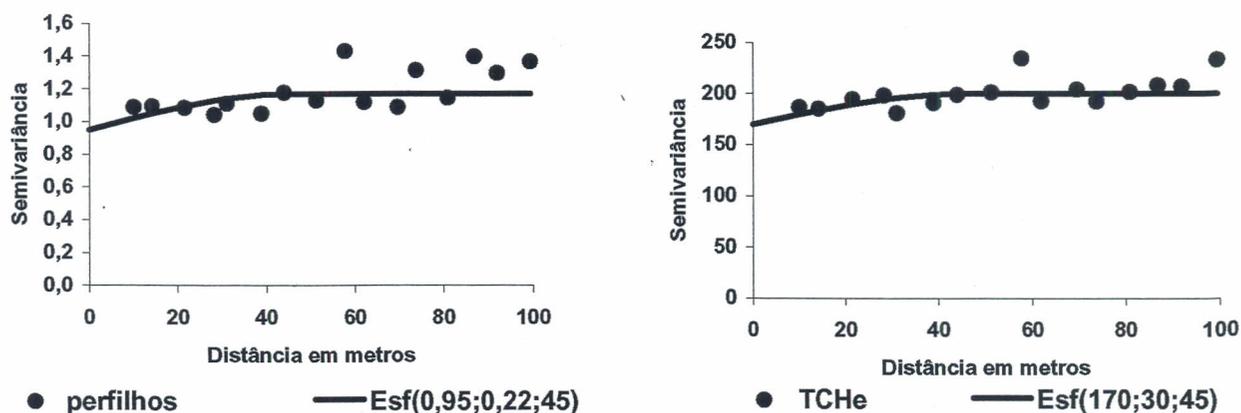


Figura 14. Semivariogramas para número de perfis por metro e produtividade (TChE, t ha<sup>-1</sup>) da variedade IACSP 93-3046.

## REFERÊNCIAS BIBLIOGRÁFICAS

CAMPBELL, J. B. Spatial variation of sand content and pH within single contiguous delineation of two soil mapping units. **Soil Science Society of America Journal**, v. 42, p. 460-464, 1978.

DINARDO-MIRANDA, L. L.; VASCONCELOS, A. C. M.; VIEIRA, S. R.; FRACASSO, J. V.; GREGO, C. R. Uso da geoestatística na avaliação da distribuição espacial de *mahanarva fimbriolata* em cana-de-açúcar. **Bragantia**, v. 66, n. 3, p. 449-455, 2007.

FARIAS, P. R. S.; BARBOSA, J. C.; BUSOLI, A. C. Distribuição espacial da lagarta-do-cartucho, *Spodoptera frugiperda* (J.E. Smith) (Lepidoptera: Noctuidae), na cultura do milho. **Neotropical Entomology**, v. 30, p. 681-689, 2001.

FARIAS, P. R. S.; ROBERTO, S. R.; LOPES, J. R. S.; PERECIN, D. Geostatistical characterization of

the spatial distribution of *Xylella fastidiosa* sharpshooter vectors on citrus. **Neotropical Entomology**, v. 33, p. 13-20, 2004.

GREGO, C. R.; VIEIRA, S. R.; ANTONIO, A. M.; ROSA, S. C. D. Geostatistical analysis for soil moisture content under the no tillage cropping system. **Scientia Agrícola**, v. 63, n. 4, p. 341-350, 2006a.

GREGO, C. R.; VIEIRA, S. R.; LOURENÇÃO, A. L. Spatial distribution of *Pseudaletia sequax* franclemont in triticale under no-till management. **Scientia Agrícola**, v. 63, n. 4, p. 321-327, 2006b.

SALVIANO, A. A. C.; VIEIRA, S. R.; SPAROVEK, G. Erosion intensity and crotalaria juncea yield on southwest brazilian ultisol. **Advances in GeoEcology**, v. 31, p. 369-374, 1998.

SNEDECOR, G. W.; COCHRAN, W. G. **Statistical methods**. 6. ed. Ames: Iowa State University Press, 1967. 593 p.

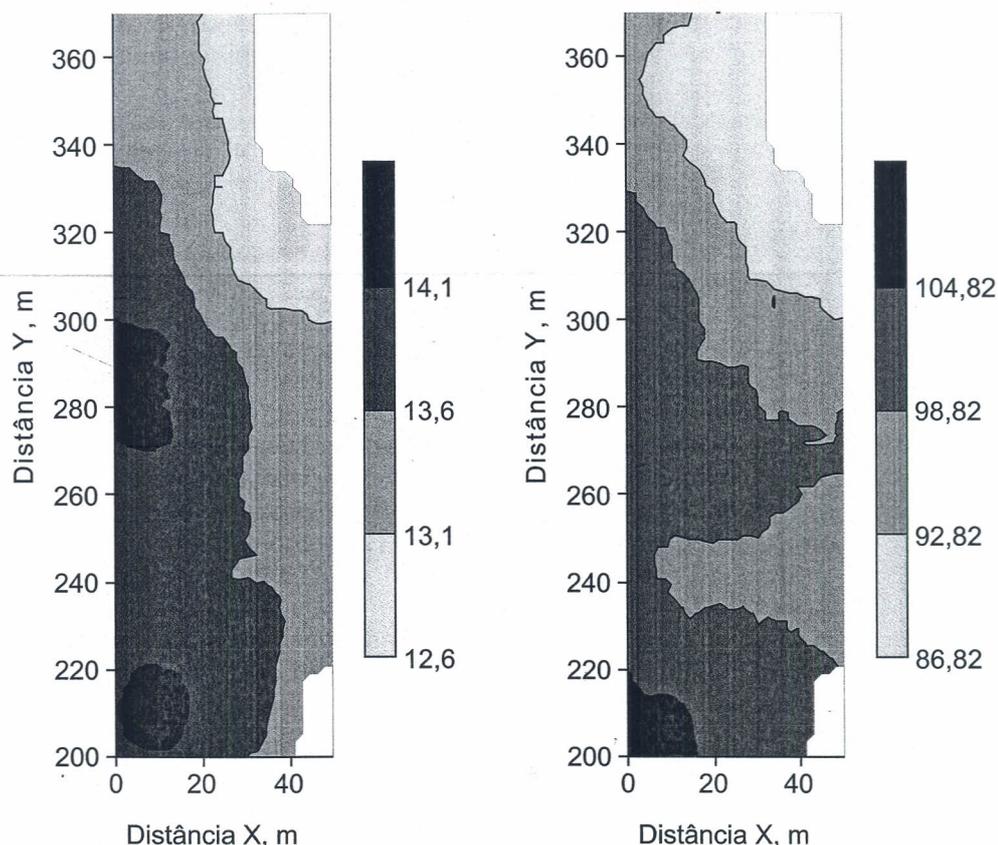


Figura 15. Mapa de número de perfis (à esquerda) e de produtividade (à direita,  $t\ ha^{-1}$ ) da variedade IACSP93-3046.

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

VIEIRA, S. R. Uso de geoestatística em estudos de variabilidade espacial de propriedades do solo. In: NOVAIS, R. F. (Ed.). **Tópicos em ciência do solo**. Viçosa: Sociedade Brasileira de Ciência do Solo, 2000. p. 1-54.

VIEIRA, S. R. Variabilidade espacial de argila, silte e atributos químicos em uma parcela experimental em um latossolo roxo de Campinas (SP). **Bragantia**, v. 57, n. 1, p. 181-190, 1997.

VIEIRA, S. R.; HATFIELD, J. L.; NIELSEN, D. R.; BIGGAR, J. W. Geostatistical theory and application to variability of some agronomical properties.

**Hilgardia**, v. 51, n. 1, p. 1-75, 1983.

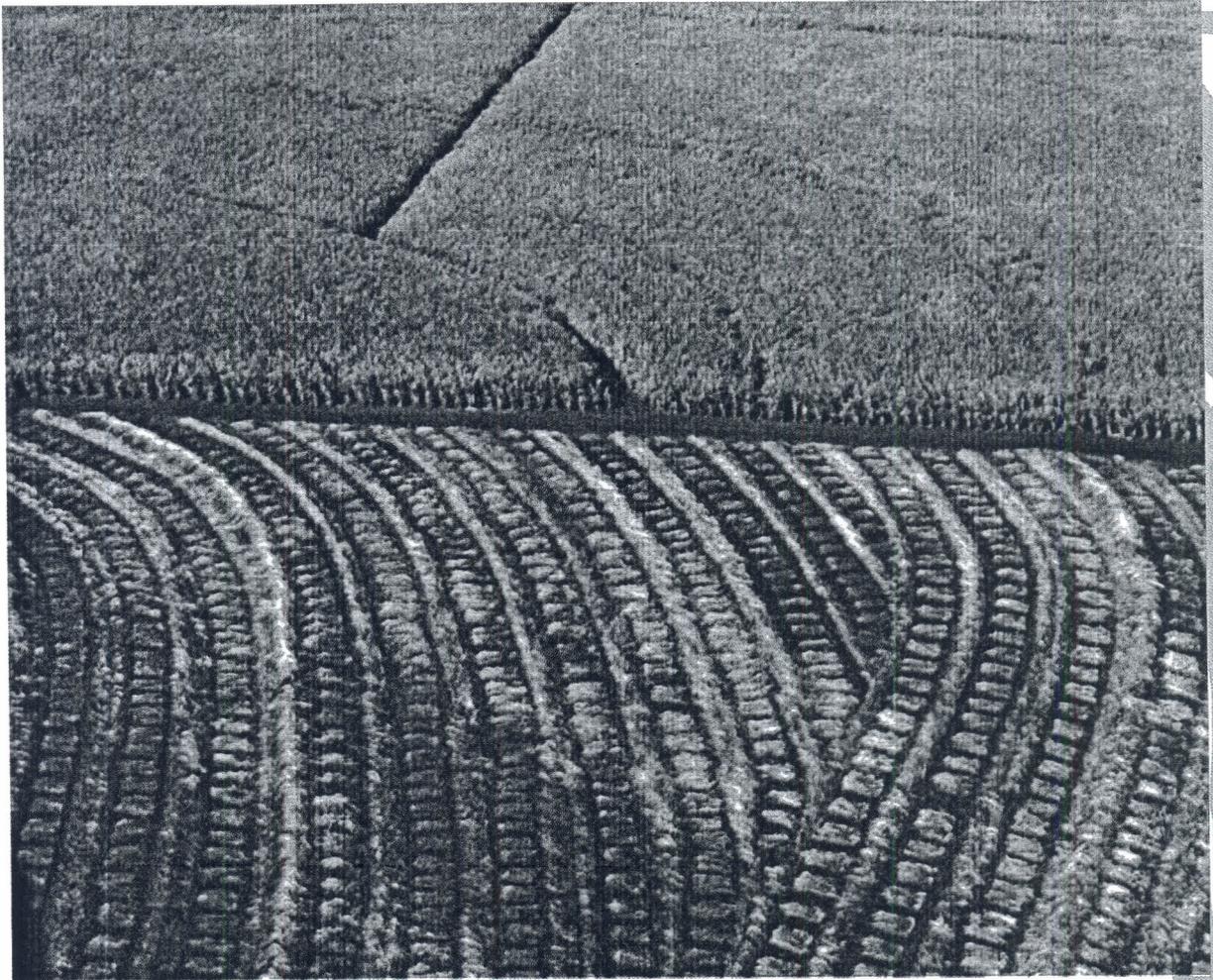
VIEIRA, S. R.; MILLETE, J.; TOPP, G. C.; REYNOLDS, W. D. Handbook for geostatistical analysis of variability in soil and climate data. In: ALVAREZ, V. V. H.; SCHAEFER, C. E. G. R.; BARROS, N. F.; MELLO, J. W. V.; COSTA, L. M. (Eds.). **Tópicos em ciência do solo**. Viçosa: SBCS, 2002, v. 2., p. 1-45.

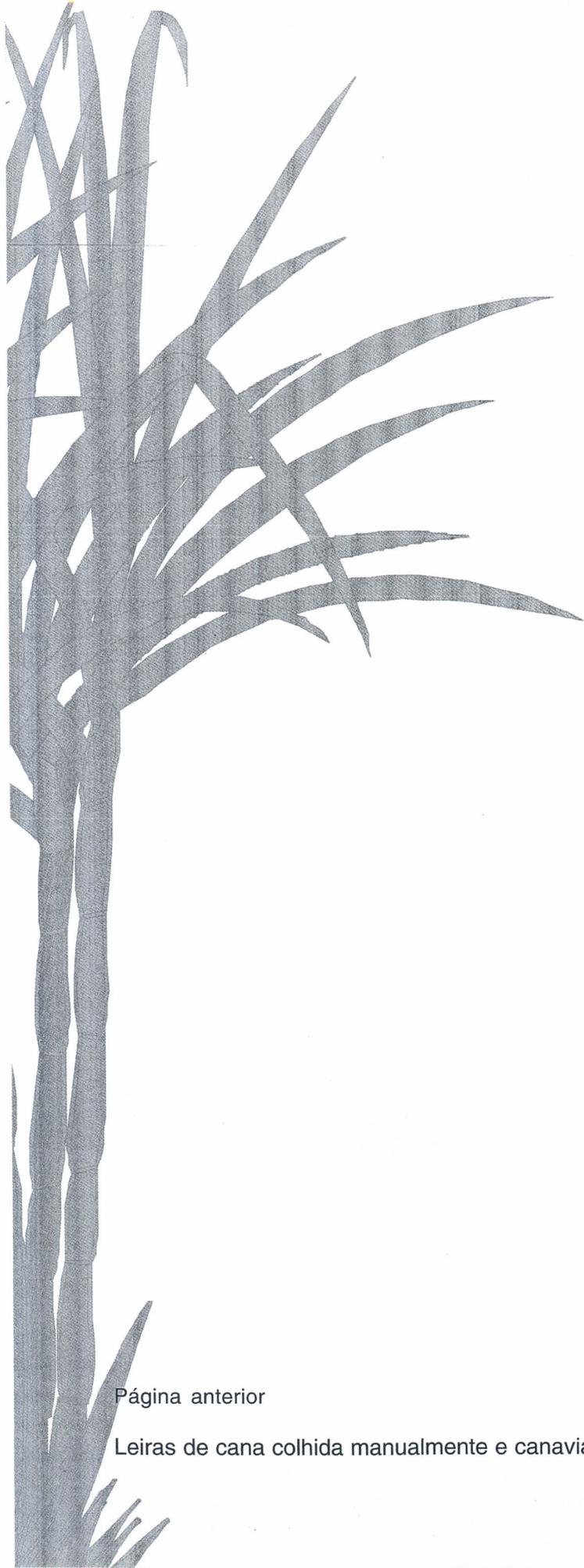
WARRICK, A. W.; NIELSEN, D. R. Spatial variability of soil physical properties in the field. In: HILLEL, D. (Ed.). **Applications of soil physics**. New York: Academic Press, 1980. p. 319-344.

ZIMBACK, C. R. L. **Análise especial de atributos químicos de solo para o mapeamento da fertilidade do solo**. 2001. 114 f. Tese (Livro Docência) - Botucatu, UNESP, 2001.

---

## PARTE 11: ASPECTOS ECONÔMICOS E AMBIENTAIS





Página anterior

Leiras de cana colhida manualmente e canavial ainda em pé (Foto: R. ROSSETTO).