

## **3 AVALIAÇÃO E MONITORAMENTO DA SALINIDADE DO SOLO**

**José Elenildo Queiroz, Antonio Carlos Gonçalves, Jacob Silva Souto e Marcos Vinicius Folegatti**

### **3.1 INTRODUÇÃO**

A salinidade do solo é um dos importantes fatores que afetam o rendimento dos cultivos. Em se tratando de regiões áridas e semi-áridas irrigadas, constitui um sério problema, limitando a produção agrícola e reduzindo a produtividade das culturas a níveis anti-econômicos. Nessas regiões, caracterizadas por baixos índices pluviométricos e intensa evapotranspiração, a baixa eficiência da irrigação e a drenagem insuficiente, contribuem para aceleração do processo de salinização.

Assim como outras de propriedades químicas e físicas do solo, a salinidade é uma propriedade bastante variável no espaço e no tempo devido a natureza dinâmica dos efeitos e interações de diversos fatores edáficos (permeabilidade do solo, nível do lençol freático etc.), climáticos (quantidade e distribuição das chuvas, umidade relativa, temperatura, etc.) e ação do homem (irrigação, práticas culturais, etc.). Dessa forma, na definição de estratégias de manejo e no estabelecimento de planos de recuperação dos solos afetados por sais, o monitoramento espacial (e temporal) dos parâmetros que definem a salinidade constitui um aspecto de fundamental importância.

O monitoramento da salinidade requer a aplicação de técnicas rápidas e apropriadas de avaliação e análise dos dados. A avaliação tem como objetivo identificar os fatores que contribuíram e/ou estão contribuindo para aumentar o grau de salinidade, definir um sistema de amostragem adequado e classificar os resultados conforme a magnitude dos dados. Devido a grande variabilidade normalmente encontrada nos parâmetros de avaliação da salinidade, a análise dos dados requer habilidade, conhecimento, muita experiência e a utilização de recursos computacionais, além de técnicas estatísticas adequadas.

Nas análises estatísticas usuais supõe-se que as amostras são espacialmente aleatórias e independentes, o que nem sempre é verdade. É de se esperar, por exemplo, que amostras de solo próximas resultem valores de salinidade mais parecidos do que amostras mais distantes, o que indica a existência de alguma relação espacial entre as medidas. Isto deve ser objeto de análise por meio da geoestatística, a qual constitui uma ferramenta apropriada para identificar a existência de estrutura de dependência espacial entre as observações. Também pode ser usado o processo de krigagem para estimar valores em pontos não medidos, permitindo, assim, o mapeamento da área estudada. Este mapeamento, feito para várias variáveis de interesse, permite estudar a correlação espacial entre elas. Por exemplo, correlação entre níveis de salinidade e elevação do lençol freático.

A aplicação das técnicas geoestatísticas deve ser orientada por uma análise descritiva geral e espacial dos dados, isto é, através de uma boa análise exploratória dos dados. Esta análise constitui uma forma de auxiliar na decisão de se assumir algum tipo de estacionariedade dos dados, o que é um pressuposto para aplicação das técnicas geoestatísticas.

Os recursos da geoestatística podem ser usados para tomada de decisão quanto ao manejo da salinidade e estudos de recuperação dos solos afetados por sais. Russo (1984), por exemplo, descreveu um método usando a geoestatística para estudo do transporte de solutos em solos heterogêneos e sua aplicação no manejo da salinidade.

Este capítulo trata da avaliação e classificação dos solos afetados por sais, do problema da amostragem e da análise de dados por meio de técnicas exploratórias e o emprego de recursos geoestatísticos. Um exemplo de aplicação a partir de um conjunto hipotético de dados de condutividade elétrica do extrato de saturação, medida de uso praticamente generalizado para avaliar a salinidade do solo, é apresentado e discutido.

## **3.2 DIAGNOSE E CLASSIFICAÇÃO DOS SOLOS SALINOS**

### ***3.2.1 Diagnose dos solos salinos***

Os solos salinos e sódicos se desenvolvem em consequência do acúmulo de sais e de sódio. Os sais encontrados mais frequentemente são formados por cátions de cálcio, magnésio, sódio, e dos ânions cloreto e sulfato; também, com menos frequência, nesses solos, encontram-se outros íons como potássio, carbonato e bicarbonato.

A formação dos solos salinos e sódicos ocorre principalmente em climas áridos e semi-áridos, onde a ausência de lixiviação, juntamente com a excessiva evaporação da água, possibilita o acúmulo de sais no solo. A situação pode ser muito agravada com a irrigação, já que a água utilizada sempre carrega sais para o solo. Este assunto é tratado com maiores detalhes no Capítulo 1.

Apesar dos solos afetados por sais serem comuns em regiões áridas e semi-áridas, os problemas de sais não são restritos a essas regiões. Os solos salinos e sódicos podem se desenvolver também em regiões úmidas e sub-úmidas sob condições apropriadas.

As três principais fontes naturais de sais do solo são o intemperismo mineral, a precipitação atmosférica e os sais fósseis (aqueles remanescentes de ambientes marinhos ou lacustres). Os sais também podem ser adicionados ao solo através de atividades humanas, incluindo o uso de águas de irrigação e salmouras altamente salinas, ou resíduos industriais (Bohn *et al.*, 1985).

### ***3.2.2 Parâmetros principais para avaliação da salinidade***

Diversas medidas de laboratório são usadas para avaliar a salinidade do solo, sendo as mais importantes o pH, a condutividade elétrica do extrato de saturação e a porcentagem de sódio trocável (Raij, 1991). Para avaliar o perigo de sodificação do solo pelo uso da água de irrigação utiliza-se um outro índice chamado relação de adsorção de sódio (RAS).

#### ***O pH do solo***

O pH de uma solução aquosa é o logaritmo negativo da atividade do íon hidrogênio. Sua determinação pode ser feita a partir de uma solução aquosa usando-se um potenciômetro, ou, colorimetricamente, mediante indicadores que mudam de cor com a atividade do íon hidrogênio. Existe uma certa dúvida com relação a determinação do pH nas soluções, se se aplica a suspensão do solo na água (Richards, 1954).

O pH do solo é influenciado pela composição e natureza dos cátions trocáveis, composição e concentração dos sais solúveis e a presença ou ausência do gesso e carbonatos de cálcio e magnésio.

Fireman & Wadleigh (1951) realizaram um estudo quanto ao efeito do teor de umidade, do nível de salinidade, da presença ou ausência de carbonatos e do gesso, na relação entre pH e porcentagem de sódio trocável (PST) em solos de regiões áridas, cujos resultados encontram-se na Tabela 3.1.

Tabela 3.1 Relação entre pH e PST.

pH (pasta de saturação)	PST
< 7	O solo contém quantidades consideráveis de H <sup>+</sup>
< 7,5	O solo quase nunca contém carbonatos de cálcio e de magnésio
< 8,5	PST pode ser maior ou menor que 15
> 8,5	PST quase sempre maior que 15 e presença de carbonatos de cálcio e de magnésio

Fonte: Fireman & Wadleigh (1951).

No item 3.2.3 deste capítulo, será mostrado o pH para as várias categorias de solos afetados por sais.

### ***Condutividade elétrica do extrato de saturação***

A condutividade elétrica (CE) expressa a habilidade que um meio apresenta em conduzir uma corrente elétrica.

Devido ao fato de que a CE de uma solução aquosa está intimamente relacionada com a concentração total de eletrólitos dissolvidos (solutos iônicos) na solução, ela é comumente usada como uma expressão da concentração total de sais dissolvidos de uma amostra aquosa, embora também seja afetada pela temperatura da amostra, pela mobilidade, valência, e concentração relativa dos íons contidos na solução (Rhoades, 1994).

A determinação da CE geralmente envolve a medição da resistência elétrica da solução, a qual é inversamente proporcional a sua área seccional e diretamente proporcional ao seu comprimento. A magnitude da resistência medida depende, contudo, das características da célula condutivimétrica usada para conter a amostra e dos eletrodos. Maiores detalhes podem ser vistos no Capítulo 2.

A CE de um solo pode ser determinada por meio de um extrato de uma pasta de solo saturado ou em uma suspensão mais diluída. Porém, para preparação da pasta de saturação, há necessidade de um bom adestramento técnico e certas precauções com a textura do solo (Filgueira & Souto, 1996). Em seu preparo, que consiste na agitação do solo (200- 400g) durante a adição de água destilada até alcançar o ponto final desejado, erros podem ocorrer quanto a observação se o solo já está saturado ou não (Rhoades, 1994).

Diferentes relações solo:água são também utilizadas para determinação da CE de um solo, de forma rápida e eficiente (por exemplo, 1:1, 1:2 e 1:5). Entretanto, a segurança das determinações depende do solo e dos sais nele presentes (Richards, 1954). Em solos

salino-sódicos e sódicos degradados da região de Patos, Estado da Paraíba, verificou-se que há possibilidade de se utilizar extratos obtidos em relações mais diluídas de solo:água (1:2 e 1:5), em substituição ao extrato de saturação, para estimar problemas de salinidade, por meio da CE (Filgueira & Souto, 1996).

### **Porcentagem de sódio trocável (PST)**

No estudo de solos com problemas de sais é de muita importância a porcentagem que o sódio representa em relação aos demais cátions adsorvidos. Esta porcentagem denomina-se porcentagem de sódio trocável. Na literatura inglesa a PST se expressa com as letras ESP, iniciais de *Exchangeable Sodium Percentage*. Em espanhol, expressa-se com as letras PSI (*Porcentaje de Sódio Intercambiable*). Seu valor é dado pela seguinte Expressão:

$$PST = \frac{Na}{(Ca + Mg + K + Na + H + Al)} 100 \quad (3.1)$$

A Tabela 3.2 mostra a classificação dos solos de acordo com a PST.

Tabela 3.2 Classificação dos solos segundo sua PST.

Classe	PST
Não sódicos	< 7
Ligeiramente sódicos	7 - 10
Mediamente sódicos	11 - 20
Fortemente sódicos	21 - 30
Excessivamente sódicos	>30

Fonte: Pizarro (1978).

### **Relação de adsorção de sódio (RAS)**

A RAS é um índice que expressa a possibilidade de que a água de irrigação provoque a sodificação do solo, no que depende da proporção do Na<sup>+</sup> em relação aos demais cátions.

Na língua inglesa, este índice recebe a denominação de SAR, iniciais de *Sodium Adsorption Ratio*. A expressão que define a RAS já foi vista no Capítulo 2 (Equação 2.20). Entre o complexo de troca e a solução do solo existe um equilíbrio no que se refere aos cátions adsorvidos e dissolvidos. Existe uma relação entre a RAS e a PST do solo, conforme a Equação 2.23 do Capítulo 2. Para Klar (1984) é possível que a Equação 2.23 não se aplique em condições de campo, devido à solução do solo ser sempre mais concentrada do que a água de irrigação. Pode-se considerar, entretanto, que, via de regra, a concentração da solução do solo não é maior que duas a três vezes a concentração da água de irrigação nos 30 cm superficiais.

### 3.2.3 Categoria dos solos afetados por sais

Existem várias classificações para os solos afetados por sais, sendo as mais importantes a russa, francesa e americana. A classificação russa combina os princípios da pedogênese, geoquímica dos sais e fisiologia vegetal. Contudo, a classificação mais simples e mais prática tem sido a americana (Richards, 1954).

Esta classificação é baseada na concentração de sais solúveis (expressa por meio da CE) extraídos da solução do solo e da percentagem de sódio trocável do solo (PST). A CE indica os efeitos da salinidade sobre as plantas, enquanto que, a PST indica os efeitos do sódio sobre as propriedades do solo.

A linha divisória entre solos salinos e solos não-salinos tem estabelecido o valor  $4 \text{ dS m}^{-1}$  para extratos de pasta saturada do solo. Entretanto, podemos encontrar plantas sensíveis à sais, que podem ser afetadas em solos cujo extrato de saturação apresenta CE entre 2 e  $4 \text{ dS m}^{-1}$ . O Comitê de Terminologia da Sociedade Americana de Ciência do Solo tem recomendado baixar o limite entre solos salinos e não-salinos para  $2 \text{ dS m}^{-1}$  (Bohn *et al.*, 1985). A classificação tradicional e a recentemente proposta são apresentadas na Tabela 3.3.

Tabela 3.3 Classificação de solos afetados por sais.

	Classificação tradicional	Classificação proposta
Solos normais	CE < 4	CE < 2
	PST < 15	PST < 15
	pH < 8,5	pH < 8,5
Solos salinos	CE > 4	CE > 2
	PST < 15	PST < 15
	pH < 8,5	pH < 8,5
Solos sódicos	CE < 4	CE < 2
	PST > 15	PST > 15
	pH > 8,5	pH > 8,5
Solos salinos-sódicos	CE > 4	CE > 2
	PST > 15	PST > 15
	pH < 8,5	pH < 8,5

Fonte: Bohn *et al.*(1985).

Embora sejam designados como sódicos os solos com  $\text{PST} > 15$ , vários resultados publicados na literatura têm mostrado efeitos do sódio sobre a estrutura do solo, mesmo sob níveis inferiores, o que parece ser mais prudente considerar como sódicos solos com  $\text{PST} > 7$  (Pizarro, 1978).

Os solos salinos correspondem aos solos “alcali-branco” descritos por Hilgard<sup>1</sup> (citado por Richards, 1954) e aos “solonchacks” dos pesquisadores russos. Nestes, o crescimento das plantas é limitado pela grande quantidade de sais solúveis e, podem ser convertidos a solos normais simplesmente pela lavagem do excesso de sais da zona radicular das plantas, o que requer um bom sistema de drenagem. Plantas crescendo em

<sup>1</sup> HILGARD, E.W. Soils. Their formation, properties, composition, and relation to climate and plant growth. New York and London, 1906. 593pp., illus.

tais solos podem apresentar-se raquíticas com folhas grossas e uma coloração verde escuro.

Solos contendo altas concentrações de sais solúveis e altos níveis de sódio trocável são chamados de solos salino-sódicos. Tais solos geralmente limitam o crescimento das plantas pelos seus altos níveis de sais solúveis.

Bohn *et al.* (1985) relatam que o principal risco quando da lixiviação dos sais solúveis dos solos salino-sódicos é que a lixiviação dos sais ocorre mais rapidamente do que a remoção do sódio trocável, o que irá causar uma conversão para solo sódico, podendo ocorrer uma redução na permeabilidade ou condutividade hidráulica do solo, com consequentes efeitos nas relações água-planta. Devido os sais solúveis prevenirem as reações de hidrólise, o pH dos solos salino-sódicos é tipicamente menor que 8,5.

Os solos sódicos correspondem aos chamados “alcali-negro” por Hilgard (citado por Richards, 1954) e, em certos casos, aos “Solonetz” de pesquisadores russos. O pH dos solos sódicos é comumente maior que 9 ou 9,5, e as frações orgânicas e argila encontram-se dispersas (Bohn, 1985). A matéria orgânica dispersa pode acumular-se na superfície do solo devido a evaporação, causando um enegrecimento, dando origem ao termo “álcali-negro”.

Nos solos sódicos a percolação da água é muito baixa e o controle da salinidade é o principal problema associado com esses solos. Devido ter uma baixa concentração de sais solúveis e, um pH alto, os solos sódicos podem apresentar toxidez direta para muitas plantas devido à ação direta do sódio.

#### **3.2.4 Monitoramento da salinidade**

A agricultura irrigada em regiões áridas e semi-áridas requer um monitoramento periódico da salinidade do solo. Para tanto, procedimentos práticos e rápidos de campo para medida da CE, que permitam identificar áreas fontes de carregamento de sais e mapear a distribuição e extensão dos solos afetados por sais, devem ser utilizados.

O nível de sais na zona radicular deve ficar abaixo do nível nocivo às plantas cultivadas. Assim, o monitoramento direto da salinidade na zona radicular é recomendado para avaliar a eficiência dos diversos programas de manejo na área irrigada (Rhoades *et al.*, 1981).

A determinação da salinidade pode ser feita tanto por métodos de laboratório como por métodos de campo. No laboratório, conforme já comentado no item 3.2.2, estima-se a salinidade a partir de medidas da CEes ou da CE em diferentes relações solo:água. Em condições de campo, basicamente três métodos são disponíveis para se determinar a CE e avaliar a salindade: a) técnica de indução eletromagnética; b) “four-electrode probes”; e c) “time domain reflectometry - TDR”. Estes três métodos são amplamente usados, especialmente onde se deseja informações imediatas visando o mapeamento ou monitoramento das variações de salinidade em escala de campo. Cada um destes métodos tem vantagens e desvantagens, de modo que a escolha do mais apropriado ou de uma combinação de métodos vai depender das necessidade específicas e dos objetivos de cada situação particular de interesse Rhoades, 1994).

### 3.3 AMOSTRAGEM PARA ESTUDOS DE SALINIDADE

#### 3.3.1 Introdução

Quando se pretende avaliar a salinidade do solo, propriedade de grande importância na produção vegetal, faz-se uso de uma metodologia padronizada, de forma a quantificá-la em alguns pontos dentro da área em estudo, selecionados para a amostragem.

A questão que se pretende discutir é: que critérios devem ser estabelecidos para se decidir como amostrar uma determinada área, de forma que os valores medidos permitam fazer inferências sobre esta propriedade do solo, com um desejado nível de precisão, em toda a área?

Para abordar esta questão, inicialmente deve-se considerar que o procedimento clássico da estatística consiste em se assumir que o valor esperado de uma propriedade do solo  $Z(x)$  qualquer, em uma posição  $x$  dentro da área em estudo, é dado por:

$$z(x) = \mu + e(x) \quad (3.2)$$

em que  $\mu$  é a média populacional ou a esperança de  $Z(x)$  e  $e(x)$  são os desvios dos valores em torno da média. Assume-se que estes desvios são aleatórios e espacialmente independentes, apresentando distribuição normal com média zero e variância representada por  $\sigma^2$  (Trangmar *et al.*, 1985).

Assumindo estas hipóteses e sendo a distribuição de frequências da propriedade  $Z(x)$  simétrica o bastante para ser suficientemente próxima da normal, a média aritmética dos dados amostrais é adotada como sendo uma boa estimativa da tendência central dos valores da população. A média é então tomada como uma estimativa da propriedade em locais não amostrados, tornando necessário identificar o nível de precisão desta média como estimador. Para tanto, tem sido usado algum parâmetro que quantifica a dispersão dos dados em torno desta média, como o coeficiente de variação ou os limites de confiança para a média. Para a comparação da variação entre unidades de amostragem, a análise de variância e subsequentes testes estatísticos são usados, segundo Trangmar *et al.* (1985).

Dentro desta abordagem do problema, o procedimento de amostragem adequado para a caracterização do comportamento de uma propriedade do solo  $Z(x)$ , em uma determinada área, consiste em se proceder a uma amostragem aleatória na área, garantindo assim a independência dos dados. A intensidade de amostragem dependerá então da variabilidade da propriedade na área, assim como do nível de precisão desejado em torno da média. Quanto maior a variabilidade da propriedade, assim como quanto maior a precisão desejada, maior será o número necessário de amostras.

De acordo com Warrick & Nielsen (1980), a partir de um conjunto de valores de uma determinada propriedade, pode-se estimar o número de amostras necessário ( $N$ ) em futuras amostragens, para se obter uma previsão com um nível de probabilidade desejado, por meio da Expressão 3.3:

$$N = \frac{t_{(n-1, 1-\alpha/2)}^2 s^2}{d^2} \quad (3.3)$$

em que  $d$  é o limite de variação aceitável em torno da média, expresso em porcentagem,  $s$  o desvio padrão dos  $n$  dados conhecidos e  $t$  o valor do parâmetro “ $t$ ” da distribuição de Student, para o nível de significância  $\alpha$  desejado. Lascano & Hatfield (1992) salientam que, caso a população siga a distribuição normal, a equação (3.3) é adequada. Em caso contrário, esta expressão fornece apenas uma aproximação do número necessário de amostras.

Segundo Bresler *et al.* (1982), o procedimento decrito acima pressupõe estacionariedade e independência dos valores amostrais. Russo & Bresler (1981), no entanto, afirmam que é usual propriedades do solo apresentarem estrutura de variação no espaço bem definida, o que ocorre, segundo Ahuja & Nielsen (1990), em função do material originário e dos fatores de formação. Desta forma, as hipóteses formuladas anteriormente são negadas, uma vez que os procedimentos clássicos da estatística assumem que os valores da propriedade se distribuem aleatoriamente na área. De acordo com Burgess & Webster (1980), uma propriedade de solo é uma variável contínua cujo valor em uma dada posição deve ser função da direção e da distância de separação em relação aos valores vizinhos amostrados. Em função desta dependência espacial, os métodos tradicionais de interpolação tornam-se inadequados por não levarem em consideração a correlação espacial e a posição relativa das amostras.

O estudo da distribuição dos valores de uma propriedade do solo em uma área deve considerar a possível correlação espacial dos seus valores, tornando necessário que não apenas os valores em uma série de locais sejam conhecidos, mas que se conheçam também as coordenadas destes locais. Em outras palavras, deve-se conhecer o valor da propriedade em cada ponto, assim como a posição de cada ponto de amostragem, constituindo assim uma amostragem sistemática, como será descrito a seguir.

### **3.3.2 Formas de amostragens**

Não é possível medir o valor de uma propriedade do solo em todos os pontos. É necessário tomar algumas amostras, as quais são úteis desde que sejam capazes de representar confiavelmente uma região. A teoria da amostragem procura estabelecer bases seguras para a definição do padrão de amostragem. De um modo geral, procura-se estabelecer uma amostragem aleatória como forma de garantir que esta será não tendenciosa. De posse de um conjunto de valores amostrais, pode-se calcular a sua média e a sua variância, assim como momentos de maior ordem. Se a amostragem for não tendenciosa, estes valores podem ser usados para estimar os parâmetros populacionais correspondentes, ou seja,  $\bar{m}$  e  $s^2$  (média e variância amostral) estimam  $\mu$  e  $\sigma^2$  (média populacional e variância populacional), respectivamente. Sendo a amostragem não tendenciosa, os valores amostrais convergem para os valores populacionais, à medida que o número de amostras cresce.

Em se tratando de uma área de solo, é usual representá-la em um primeiro quadrante de um sistema de eixos cartesianos, de forma que cada ponto tenha duas coordenadas,  $x$  e  $y$ . Busca-se então valores aleatórios para estas coordenadas, definindo assim os pontos que serão amostrados. Para tanto, pode-se fazer uso das tabelas de números aleatórios ou pode-se usar um gerador de números pseudo-aleatórios disponível em calculadoras e computadores. Este procedimento constitui uma amostragem aleatória simples, o qual é mais apropriado para solos homogêneos. Um grave inconveniente deste padrão de amostragem é que algumas regiões tendem a ser intensamente

amostradas, enquanto outras o são precariamente. A aleatoriedade não garante uma boa distribuição dos pontos de amostragem em toda a área. Assim, pode-se fazer uso da amostragem aleatória estratificada, técnica mais apropriada para solos heterogêneos, que consiste em se subdividir a área total em células relativamente homogêneas, usualmente quadradas, dentro das quais é feito a casualização dos pontos amostrais, como descrito anteriormente. É necessário que cada célula seja de igual tamanho, sem haver superposição, sendo a seleção de cada uma baseada nas diferenças em termos de aparência das culturas, textura do solo, acumulação de sais na superfície do solo ou práticas de manejo.

Em estudos de salinidade do solo, por ser esta uma propriedade altamente variável e dinâmica, esta técnica de amostragem proporciona uma boa caracterização da região estudada. Uma grande região heterogênea é subdividida em sub-regiões menores que apresentam níveis de salinidade relativamente homogêneos, associando-se a estes, fatores edáficos, climatológicos e operações de manejo dentro de cada sub-região.

O número de amostras por célula deve ser proporcional a porcentagem da área total representada pela célula. Com a amostragem aleatória estratificada elimina-se o erro amostral entre células, sendo o erro amostral proveniente apenas de dentro de cada célula.

De acordo com Webster & Olivier (1990), sendo assegurado que todas as células sejam iguais e contenham o mesmo número de pontos, a média de todas as observações é um estimador não tendencioso da média populacional. Contudo, o desvio padrão deve ser estimado de acordo com a expressão:

$$S_{est} = \left( \frac{1}{m} \right) \left[ \sum_{k=1}^m S_k^2 / n_k \right]^{1/2} \quad (3.4)$$

em que  $m$  é o número de células,  $S_k^2$  é a variância das  $n_k$  observações na célula  $k$ .

Outro padrão de amostragem é a amostragem sistemática, na qual os pontos de amostragem são localizados em intervalos regulares em um “grid” ou ao longo de um “transect”. Isto garante uma cobertura total da área, sendo a forma mais simples de se identificar os pontos e demarcá-los na área. Para se determinar a localização dos pontos, inicialmente, de forma aleatória, escolhe-se as coordenadas  $x$  e  $y$  do primeiro ponto e, em seguida, adota-se o espaçamento constante nas duas direções, identificando-se os demais pontos. Reichardt *et al.* (1986) salientam que a distância constante entre as amostras não é estritamente necessária, desde que a posição relativa de cada amostra seja conhecida.

Neste tipo de amostragem como normalmente é obtida apenas uma observação por intervalo, não há como estimar o desvio padrão da média amostral. Peterson & Calvin (1967) e Yates (1981) recomendam estratificar a amostra sistemática em blocos, com um mínimo de duas amostras por bloco, assumindo que a variação dentro do bloco é apenas a variação amostral. Com isto, a estimativa do desvio padrão pode ser feita de modo similar ao caso da amostragem aleatória estratificada.

Quando a propriedade em estudo não se distribui aleatoriamente na área, mas segundo um padrão definido, a amostragem sistemática pode conduzir a erros de estimativa. De acordo com Webster & Olivier (1990), no entanto, existem evidências de que esta forma de amostragem é mais precisa que a amostragem aleatória simples e mesmo que a estratificada.

Além destas formas de amostragem, pode-se citar também a amostragem sistemática estratificada não alinhada, na qual a área é subdividida em células e dentro destas, de forma não alinhada, os pontos amostrais são distribuídos sistematicamente. Outra forma é a amostragem desigual, onde subáreas são mais intensamente amostradas. Outras formas de amostragem se adequam a questões específicas e não são de interesse neste texto.

### ***3.3.3 Estratégia de amostragens***

Na estratégia de amostragem para estudos de salinidade vários aspectos como a seleção da forma de amostragem, a distribuição espacial (e temporal) da salinidade, a independência das observações, o volume da amostra, o tamanho da área, entre outros, devem ser considerados.

#### ***Seleção da forma de amostragem***

A seleção de uma determinada forma de amostragem depende dos objetivos e das fontes de variação. Para problemas de salinidade em áreas relativamente pequenas, um critério de amostragem composta pode ser usada. Se, entretanto, necessita-se avaliar a salinidade em uma grande área, uma técnica de amostragem não tendenciosa e precisa pode ser requerida, principalmente quando o monitoramento da salinidade com o tempo é um dos objetivos. Para este propósito, os erros amostrais devido a variabilidade espacial devem ser mínimos para detectar diferenças na salinidade (Tanji, 1990).

As causas comuns de variação estão associadas a tendência, descontinuidade, periodicidade e aleatoriedade. A tendência em salinidade do solo pode ser resultante de variações na textura ou infiltração da água no solo, de forma que se esta é a maior causa de variabilidade, pode ser melhor realizar uma amostragem aleatória estratificada ou sistemática. Em áreas com descontinuidade ou mudanças relativamente abruptas na salinidade, deve-se usar também a técnica de amostragem aleatória estratificada. Se a maior fonte de variabilidade é devida ao acaso e o extenso campo populacional parecer homogêneo, então a amostragem aleatória simples seria satisfatória.

No caso da periodicidade ser substancial, a estratégia de amostragem deve considerar esta fonte de variação. Amostragem em blocos de observações de tamanhos iguais pode ser necessária para períodos relativamente curtos (alta frequência de amostragem). Os blocos poderiam ser distribuídos aleatoriamente ou sistematicamente sobre todo o campo, dependendo da forma de amostragem desejada. Em áreas cultivadas, o período e o tamanho dos blocos podem ser estimados a partir de informações das práticas culturais.

#### ***Distribuição espacial***

Se os parâmetros que caracterizam a salinidade variam espacialmente e não se ajustam a uma distribuição normal de probabilidade, a média não pode ser usada como um parâmetro representativo para recomendações quanto as técnicas de recuperação e manejo. O conhecimento da distribuição espacial dos dados permite o planejamento quanto a aplicação de tratamentos adequados nas áreas com problemas mais severos e evita despesas com tratamentos além das necessidades em outras áreas. Este tipo de estudo permite subdividir a área em zonas de menor variabilidade onde valores médios podem ser usados como representativos, se a distribuição for normal. Nesse caso, o

número necessário de amostras pode ser estimado pela Equação 3.3, apresentada no item 3.3.1.

### ***Independência das observações***

A técnicas de amostragens comumente utilizadas assumem independência entre as observações, o que implica dizer que cada observação proporciona completamente uma nova informação acerca do parâmetro populacional. Entretanto, até uma determinada distância, as propriedades do solo podem ser autocorrelacionados, indicando uma dependência espacial dos dados, isto é, cada observação não fornece uma nova informação à respeito do parâmetro populacional. Neste caso, a estratégia de amostragem deve considerar a distância amostral para satisfazer independência espacial das amostras para uso das técnicas clássicas da estatística.

### ***Volume da amostra***

O volume da amostra é um fator que deve ser considerado no planejamento amostral, pois os estudos mostram que existe uma relação direta entre o volume amostrado por observação e a variância ou coeficiente de variação dos dados. Vários autores (Hawley *et al.*, 1982; Hassan *et al.*, 1983; entre outros) têm mostrado que quando o volume amostrado aumenta, a variância diminui. Portanto, aumentando o volume da amostra por observação, um menor número de observações é requerido.

### ***Tamanho da área***

O número necessário de amostras depende da variância dos dados. Assim, o número de observações aumenta com o aumento da área apenas se a variância aumentar.

A relação entre o coeficiente de variação e o tamanho da área amostrada varia muito. O efeito do tamanho da área sobre a variabilidade depende da distribuição das observações na área de interesse.

### ***Variabilidade local***

Quando existe uma substancial variabilidade na salinidade do solo em pequena escala, mais cuidados são necessários na amostragem. Os fatores locais que estão contribuindo para esta variabilidade, devem ser observados. O sistema de irrigação utilizado, conforme a distribuição da água no solo, é um dos fatores que pode proporcionar diferenças na variabilidade da salinidade. Amostragens em um particular local podem ser necessárias para considerar esta variabilidade. Para os sistemas de irrigação por aspersão e gotejamento, por exemplo, amostras obtidas no espaço médio entre os emissores resultaria em níveis diferentes de salinidade em relação a amostras próximas aos emissores.

Quando o monitoramento da salinidade é feito em função do tempo, cada observação deve ser feita na mesma posição relativa.

### ***Amostragem composta***

As amostras compostas são frequentemente utilizadas com o objetivo de reduzir o custo de análises. Cline\* (citado por Tanji, 1990) condiciona que a amostra composta é válida se:

1. O volume amostrado representa uma população homogênea. Para populações heterogêneas, o conhecimento da distribuição espacial é importante, no qual uma amostra composta resultaria em erros.

2. Quantidades iguais de cada unidade amostral contribuem para a amostra composta.

3. Não ocorrem interações dentro da amostra composta como resultado da composição.

4. Uma estimativa não tendenciosa da média é o único objetivo. Amostragem composta não é satisfatória se qualquer outra estatística é requerida.

\* **Rodapé: Cline ...**

### 3.3.4 Amostragem hierárquica

A amostragem hierárquica pode ser usada quando se deseja fazer uma avaliação preliminar da variação em uma população, permitindo identificar a variação devida a diferentes níveis de subdivisão da população. Em ciência do solo, é de grande interesse a adaptação desta técnica realizada por Youden & Mehlich (1937), os quais associaram as diferentes componentes da variância total de um conjunto de valores de propriedades do solo à distância de separação entre os respectivos valores.

Uma vez amostrado o solo em alguns subníveis, correspondentes a regiões ou mais apropriadamente a diferentes distâncias de separação entre pontos amostrais, a variância total pode ser particionada por meio de uma análise de variância hierárquica. O programa computacional para análises estatísticas SAS apresenta recursos para a execução desta análise de variância, que pode ser esquematizada como apresentado por Webster (1985) por meio de um exemplo onde se tem dados hierarquizados em três estágios.

O modelo de variação é dado por:

$$Z_{ijkl} = \mu + A_i + B_{ij} + C_{ijk} + \varepsilon_{ijkl} \quad (3.5)$$

em que  $Z_{ijkl}$  é o valor da propriedade no ponto  $l$  no  $k$ -ésimo sub-nível, dentro do  $j$ -ésimo subnível, dentro do nível  $i$ .  $\mu$  é a média de  $z$  na área.  $A_i$ ,  $B_{ij}$  e  $C_{ijk}$  são variáveis aleatórias associadas aos níveis 1, 2 e 3, com variâncias  $s_1^2$ ,  $s_2^2$  e  $s_3^2$ , respectivamente.  $\varepsilon_{ijkl}$  é o termo residual com variância  $s_4^2$ . Sendo  $n$  o número de subdivisões em cada nível, o quadro de análise de variância é apresentado na Tabela 3.4.

Tabela 3.4 Graus de liberdade e quadrados médios para análise de variância hierárquica de três níveis, balanceados.

Fonte (nível)	Graus de Liberdade	Componentes do quadrado médio
1	$n_1 - 1$	$s_3^2 + n_3 s_2^2 + n_2 n_3 s_1^2$
2	$n_1(n_2 - 1)$	$s_3^2 + n_3 s_2^2$
3	$n_1 n_2 (n_3 - 1)$	$s_3^2$
Total	$n_1 n_2 n_3 - 1$	

**Exemplo de aplicação:**

Para exemplificar o uso da amostragem hierárquica, vamos tomar uma situação semelhante à descrita por Youlden & Mehlich (1937), porém com dados hipotéticos. Estes autores mediram o valor de pH do solo em nove (9) estações, espaçadas entre si de 1,6 Km. Em cada estação, foram definidas duas subestações espaçadas de 305 m (nível 2), em cada uma, duas áreas de amostragem espaçadas de 30,5 m (nível 3) e, em cada uma, foram definidos dois pontos de amostragem defasados de 3,05 m (nível 4). Este é um esquema hierárquico balanceado, com uma progressão geométrica dos espaçamentos. Um quadro de análise de variância com valores hipotéticos de variância é apresentado a seguir, com o propósito de exemplificar o uso das informações resultantes.

Na Figura 3.2 é mostrado o gráfico da distribuição da variância acumulada em função da distância de separação entre os pontos, ou seja, da distância correspondente aos níveis adotados (Tabela 3.5).

Tabela 3.5 Análise de variância hierárquica para amostragem balanceada com 4 níveis.

Fonte Espaçamento - m	Graus de Liberdade	componente estimado da variância	percentagem da variância total
1 - 1600	8	0.00368	5
2 - 305	9	0.004786	6.5
3 - 30,5	18	0.05007	68
4 - 3,05	36	0.01509	20.5

A análise da Figura 3.2 permite avaliar a utilidade desta técnica. Inicialmente, pode-se observar que para o menor espaçamento, a variância assume um valor de 20,5 % do total, o que é usualmente denominado “nugget effect” e corresponde à variação da propriedade em distâncias inferiores à amostrada, bem como a erros experimentais. À medida que a distância cresce, a variância acumulada também cresce. No entanto, 88,5 % da variância total está concentrada nas distâncias de até 30,5 m, o que permite dizer que a contribuição restante é muito pequena. Assim, pode-se admitir que para distâncias maiores que 30,5 m, o gráfico assume um patamar praticamente constante, denominado “Sill”. Assim, pode-se concluir que a variância desta propriedade é função da distância de separação entre os pontos, para distâncias de até 30,5 metros. A partir daí, a distância de separação entre pontos amostrais não mais altera a variância dos dados. Evidentemente, estudos futuros da variabilidade espacial desta propriedade deve concentrar a atenção em distâncias desta ordem.

Muito embora as ferramentas da geoestatística permitam um estudo muito mais bem elaborado da variabilidade espacial, a amostragem hierárquica pode ser de grande utilidade para uma análise exploratória em áreas onde nada se conhece sobre a variação da propriedade no espaço. De fato, como ponto de partida, pode-se realizar uma amostragem hierárquica com diferentes espaçamentos na área em estudo e, com base nas informações resultantes da análise de variância hierárquica, estabelecer espaçamento adequado para uma futura amostragem sistemática com vistas ao estudo usando as ferramentas da geoestatística, discutidas a seguir.

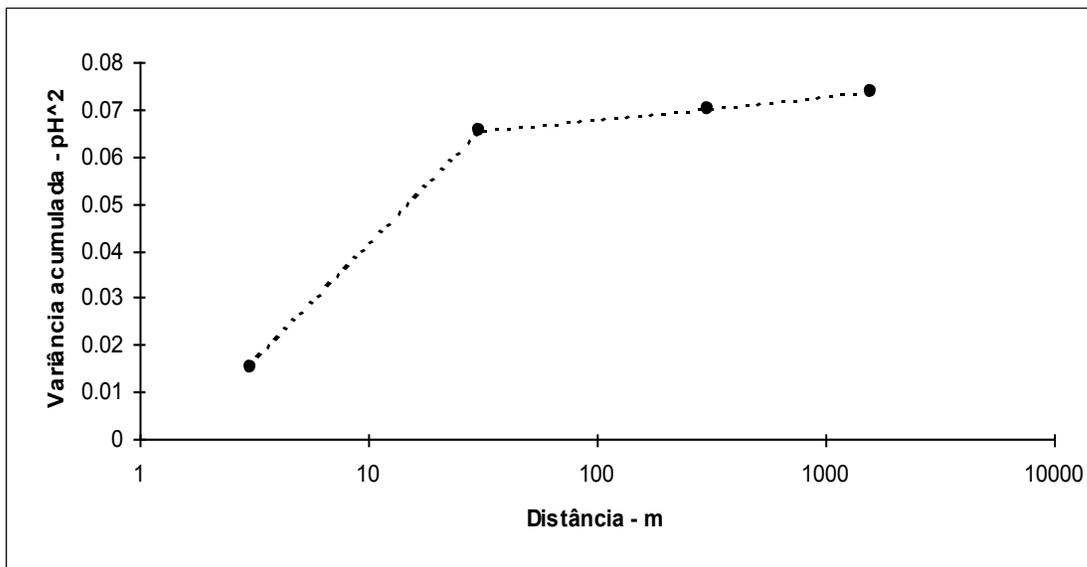


Figura 3.2 Distribuição da variância acumulada com a distância de separação.

### 3.4 VARIABILIDADE DA SALINIDADE DO SOLO

#### 3.4.1 *Análise exploratória de dados*

##### **Introdução**

A estatística atualmente é empregada em praticamente todas as áreas do conhecimento humano, tendo se tornado uma poderosa ferramenta no planejamento de experimentos, organização, apresentação e análise de dados, permitindo a dedução de conclusões válidas, para um dado nível de confiança, e auxiliando na tomada de importantes decisões.

Em agricultura irrigada, os parâmetros envolvidos, sejam relativos a água, ao solo, a planta ou a atmosfera, exibem um comportamento variável no espaço e no tempo, requerendo uma boa análise estatística para interpretação dos dados. Uma análise exploratória adequada dos dados constitui o ponto de partida para observar o comportamento das variáveis de interesse, o que pode ser feito por meio de técnicas estatísticas descritivas gerais e espaciais. Esta análise servirá de base para orientar na decisão de se assumir algum tipo de estacionariedade dos dados, o que é de fundamental importância nas análises geoestatísticas.

A salinidade, assim como outras propriedades físicas e químicas do solo, apresentam uma variabilidade espacial (e temporal) natural devido a influências das práticas de manejo utilizadas, profundidade do lençol freático, permeabilidade do solo, taxa de evapotranspiração, chuvas, salinidade da água subterrânea e outros fatores geohidrológicos.

A análise de dados de salinidade deve ser realizada de modo a permitir a identificação de áreas problema e o monitoramento de variações temporais, tornando

possível planejar o estudo de recuperação, definir estratégias de manejo, bem como identificar que fatores estão contribuindo para o problema.

### ***Análise descritiva geral***

A análise descritiva geral tem por objetivo fazer uma identificação inicial do comportamento dos dados, tratando os valores de cada variável como independentes espacialmente, isto é, neste tipo de análise a posição onde a variável é medida não é levada em consideração. Os procedimentos utilizados permitem interpretar os dados por meio de medidas sínteses, conhecidas como estatísticas descritivas.

Assumir independência entre as medidas de salinidade do solo, por exemplo, significa considerar que a obtenção de um determinado valor não tem nenhum tipo de influência na obtenção de qualquer outro.

Assim como todas as outras técnicas estatísticas, a descritiva exige muita experiência, bom senso e precisão para interpretação dos resultados.

Neste ítem comentaremos, de forma resumida, os seguintes recursos da estatística descritiva: a) principais estatísticas descritivas; b) histogramas; c) gráficos de probabilidade e d) “box-plot” ou gráficos de caixa. O objetivo é fazer uma apresentação didática, incluindo um exemplo de aplicação com dados hipotéticos de Cees.

***Principais estatísticas descritivas:*** As estatísticas descritivas resumem o comportamento geral dos dados, sendo representadas por meio de medidas de posição e de variabilidade. As medidas de posição são utilizadas para representar um determinado fenômeno por meio de valores em torno dos quais tende a haver uma maior concentração dos dados observados (Fonseca & Martins, 1993).

A média aritmética, a moda e a mediana, representam as três medidas de posição mais frequentemente utilizadas para descrever a tendência central de um conjunto de dados. As médias geométrica, quadrática, biquadrática, harmônica e cúbica, são de uso menos frequente.

Os quartís, embora não sendo medidas de tendência central, são medidas de posição que apresentam concepção semelhante a mediana. Enquanto a mediana divide o conjunto de dados em duas partes iguais quanto ao número de elementos de cada parte, os quartís divide-os em quatro partes iguais. Ao leitor interessado numa leitura detalhada sobre estas medidas, segure-se consultar literaturas específicas de estatística (Toledo & Ovalle, 1982; Spiegel, 1985; Costa Neto, 1990; Fonseca & Martins, 1993; Beiguelman, 1994; entre outros).

Numa distribuição normal, como os dados são simétricos em torno de um valor central, a média, a moda e a mediana, são coincidentes, podendo-se utilizar a média como medida de posição representativa dos dados, caso seja necessário obter outros parâmetros como o desvio padrão e a variância. Por outro lado, sendo a distribuição assimétrica, como é o caso de dados com distribuição do tipo log-normal, por exemplo, é preferível usar a mediana, por tratar-se de uma medida resistente a influência de valores extremos (Parkin & Robinson, 1992).

A verificação da normalidade dos dados pode ser baseada nos coeficientes de assimetria ( $C_s$ ) e curtose ( $C_r$ ), definidos conforme as Equações 3.6 e 3.7, ou através de testes de aderência como o qui-quadrado ( $X^2$ ) ou Kolmogorov-Smirnov (Spiegel, 1985; Costa Neto, 1990).

$$C_s = \frac{\hat{m} - M_o}{S} \quad (3.6)$$

$$C_r = \frac{M_4}{S^4} \quad (3.7)$$

onde  $\hat{m}$  é a estimativa da média da população,  $M_o$  é a moda,  $S$  é a estimativa do desvio padrão,  $M_4$  é o momento da quarta ordem centrado na média e  $S^2$  é o quadrado da variância. Valores de  $C_s = 0$  e  $C_r = 3$ , caracterizam uma distribuição normal, sendo utilizado a teoria estatística da decisão (testes de hipóteses) para testar se os valores encontrados para um conjunto de dados são estatisticamente iguais a 0 e 3, respectivamente. Segundo Spiegel (1985), muitas vezes, utiliza-se apenas a raiz quadrada do coeficiente momento de assimetria ( $b_1$ ), definido conforme a Equação 3.8, para representar o coeficiente de assimetria da distribuição de probabilidade:

$$b_1 = \frac{m_3^2}{m_2^3} \quad (3.8)$$

onde  $m_2$  e  $m_3$  são os momentos de segunda e terceira ordem, centrados na média. Para uma distribuição perfeitamente simétrica,  $\sqrt{b_1} = 0$ .

No estudo da variabilidade dos dados, as medidas clássicas mais utilizadas da estatística são: a variância, o desvio padrão e o coeficiente de variação. A amplitude total e a amplitude interquartílica, também são medidas frequentemente utilizadas para analisar a dispersão de um conjunto de dados. A amplitude total, entretanto, por ser bastante influenciada pela ocorrência de observações perturbadoras, isto é, observações que produzem efeito adverso substancial nas medidas estatísticas não resistentes, constitui uma medida que apresenta grande instabilidade, podendo conduzir a interpretações errôneas à respeito do fenômeno estudado. Por outro lado, a amplitude interquartílica representa uma medida insensível a estas observações, constituindo uma medida considerada resistente.

**Histogramas:** Uma vez coletado um conjunto de dados, o passo seguinte consiste em fazer uma apuração dos mesmos, para posterior tabulação. A apuração diz respeito a ordenação (ordem crescente ou decrescente) e contagem. Após a ordenação, tomando-se os valores máximo e mínimo, por diferença se obtém a amplitude total, a qual é dividida pelo número total de classes para determinação do intervalo de classes. Tendo-se os intervalos de classe e a frequência de ocorrência dos dados em cada intervalo, constrói-se o histograma.

Os histogramas constituem uma ferramenta gráfica muito importante da estatística descritiva, os quais dão uma primeira visão sobre o comportamento do fenômeno como um todo e permitem observar características como: simetria, distribuição, observações perturbadoras (valores discrepantes em relação ao conjunto de dados), concentração e lacunas nos dados. Uma predominância de valores superiores à moda revela uma distribuição assimétrica à direita (assimetria positiva). Neste caso, ocorre uma maior

concentração dos dados na escala inferior do gráfico, sendo a média aritmética superior a mediana e a moda. Por outro lado, uma predominância de valores inferiores à moda indica uma distribuição assimétrica à esquerda (assimetria negativa), tendo-se uma maior concentração de dados na escala superior do gráfico. A média é menor que a mediana e a moda.

Atualmente a construção de histogramas, inclusive com o ajuste da melhor distribuição de frequência, é feito de forma eficiente e rápida através de inúmeros programas computacionais disponíveis (EXCEL, STATÍSTICA, SANEST, GeoEAS, GRAPH, etc.).

**Gráficos de probabilidade:** Os gráficos de probabilidade ou “probability plot” constituem um instrumento bastante útil na análise descritiva geral, permitindo verificar se os dados são provenientes ou não de uma população de distribuição normal, como também detectar a existência de valores discrepantes. Quando o gráfico apresenta um comportamento linear aceita-se a hipótese de que os dados se ajustam a uma distribuição normal, podendo-se usar a média como medida de tendência central representativa do conjunto de dados.

Ao se usar o logaritmo dos dados, um comportamento linear no gráfico de probabilidade indica ser razoável aceitar a hipótese de que os dados originais ( $X_i$ ) são provenientes de uma população log-normalmente distribuída e que os dados transformados,  $Y_i = \ln X_i$ , se ajustam a uma distribuição normal. Os programas Geo-EAS (Geostatistical Enviromental Assesment Software) e o SAS (Statistical Analysis Sistem) permitem a obtenção dos gráficos de probabilidade com relativa facilidade.

**“Box-plot” ou gráfico de caixa:** O gráfico “box-plot” representa o resumo de 5-números de um conjunto de dados (limite inferior, limite superior, quartil inferior, quartil superior e mediana), podendo também apresentar os valores acima dos limites considerados “críticos”, isto é, os valores discrepantes ou “outliers”. Constitui um gráfico especialmente útil para comparação de vários grupos de dados. Quando feito com o dados de cada linha e de cada coluna permite uma visualização do comportamento espacial dos dados em duas dimensões.

A caixa tem comprimento igual a dispersão dos quartís ou amplitude interquartílica. Duas linhas contínuas saindo da caixa, uma da base inferior e outra da base superior, indicam os limites críticos inferior e superior das das observações não “outliers”. Normalmente acima dos limites críticos são destacados os valores considerados “outliers”. De acordo com Hoaglin *et al.* (1983) os limites críticos inferior ( $L_i$ ) e superior ( $L_s$ ) são dados pelas Equações 3.9 e 3.10, respectivamente:

$$L_i = Q_i - 1,5D_q \quad (3.9)$$

$$L_s = Q_s + 1,5D_q \quad (3.10)$$

onde,

$Q_i, Q_s$  - quatís inferior e superior, respectivamente;

$D_q$  - dispersão dos quartís.

**Exemplo:**

A partir de um conjunto de dados hipotéticos de CEes, serão utilizados alguns recursos da estatística descritiva com o objetivo de mostrar a utilização de tais recursos na análise de dados de salinidade, sem, no entanto, ter a pretensão de discutir detalhadamente os resultados. Para tanto, será adotado um esquema de amostragem sistemática, numa área retangular de 45 m x 120 m, quadriculada com espaçamento de 5 m, totalizando 10 linhas e 25 colunas, com um total de 250 pontos.

A Tabela 3.6 resume as principais estatísticas dos dados de CEes e dos valores transformados na forma logarítmica,  $Y_i = \text{LnCEes}$ , obtidas com o programa **STATÍSTICA**. O **GeoEAS** permite o cálculo de todos os momentos estatísticos apresentados, assim como vários outros programas computacionais. Pode-se verificar pela tabela que os valores de CEes não se ajustam a uma distribuição normal, enquanto os valores transformados se ajustam bem a este tipo de distribuição. Os histogramas das Figuras 3.3 e 3.4 e os gráficos de probabilidade das Figuras 3.5 e 3.6, construídos no módulo STAT1 do programa GeoEAS e melhorados em um programa de tratamento de imagens (o Paint de Windows 95), ilustram e reforçam esta análise. Observa-se claramente um comportamento assimétrico dos dados de CEes e simétrico para LnCEes, com média igual a mediana, o que é uma das características de dados com distribuição normal.

Tabela 3.6 Momentos estatísticos para os dados originais (CEes) e transformados ( $Y_i = \text{LnCEes}$ ) de condutividade elétrica do estrato de saturação.

Estatísticas	CEes	Ln(CEes)
	(dS m <sup>-1</sup> à 25°C)	
Número de dados	250	250
Média	9,90	2,17
Mediana	8,79	2,17
Variância	33,52	0,23
Desvio padrão	5,79	0,48
C.V. (%)	58,46	21,96
Mínimo	2,05	0,72
Máximo	41,98	3,74
Amplitude total	39,93	3,02
Quartil inferior	6,62	1,89
Quartil superior	11,42	2,44
Amplitude interquartilica	4,80	0,55
Assimetria	2,80	0,34
Curtose	11,10	1,17

Observa-se por meio das medidas de dispersão uma elevada variabilidade dos dados. A amplitude total (39,93 dS m<sup>-1</sup>) apresenta uma elevada diferença quando comparada com a amplitude interquartilica (4,80 dS m<sup>-1</sup>), o que se deve ao fato da primeira ser uma medida estatística de dispersão muito influenciada por observações

discrepantes, o que não ocorre com a segunda, a qual é considerada uma medida estatística de dispersão resistente.

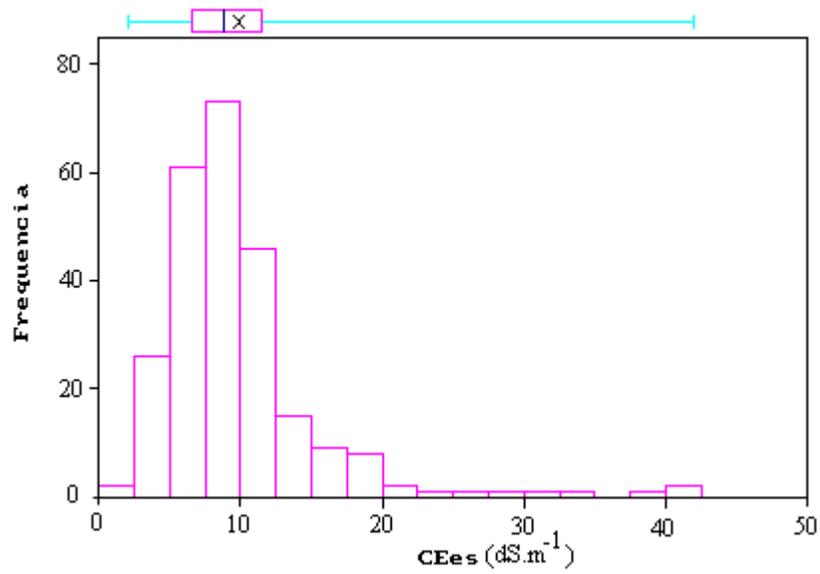


Figura 3.3 Histograma dos dados de condutividade elétrica do extrato de saturação (CEes).

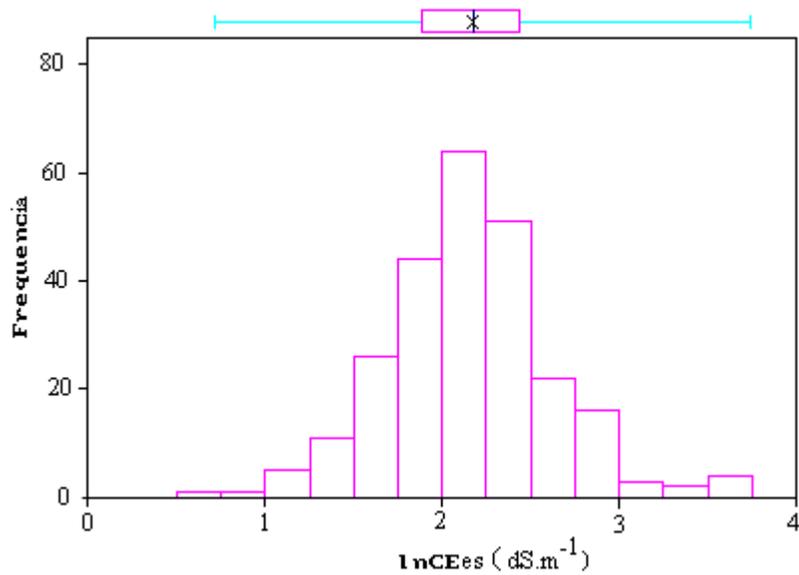


Figura 3.4 Histograma dos dados transformados de condutividade elétrica do extrato de saturação ( $Y = \text{LnCEes}$ ).

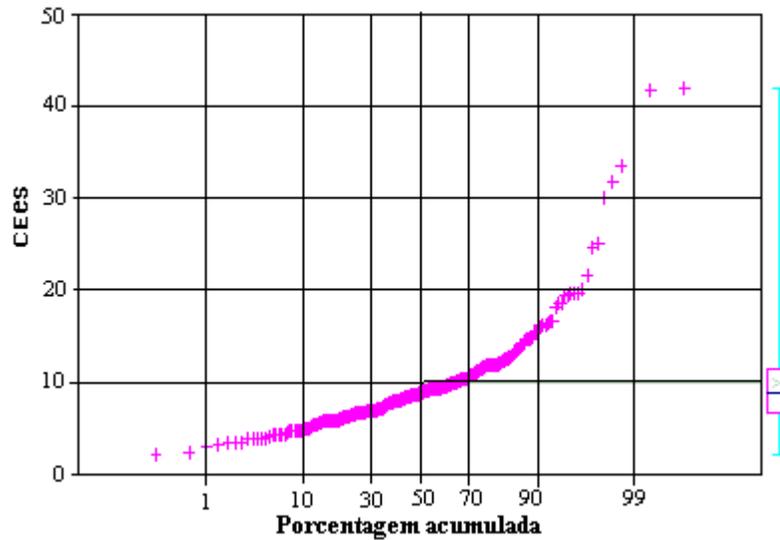


Figura 3.5 Gráfico de probabilidade normal dos dados de condutividade elétrica do estrato de saturação (CEEs).

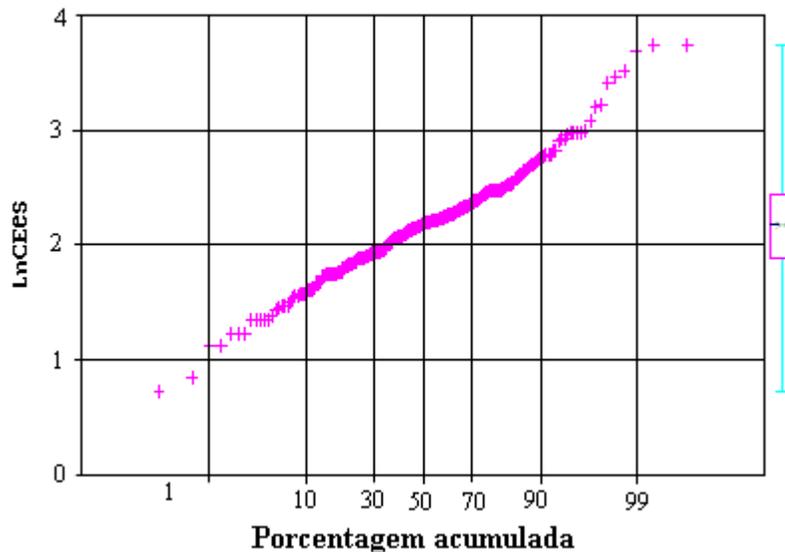


Figura 3.6 Gráfico de probabilidade normal dos dados transformados de condutividade elétrica do estrato de saturação ( $Y = \text{LnCEEs}$ ).

### ***Análise descritiva espacial***

As características espaciais de um conjunto de dados, tais como a localização de valores extremos, a tendência, ou o grau de continuidade, são frequentemente de considerável interesse (Isaaks & Srivastava, 1989).

A análise descritiva espacial, por considerar de alguma forma a posição espacial em que as variáveis são medidas, constitui uma alternativa para orientar na tomada de decisões quanto a eliminação ou não de dados, bem como a remoção de tendências, o que é de fundamental importância nas análises geoestatísticas.

Os principais recursos gráficos para análise descritiva espacial são: mapa de localização e valores de cada observação ou “datapost”, “postplot”, mapa de contorno, gráficos por linhas e colunas ( dispersão dos dados brutos, médias e desvios padrão, box-plot, etc.) e janelas móveis. A título de exemplo de aplicação, somente alguns desses recursos serão utilizados, para o mesmo conjunto de dados usados na análise descritiva geral. O objetivo é verificar a existência de tendências, heterogeneidades de variâncias e comportamentos de dados isolados. Para maiores detalhes sobre estes recursos sugere-se aos interessados consultar Isaaks & Srivastava (1989) e Ribeiro Júnior (1995). Alguns desses recursos também foram utilizados por Queiroz (1995) na análise da variabilidade espacial de parâmetros hidrodinâmicos de um solo de várzea para fins de drenagem subterrânea.

O “datapost” é um recurso gráfico que, além de revelar possíveis erros quanto a localização dos dados, permite chamar atenção quanto a valores possivelmente errôneos ou que estão associados a algum fenômeno localizado que precisa ser investigado. Algumas características dos dados aparecem de forma clara no “datapost”, podendo-se destacar zonas de valores máximos e mínimos. Para o mesmo conjunto de dados utilizado na análise descritiva geral, obteve-se o “datapost” apresentado na Figura 3.7. Observa-se uma concentração de maiores valores no canto inferior direito da área, o que está associado a algum efeito localizado.

O “postplot” é um recurso gráfico muito parecido com o “datapost”, onde o valor da propriedade analisada é distribuído em classes, em vez de ser assinalado o seu valor bruto. No caso de uso do programa GoeEAS, as seguintes classes são consideradas:  $C_1$  - entre o mínimo e o quartil inferior,  $C_2$  - entre o quartil inferior e a média,  $C_3$  - entre a média e o quartil superior e  $C_4$  - entre o quartil superior e o valor máximo. Este gráfico dar uma idéia inicial sobre a continuidade do fenômeno. Variações gradativas numa dada direção indicam a existência de tendência, o que é incompatível com as hipóteses de estacionariedade assumidas nas análises geoestatísticas.

A Figura 3.8 mostra o “postplot” dos valores dos logaritmos dos dados de condutividade elétrica do estrato de saturação ( $Y = \ln CE_s$ ), gerado no módulo POSTPLOT do GeoEAS. Observa-se, conforme já confirmado pelo “datapost”, uma tendência de concentração de maiores valores no canto inferior direito da área, o que não chega a caracterizar uma tendência global.



Os gráficos por linhas e por colunas permitem identificar variações da propriedade de interesse em relação às direções da malha experimental, podendo-se também utilizar para esta finalidade janelas móveis, as quais consistem na divisão da área em várias subáreas vizinhas e de igual tamanho, englobando um certo número de linhas e de colunas, podendo haver ou não superposição. Através das janelas móveis é possível identificar regiões onde os dados são mais variáveis do que outras, indicando heterocedasticidade nos dados, o que tem sérias implicações práticas que devem ser analisadas.

Para o conjunto de dados de CEes que vem sendo utilizado como exemplo, foram construídos os gráficos de “box-plot” múltiplos por colunas (25 dados por coluna) e por linhas (10 dados por linha), sem delimitação de janelas móveis, apresentados através das Figuras 3.9 e 3.10. As colunas e as linhas com valores elevados ficam bem evidenciadas nestas figuras. Observa-se, entretanto, não haver uma tendência global nas direções X e Y da malha experimental. De um modo geral as análises descritivas mostram que a estacionariedade intrínscica pode ser assumida, indicando não haver restrição quanto a construção do semivariograma.

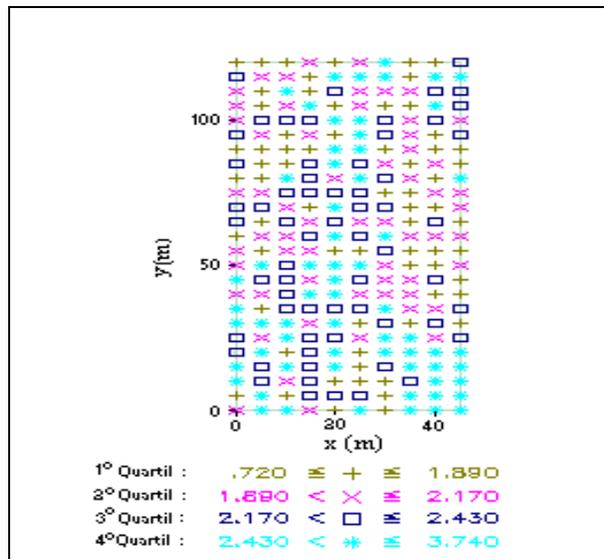


Figura 3.8 “Postplot” dos dados transformados de CEes,  $Y_i = \text{LnCEes}$ , na área amostrada.

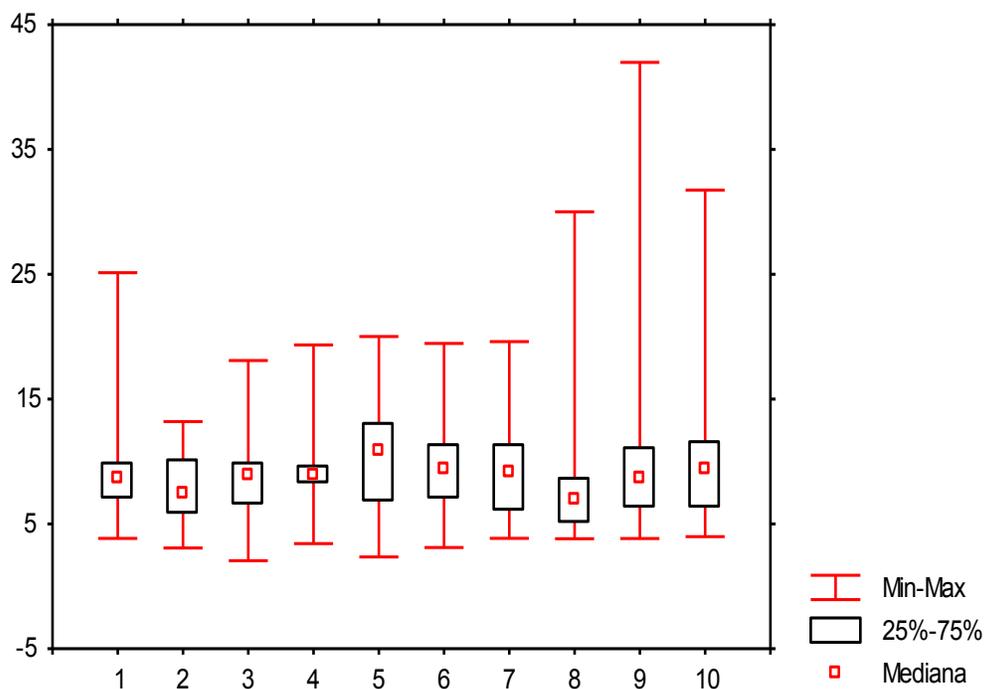


Figura 3.9 Gráficos “box-plot” por coluna dos dados de condutividade elétrica do estrato de saturação (CEEs).

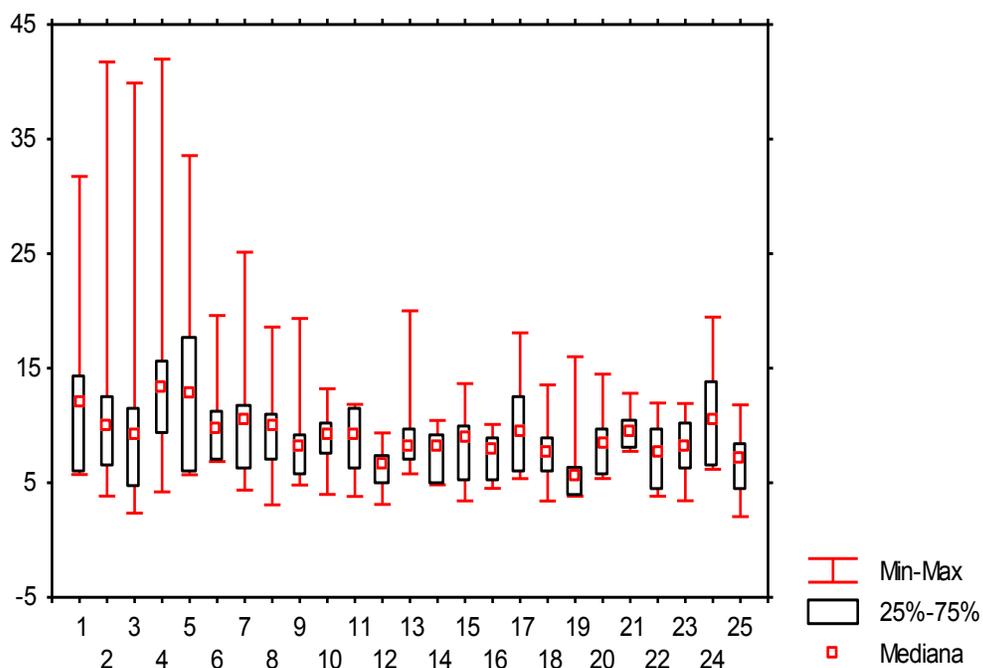


Figura 3.10 Gráficos “box-plot” por linha dos dados de condutividade elétrica do estrato de saturação (CEEs).

### 3.4.2 Técnicas geoestatísticas

#### ***Dependência espacial***

Os métodos clássicos da estatística assumem que o valor médio de uma propriedade de solo em uma região amostrada é igual ao valor da esperança desta propriedade em qualquer ponto dentro da região, com um erro de estimativa correspondente à variância dos dados amostrais, conforme Trangmar *et al.* (1985). Isto corresponde a assumir que a variabilidade em torno da média é aleatória e independente da posição espacial dos valores amostrais. No entanto, muitos trabalhos, como o de Vieira *et al.* (1981), mostram que a variabilidade de propriedades do solo é espacialmente dependente, ou seja, dentro de um certo domínio, as diferenças entre valores de uma propriedade do solo podem ser expressas como uma função da distância de separação entre estes pontos medidos.

Webster & Olivier (1990) afirmam que muitas propriedades do solo variam continuamente no espaço e, conseqüentemente, os valores em locais mais próximos entre si tendem a ser mais semelhantes, até um determinado limite, correspondente ao domínio destas propriedades, que aqueles tomados a maiores distâncias. Caso isto ocorra, os dados não podem ser tratados como independentes e um tratamento estatístico mais adequado é necessário.

As ferramentas estatísticas que consideram a posição de cada valor no espaço tem como base trabalhos desenvolvidos por Krige (citado por WEBSTER & OLIVIER, 1990), em mineração, os quais foram seguidos pelos trabalhos de MATHERON (1971). Estes trabalhos formalizam a teoria das variáveis regionalizadas, fundamento da geoestatística. Os conceitos de variável aleatória, variável regionalizada, funções aleatórias e estacionariedade são fundamentais para a geoestatística.

A geoestatística possibilita a descrição quantitativa da variação espacial no solo, a estimativa não tendenciosa e de variância mínima de valores da propriedade em locais não amostrados, permitindo a construção de mapas de valores e também a identificação de esquemas de amostragem eficientes.

#### ***Hipóteses de estacionariedade***

Uma medida de uma propriedade do solo em uma posição qualquer pode ser entendida como uma realização de uma **variável aleatória (v.a.)**, a qual deve variar segundo alguma lei de distribuição de probabilidade que pode ser descrita pelos seus parâmetros. De acordo com Journel & Huijbregts (1978), uma **variável regionalizada**  $z(x)$  é uma variável aleatória que assume diferentes valores  $z$  em função da posição  $x$  dentro de uma certa região. De acordo com Trangmar *et al.* (1985), uma variável regionalizada  $z(x)$  pode ser considerada como uma realização particular de uma variável aleatória  $Z$  para uma dada localização  $x$ . O conjunto de variáveis  $z(x)$  medidas em todos os pontos  $x$  pode ser considerado uma função aleatória  $Z(x)$ , uma vez que, segundo Isaaks & Srivastava (1989), são variáveis aleatórias, regionalizadas e assume-se que a dependência entre elas é especificada por algum mecanismo probabilístico.

Segundo Journel & Huijbregts (1978), a interpretação probabilística de que a variável regionalizada  $z(x)$  é uma particular realização de uma certa função aleatória  $Z(x)$  é consistente quando se pode inferir toda ou pelo menos parte da lei de distribuição de probabilidade que define esta função aleatória. Em problemas práticos no entanto, em

cada ponto  $x$  tem-se apenas uma realização  $z(x)$  e o número de pontos  $x$  é sempre finito. Isto torna usualmente impossível inferir sobre a distribuição de  $Z(x)$ . Em vista disto, certas hipóteses são necessárias, as quais envolvem diferentes graus de homogeneidade espacial, sendo comumente denominadas hipóteses de estacionariedade.

Assumir que em cada ponto  $x$  tem-se uma v.a. e que todas elas apresentam a mesma distribuição de frequência seria uma hipótese de estacionariedade extremamente forte, denominada estacionariedade estrita por Journel & Huijbregts (1978). O que se adota usualmente então são hipóteses menos rígidas, admitindo-se que apenas alguns momentos das distribuições das v.a. são iguais, uma vez que isto é o bastante para a geoestatística.

A hipótese de estacionariedade de primeira ordem é definida por Trangmar *et al.* (1985) como sendo a hipótese de que o momento de primeira ordem da distribuição da função aleatória  $Z(x)$  é constante em toda a área, ou seja:

$$E[Z(x)] = m \quad (3.11)$$

em que  $m$  é a média dos valores amostrais.

Decorre desta definição que se for tomado um vetor  $h$  de separação entre dois pontos, o qual apresenta módulo e direção, para qualquer  $h$  tem-se:

$$E[Z(x) - Z(x+h)] = 0 \quad (3.12)$$

A estacionariedade de segunda ordem é definida por Journel & Huijbregts (1978) quando a função aleatória, além de atender à estacionariedade de primeira ordem, apresenta a característica de, para cada par de valores  $\{Z(x), Z(x+h)\}$ , a covariância existe e depende apenas da distância de separação  $h$ :

$$C(h) = E[Z(x+h) - m][Z(x) - m] = E[Z(x+h).Z(x)] - m^2 \quad (3.13)$$

Usualmente a geoestatística não necessita do atendimento a esta estacionariedade, bastando uma estacionariedade menos forte, definida como sendo a hipótese intrínseca que, segundo Journel & Huijbregts (1978), corresponde ao fato de que para todo vetor  $h$ , a variância da diferença  $[Z(x) - Z(x+h)]$  é finita e independente da posição na região, dependendo apenas do valor de  $h$ , ou seja:

$$\text{VAR}[Z(x) - Z(x+h)] = E[Z(x) - Z(x+h)]^2 = 2 G(h) \quad (3.14)$$

Nesta expressão é definida a função de semivariância,  $G(h)$ , igual a esperança do quadrado da diferença entre dois valores defasados de uma distância  $h$ .

A determinação do valor do semivariograma poderia ser feita para cada valor de  $h$ , caso houvesse disponível várias realizações da v.a.  $z(x)$  em cada posição  $x$ . Como em cada posição  $x$  usualmente tem-se apenas uma realização da v.a., apenas um par de valores  $[Z(x), Z(x+h)]$  existe para cada posição  $x$ . Para contornar este problema, segundo Journel & Huijbregts (1978), é introduzida a hipótese intrínseca, segundo a qual a função de semivariância é função apenas do vetor de separação  $h$ , e não da localização na região. Isto torna possível estimar a semivariância  $G(h)$  a partir dos dados disponíveis, definindo-

se um estimador  $G^*(h)$  igual à média aritmética das diferenças quadradas entre dois valores experimentais  $[Z(x_i), Z(x_i + h)]$ , em todos os pontos separados pela distância  $h$ :

$$G^*(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i + h)]^2 \quad (3.15)$$

em que  $N(h)$  é o número de pares de pontos defasados pelo vetor  $h$ .

Journel & Huijbregts (1978) salientam ainda que a hipótese intrínseca corresponde à hipótese de estacionariedade de segunda ordem da v.a. constituída pela diferença  $[Z(x) - Z(x+h)]$ . Esta v.a. está sempre presente na geoestatística. Em termos físicos, assume-se que a estrutura de variabilidade entre dois pontos, dentro da região de estudo, independe da posição. Quando a propriedade se distribui de forma homogênea, isto é verdadeiro.

Na prática, contudo, muitas vezes uma propriedade nega a estacionariedade dentro da região estudada, apresentando por exemplo uma tendência de crescimento de valores numa determinada direção. No entanto, Journel & Huijbregts (1978) definem uma vizinhança de quase-estacionariedade ( $b$ ), dentro da qual a propriedade pode ser considerada estacionária. Caso o semivariograma, como definido a seguir, seja usado em distâncias inferiores a este valor  $b$ , pode-se assumir estacionariedade da propriedade.

### ***Análise da dependência espacial***

A partir dos conceitos apresentados para variáveis regionalizadas, funções aleatórias, estacionariedade, assim como a definição da função de semivariância e do seu estimador clássico, pode-se estudar a dependência espacial em um conjunto de valores de uma determinada propriedade, em uma região onde se procedeu uma amostragem, usualmente sistemática, mas que, sobretudo, sejam conhecidas as posições correspondentes aos valores.

A função de semivariância pode ser estimada para um dado valor de  $h$ , vetor de separação entre dois pontos, por meio do estimador clássico definido. Se na região amostrada a máxima distância de separação for representada por  $L$  e os pontos estiverem espaçados entre si de uma distância “ $e$ ”, então, pode-se definir um conjunto de valores de  $h$ , variando de zero até à metade de  $L$ , com incremento igual a “ $e$ ” e, em seguida, calcular o valor da função de semivariância para cada valor de  $h$ . Inicialmente, tomam-se todos os pares de pontos defasados entre si do valor de  $h$ , independente da direção. O mesmo procedimento pode ser adotado para diferentes direções dentro da área. O gráfico que relaciona o valor de semivariância obtido com o correspondente valor de  $h$ , é denominado semivariograma. A Figura 3.11 representa um esquema de um semivariograma, mostrando os seus elementos constituintes.

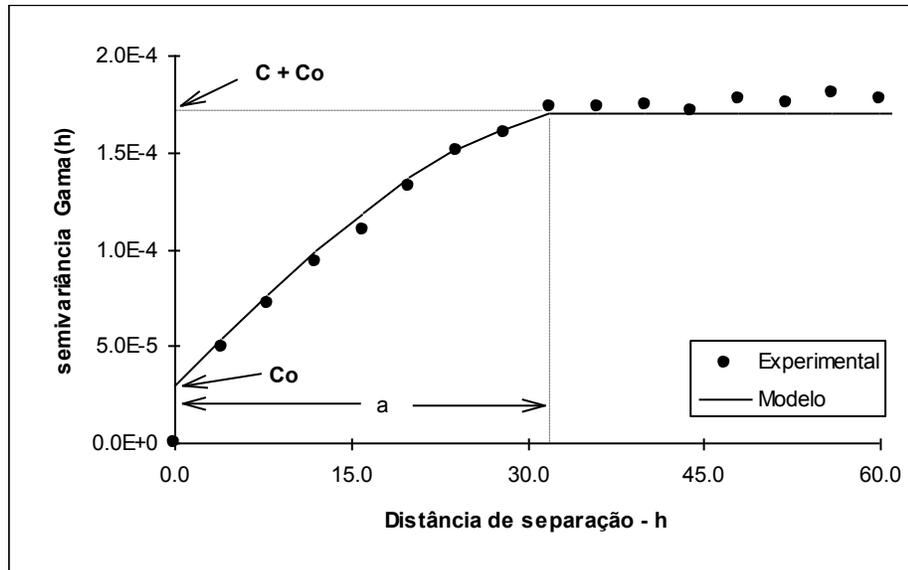


Figura 3.11 Semivariograma experimental e modelo matemático ajustado.

Na Figura 3.11 é mostrado o comportamento típico de um semivariograma experimental de uma propriedade que apresenta dependência espacial. Pontos próximos entre si são mais semelhantes que pontos mais afastados. Como a função de semivariância quantifica a dessemelhança entre os pontos, ao contrário da correlação, então o semivariograma começa com um baixo valor, denominado efeito pepita ou “nugget effect” e representado por  $C_0$  e cresce à medida que  $h$  cresce, até uma distância “ $a$ ”, denominada alcance ou “range” do semivariograma, a qual determina a distância até onde a propriedade se apresenta espacialmente dependente. A partir desta distância, os dados podem ser considerados independentes, condição assumida pela estatística. Para distâncias maiores que o alcance, o semivariograma tende a se estabilizar em torno de um valor denominado patamar, representado por  $C + C_0$ , onde  $C$  é denominado “Sill”, nos textos de geoestatística. A curva apresentada na Figura 3.11 corresponde a um modelo matemático ajustado aos valores experimentais, o que será tratado a seguir.

### ***Modelos de semivariogramas***

Uma vez que as propriedades do solo variam continuamente no espaço, os seus semivariogramas são funções contínuas. O semivariograma experimental consiste em alguns pontos estimados ao longo desta função, a partir dos pontos experimentais e, portanto, sujeitos a erros. De acordo com Webster (1985), embora os pontos em um semivariograma experimental bem estimado se apresentem de forma ainda irregular, usualmente é possível ajustar funções simples a eles. Basicamente, um modelo matemático a ser ajustado ao semivariograma precisa incluir os três parâmetros descritos anteriormente: um **intercepto** ou efeito pepita, um **patamar** ou “Sill”, atingido após uma distância correspondente ao **alcance**. Além disto, a forma da curva deve se ajustar aos pontos experimentais na região de crescimento da função, ou seja, para  $h$  entre zero e o alcance. Diversos modelos matemáticos podem ser usados, deste que algumas condições

sejam atendidas, como descrito por Mcbratney & Webster (1986). Estes autores apresentam os modelos mais comumente usados, ou seja:

1) modelo linear com patamar:

$$\begin{aligned} G(h) &= C_0 + C(h/a) && \text{para } 0 < h < a \\ G(h) &= C_0 + C && \text{para } h > a \end{aligned} \quad (3.16)$$

2 ) modelo esférico:

$$\begin{aligned} G(h) &= C_0 + C \{ [3h/2a] - [(1/2) (h/a)^3] \} && \text{para } 0 < h < a \\ G(h) &= C_0 + C && \text{para } h > a \end{aligned} \quad (3.17)$$

3 ) modelo esférico:

$$G(h) = C_0 + C \{ 1 - \exp(-3h/a) \} \quad \text{para } h > 0 \quad (3.18)$$

Esta função cresce assintoticamente em relação ao patamar. Outros modelos podem ser encontrados na literatura de geoestatística. Webster (1985) apresenta uma discussão detalhada sobre o assunto.

Obtido o semivariograma experimental para um conjunto de valores de uma propriedade e ajustado um modelo matemático a este, tem-se uma função contínua que descreve a dependência espacial da propriedade, sendo de grande utilidade para a compreensão de vários aspectos da variabilidade do solo, da sua formação, implicações no manejo, etc. Trangmar *et al.* (1985) apresentam discussão muito útil sobre estas aplicações. Pode-se dizer que a função de semivariograma descreve uma característica intrínseca àquela propriedade, nas condições estudadas. Esta característica, quantificada e descrita por um modelo matemático, pode ser usada para, talvez, o mais importante recurso da geoestatística: a interpolação por Krigagem, descrita a seguir.

### ***Interpolação por krigagem***

A krigagem é uma técnica de interpolação para estimativa dos valores de uma propriedade em locais não amostrados, a partir de valores vizinhos resultantes da amostragem realizada. Diversas outras técnicas estão disponíveis para este propósito. A krigagem, no entanto, faz uso de um interpolador linear não tendencioso e de variância mínima que assegura a melhor estimativa. Este estimador tem como base os dados amostrais da variável regionalizada e as propriedades estruturais do semivariograma obtido a partir destes dados. A teoria pertinente a esta técnica está descrita em Journel & Huijbregts (1978), Isaaks & Srivastava (1989), Vieira *et al.* (1983) e outros, não sendo o propósito deste texto. Diversas formas de krigagem são utilizadas, cada qual compatível com as características do problema em questão. Para fins de estudo de salinidade de solo, a forma mais simples de krigagem, a krigagem ordinária puntual, pode fornecer resultados muito úteis, permitindo o mapeamento dessa propriedade em áreas amostradas.

Para o uso da krigagem em mapeamento de uma propriedade do solo, o mais conveniente é a estimativa de valores nos vértices de um grid suficientemente fino para os propósitos em questão. Diversos programas computacionais permitem a realização da krigagem desta forma, a partir do grid de dados originais e do modelo matemático

ajustado ao semivariograma experimental. Veremos a aplicação destes recursos computacionais com um exemplo de aplicação, apresentado a seguir.

### ***Exemplo de aplicação - uso de programas computacionais***

Para aplicação da teoria geoestatística em estudos de salinidade, vamos considerar o mesmo conjunto de dados utilizado nas análises descritivas. De um modo geral, estas análises mostraram não haver incompatibilidade em aceitar a hipótese intrínseca, indicando que a estrutura de dependência espacial entre os valores de CEes pode ser analisada através da construção do semivariograma.

Para obtenção do semivariograma pode-se usar o módulo PREVAR do GeoEAS para construir todos os pares de valores possíveis, usando-se como FRACTION o valor 0,6, de forma a que o número de pares não supere o máximo possível. Após o uso deste módulo, usa-se o módulo VARIO, para a construção do semivariograma, com distância máxima de 45 metros e lag de 5 metros. O gráfico resultante, bem como o modelo ajustado, é mostrado na Figura 3.12.

O modelo que melhor se ajustou ao semivariograma experimental foi o modelo exponencial, com efeito pepita igual a 0,062, Sill igual a 0,154, ou seja o patamar do semivariograma igual a 0,216 e um alcance de 18 metros. A curva correspondente a este modelo é mostrada na Figura 3.12. O ajuste do modelo pode ser feito dentro do módulo VARIO do GeoEAS ou utilizando-se o software VARIOWIN. Também pode ser feito um ajuste por tentativas em uma planilha eletrônica qualquer.

Os programas computacionais elaborados por Vieira (1995) também podem ser usados para o estudo descritivo e para o estudo da dependência espacial dos dados. O uso destes programas, elaborados em FORTRAN, é descrito na referência citada. Apresentam a conveniência de não serem programas fechados, mas sim rotinas que podem ser seguidas passo a passo, possibilitando a compreensão do que está sendo executado.

De posse da função de semivariograma pode-se proceder à Krigagem, ou seja, a partir dos pontos experimentais e do semivariograma, estimar o valor da propriedade em locais não amostrados. Diversas formas de krigagem podem ser usadas. A mais simples é a Krigagem ordinária pontual, que atende aos nossos propósitos e será usada neste exemplo. Inicialmente deve-se definir espaçamento do grid fino para os vértices dos quais será feita a estimativa.

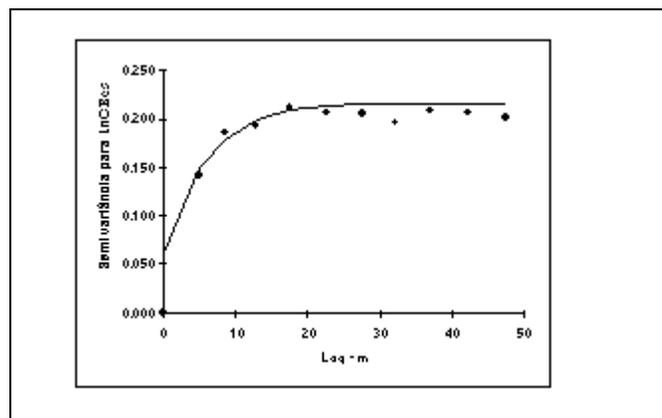


Figura 3.12 Semivariograma experimental e modelo ajustado para LnCEes.

Para o exemplo, adotar-se-á um espaçamento de 2 metros em ambas as direções, X e Y. O módulo KRIGE do programa GeoEAS permite a realização desta krigagem. Outra forma é pelo uso do programa SURFER para Windows, onde os dados amostrais são fornecidos, o modelo do semivariograma e o espaçamento para as duas direções. O programa retorna um conjunto de dados estimados, além dos dados experimentais. Estes dados podem ser usados para se traçar uma superfície de valores, assim como um mapa de curvas de nível. Para o exemplo, ambos são mostrados na Figura 3.13.

A observação da Figura 3.13 anterior fornece uma idéia da utilidade das técnicas da geoestatística na ciência do solo, bem como na irrigação e na drenagem, por permitir o mapeamento de uma propriedade dentro da área em estudo, de forma a se ter um conhecimento detalhado do seu comportamento, não se limitando a um valor único, considerado como representativo da área, qualquer que seja a forma de sua obtenção.

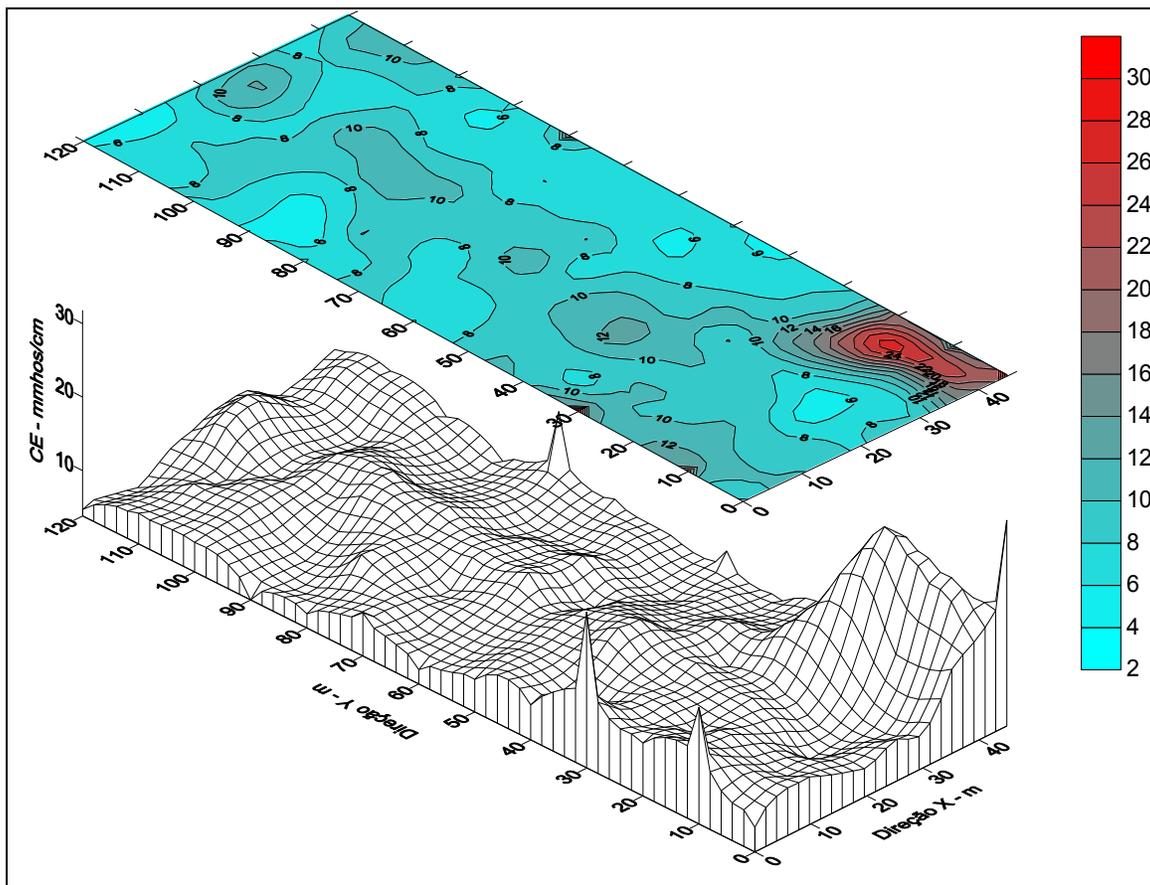


Figura 3.13 Superfície de valores de CEes e mapa de isoCEes na área em estudo.

### 3.5 CONCLUSÕES

A salinidade do solo é uma propriedade que varia no tempo e no espaço, o que requer processos adequados de amostragens, determinações e análises dos dados, que permitam obter informações de seu comportamento.

O monitoramento da salinidade é de fundamental importância para orientar as práticas de manejo da irrigação e as técnicas de recuperação de áreas afetadas. A utilização de recursos da estatística clássica e da geoestatística se complementam na análise e interpretação dos dados de salinidade. Porém, somente através da geoestatística é possível caracterizar a estrutura de dependência espacial entre os parâmetros que definem a salinidade do solo. Uma vez detectada a dependência espacial entre as observações, o processo de krigagem permite estimar valores em locais não amostrados, sem tendência e com variância mínima, assegurando a melhor qualidade das estimativas. Por outro lado, caso as observações sejam independentes espacialmente, os procedimentos clássicos da estatística podem ser utilizados para avaliar a distribuição de probabilidade e os momentos estatísticos dos parâmetros estudados.

O mapeamento pelo processo de krigagem é de grande importância para orientação do manejo e controle da salinidade, como também o planejamento de estudos quanto a resposta de cultivos em diferentes setores. O mapeamento permite uma visualização espacial da salinidade na área de interesse, o que constitui um aspecto de inquestionável importância no manejo da agricultura irrigada em zonas áridas e semi-áridas. Em resumo, neste capítulo procurou-se apresentar um texto didático quanto a importância de uma boa análise exploratória dos dados de salinidade, como base para orientação de posteriores análises geoestatísticas, e das técnicas geoestatísticas, como recursos complementares na análise e monitoramento da salinidade do solo.

### REFERÊNCIAS BIBLIOGRÁFICAS

- AHUJA, L.R. & NIELSEN, D.R. Field soil-water relations. In: **Irrigation of Agricultural Crops**. Agronomy Monograph nº 30. 143 - 189. 1990.
- BEIGUELMAN, B. **Curso prático de bioestatística**. 3.ed. Ribeirão Preto: Sociedade Brasileira de Genética, 1994. 244p.
- BRESLER, E. G.; DAGAN, G. & HANKS, R. J. Statistical analysis of crop yield under controlled line-source irrigation. **Soil Science Society of America Journal**, v.46, n.4, p.841-47, 1982.
- BOHN, H.L.; McNEAL, B.L.; O'CONNOR, G.A. **Soil chemistry**. New York, John Wiley & Sons, 1985. 341p.
- BURGESS, T.M. & WEBSTER, R. Optimal interpolation and isarithmic mapping of soil properties.4r I. The semi-variogram and punctual kriging. **Journal of Soil Science**, v.31, p. 315-331, 1980.
- COSTA NETO, P.L.O. **Estatística**. 10.ed. São Paulo: Edgard Blücher, 1990. 264p.
- FILGUEIRA, H.J.A.; SOUTO, J.S. Avaliação de quatro níveis de relação solo:água para a caracterização da condutividade elétrica em solos da região de Patos, PB. In:

- REUNIÃO BRASILEIRA DE FERTILIDADE DO SOLO E NUTRIÇÃO DE PLANTAS, 22. Manaus, 1995. **Anais**, 1995. p.p.282-283.
- FIREMAN, M.; WADLEIGH, C.H. A statistical study of the relation between pH and the exchangeable-sodium-percentage of western soils. **Soil Science**, v.71, p.273-85, 1951.
- FONSECA, J.S.; MARTINS, G.A. **Curso de estatística**. São Paulo: Atlas, 1993. 317p.
- HASSAN, H.M.; WARRICK, A.W. & AMOOZEGAR-FARD, A. Sampling volume effects on determining salt in a soil profile. **Soil Sci. Soc. of Am. Proc.** v.47, p.1265-67, 1983.
- HAWLEY, M.E.; McCUEN, R.H. & JACKSON, T.J. Volume-accuracy relation in soil moisture sampling. **Proc. Am. Soc. of Civil Eng. Irrig. and Drain. Div.** v.108, n. IR1, p.1-11, 1982.
- HOAGLIN, D.C.; MOSTELLER, F.; TYKEY, J.W. **Análise exploratória de dados: técnicas robustas; um guia**. Lisboa: Ed. Salamandra, 1983, 446p.
- ISAAKS, E.H. & SRIVASTAVA, R.M. **An introduction to applied geostatistics**. Oxford. Oxford University Press, 1989. 561p.
- JOURNEL, A.G, HUIJBREGTS, CH.J. **Mining geostatistics**. Academic Press, London, 1978. 600 p.
- KLAR, A.E. **A água no sistema solo-planta-atmosfera**. São Paulo, Nobel, 1984. 408p.
- KRIGE, D.G. Two-dimensional weighted moving average trend surfaces for ore-evaluation. **Journal of the south african institute of mining and metallurgy**, v.66, p.13-38, 1966.
- LASCANO, R. J. & HATFIELD, J. L. Spatial variability of evaporation along two transects of a bare soil. **Soil Sci. Soc. Am. J.**, v.56, p.341 -346, 1992.
- MATHERON, G. **The theory of regionalized variables and its applications**. Cahiers du Centre de Morphologie Mathématique de Fontainebleau n° 5. 1971.
- McBRATNEY, A.B. & WEBSTER, R. Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates. **Journal of Soil Science**, v.37, p.617-39. 1986.
- PARKIN, T.B.; ROBINSON, J.A. Analysis of lognormal data. **Advances in Soil Science**, v.20, p.193-235, 1992.
- PIZARRO, D. **Drenaje agrícola y recuperacion de suelos salinos**. Madrid, Ed. Agrícola Española, 1978. 520p.
- QUEIROZ, J.E. Parâmetros hidrodinâmicos de um solo de várzea para fins de drenagem subterrânea. Piracicaba, 1995. 167 p. Tese (Doutorado) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo.
- RAIJ, B. Van. **Fertilidade do solo e adubação**. São Paulo, Ed. Agrônômica Ceres, 1991. 343p.
- REICHARDT, K.; VIEIRA, S.R. & LIBARDI, P.L. Variabilidade espacial de solos e experimentação de campo. **Revista brasileira de ciência do solo**, v.10, p.1-6, 1986.
- RIBEIRO JÚNIOR, P. J. Métodos geoestatísticos no estudo da variabilidade espacial de parâmetros do solo. Piracicaba, 1995. 99p. Dissertação (Mestrado) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo.
- RICHARDS, L.A. **Diagnostico y Reabilitacion de Suelos Salinos y Sodicós**. Ed. Limusa. México, 1954. 172p.

- RHOADES, J.D. Electrical conductivity methods for measuring and mapping soil salinity. **Advances in agronomy**, v.49, p.201-51, 1994.
- RUSSO, D. A geostatistical approach to the solute transport in heterogeneous fields and its applications to salinity management. **Water Resour. Res.**, v.20, p.1260-70, 1984.
- RUSSO, D.& BRESLER, E. Soil hydraulic properties as stochastic processes: I. An analysis of field spatial variability. **Soil Science Society of America Journal**, v.45, n.4, p.682-87, 1981.
- SPIEGEL, M. R. **Estatística**. São Paulo: McGraw-Hill do Brasil, 1985. 454p.
- TANJI, K.K. **Agricultural salinity assessment and management**. New York. American Society of Civil Engineers, 1990. 619p.
- TOLEDO, G. R.; OVALLE, I. I. **Estatística básica**. 2.ed. São Paulo: Atlas, 1982. 459p.
- TRANGMAR, B. B.; YOST, R.S., & UEHARA, G. Applications of geostatistics to spatial studies of soil properties. **Advances in Agronomy**, 38, Academic Press, 45-94. 1985.
- VIEIRA, S.R. Geoestatística em estudos de variabilidade espacial. In: **Curso de atualização em conservação do solo**, Campinas, 1995. Campinas, IAC, 1995.
- VIEIRA, S.R.; NIELSEN, D.R.; BIGGAR, J.W. Spatial variability of field-measured infiltration rate. **Soil Science Soc. Am. Journal**, v.45, p.1040-48, 1981.
- VIEIRA, S.R.; HATFIELD, J.L.; NIELSEN, D.R. & BIGGAR, J.W. Geostatistical theory and application to variability of some agronomical properties. **Hilgardia**, v.31, n.3, 1983, 75p.
- WARRICK, A. W. & NIELSEN, D. R. Spatial variability of soil physical properties in the field. In: HILLEL, D.(ed.), **Applications of soil physics**. Academic press, New York, 1980.
- WEBSTER, R. & OLIVIER, M.A. **Statistical methods in soil and land resource survey**. Oxford University Press. Oxford. 1990. 316 p.
- WEBSTER, R. Quantitative spatial analysis of soil in the field. **Advances in Soil Science**, v.3, p.1-70, 1985.
- YOU DEN, W.J. & MEHLICH, A. Selection of efficient methods for soil sampling. **Contr. Boyce Thompson Inst. Plnt. Res.**, v.9, p. 59 - 70, 1937.