APLICAÇÃO DE MODELOS GEOESTATÍSTICOS UNIVARIADOS E BIVARIADOS NA ESTIMAÇÃO DE PRODUTIVIDADE DE SOJA, COM CONTRIBUIÇÃO DA VARIABILIDADE DO ÍNDICE DE CONE SOB DIFERENTES ESQUEMAS AMOSTRAIS

Edson Antonio Alves da Silva, Paulo Justiniano Ribeiro Jr. e Miguel Angel Uribe Opazo

RESUMO: O desenvolvimento de cultivares pode ser determinado por características físicas e químicas do solo que controlam o desenvolvimento radicular da planta, favorecendo ou restringindo a absorção de nutrientes. A produtividade da cultura pode ser correlacionada com qualquer dessas propriedades ou características pedológicas. Neste trabalho escolhemos correlacionar a produtividade de soja com o parâmetro físico Índice de Cone (resistência mecânica à penetração) por caracterizar o solo em várias propriedades, como porosidade, compactação, permeabilidade, dentre outros. A correlação serviu para aplicar um modelo geoestatístico bivariado visando, juntamente com informação de rendimento de soja em várias situações amostrais, obter estimativas por krigagem da produção total da área. Essas estimativas, quando comparadas com o rendimento conhecido, não mostraram grandes diferenças entre si, sendo todas muito próximas da produção obtida. Esse resultado levou a supor que a área escolhida para o experimento foi muito pequena e as variáveis não mostraram variabilidade suficiente para identificar estruturas acentuadas de dependência espacial. Todavia, a utilização de uma segunda variável no modelo de dependência espacial mostrou ser útil, principalmente em casos que essa variável secundária for de fácil obtenção.

INTRODUÇÃO: Em 1997 foi aprovado pelo Centro Nacional de Pesquisa – CNPq, um projeto em Agricultura de Precisão proposto por Souza *et al*. (1998) junto à Universidade Estadual do Oeste do Paraná – Unioeste, onde se desenvolveu, durante 5 anos, uma pesquisa que teve, dentre outros objetivos, aplicar metodologia geoestatística para produção de mapas temáticos de variáveis químicas e físicas do solo, bem como mapas de produtividade de soja cultivada na área de estudo. Várias dissertações de mestrado foram produzidas pelo programa de pós-graduação em Engenharia de Sistemas Agroindustriais do Centro de Ciências Exatas e Tecnológicas desta instituição.

Os primeiros trabalhos desenvolvidos pelo programa envolviam o uso de semivariogramas, tentando caracterizar uma estrutura de variação espacial para uma única variável, representando-a através de um mapa temático. Silva (2000), Kavanagh (2001), Carvalho (2004), Assumpção (2004) foram as primeiras dissertações do programa envolvendo a geoestatística. A metodologia aplicada foi baseada em autores como Vieira (1996), Journel e Huijbregts (1978), Cressie (1993), dentre outros. Esses autores caracterizaram dados espaciais como sendo aqueles que apresentavam uma única informação em cada localização e com dependência entre as localizações, o que não constituia em si uma situação realística, pois em experimentos dessa natureza era amostrado de um conjunto de variáveis em cada localização. Ver Hoef e Cressie (1993) alegam que a cokrigagem em modelos bivariados com o uso do semivariograma cruzado pode resultar em predições inferiores. Neste caso os autores querem se referir à predição de uma variável tendo como informação adicional a influência da variabilidade espacial de outra variável tomada na mesma localização. Já uma predição multivariada produzirá uma matriz do erro médio quadrático de predição que permitirá construir regiões de predição multivariada, o que será de difícil interpretação física. Todavia, a distribuição marginal de qualquer das variáveis poderá fornecer informações de interesse específico, com interpretação física fácil.

Definindo-se “suporte” como a geometria da unidade amostral de coleta de dados atribuída a um ponto amostral, autores de grande parte de trabalhos científicos com métodos geoestatísticos em estudos de solo e planta, mediram variáveis químicas e físicas em suporte pequeno, muito próximo de um ponto amostral. Já para os dados de produção tomaram medidas em suporte maior, envolvendo a quantidade do produto colhido em uma área correspondendo, muitas vezes, a uma parcela de tamanho significativo. Tal procedimento leva à necessidade de transformar essas medidas em um indicador de produtividade associando a um ponto (o centróide da parcela, por exemplo). Nesta pesquisa os pontos amostrais foram planejados de forma desalinhada, sistemática e estratificada (WOLLENHAUPT e

WOLKOWSKI, 1994). Cada ponto representou um conjunto de variáveis e a parcela correspondeu a uma área de 25 m^2.

Já se passaram cerca de 10 anos até a presente data e durante esse período muito se aprimorou na área da estatística, sobretudo no que se refere aos recursos computacionais. Foram desenvolvidos novos programas capazes de executar algoritmos mais atualizados e com maior rapidez, que, para a análise estatística, forneceram mais elementos de tomada de decisões. Na geoestatística foram implementados métodos nem tão recentes, envolvendo a função de verossimilhança (CURRIERO e LELE, 1999), que permitiram estimar melhor parâmetros de modelos representantes da variabilidade espacial.

O objetivo geral desse trabalho foi avaliar o efeito de diferentes planejamentos amostrais na recuperação da informação disponível da produção de uma área agrícola pequena, utilizando ferramentas geoestatísticas baseadas em modelos univariados e bivariados, com estimativa de parâmetros por maximização da função log-Verossimilhança e estimativas por krigagem convencional (ordinária).

MATERIAL E MÉTODOS: A pesquisa que gerou os dados estatísticos foi realizada em área de 151,2 x 115,2 m, localizada no Centro de Pesquisa Eloy Gomes, da Cooperativa Central Agropecuária de Desenvolvimento Tecnológico e Econômico Ltda, situado no município de Cascavel-PR. Nessa área, no final do ano de 1997, cultivou-se soja em sistema de semeadura direta e em abril de 1998, após serem demarcadas na área cultivada 256 parcelas de 7,20 x 7,20 m com o uso de trena e estacas de madeira, a soja produzida em área de 25 x 25 m dentro de cada parcela foi colhida e pesada. O histórico de manejo foi considerado desconhecido até esse plantio. As amostras para determinação dos atributos químicos e físicos do solo foram obtidas em pontos com 7 cm de diâmetro e 15 cm de profundidade dentro de cada uma das 256 parcelas com pontos amostrais marcados pelo sistema desalinhado, sistemático estratificado (*op. cit.*).

Para a análise geoestatística utilizou-se a metodologia baseada em modelos, onde o conjunto de dados experimentais foi definido como $\{(x_i, y_i): x_i \in \mathbb{R}^2, y_j \in \mathbb{R}, i=1,2, ..., n\}$ para x_i

indicando a localização espacial da i-ésima coordenada em uma região do espaço bidimensional e y_i indicando uma medida da variável aleatória Y tomada na coordenada. Essa variável aleatória Y é a representação mensurável de um processo estocástico $\{S(x_i)$: $x_i \in \mathbb{R}^2$, $i=1,2, \ldots, n\}$ que, embora desconhecido, julgamos descrever de maneira teórica o comportamento do fenômeno na área de estudo (DIGGLE e RIBEIRO JR, 2007). Suporemos aqui que o processo S seja gaussiano, com função de distribuição contínua e que o evento Y ocorra devido a sua lei de probabilidades. Uma realização do evento Y corresponderá a um conjunto de n observações tomadas cada uma em uma localização distinta e fixa. O modelo adequado para representar esse processo estocástico Y é dado por: $Y(x_i) = \mu(x_i) + S(x_i) + \delta_i$, $i=1,2, \ldots, n$, onde $Y(x_i)$ terá distribuição normal multivariada com média $\mu(x_i)$ e variância τ^2, ambas condicionadas a $S(x_i)$, $\mu(x_i) = \beta_0 + \beta_1 d_1(x_i) + \beta_2 d_2(x_i) + \ldots + \beta_p d_p(x_i)$ é um efeito espacial associado a p variáveis externas diferentes de $Y(x_i)$ mas que dependerão das localizações (x_i) e β são coeficientes a serem determinados. $S(x_i)$ tem média zero e variância $\sigma^2\rho(u_{ij})$ sendo $\rho(u_{ij}) = \mathrm{Corr}\{S(x_i),S(x_j)\}$ e $u_{ij}=\|x_i - x_j\|$ a distância euclidiana que separa duas coordenadas quaisquer. δ_i são erros aleatórios mutuamente independentes, tais que $\delta_i \sim N(0;\tau^2)$. Na forma matricial, representamos o modelo como:

$$\mathbf{Y} \sim N(\mathbf{D}\beta\ ;\ \sigma^2\mathbf{R} + \tau\mathbf{I})$$

onde **R** é a matriz cujos elementos representam as autocorrelações entre as localizações amostrais, **I** é a matriz identidade, ambas de tamanho $n \times n$ e **D** é a matriz de covariáveis de efeitos externos, com dimensão $n \times p$.

A função de correlação que adotamos neste trabalho pertence a família Matérn (MATÉRN, 1986) dada por:

$$\rho(u, \phi, \kappa) = \frac{1}{2^{\kappa-1}\Gamma(\kappa)}\left(\frac{u}{\phi}\right)^{\kappa} K_{\kappa}\left(\frac{u}{\phi}\right) \tag{1}$$

onde K é a função modificada de Bessel e Γ a função Gamma (ABRAMOWITZ e SEGUN, 1965). Essa função de correlação é conveniente pela equivalência a outras funções de correlação adotadas por autores como Isaaks e Srivastava (1989). A escolha do parâmetro de diferenciabilidade κ que determina a suavidade de decaimento da função e o parâmetro ϕ, dá a taxa de decaimento até a função atingir um valor que corresponderá à distância de separação das coordenadas onde a correlação a correlação entre elas é considerada desprezível.

Sendo Y um processo gaussiano correlacionado, sua função de verossimilhança será composta pela sua distribuição conjunta de probabilidades dada por:

$$f(y;\theta)=L(\theta)=\frac{1}{(2\pi)^{n/2}(|\sigma^2 R+\tau^2 I\ |)^{1/2}}\ \exp\left\{-\frac{1}{2}(y-D\beta)'(\sigma^2 R+\tau^2 I)^{-1}(y-D\beta)\right\}$$

ou ainda:

$$\log L(\theta)=l(\theta)=-\frac{1}{2}\left[n\log(2\pi)+\log(|\sigma^2 R+\tau^2 I|)+(y-D\beta)'(\sigma^2 R+\tau^2 I)^{-1}(y-D\beta)\right]$$

Fazendo $\frac{\tau^2}{\sigma^2}=\nu^2$ e $\sigma^2(R+\frac{\tau^2}{\sigma^2}I)=\sigma^2 V$ e substituindo em *l(θ)* acima vem:

$$l(\theta)=-\frac{1}{2}\left[n\log(2\pi)+\log(|\sigma^2 V|)+y'(\sigma^2 V)^{-1}y-2y'(\sigma^2 V)^{-1}D\beta+\beta' D'(\sigma^2 V)^{-1}D\beta\right] \quad (2)$$

Derivando-se a expressão do log-Verossimilhança em relação a β e igualando o resultado a zero, obteremos o estimador para esse parâmetro dado por:

$$\hat{\beta}=(D'(\sigma^2 V)^{-1}D)^{-1}D'(\sigma^2 V)y \quad (3)$$

Considerando o fato de que $y'(\sigma^2\ V)^{-1}D\beta$ é um escalar e derivando-se novamente *l(θ)* em relação a σ^2 e igualando-se a zero obtemos o estimativa para esse parâmetro dada por:

$$\hat{\sigma}^2_{\phi,\nu^2}=\frac{(y-D\hat{\beta})'V^{-1}_{\phi,\nu^2}(y-D\hat{\beta})}{n} \quad (4)$$

Inserindo-se essa estimativa na equação (3) obtém-se:

$$\hat{\beta}=(D'V^{-1}_{\phi,\nu^2}D)^{-1}DV^{-1}_{\phi,\nu^2}y$$

que irá depender somente dos parâmetros ϕ e ν. Substituindo-se ainda a estimativa dada pela equação (4) na equação (2), a função log-Verossimilhança fica:

$$l(\phi;\nu^2)=-\frac{1}{2}\left\{n\log(2\pi)+n\ \log\left[(y-D\hat{\beta})'V^{-1}_{\phi,\nu}(y-D\hat{\beta})\right]-n\ \log n+\log|V|+n\right\} \quad (5)$$

A função dada pela equação (5) recebe como argumentos o vetor das observações de Y e a matriz de distâncias de cada coordenada com as demais, o que permitirá obter V pela escolha adequada dos parâmetros ϕ e κ da função de correlação (1). Essa função, em seu ponto de máximo, fornecerá estimativas dos seus parâmetros. Empregamos o software geoR para se obter essas estimativas (RIBEIRO JR e DIGGLE, 2001) pois possui algoritmo e implementação computacional eficientes.

Existe ampla discussão na literatura das razões para se transformar dados estatísticos buscando obter uma forma de distribuição próxima da distribuição gaussiana de probabilidades, principalmente devido às desejáveis características de simetria e de uma curtose mesocúrtica. Empregamos o método paramétrico de transformação proposto por Box e Cox (1964) tanto para verificar se os dados se ajustam a uma distribuição Normal de probabilidades quanto para definir o parâmetro de transformação.

Um aspecto importante da modelagem estatística é a utilização do modelo obtido para efetuar predições. No caso geoestatístico o interesse é estimar valores do processo gaussiano *S(x)* em locais não cobertos pela malha amostral estabelecida para *Y*. Krige (1951) introduziu os primeiros conceitos e Matheron (1963) usou o termo "krigagem" para esse tipo de predição linear espacial. O método estima um valor em um ponto arbitrário de uma região "fechada" onde a função de correlação do processo e um conjunto de pontos amostrais são conhecidos. Para Isaaks e Srivastava (1989) a krigagem ordinária é tida como um método BLUE, acrônimo do inglês *Best Linear Unbiased Estimator*, ou seja, melhor estimador não viciado e de variância mínima.

Shabenberger e Gotway (2005) fizeram distinção entre estimação e predição, por serem expressões muitas vezes tidas como equivalentes. Em um modelo clássico de regressão linear simples, os erros não são correlacionados e os coeficientes são estimados usualmente pelo método dos mínimos quadrados. Com isso se prediz um valor $\hat{Y}(x_0)$ para x_0. Não fica claro se esse valor é um preditor de $Y(x_0)$ como uma "resposta" em x_0 ou é um estimador de E[$Y(x_0)$]. Apesar de que estimar uma quantidade fixa ou predizer uma quantidade aleatória ser uma questão menor, sua importância fica clara ao se considerar uma incerteza associada a essas quantidades. No caso da geoestatística, apesar do desconhecimento do processo estocástico S(x), aplicações com predição são freqüentemente mais empregadas do que aquelas que buscam a estimação de uma média. Esses autores (*op. cit.*) apresentam como modelo de predição linear (como sinônimo de krigagem) o seguinte preditor: $\hat{Y}(x_0)=\hat{\mu}+r'\Sigma^{-1}[Y(x)-\hat{\mu}\mathbf{1}]$ onde r é a covariância entre a coordenada x_0 onde se deseja predizer e as demais localizações x. $\Sigma = \sigma^2 V$ é a matriz de covariâncias relativa às coordenadas amostrais e $\hat{\mu} = (\mathbf{1}' V^{-1} \mathbf{1}')^{-1} \mathbf{1}' V^{-1} Y$ é o estimador de μ. Neste trabalho empregaremos a função `krige.conv`, presente no software geoR (*op.cit.*) por permitir, dessa forma, predições espaciais com parâmetros de covariância fixos, utilizando a vizinhança total da coordenada a predizer.

Muitos trabalhos predizem valores sob o cenário de uma única variável, e em alguns casos utilizam informações de outras variáveis nas mesmas coordenadas (co-variáveis) visando controlar o componente de $\mu(x)$ do modelo proposto, entretanto, Ver Hoef e Cressie (1993) e outros autores afirmam que em ciências da terra é freqüente o interesse em predizer conjuntamente uma grande quantidade de variáveis. Predições espaciais multivariadas permitirão construir regiões de predição multivariada. Estamos aqui interessados em modelos multivariados que digam respeito a um conjunto de processos estocásticos gaussianos diferentes; especificamente dois, tomados em uma mesma região, todas com igual interesse científico. Parece ser uma situação pouco realística pois esta descrição não leva a uma interpretação física de sentido prático. Entretanto, se elegermos uma variável de interesse primário e descrevermos sua distribuição de probabilidades condicionada a outras variáveis então encontraremos facilmente aplicações de interesse. Neste caso, exige-se que todas as variáveis sejam tomadas nas mesmas coordenadas e que haja uma certa correlação entre elas. Outra situação prática ocorre quando uma variável de interesse primário for de difícil e/ou onerosa aquisição. Nesse caso, toma-se um conjunto de variáveis preditoras complementares, convenientemente modeladas para se fazer estimativas da variável primária. O conjunto de variáveis complementares poderá ser em quantidades, tipos e localizações diferentes.

Sendo $\{S(x) = (S_1(x_i), S_2(x_j)): x_i, x_j \in \mathbb{R}^2, i=1,2, \ldots, m; j=1,2, \ldots, n\}$ um processo gaussiano bivariado. Uma proposta de modelagem para $Y = (Y_1;Y_2)$ nesse processo pode ser dado por:

$$\begin{aligned} Y_{1,j} &= \mu_1(x_i) + \sigma_{0;1}R_0(x_i) + \sigma_1 R_1(x_i) + \varepsilon_1;\ i=1, 2, \ldots, m. \\ Y_{2,j} &= \mu_2(x_j) + \sigma_{0;2}R_0(x_j) + \sigma_2 R_2(x_j) + \varepsilon_2;\ j=1, 2, \ldots, n. \end{aligned} \tag{6}$$

Nesse modelo, Y_1 estará correlacionado com Y_2 devido à presença de uma mesma matriz de correlação $R_0(x_i)$. A distribuição de probabilidades conjunta de Y_1 e Y_2 será:$(Y_1 ; Y_2) \sim N(\mu^*(x) ; \Sigma^*)$ onde $\mu^*(x) = \mu_1(x) + \mu_2(x)$ e $\Sigma^* = (\sigma_{0;1}R_0(x) + \sigma_1 R_1(x) + \tau_1) + (\sigma_{0;2}R_0(x) + \sigma_2 R_2(x) + \tau_2)$. A função log-

Verossimilhança será obtida com procedimento análogo aos adotados para a obtenção da equação (5). Neste trabalho utilizamos a função geoR (*op.cit.*) `likfitBGCCM` que forneceu estimativas de máxima-verossimilhança para os parâmetros da equação (6). Para a predição multivariada, empregamos a função `predict` a qual forneceu estimativas da variável Y_1 (variável de interesse primário) em localizações compatíveis com as predições univariadas.

O problema foi estruturado da seguinte maneira: A produção pseudo-real foi inferida das quantidades efetivamente colhidas e pesada nas 256 parcelas de $25m^2$. Como o planejamento inicial do experimento intercalou na área dois grupos de 128 amostras, designando um grupo para manejo localizado (AP) e outro grupo para manejo convencional (CV), utilizamos esses dois arranjos como amostras dos 256 pontos planejados. Adicionalmente, sorteamos dos 256 pontos, uma amostra aleatória de 128 pontos e outra de 64 pontos, irregularmente localizadas. Essas amostras forneceram modelos univariados e bivariados (pela inclusão da variável resistência mecânica à penetração) para predição espacial da produção, comparando-se com o valor pseudo-real.

RESULTADOS E DISCUSSÃO: Nesse estudo denominamos as variáveis como Soja256 para nos referirmos a todo os pontos amostrais onde foi colhida e pesada a soja plantada, SojaAP para as 128 parcelas intercaladas, reservadas para futuro manejo localizado (Agricultura de Precisão), SojaCV para as 128 parcelas intercaladas, reservadas para futuro manejo convencional, SojaS128 para 128 parcelas tomadas aleatoriamente dentre as 256 disponíveis, SojaS64 para 64 parcelas tomadas aleatoriamente dentre as 256 disponíveis, iConeCV para as medidas dessa variável nas 128 parcelas destinadas ao manejo localizado, iConeAP para as medidas nas 128 parcelas destinadas ao manejo convencional, iConeS128 para medidas em 128 parcelas sorteadas aleatoriamente e iConeS64 para medidas em 64 parcelas sorteadas aleatoriamente nas 256 parcelas disponíveis.

Na tabela 1 são apresentados resultados descritivos do comportamento da produtividade de soja de acordo com diferentes desenhos de amostragem. Nela observa-se uma produtividade média que varia entre 2.683 a 2.785 Kg. por hectare e Coeficiente de Variação - CV variando entre 15,7% e 19,2% sugerindo uniformidade do rendimento das parcelas. Esses resultados são compatíveis com a produção dessa cultivar no Paraná segundo levantamento da CONAB - Companhia Nacional de Abastecimento, que foi de 2.550 Kg/ha. Destaca-se ainda nessa tabela uma fraca correlação dos resultados em sentido ao eixo vertical das coordenadas nos desenhos amostrais envolvendo os 256 pontos (-0,30), nos 128 pontos das parcelas destinadas a manejo convencional (-0,45) e nas 128 parcelas escolhidas aleatoriamente (-0,35). A mesma tabela apresenta o Índice de Cone (resistência mecânica à penetração), medida em Kg/cm^2 na profundidade 0-5 cm. Constata-se ali uma média variando de 20,35 a 20,73 Kg/cm^2 o que equivale a 1,997 e 2,034 Mpa, valores estes considerados limítrofes para crescimento radicular das principais cultivares (LIMA *et. al*., 2007). O Coeficiente de Variação segundo os diferentes desenhos amostrais ficou entre 20,0 e 21,8% sugerindo uniformidade na distribuição das medidas dessa variável. A Correlação Linear de Pearson para a variabilidade das medidas de produção de soja nas parcelas comparadas com a variabilidade do Índice de Cone só se mostrou significativo para as parcelas marcadas para menejo localizado (AP), mesmo assim o adotamos como variável de suporte para os modelos de predição bivariada da soja. A baixa correlação encontrada para essas duas variáveis se deu devido à baixa variabilidade.

Tabela 1: Informações descritivas da variável Soja safra 97/98 e Índice de Cone.

Desenho	N	Média	Min	Max	Cor_X	Cor_Y	DP	CV%
Soja256	256	2.746	1.190	4.140	-0,02	-0,30	0,490	17,8
SojaCV	128	2.785	1.390	4.140	-0,05	-0,45	0,535	19,2
Soja AP	128	2.708	1.190	3.700	0,02	-0,14	0,440	16,3
SojaS128	128	2.754	1.710	4.140	0,10	-0,35	0,433	15,7
SojaS64	64	2.683	1.190	3.780	-0,09	-0,02	0,494	18,4
iConeAP	128	20,73	10,50	32,50	--	--	4,346	21,0
iConeCV	128	20,35	10,00	30,00	--	--	4,302	21,1
iConeS128	128	20,48	10,00	32,50	--	--	4,463	21,8
iConeS64	64	20,70	13,00	31,50	--	--	4,139	20,00

Média, Min., Max e DP dados em Kg/hectare. N: Número de parcelas; Cor_X e Cor_Y: Coef. de correlação.

Na tabela 2 estão os resultados para um intervalo de 95% de confiança para o parâmetro λ de transformação dos valores de Y (BOX e COX, 1964). Nessa tabela observa-se que o valor 1 está presente em todos os intervalos. Optamos assim por não transformar a variável resposta em nenhum dos planejamentos amostrais.

A área total de cultivo foi de 17.418,24 m^2 (115,2 m x 141,2 m). Cada uma das 256 parcelas mediu 25 m^2 perfazendo um total de 6.400 m^2, correspondendo a 36,7% da área total cultivada. O rendimento das 256 foi de 1.757,7 Kg de soja que, projetada para a área total cultivada resultou em 4.789,4 Kg. Essa produção corresponde a 2.749,6 Kg/ha. Melo e Fontana (2007) alegaram que o IBGE registrou 2.743,0 Kg/ha na safra de soja de 2003. Segundo levantamento feito pela CONAB, a produtividade obtida nessa pesquisa é compatível com produção no Paraná para a safra 97/98 que foi de 2.550,0 Kg/ha. Tomamos então como valor referência a produtividade projetada de 2.749,6 Kg/ha.

Tabela 2: Intervalo de 95% de confiança para o parâmetro λ de transformação de Box-Cox.

Desenho	N	L.I.	L. S.
Soja256	256	0,30	1,32
SojaCV	128	-0,10	1,30
SojaAP	128	0,49	2,10
SojaS128	128	-0,10	1,35
SojaS64	64	0,70	2,70

L.I.: Limite inferior do intervalo; LS: Limite superior do intervalo

O ajuste de modelos com dependência espacial indicou, conforme tabela 3, estruturas de correlação com Coeficiente de Efeito Pepita – CEP elevados, sugerindo ausência de uma estrutura de dependência espacial. Exceto para o caso de uma amostragem de 64 das 256 parcelas, obtivemos CEP acima de 57%. Esses resultados estão de acordo com os resultados de coeficiente de variação obtidos na tabela 1. As estimativas de produtividade comparadas com o valor referência são superestimadas em 7,8% com base em amostragem densa (147 parcelas por hectare), em 16,6% em 128 parcelas de manejo convencional (73/hectare), em 9,5% em 128 parcelas tomadas aleatoriamente e subestimadas em 5,2% em 128 parcelas para manejo localizado e em 2,5% em 64 das 256 parcelas (37/hectare), tomadas aleatoriamente. Esses resultados sugerem a tendência a se obter um adensamento amostral ótimo, objetivo este não coberto por esse estudo.

Tabela 3: Estimativas de parâmetros com ajustes de otimização numérica da função de verossimilhança para modelos de correlação de Matérn

Amostra	κ	β_0	β_1	τ^2	σ^2	ϕ	alcance	CEP	log-L
Soja256	1,5	2,9693	-0,0045	0,1960	0,0263	15,9111	63,64	0,8817	-164,5
SojaCV	0,5	3,2121	-0,0074	0,1301	0,0988	12,0928	36,28	0,5684	-83,55
SojaAP	2,5	2,6068	--	0,1363	0,0703	12,4892	67,07	0,6597	-68,12
SojaS128	0,5	3,0145	-0,0046	0,1378	0,0258	18,6597	55,98	0,8423	-64,33
SojaS64	0,5	2,6819	--	0	0,2394	2,8902	8,67	0	-45,01

SojaS64: 64 parcelas e SojaS128: 128 parcelas (aleatórias); CEP: Coeficiente de Efeito Pepita.

Na tabela 4 são apresentadas as estimativas dos parâmetros do modelo bivariado conforme equação 6. Devido ao histórico de correlação da variável Índice de Cone com produtividade, adotamo-la como a variável de interesse secundário, visando utilizar sua variabilidade para melhorarmos as predições da variável Soja. Foram construídos vários modelos bivariados, sempre procurando completar as informação da variável principal com informações da variável secundária tomada em locais diferentes. Assim, denominamos os modelos como M1: SojaCV x iConeAP; M2: SojaAP x iConeCV; M3: SojaS128 x iConeCV; M4: SojaS128 x iConeAP; M5: SojaS64 x iConeCV; M6: SojaS64 x iConeAP; M7: SojaS64 x iConeS128; M8: SojaCV x iConeS128; M9: SojaAP x iConeS128. Nessa tabela observamos que as estimativas da produtividade média da Soja (μ_1) são compatíveis com a produção do Paraná, segundo todos os modelos ajustados.

Tabela 4: Estimativas de parâmetros com ajustes de otimização numérica da função de verossimilhança para modelos bivariados e função de correlação de Matérn com $\kappa = 0,5$.

Mod.	μ_1	μ_2	$\sigma^2_{0;1}$	σ^2_1	$\sigma^2_{0;2}$	σ^2_2	ϕ_0	ϕ_1	ϕ_2	η	ν_1	ν_2	log-L
M1	2,7848	20,8388	0,151	0,172	16,977	1,868	0,0000	42,4302	24,7766	10,619	1,069	3,522	-222
M2	2,5654	20,3524	0,123	0,094	18,374	0,002	1,1607	40,2915	30,6323	12,200	0,872	0,131	-203
M3	2,7561	20,3339	0,061	0,113	0,967	19,800	10,8876	2,4106	0,1343	3,991	1,363	18,054	-207
M4	2,7331	20,8907	0,055	0,527	3,671	15,943	24,5093	29,3775	1,7622	8,154	3,089	16,992	-222
M5	2,5463	20,3514	0,057	0,683	0,893	18,332	7,8521	47,0128	0,0008	3,964	3,466	17,963	-247
M6	2,6219	20,7249	0,061	0,347	0,030	20,332	26,2107	19,1828	0,2840	0,698	2,376	18,191	-247
M7	2,5494	20,4974	0,060	0,697	1,150	19,177	0,0008	51,3994	3,3566	4,394	3,422	17,945	-251
M8	2,7849	20,6467	0,060	0,292	5,282	17,256	28,6989	12,5374	3,1931	9,403	2,210	16,995	-231
M9	2,6528	20,3338	0,057	0,204	5,228	16,709	37,8722	8,3863	2,2903	9,586	1,896	17,138	-217

Modelos: M1: SojaCV x iConeAP; M2: SojaAP x iConeCV; M3: SojaS128 x iConeCV; M4: SojaS128 x iConeAP; M5: SojaS64 x iConeCV; M6: SojaS64 x iConeAP; M7: SojaS64 x iConeS128; M8: SojaCV x iConeS128; M9: SojaAP x iConeS128.

A tabela 5 mostra as estatísticas da predição sobre uma malha de 4.466 pontos regularmente distribuídos dentro da área de cultivo. Tanto as predições univariadas (parte superior da tabela) quanto as predições bivariadas apresentam médias muito próximas do valor projetado para a área. O erro de predição de que trata a última coluna da tabela é obtido pela comparação da coluna da média dos valores preditos pela krigagem convencional (segunda coluna) com o valor calculado de 2.749,6 Kg/ha. Nota-se que o valor absoluto não supera 1,6%, resultado esse considerado como evidência da eficiência do método.

Tabela 5: Média da predições por krigagem comparada com o valor de referência com base nas 256 parcelas colhidas e pesadas.

Modelo	M. P.	D.P.	CV	E. Pr.
Soja256	2,7471	0,1834	6,7%	-0,1%
SojaCV	2,7878	0,2418	8,7%	1,4%
SojaAP	2,7066	0,1952	7,2%	-1,7%
SojaS128	2,7631	0,1755	6,4%	-0,5%
SojaS64	2,7774	0,1300	4,7%	1,0%
M1	2,7886	0,2936	10,5%	1,4%
M2	2,7066	0,2040	8,9%	-1,6%
M3	2,7596	0,1783	6,5%	0,4%
M4	2,7758	0,3176	11,4%	1,0%
M5	2,7941	0,2087	7,5%	1,6%
M6	2,7926	0,2004	7,2%	1,6%
M7	2,7514	0,0080	0,3%	-0,3%
M8	2,7903	0,4005	14,4%	1,5%
M9	2,7075	0,2890	10,7%	-1,5%

M.P.: Média dos valores preditos por krigagem convencional (ton/ha); D.P.: Desvio Padrão (ton/ha); C.V.: Coeficiente de Variação; E.Pr: Diferença percentual entre o valor médio da predição de produtividade por krigagem e valor projetado para a área com base nas parcelas de soja colhido e pesado.

CONCLUSÕES: O efeito de diferentes planejamentos amostrais em modelos geoestatísticos univariados e bivariados na recuperação da informação disponível da produção de uma área agrícola pequena e com pouca variabilidade espacial, utilizando estimativas com interpolação por krigagem convencional (ordinária), não mostraram diferenças, mantendo o erro relativo abaixo de 1,7%. Devido à uniformidade das variáveis tomadas na área não foi possível atribuir vantagem em se utilizar métodos bivariados, todavia o método manteve a robustez e eficiência nas estimativas.

REFERÊNCIAS:

ABRAMOWITZ M. and STEGUN, I. A. Handbook of Mathematical Functions. New York: Dover, 9.ed, 1972. 1.046p.

ASSUMPÇÃO, R. A. B. **Uso de tipos de krigagem no estudo da produtividade de soja (*Glycine max L. Merrill*) em uma área espacialmente referenciada**. Cascavel: Edunioeste, 2004, 124p. (Dissertação de Mestrado)

BOX, G. E. P., COX, D. R. An analysis of transformations (with discussion), **Journal of the Royal Statistical Society**, Serie N, n.26, pp.211-52, 1964

CARVALHO, L. Pagliosa C. **Estudo da Anisotropia em uma Área Espacialmente Referenciada**. Cascavel: Edunioeste, 2004. 80p. (Dissertação de Mestrado)

CRESSIE, N. A. C. **Statistics for spatial data**. New York: John Wiley e Sons, 1993. 900p.

DIGGLE, P. J.; RIBEIRO-JR, P. J. **Model-based Geostatistics**. USA: Springer . Series in Statistics, 2007. 228p.

CURRIERO, F. C. and LELE, S. A composite likelihood approach to semivariogram estimation. **American Statistical Association and the International Biometric Society Journal of Agricultural, Biological, and Environmental Statistics**. v.4, n.1, pp.9-28, 1999.

ISSAKS, E. H. e SRIVASTAVA, R. M. **Applied geostatistics**. New York: Oxford University Press, 1989. 561p.

JOURNEL, A. G. and HUIJBREGTS, Ch. J.**Mining geostatistics.** New Jersey: Blackburn Press, 1978. 600p.

KAVANAGH, C. R. **Estimação de parâmetros no ajuste de modelos teóricos a semivariogramas experimentais**. Cascavel: Edunioeste, 2001. 85p. (Dissertação de Mestrado)

KRIGE, D. G. A. **Statistical Approach to Some Mine Valuations and Allied. Problems at Witwatersrand**. England: University of Witwatersrand, 1951. (Tese de Doutorado)

LIMA, C. L. R.; REICHERT, J.M.; REINERT, D. J.; SUZUKI, L. E. A. S.; DALBIANCO, L. Densidade crítica ao crescimento de plantas considerando água disponível e resistência à penetração de um Argissolo Vermelho distrófico arênico. **Ciência Rural**. Santa Maria: v.37, n.4, p.1166-69, jul-ago, 2007.

MATERN, B. **Spatial Variation Analysis**. 2.ed, Berlin: Springer Verlag, 1986.

MATHERON, G. Principles of geoestatistics. Economic Geology, pp.1246-66, n.58, 1963.

MELO, R. W., FONTANA, D. C. Estimativa do rendimento de soja usando dados do modelo do ECMWF em um modelo agrometeorológico-espectral no Estado do Rio Grande do Sul . XIII Simpósio Brasileiro de Sensoriamento Remoto. **Anais...** Florianópolis: INPE, p. 279-286, 21-26/04, 2007.

RIBEIRO Jr., P. J. and DIGGLE, P. J. geoR: A package for geostatistical analysis. **R-NEWS**, v.1, n.2, pp.15-18, 2001.

SCHABENBERGER, O.; GOTWAY, C. A. **Statistical Methods for Spatial Data Analysis**. New York: Chapman-Hall, 2005. 488p,

SILVA, E. A. A. **Aplicação de um estimador robusto na análise de variabilidade espacial de parâmetros de Agricultura de Precisão**. Cascavel: Edunioeste, 2000, 85p. (Dissertação de Mestrado)

SOUZA, E. G.; RIBEIRO, S. R. A.; URIBE-OPAZO, M. A.; VILAS-BOAS, M.; SILVA, M. S.; JOHANN, J. A.; MOLIN, J. P.; NÓBREGA, L. H. P.; OLIVEIRA, E. F.; CARRARO, I. Metodologia para análise da variabilidade espacial dos atributos do solo e da produtividade em uma área piloto de agricultura de precisão. In: XXVII Congresso Brasileiro de Engenharia Agrícola, **Anais...** v.3, pp. 109-11, 1998.

Ver HOEF, J. M., CRESSIE, N. **Multivariate Spatial Prediction. Mathematical Geology**, v. 25, n.2, pp.219-40, 1993.

VIEIRA, S. R. **Geoestatística**: curso de extensão universitária. Botucatu: UNESP, 1996. 166p.

WOLLENHAUPT, N. C. and WOLKOWSKI, R. P. Grid soil sampling. **Better Crops with plant food**. NORCROSS, v.78, n.4, pp.6-9, 1994.