

# Uso da krigagem indicatriz na avaliação espacial de percevejos da espécie *Piezodorus guildinii* na Safra 2011/12.

Farias, C.H.<sup>1</sup>; Oliveira, M.C.N. de<sup>2</sup>; Roggia, S.<sup>2</sup>;

<sup>1</sup>Universidade Norte do Paraná (Unopar), <sup>2</sup>Embrapa Soja, carlos@cnpso.embrapa.br

## Introdução

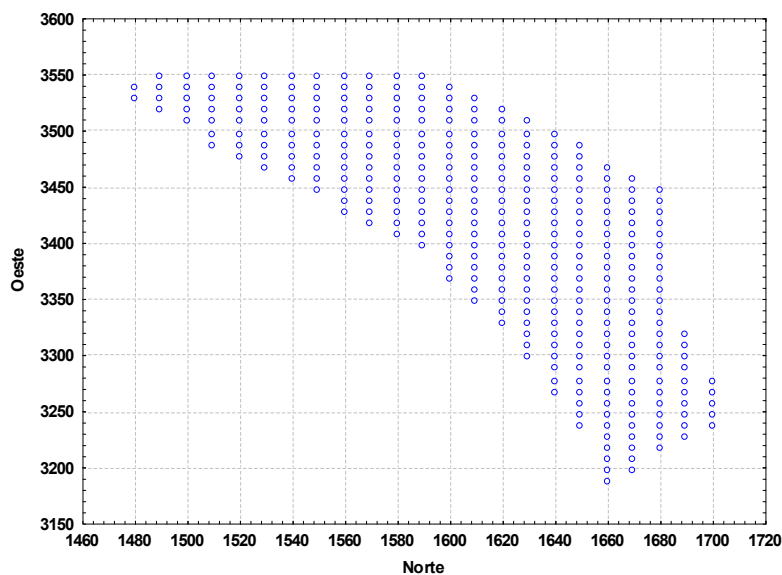
A cultura da soja apresenta destaque no cenário agrícola nacional onde está sendo cultivada em uma área de 27,7 milhões de hectares e é responsável por mais de 81,94 milhões de toneladas de grãos, o que equivale à aproximadamente 37% do volume total de grãos produzidos no Brasil. O grão é componente essencial na fabricação de rações animais, com uso na alimentação humana, e encontra-se em contínua expansão, pois tanto o mercado doméstico quanto o internacional é francamente comprador. Em função da extensa área cultivada, a cultura está sujeita à infestação de vários insetos-pragas, sendo um dos principais o percevejo *Piezodorus guildinii* (HOFFMANN-CAMPO et al, 2000). A pesquisa deve vir sempre à frente para a solução dos problemas e, por sua vez, deve ser realizada de forma a gerar conclusões confiáveis. Neste sentido é fundamental a obtenção de amostras que representem com confiabilidade o fenômeno estudado, sendo esta, muitas vezes, uma das etapas mais trabalhosas do estudo. Os métodos estatísticos são também fundamentais e auxiliam na tomada de decisões sobre as respostas de pesquisas (BANZATO & KRONKA, 1995). Esses métodos também podem melhorar o processo da amostragem, como os métodos geoestatísticos, os quais permitem simular amostras não realizadas pelo método da *krigagem* indicatriz. A *krigagem* é um dos melhores métodos de interpolação de dados em que os valores intermediários dos dados são preservados, obtendo-se como resultado final uma superfície contínua de dados mais suavizados.

Nesse estudo foi estimado a flutuação populacional de *Piezodorus guildinii* na cultura da soja utilizando o método da *krigagem* indicatriz.

## Materiais e Métodos

### 1-Local experimental e mapa de amostragem

Nesse estudo foram utilizados os dados obtidos no campo experimental da Embrapa Soja no Distrito de Maravilha (Londrina-PR), segundo as coordenadas" 23° 28' 44" S; 50° 59' 03" O, e altitude média de 465 m, com área de aproximadamente 3,38 ha. Em cada ponto amostral foram realizadas quatro amostragens com pano de batida para a coleta de percevejos da espécie *Piezodorus guildinii* e nessa amostragem foram consideradas somente as ninfas grandes. Na área experimental foram amostrados e georreferenciados 338 pontos numa grade de 10 X 10m, utilizando o aparelho GPS de navegação marca Garmin, modelo MAP 62. Abaixo o mapa amostral da área pesquisada na presença de *outliers*. (Figura 1).



**Figura 1.** Mapa da área experimental em que foi realizada a coleta de dados na presença de *outliers* em função dos pontos cardeais Norte e Oeste.

## 2- Métodos Exploratórios

As estatísticas deste diagnóstico foram realizadas pelo programa SAS - Statistical Analysis System (SAS Institute, 2001) tais como: a média, a mediana, os coeficientes de assimetria, de curtose, de variação e o teste de normalidade de Shapiro-Wilk (1965). Os gráficos como: o *boxplot*, o *scatterplot* e a carta de controle pelo software Statística versão 7.1 (2001) para detectar dados discrepantes (*outliers*) que podem invalidar as respostas da pesquisa. No *boxplot* são apresentadas as estatísticas como: a mediana (q2), os quartis inferior e superior (q1, q3) e a diferença entre estes quartis (intervalo interquartil). O ponto central do *boxplot* é a mediana do conjunto de dados. Valores que estão acima e abaixo das linhas do *boxplot* são determinados como *outliers*.

## 3- Métodos Geoestatísticos

Para a aplicação dos métodos geoestatísticos foi utilizado o Geoest (VIEIRA, S.R., 2002) e o Surfer v.9 para gerar as superfícies obtidas pelo método de *krigagem* indicatriz. Após o diagnóstico exploratório e a retirada dos pontos discrepantes (*outliers*) foi aplicado a *krigagem* indicatriz, calculado os semivariogramas ajustados pelos modelos esférico, exponencial e gaussiano e mapa da *krigagem* da área amostrada. O semivariograma é usado para modelar dois valores correlacionados no espaço ou no tempo e é pré-requisito para a *krigagem*. Essa ferramenta mede a variação ou discrepância entre pares de valores separados por uma distância determinada  $h$ , sendo possível observar que essa dependência será tanto maior quanto menor for a distância entre os pontos, e ainda, resume as nuvens de pontos ( $z(x)$ ,  $z(x+h)$ ) pela média do quadrado das diferenças entre  $z(x)$  e  $z(x+h)$ , para vários valores de  $h$ . O semivariograma, pode ser definido matematicamente como:

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

em que:

$Z(x_i)$  é o valor observado da variável  $x$  no ponto  $i$ ;

$Z(x_{i+h})$  é o valor observado da variável  $x$  no ponto  $i+h$ ;

$n(h)$  é o número de pares distanciados a uma distância  $h$ .

No estudo geoestatístico o número de percevejos da espécie *Piezodorus guildinii* foram avaliados pelos semivariogramas com as estruturas de covariância exponencial (1), esférico (2) e gaussiano (3) e podem ser apresentados abaixo:

$$\gamma(h) = \begin{cases} C(0) + C \left[ \frac{3}{2} \frac{h}{a} - \frac{1}{2} \left( \frac{h}{a} \right)^3 \right], & h \leq a \\ C(0) + C, & h > a \end{cases} \quad (1) \quad \gamma(h) = \begin{cases} C(0) + C \left[ 1 - e^{-\frac{h}{a}} \right], & h \leq a \\ C(0) + C, & h > a \end{cases} \quad (2)$$

$$\gamma(h) = \begin{cases} C(0) + C \left[ 1 - e^{-\left(\frac{h}{a}\right)^2} \right], & h \leq a \\ C(0) + C, & h > a \end{cases} \quad (3)$$

A *krigagem* indicatriz, técnica de inferência estatística não linear baseia-se na premissa de que a propriedade em estudo é considerada uma variável aleatória em cada posição não amostrada. O método é baseado em uma transformação binária dos dados, sendo cada dado transformado em um indicador antes de ser submetido às análises geoestatísticas e à krigagem. Por convenção, os dados são codificados em 0 ou 1, se estiverem acima ou abaixo de determinado valor de corte, respectivamente, e segundo Cressie (1993) segue a função :

$$\text{Função } (X_i, Y_k) = \begin{cases} 0 & \text{se } Y_k \geq (X_i) \\ 1 & \text{se } Y_k < (X_i) \end{cases}$$

Em que:  $Y_k$  são valores do conjunto de dados e  $X_i$  é o valor de corte, em que é o valor do número de percevejo.

O valor de corte é arbitrário e depende do objetivo do estudo. O resultado da transformação é um novo conjunto de dados, composto de 0 e 1, o qual é então submetido às análises geoestatísticas, obtendo-se, assim, o semivariograma indicador, que reflete o modelo de continuidade espacial para o valor de corte preestabelecido. Nesse trabalho o ponto de corte será considerado o nível de dano da praga da soja que é igual a dois. Valores iguais ou superiores a dois são codificados com o valor um (1) e abaixo de dois o valor zero (0). Os estimadores geoestatísticos da autocorrelação espacial são usados como ferramentas de continuidade espacial, como o semivariograma. Essa ferramenta é usada para investigar a magnitude da correlação entre as amostras e sua similaridade ou não, com a distância.

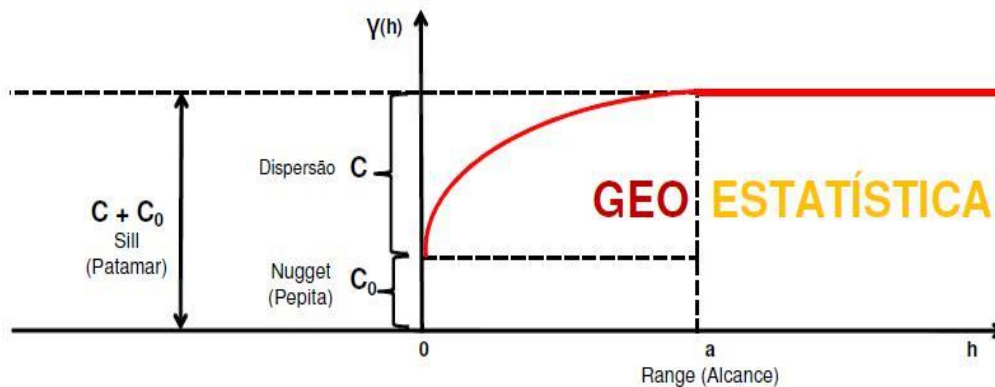
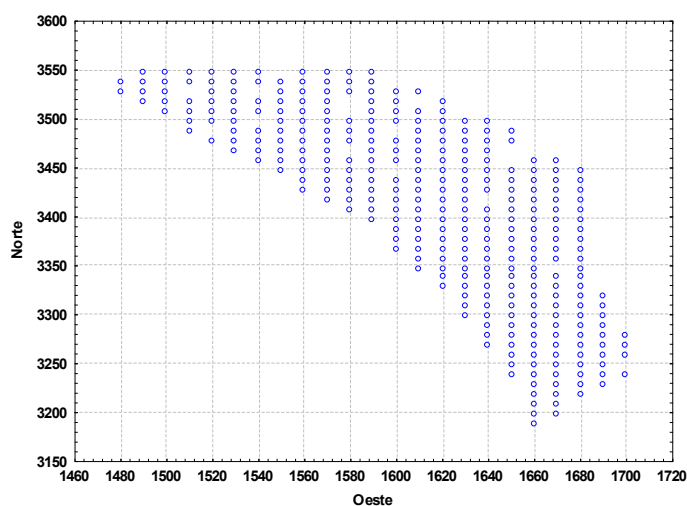


Figura 2. Esquema representando a estrutura de um semivariograma teórico, adaptado de Rosa Neto (2009).

## Resultados e Discussão

Na Figura 3 observam-se os locais onde foram retirados os outliers (Figura 3).

Pela Tabela 1 pode-se verificar que inicialmente com 338 observações foram detectados vários valores discrepantes e os resultados para a curtose e assimetria foram (18,07 e 3,65), coeficiente de variação igual a 189,30, antes da transformação dos dados para uma variável categórica. Esses resultados apresentam *outliers* e pode ser verificado pelas Figuras 4 e Figura 5 A,B.



**Figura 3.** Mapa da área experimental em que foi realizada a coleta de dados na ausência de *outliers* em função dos pontos cardeais Norte e Oeste.

**Tabela 1.** Estatísticas descritivas para contagem de insetos (*Piezodorus guildinii*) em campos de soja. Safra 2011/2012.

Testes estatísticos <sup>1</sup>	Dados Originais	Sem 19 outliers
Tamanho da Amostra	338	319
$g_1$ = Coeficiente de Assimetria	3,657	1,586
$g_2$ = Coeficiente de Curtose	18,07	1,748
Média	1,289	0,836
Mediana	0	0
Moda	0	0
Máximo	20	5
Variância	5,963	1,658
CV= Coeficiente Variação (%)	189,30%	153,88%
Shapiro-Wilk	W =0,5740 Pr < W = 0,00	W=0,694 Pr < W = 0,00

<sup>1</sup> As estatísticas W devem ser não significativas para obtenção da distribuição normal.

O *Boxplot* identificou a presença de *outliers* (Figura 3), cujos valores variaram de seis a vinte percevejos. Após a retirada dos *outliers* foram detectados pelos gráficos do *boxplot* e carta de controle, que os valores dos coeficientes de assimetria e curtose reduziram para 1,586 e 1,748, respectivamente e coeficiente de variação para 153,88 (Tabela 1). Pode-se verificar que mesmo após a retirada de dados discrepantes o conjunto de dados ainda apresenta variabilidade, bastante comum em estudos com população de insetos.

Com a Figura 5A foram detectados os mesmos pontos discrepantes já indicados na Figura 4. Esse gráfico em geral é mais rigoroso que os demais gráficos que detectam *outliers*. Na retirada dos pontos discrepantes foi observado que na carta de controle (Figura 5B), embora ainda apresentem alguns valores distantes da massa de dados, houve maior uniformidade para os demais dados entre os limites superiores e inferiores. Os métodos exploratórios numéricos e gráficos são essenciais antes de qualquer análise conclusiva. Por esse processo que avaliamos se existem valores discrepantes no conjunto de dados, e sua retirada contribui para aumentar a eficácia da análise estatística (Hoaglin et al,1992). Salienta-se que tanto para dados sistematizados (georreferenciados) quanto aos dados casualizados a análise exploratória é uma ferramenta importante na análise de dados.

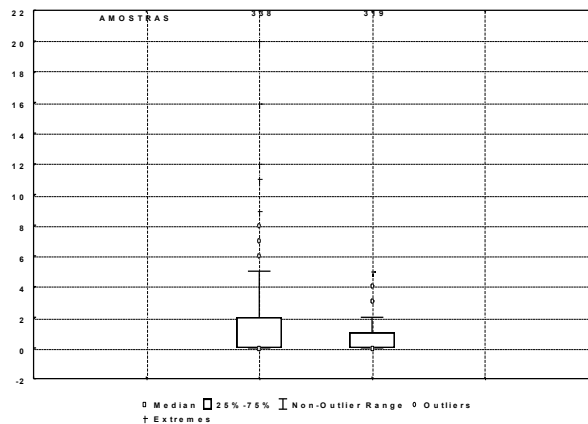


Figura 4. Boxplot antes (338 observações) e após a retirada dos outliers (319 observações) da variável ninfa grande de *Piezodorus guildinii*.

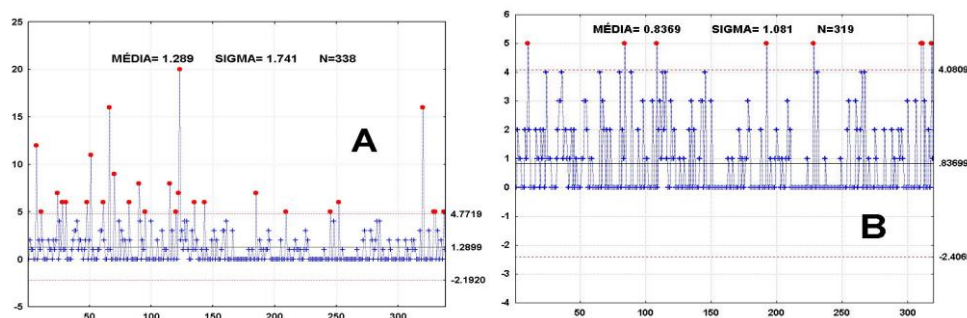


Figura 5. Cartas de Controle para as mesmas situações dos boxplot, com 338 observações (A) e 319 observações (B) após a retirada dos outliers.

Inicialmente, para o estudo geoestatístico é necessário avaliar se os dados apresentam tendência, ou seja, se não indicam patamar bem definido. Em caso afirmativo, após a determinação da semivariância experimental em conjunto com a definição do tamanho do passo a ser dado, os dados foram ajustados por meio da análise de regressão cúbica. Nesse método houve a retirada da tendência tornando-se o modelo mais adequado a esses dados com o passo de 15 metros (Tabela 2). No estudo da dependência espacial pelos modelos exponencial, gaussiano e esférico, foram obtidas as estimativas dos parâmetros e indicado como melhor modelo o gaussiano que se ajustou ao padrão espacial da distribuição de *Piezodorus guildinii*. Com esse modelo foi obtido o menor valor do quadrado médio do erro, maior valor de  $R^2$  e com esse ajuste os primeiros valores foram contemplados pelo modelo gaussiano, pré-requisito ideal para avaliação da dependência espacial (Tabela 3 e Figura 5 B).

Tabela 2. Estimativas dos parâmetros geoestatísticos, coeficiente de determinação e quadrado médio para a variável número de insetos retirando tendência cúbica.

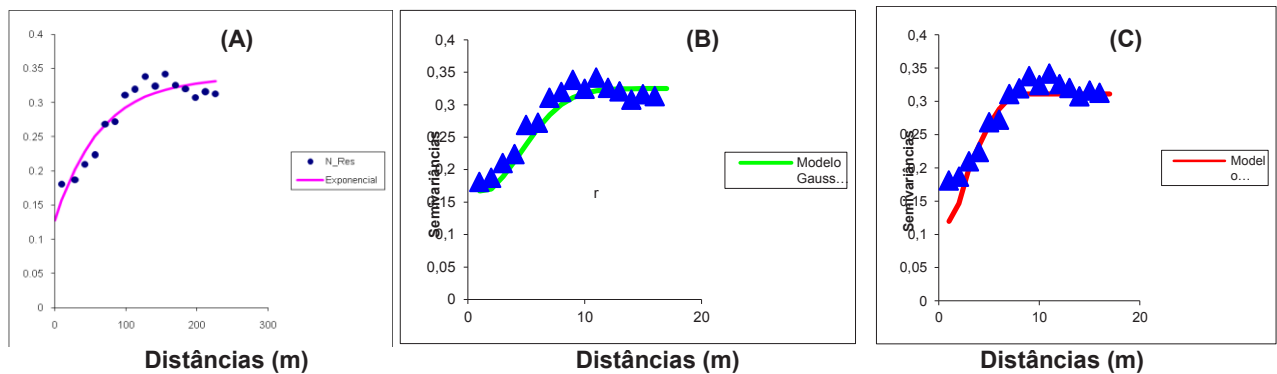
Fonte de Variação	Graus de Liberdade	Soma dos Quadrados	Quadrado Médio	Teste F
Regressão	9	29,81	3,062	34,33
Desvio	309	27,55	0,0891	
<b>Total</b>	<b>318</b>	<b>57,37</b>		

Coeficiente de Determinação do Ajuste ( $R^2$ )= 0,5197      Coeficiente de Correlação (R) = 0,7209

Observou-se nesta tabela que a dependência espacial no modelo gaussiano ocorreu até 125,8049 metros. O modelo Gaussiano apresenta também uma das exigências para que seja considerado como melhor modelo que representa dependência espacial, que o ajuste passe pelos primeiros pontos do semivariograma (Figura 6).

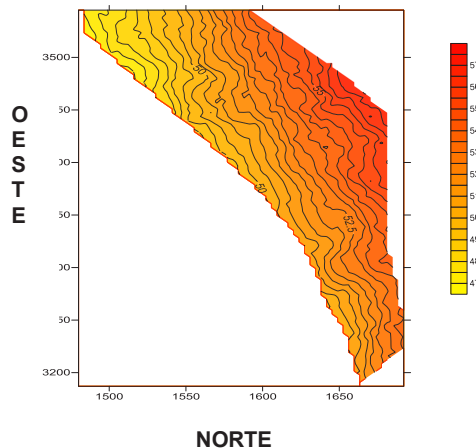
**Tabela 3.** Estimativas dos Parâmetros geoestatísticos, coeficiente de determinação e quadrado médio para a variável número de insetos da espécie *Piezodorus guildinii*.

Modelos	Efeito Pepita ( $\tau^2$ )	Patamar ( $\sigma^2$ )	Alcance ( $\phi$ )	R <sup>2</sup>	Quadrado Médio do Erro
Exponencial	0,12799	0,20934	193,1995	0,8857	0,001191
Gaussiano	0,16730	0,15769	125,8049	0,9548	0,000667
Esférico	0,11960	0,19122	100,3705	0,8393	0,001556



**Figura 6.** Semivariogramas ajustados aos modelos: Exponencial (A); Gaussiano (B); Esférico (C) utilizando os dados de contagem de *Piezodorus guildinii*.

Com a obtenção do mapa pela *krigagem* indicatriz foram observados que o maior número de percevejos ajustados pelo modelo gaussiano entraram na área experimental pelo lado superior leste. Essas informações são importantes para que o pesquisador escolha estratégias no controle do dano dos percevejos monitorando a entrada dos mesmos na área experimental (Figura 7).



**Figura 7.** Mapa realizado com a krigagem indicatriz ajustado pelo modelo gaussiano aos dados de contagem de *Piezodorus guildinii*.

## Conclusões

- o método exploratório é eficiente para detectar valores discrepantes;
- os dados de contagem de percevejos da espécie *Piezodorus guildinii* apresentam dependência espacial;
- por meio de mapa obtido pela *krigagem* indicatriz é possível detectar a entrada de percevejos na área experimental e,
- com a aplicação dos métodos geoestatísticos é possível reduzir o número de amostras em áreas georreferenciadas na pesquisa com insetos-pragas.

## Referências

- HOFFMANN-CAMPO, C.B.; MOSCARDI, F.; CORRÊA-FERREIRA, B.S.; SOSA-GÓMEZ, D.R.; PANIZZI, A.R.; CORSO, I.C.; GAZZONI, D.L.; OLIVEIRA, E.B. de. **Pragas da soja no Brasil e seu manejo integrado**. Londrina: Embrapa Soja. 2000. 70p. (Embrapa Soja. Circular Técnica, 30).
- BANZATO, D. A.; KRONKA, S. N. **Experimentação agrícola**. 4. ed. Jaboticabal: Funep,2006. 237 p.
- SAS INSTITUTE. **SAS/STAT: user's Guide**. Version 9.2. Cary: SAS Institute, 2009. 7869p.
- SHAPIRO, S.S. & WILK, M.B. "An analysis of variance test for normality".*Biometrika* 52:591-611. 1965.
- STATSOFT, Inc. **STATISTICA for windows** (Computer program manual).Tulsa, Oklahoma, USA, 2001.
- SURFER. **Surfer version 9.9.785 – mar 26.2010, surface mapping system**. [www.Goldensoftware.com](http://www.Goldensoftware.com).
- VIEIRA, S. R. **Curso de Geoestatística**. Curso ministrado na Embrapa Soja, Dezembro, 2010.
- CRESSIE, N. A. C. **Statistics for spatial data**. New York: John Wiley, 1993. 900p.
- HOAGLIN, D. C.; MOSTELLER, F.; TUKEY, J.W. **Análise exploratória de dados: técnicas robustas - um guia**. Edições Salamandra: Lisboa, 1992. 446 p