GEOSTATISTICAL ANALYSIS OF CROP YIELD MAPS IN A LONG TERM NO TILLAGE SYSTEM (1)

OSVALDO GUEDES FILHO (2*); SIDNEY ROSA VIEIRA (3);
MARCIO KOITI CHIBA (3); CÉLIA REGINA GREGO (4)

ABSTRACT

It is known, for a long time, that crop yields are not uniform at the field. In some places, it is possible to distinguish sites with both low and high yields even within the same area. This work aimed to evaluate the spatial and temporal variability of some crop yields and to identify potential zones for site specific management in an area under no-tillage system for 23 years. Data were analyzed from a 3.42 ha long term experimental area at the Centro Experimental Central of the Instituto Agronômico, located in Campinas, Sao Paulo State, Brazil. The crop yield data evaluated included the following crops: soybean, maize, lablab and triticale, and all of them were cultivated since 1985 and sampled at a regular grid of 302 points. Data were normalized and analyzed using descriptive statistics and geostatistical tools in order to demonstrate and describe the structure of the spatial variability. All crop yields showed high variability. All of them also showed spatial dependence and were fitted to the spherical model, except for the yield of the maize in 1999 productivity which was fitted to the exponential model. The north part of the area presented repeated high values of productivity in some years. There was a positive cross correlation amongst the productivity values, especially for the maize crops.

Key words: Spatial variability, cross semivariogram, management zones.

RESUMO

ANÁLISE GEOESTATÍSTICA DE MAPAS DE COLHEITA SOB SISTEMA DE PLANTIO DIRETO DE LONGA DURAÇÃO

Há muito tempo é reconhecido que as produções das culturas não são uniformes no campo. Em alguns locais há baixas ou altas produtividades dentro de uma mesma área. Dessa forma, o presente trabalho teve como objetivo avaliar a variabilidade espacial e temporal da produtividade de culturas e identificar áreas potenciais para manejo específico para o sistema de plantio direto com 23 anos de duração. Os dados de produtividade são provenientes de um experimento de longa duração desenvolvido no Centro Experimental Central do Instituto Agronômico, Campinas, SP, Brasil, em área de 3,42 ha. As produtividades avaliadas são das seguintes culturas: soja, milho, labelabe e triticale, cultivadas ao longo desse período e amostradas em uma grade regular de 302 pontos. As produtividades foram normalizadas e analisadas utilizando-se da estatística descritiva, matriz de correlação e ferramentas de geoestatística visando determinar e modelar a estrutura da variabilidade. As produtividades das culturas tiveram grande variabilidade na área. Todas as produtividades apresentaram dependência espacial e se ajustaram ao modelo matemático esférico, com exceção da produtividade milho de 1999 que se ajustou ao modelo exponencial. Na parte norte da área ocorreu repetição de altos valores de produtividade em alguns anos. Houve uma correlação cruzada positiva entre produtividades, especialmente para o milho.

Palavras-chave: variabilidade espacial, semivariograma cruzado, zonas de manejo.

---

(1) Received for publication in August 28, 2008 and accepted in October 6, 2009.
(2) ESALQ/USP. Av. Pádua Dias, 11. 13418-900 Piracicaba (SP). E-mail: osvaldoguedes@yahoo.com.br. (*) Corresponding author.
(3) Instituto Agronômico (IAC), Caixa Postal 28, 13020-902 Campinas (SP). E-mail: sidney@iac.sp.gov.br; mkchiba@iac.sp.gov.br.
(4) Embrapa Monitoramento por Satélite. Av. Soldado Passarinho, 303, Jardim Chapadão, 13070-115 Campinas (SP).
E-mail: crgrego@cnpm.embrapa.br.

1. INTRODUCTION

The main objective of site-specific management systems is the optimization of applied resources in order to increase crop production. Mapping crop productivities is a useful tool concerning to this type of management (Clark, 1999). Indeed, there has been an increasing interest on it in order to getting information related to the factors that cause spatial variation of crop yield.

The analysis of crop productivity map series is a technique that can be used to understand and manage its spatial variability (Lark et al., 1999). However, this variation is not only related to soil attributes, but it also may present some relation with climatic conditions, which indicates the association with time variation. Thus, this concept can be used to study the crop productivity considering its spatial distribution location.

Temporal and spatial analysis of productivity maps leads to the comprehension of the main factors that affect yield, so the application of resources can be optimized to plant maximum development with better productivity efficiency and without any environmental hazard (Searcy, 2000). The yield maps represent the crop responses to soil and weather conditions and their study can result on the identification of sites with varied yield levels (Gimenez and Molin, 2004). Mapping these sites allows the establishment of *local management zones* or locations with different productivity levels comparing to others that must be treated in a specific manner.

Yield maps may be used as the starting point to search spatial variations and to describe their causes. The geostatistical analysis is then a powerful tool to support the decisions made concerning site specific management, considering either the spatial or the temporal variability structure (Vieira, 2000).

This work aimed to evaluate the spatial and the temporal variability of crop yields and to identify potential specific management sites within an area under no-tillage system for 23 years.

2. MATERIAL AND METHODS

The dataset used in this paper were obtained from an experiment established in 1985 at the Centro Experimental Central of the Instituto Agronômico (IAC), located in Campinas, Sao Paulo State, Brazil (22°53′ S, 47°04′ W). The area, with 3.42 ha, is at 600 m above sea level and has 6.5% of slope. The soil was identified as a clayey Latossolo Vermelho (Embrapa, 2006), which is similar to the Rhodic Hapludox (United States, 1975). According to Köeppen classification, Campinas region is in the transition between the climatic zones Cwa and Cfa, which is a tropical climate with dry winter and rainy summer.

The studied area has been cultivated since 1985 on no-tillage system with mainly grains in the summer (soybean and maize) and lablab and triticale as winter crops. Yield measurements were taken at each crop season. The samples were collected from a regular 10 x 10 meters grid, with 302 sampling points. Yield data comprise the following crops: soybean [*Glycine max* (L.) Merr.] in 1987, 1991, 1994 and 2008; lablab (*Dolichos lablab* L.) in 1992 and 2002; maize (*Zea mays* L.) in 1989, 1999, 2003 and 2006; and triticale (*Triticum secale* L.) in 2004 and 2007. The yield of these crops was measured at sampling plots of 2.0 x 2.5 m and later transformed in kg ha^{-1}. All data were normalized according to the equation (1), which was adapted from Molin (2002) and allows the comparison amongst different species.

$$VN = \left(\frac{VP - VMin}{VMax - V\min} \right) * 100 \qquad (1)$$

where, *VN* is the normalized value, *VP* is the value at the point, *VMin* is the minimum value, and *VMax* is the maximum value.

The monthly rain recorded in Campinas, from 1987 to 2008, was also considered (Figure 1).

Data were analyzed by descriptive statistics (mean, variance, coefficients of variation, skewness and kurtosis) using the software STAT (Vieira et al., 2002). A correlation analysis amongst crop yields at different harvests was also done by Student's t-test ($p = 0.01$) (Fisher, 1970). If the correlation was significant ($p < 0.01$) then spatial correlations were tested.

In order to studying the spatial and the temporal variability of crop yields, dataset were analyzed using geostatistical methods, such as semivariogram calculations, according to Vieira (2000). The stationarity of the intrinsic hypothesis was assumed. Spatial autocorrelation amongst neighbouring points was calculated using the semivariance γ(h), as in the equation 2:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{I=1}^{N} \left[Z(x_i) - Z(x_i + h) \right]^2 \qquad (2)$$

where, *N(h)* is the number of pairs of measured values $Z(x_i)$ and $Z(x_i+h)$, separated by a vector h, which is the distance set from the coordinates of $Z(x_i)$ and $Z(x_i + h)$.

The calculation of equation (2) results in values of γ(h) corresponding to distances h. According to Vieira (2000), measurements made at neighboring points should be more similar to from each other than the ones

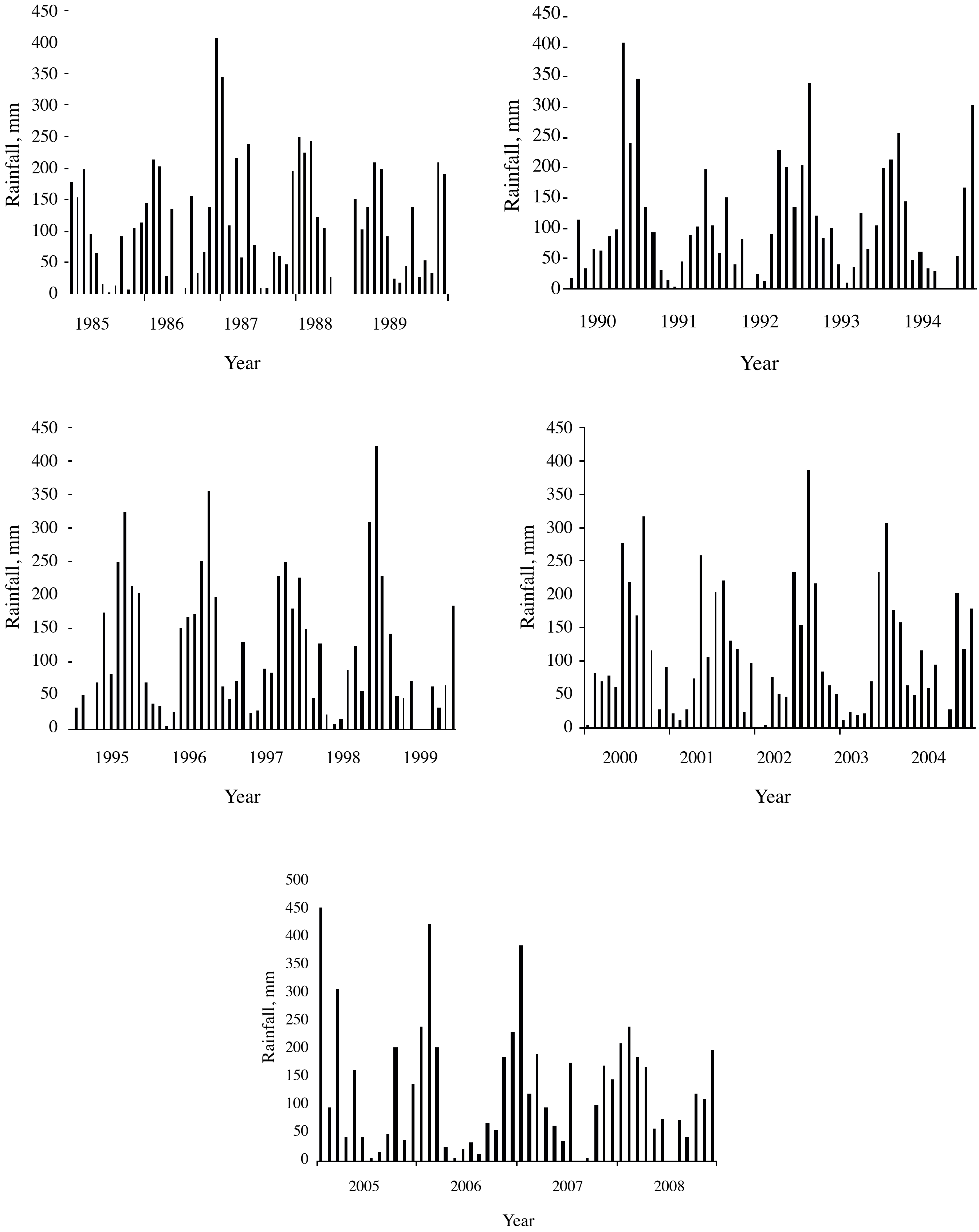

Figure 1. Monthly rain amount that occurred throughout 23 years of evaluation (1985 to 2008).

which are separated by greater distances. In this case γ(h) increases with distance until a maximum value known as sill, where it becomes stable corresponding to the range of the spatial dependence. Also, measurements at distances further than the range value, show random distribution and, in this case, classical statistics approaches may be applied.

Cross semivariograms were used to verify the spatial correlation between two sets of crop data, but only for the variables that showed spatial dependence on its individual semivariogram and significant correlation with each other. Cross semivariograms are useful to describe the spatial correlation dependence between two variables at the same location, but they are not necessary at the same sampling density (Vieira, 2000).

Mathematical models were fitted to the semivariograms according to Vieira et al. (1983). From the fitted models, the following semivariogram parameters were taken: a) nugget effect (C_0), which is the value of γ when h = 0; b) range (A), which is the distance at which γ(h) remains constant; c) sill (C_0+C1), which is the value of γ(h) at the range and is approximately equal to the data variance, if it exists.

The degree of spatial dependence (GD), which measures the proportion of the nugget effect (C_0) related to the sill ($C_0 + C_1$), was used to express it, and was calculated using the equation (3):

$$GD = \left(\frac{C_0}{C_0 + C_1} \right) \cdot 100 \quad (3)$$

According to Cambardella et al. (1994), the GD can be used to classify the spatial dependence in strong (GD < 25%), moderate (26% < GD < 75%) and weak (GD > 75%).

When the spatial autocorrelation was confirmed by the semivariogram analysis, kriging was used to interpolate values at unsampled places with no trends and with minimum variance, as described by Vieira (2000).

In order to make management decisions, it is necessary to establish criterion based on the yield values. Thus, the crop productivities were classified as low, medium-low, medium-high and high, according to the normalized productivity rank of 0-25, 26-50, 51-75 e > 75, respectively. The maps of spatial distribution were delineated using the software SURFER 7.0 (1999).

3. RESULTS AND DISCUSSION

Descriptive statistics analysis showed that crop yield presented high values of coefficients of variation (CV), between 26 and 58% (Table 1), which are higher than the ones cited by Amado et al. (2007) and Milani et al. (2006) when they evaluated the productivity of different crops.

According to the proposition of Gomes (2000), CV values larger than 20% must be considered with high variation as it was observed for all crops described here. For the same plant species a possible explanation for these high CV values across time is the existing the correlation between biomass production and the rain volume throughout the 23 years period. In rainy years the biomass production improves in comparison to dry seasons due to soil water availability. Most of the values for the coefficients of skewness and kurtosis were close to zero, indicating that the data were normally distributed, except for soybean 1991 and 2008, and maize 1999 and 2003.

The soybean crop showed high variation of mean yield values, demonstrating that this variable was not stable in time. The same was observed for triticale and maize, but, when compared to the others, the crop that presented the highest yields was maize. The winter crops presented, in general, productivity classified as medium-low, probably due to the drought season.

All yields, considering the crops from 1987 to 2008, showed spatial dependence and temporal variability, as it can be observed by the semivariogram parameters (Table 2). The spatial dependence is indicated by the ranges of the semivariograms with values larger than the sampling distance (10 m). The temporal variability is verified by the large difference in semivariogram parameters for different years as this is an indication of different variabilities according to Vieira et al. (1997). The model that best fitted to the semivariograms was the spherical, except for maize 1999 which fitted to the exponential model, confirming that the spherical model is the one that fits most soil and plant attributes (Siqueira et al., 2008). The degree of spatial dependence for yields varied between strong and moderate according to the classification rank of Cambardella et al. (1994), and it was similar to the results reported by Amado et al. (2007) to maize, soybean and wheat.

The range is an important parameter for the interpretation of the semivariogram, as it indicates the maximum distance where the sampling points are spatially correlated amongst each other. With range values varying from 15 to 77 m, it is reasonable to consider that the sampling grid adopted was adequate enough to describe the spatial variability. The highest range value was equal to 77 m for soybean 1994.

According to Carvalho et al. (2001), the nugget effect (C_0) represents an indication of the spatial discontinuity in the data. The highest values of C_0 were

Table 1. Statistical parameters for the crop yields

Productivity*	Mean	Variance	CV	Skewness	Kurtosis
Soybean 1987	47.14	550.30	49.76	0.23	-0.47
Maize 1989	59.94	325.70	30.11	-0.46	0.96
Soybean 1991	36.59	277.90	45.56	1.07	2.74
Lablab 1992	44.67	662.70	57.63	0.13	-0.48
Soybean 1994	55.38	547.50	42.25	-0.29	-0.16
Maize 1999	70.57	340.50	26.15	-1.55	4.04
Lablab 2002	36.44	367.90	52.64	0.40	0.00
Maize 2003	54.73	270.90	30.08	-0.61	1.71
Triticale 2004	44.93	321.40	39.90	0.29	0.00
Maize 2006	37.97	308.90	46.29	0.37	-0.14
Triticale 2007	43.78	305.70	39.93	0.12	0.00
Soybean 2008	42.62	344.90	43.57	0.64	0.00

*Normalized productivity. CV = Coefficient of variation.

Table 2. Fitting parameters of the experimental semivariogram for the crop yields

Productivity	Model	C_0	C_1	A	GD
Soybean 1987	Spherical	180	430	63	30
Maize 1989	Spherical	80	240	58	25
Soybean 1991	Spherical	95	150	38	39
Lablab 1992	Spherical	100	600	59	14
Soybean 1994	Spherical	172	443	77	28
Maize 1999	Exponential	160	175	28	48
Lablab 2002	Spherical	120	228	41	34
Maize 2003	Spherical	100	95	35	51
Triticale 2004	Spherical	205	100	15	67
Maize 2006	Spherical	100	140	33	42
Triticale 2007	Spherical	217	85	49	72
Soybean 2008	Spherical	143	167	57	46

C_0 = Nugget effect; C_1 = Structured variance; A = Range; GD = Spatial dependency degree.

found for soybean 1987, soybean 1994, maize 1999, triticale 2004 and triticale 2007 (Table 2).

Yield map analysis showed a high variation during the 23 years of no tillage system. The soybean yield maps presented a high variation of the spatial distribution for the four years analyzed, and the highest variability was in 1994, followed by 1987, 2008 and 1991 (Figure 2). The area presented, in general, the productivity classified as medium-low, ranging between 25 and 50. However, it is worth mentioning that there were regions that showed a high response of productivity, that is, regions with yields classified as medium-high. The maize maps also showed high spatial variability, with the highest in 2003, followed by 2006, 1989 and 1999. The maize productivity was ranked as medium-high, with some areas classified as high, although there were also small regions classified as medium-high and low. The predominance of the regions classified as medium-high may be due to the fact that maize has a high response capacity to the environmental conditions as demonstrated by Molin (2002). In a similar way, the yield of the winter crop lablab had a high spatial variability (Figure 2), with the major part of the area for the years 1992 and 2002 (especially 1992) classified as medium-low, despite of the lablab 1992 has shown some regions classified as medium-high and high. Triticale was the crop which presented the lowest spatial variability amongst all

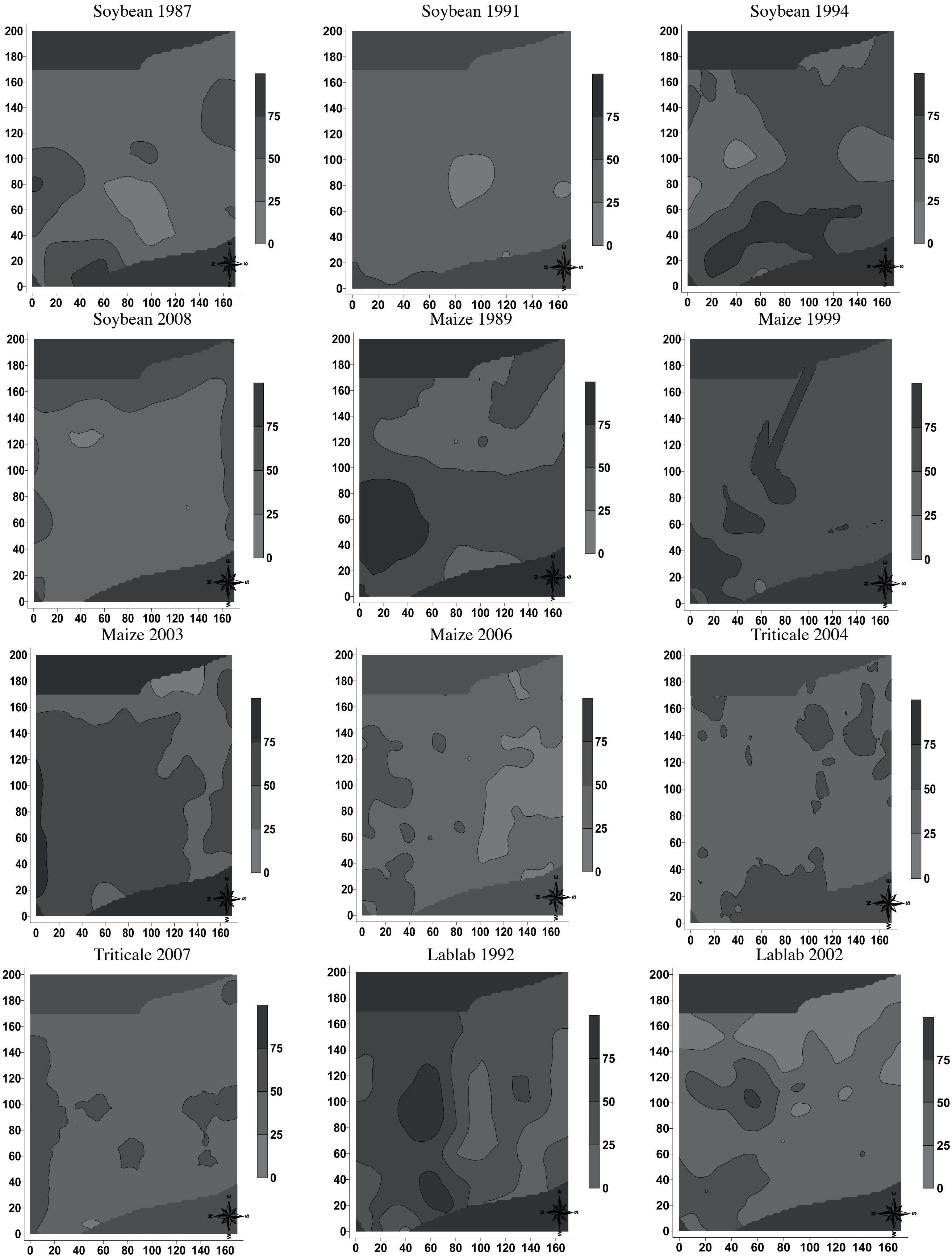

Figure 2. Maps of the spatial variability of the crop yields.

Table 3. Coefficients of correlation amongst the crop yields

Productivity	Soybean1987	Maize1989	Soybean1991	Lablab1992	Soybean1994	Maize1999	Lablab2002	Maize2003	Triticale2004	Maize2006	Triticale2007	Soybean2008
Soybean1987	1.00											
Maize1989	0.20	1.00										
Soybean1991	0.24	0.08	1.00									
Lablab1992	0.01	0.16	0.07	1.00								
Soybean1994	-0.23	-0.10	-0.23	-0.22	1.00							
Maize1999	0.00	0.15	-0.25	0.03	0.03	1.00						
Lablab2002	-0.14	0.34*	0.18	0.22	0.17	0.05	1.00					
Maize2003	0.01	0.35*	0.07	0.15	-0.05	0.35*	0.35*	1.00				
Triticale2004	-0.15	0.00	-0.14	0.09	0.02	0.03	0.02	-0.01	1.00			
Maize2006	0.05	0.25	0.13	0.31*	-0.11	0.17	0.19	0.35*	-0.07	1.00		
Triticale2007	-0.02	0.28	0.05	-0.01	-0.38 *	0.07	0.10	0.20	0.05	0.13	1.00	
Soybean2008	0.15	0.31*	0.06	-0.14	-0.08	0.06	-0.25	-0.22	-0.05	0.04	0.11	1.00

*Significant at 1% of probability according to the t test (Student).

evaluated crops, and it was classified, in general, as medium-low.

Analyzing all the maps for the last five years, except the one made for triticale 2004, it is possible to see that at the north side of the area (left side of the figure 2) there is a region with medium-high productivity, a level that is not reached at different parts of this area. Inferences concerning the triticale 2004 crop are not completely realistic due to a severe attack of wheat caterpillars (*Pseudaletia sequax* Franclemlont) (Grego et al., 2006) on that side of the area. Thus, it is reasonable to think that site-specific management, especially related to the physical and chemical soil atributtes, should be applied to other parts of this area, in particular in medium-low rank ones. There is still a need to verify what factors are more restrictive to the productivity, whether physical, chemical, or both, in order to improve its production potential.

The correlation results amongst the crop yields are shown in the table 3. The crops that presented productivity values which were significantly correlated, based on the Student t test at 1% of probability level were: maize 2006 vs. lablab 1992, maize 2003 vs. maize 1999, maize 2003 vs. lablab 2002, maize 2003 vs. maize 2006, lablab 2002 vs. maize 1989, maize 2003 vs. maize 1989, triticale 2007 vs. soybean 1994 and soybean 2008 vs. maize 1989. Almost all the significant correlations were positive, except the triticale 2007 vs. soybean 1994, which showed a negative correlation. The positive correlation indicates that these data presented a similar

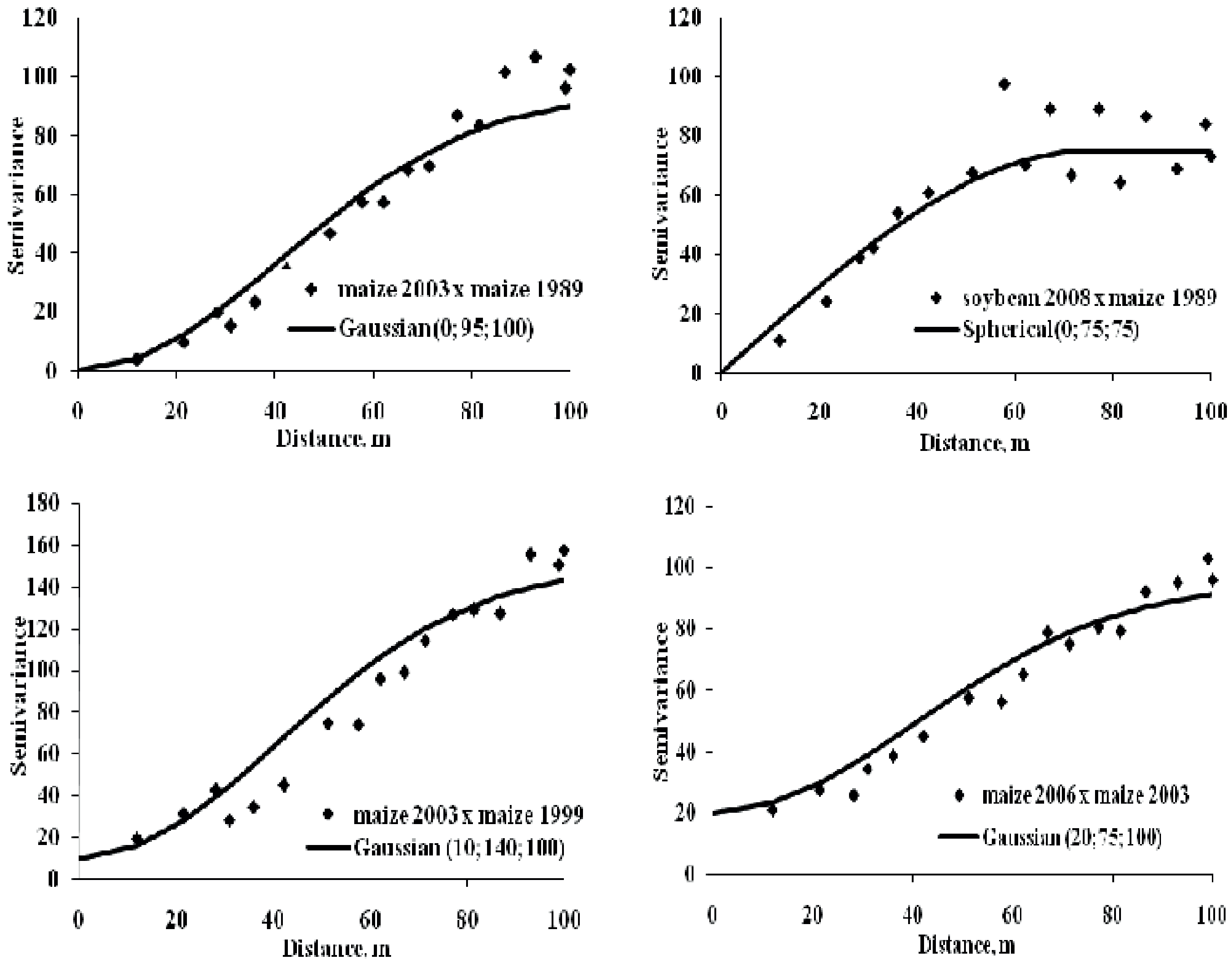

Figure 3. Cross-semivariograms for the correlated yields.

spatial behavior, especially for the maize, which was positively correlated to data belonging to the crops recorded in 1989, 1999, 2003 and 2006.

The spatial dependence among different plant species could only be clearly verified by the cross-semivariogram modeling. The spherical model was fitted for soybean 2008 vs. maize 1989, and the gaussian models were used for maize 1989 vs. maize 2003, maize 2003 vs. maize 1999 and maize 2006 vs. maize 2003. The range of the cross-semivariograms varied between 75 and 100 meters (Figure 3). Modeling cross-semivariograms is critical for the cokriging interpolation because it needs both individual and cross-semivariograms fitting in order to representing the spatial correlation of the variables (Vauclin et al., 2003; Ortiz et al., 2002).

When a cross correlation is established between two variables, as it was verified in this study, the spatial correlation analysis using cross-semivariograms is valid, as it also helps on the visualization of areas with similar spatial behavior throughout the years.

4. CONCLUSIONS

1. High variability was detected for productivity at the experimental area from 1987 to 2008 under no-tillage system. Maize was the crop which showed the highest response to this variability.

2. The north part of the area presented repeated high values of productivity in some years (2002, 2003, 2006, 2007 and 2008). This shows that, in this case, these five yield maps, made from consecutive data, were enough to evidence site-specific management zones.

3. There was a positive cross correlation amongst crop yields, especially for the maize crops, throughout time allowing the visualization of areas with similar spatial behavior.

ACKNOWLEDGEMENTS

To Agrisus Foundation for the MSc scholarship granted to the first author.

REFERENCES

AMADO, T.J.C.; PONTELLI, C.B.; SANTI, A.L.; VIANA, J.H.M.; SULZBACH, L.A.S. Variabilidade espacial e temporal da produtividade de culturas sob sistema de plantio direto. **Pesquisa Agropecuária Brasileira**, v.42, p.1101-1110, 2007.

CAMBARDELLA, C.A.; MOOMAN, T.B.; NOVAK, J.M.; PARKIN, T.B.; KARLEN, D.L.; TURV, R.F.; KONOPA, A.E. Field scale variability of soil properties in central Iowa soil. **Soil Science of Society America Journal**, v.47, p.1501-1511, 1994.

CARVALHO, J.R.; VIEIRA, S.R.; MARINHO, P.R.; DECHEN, S.C.F.; MARIA, I.C.; POTT, C.A.; DUFRANC, G. **Avaliação da variabilidade espacial de parâmetros físicos do solo sob semeadura direta em São Paulo – Brasil**. Campinas: EMBRAPA, 2001, p.1-4. (Comunicado Técnico)

CLARK, R.L. **Evaluation of the potential to develop soil strength maps using a cone penetrometer**. St. Joseph: ASAE,1999.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA-EMBRAPA. Centro Nacional de Pesquisa de Solos. **Sistema Brasileiro de Classificação de Solos**. 2.ed. Rio de Janeiro, 2006. 306p.

FISHER, R.A. **Statistical methods for research workers**. Edinburgo: Oliver & Boyd, 1970. 362p .

GIMENEZ, L. M.; MOLIN, J. P. Mapeamento da produtividade. In: MACHADO, P.L.O. de A.; BERNARDI, A.C. de C.; SILVA, C.A. **Agricultura de precisão para o manejo da fertilidade do solo em sistema plantio direto**. (Ed.). Rio de Janeiro: Embrapa Solos, 2004. p.77-92.

GOMES, F.P. **Curso de estatística experimental**. 6 ed. Piracicaba: Binetti, 1976. 430p.

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

LARK, R.M.; OLAM, H.C.; MAYR, T.; BRADLEY, R.I.; BURTON, R.G.O.; DAMPNEY, P.M.R. Analysis of yield maps in support of field investigation. In: STAFFORD, J. **Precision Agriculture** 99. Sheffield: Sheffield Academic Press, 1999. Part I, p.151-162.

MILANI, L.; SOUZA, E.G. de; URIBE-OPAZO, M.A.; GABRIEL FILHO, A.; JOHANN, J.A.; PEREIRA, J.O. Unidades de manejo a partir de dados de produtividade. **Acta Scientiarum Agronomy**, v.28, p.591-598, 2006.

MOLIN, J.P. Definição de unidades de manejo a partir de mapas de produtividade. **Engenharia Agrícola**, v.22, p.83-92, 2002.

ORTIZ, G.C. **Aplicação de métodos geoestatísticos para identificar a magnitude e a estrutura da variabilidade espacial de variáveis físicas do solo**. 2002. 91p. Tese (Mestrado) Escola Superior de Agricultura Luiz de Queiroz, Piracicaba.

SEARCY, S.W. A agricultura de precisão: um desafio para a extensão e educação. In: BALASTREIRE, L.A. **O estudo da arte da agricultura de precisão no Brasil**. Piracicaba, 2000. p.17-24.

SIQUEIRA, G.M.; VIEIRA, S.R.; CEDDIA, MB. Variabilidade espacial de atributos físicos do solo determinados por métodos diversos. **Bragantia**, v.67, p.693-699, 2008.

SURFER for Windows. Release 7.0. **Contouring and 3D surface mapping for scientist's engineers**. User's guide. New York: Golden Software, 1999. 619p.

UNITED STATES. Departament of Agriculture. Soil Survey Division. Soil Conservation Service. Soil Survey Staff. **Soil Taxonomy**: a basic system of soil classification for making and interpreting soil surveys. Washington, D.C., 1975. 754 p. (USDA. Agriculture Handbook, 436)

VAUCLIN, M.; VIEIRA, S.R.; VAUCHAUD, G.; NIELSEN, D.R. The use of cokriging with limited field soil observation. **Soil Science Society of America Journal**, v.47, p.175-84, 1983.

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

VIEIRA, S.R.; MILLETE, J.; TOPP, G.C.; REYNOLDS, W.D. Handbook for geoestatistical analysis of variability in soil and climate data. In: ALVAREZ, V.V.H.; SCHAEFER, C.E.G.R.; BARROS, N.F.; MELLO, J.W.V.; COSTA, J.M. **Tópicos em Ciência do Solo**. Viçosa: Sociedade Brasileira de Ciência do Solo, v.2, 2002. p.1-45.

VIEIRA, S.R.; NIELSEN, D.R.; BIGGAR, J.W.; TILLOTSON, P.M. The Scaling of semivariograms and the kriging estimation. **Revista Brasileira de Ciência do Solo**, v.21, p.525-533, 1997.

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