

SPATIAL STATISTICS APPLIED TO SOYBEAN PRODUCTION DATA FROM PARANÁ STATE FOR 2003-04 TO 2009-10 CROP-YEARS

VICTOR H. R. PRUDENTE¹, CARLOS H. W. DE SOUZA², ERIVELTO MERCANTE³,
JERRY A. JOHANN⁴, MIGUEL A. URIBE-OPAZO⁵

ABSTRACT: In the current study, we performed a soybean production spatial distribution analysis in Paraná State. Seven crop-year data, from 2003-04 to 2009-10, obtained from the Paraná Department of Agriculture and Supply (SEAB) were used to develop a *Boxmap* for each crop-year, show soybean production throughout this time interval. Moran's index was used to measure spatial autocorrelation among municipalities at an aggregate level, while LISA index local correlation. For each index, different contiguity matrix and order were used and there was a significance level study. As a result, we have showed spatial relationship among cities regarding the production, which allowed the indication of high and low production clusters. Finally, identifying main soybean-producing cities, what may provide supply chain members with information to strengthen the crop production in Paraná.

KEYWORDS: spatial correlation, Moran's Index, LISA index.

ESTATÍSTICA ESPACIAL DE ÁREA APLICADA A DADOS DE PRODUÇÃO DE SOJA NO ESTADO DO PARANÁ, NOS ANOS AGRÍCOLAS DE 2003/2004 A 2009/2010

RESUMO: No presente trabalho, realizou-se a análise da distribuição espacial de área dos valores de produção municipal para a cultura da soja, no Estado do Paraná. Os dados dos sete anos-safras, entre 2003/2004 e 2009/2010, foram obtidos da Secretaria da Agricultura e do Abastecimento do Paraná (SEAB). Elaborou-se um *Boxmap* para cada ano-safra, com o intuito de demonstrar como ocorreu a alternância dos valores de produção de soja neste intervalo de tempo. O Índice de Moran foi utilizado para apresentar os níveis de autocorrelação espacial a nível global entre os municípios e o índice de autocorrelação LISA, para estudar a autocorrelação espacial a nível local, com diferentes matrizes de contiguidade e com diferentes ordens de contiguidade, assim como o estudo dos níveis de significância. Através destes métodos, foi possível demonstrar como os municípios se relacionam espacialmente quanto à produção da soja, o que permitiu a indicação de clusters de alta-baixa produção, possibilitando assim que se conheçam os principais municípios produtores de soja, subsidiando os membros da cadeia produtiva com informações para melhorar seus planos no fortalecimento da cultura no Paraná.

PALAVRAS-CHAVE: correlação espacial, Índice de Moran, LISA.

INTRODUCTION

Brazil is one of the greatest world powers in soybean production, second only to the United States. From 2000 to 2010, Brazilian production accounted for about a quarter of global output against 39% of U.S. (FAOSTAT, 2012). In the Brazilian crop-year of 2008/2009, a total of 57,165,500 tons of soybeans was produced; the largest production was 17,962,500 tons (31.42%) by Mato Grosso State, followed by Paraná with 9,509,700 tons (16.64%). Regarding the Brazilian

¹ Graduando em Engenharia Agrícola, Bolsista Iniciação Científica do CNPq, Estudante de Engenharia Agrícola, UNIOESTE, Cascavel- PR, victor.rohden@yahoo.com

² Engenheiro Agrícola, Ms. Eng. Agrícola do PGEAGRI, UNIOESTE/Cascavel – PR, carlos_hws@hotmail.com

³ Engenheiro Agrícola, Dr. Eng. Agrícola, Pesquisador Produtividade do CNPq, Prof. Adjunto do PGEAGRI, UNIOESTE, Cascavel- PR, erivelto.mercante@unioeste.br

⁴ Engenheiro Agrícola, Dr. Eng. Agrícola, Prof. Adjunto do PGEAGRI, UNIOESTE, Cascavel- PR, jerry.johann@unioeste.br

⁵ Estatístico, Dr. em Estatística, Pesquisador de Produtividade do CNPq, Prof. Associado do PGEAGRI, UNIOESTE, Cascavel- PR, miguel.opazo@unioeste.br

Recebido pelo Conselho Editorial em: 10-4-2013

Aprovado pelo Conselho Editorial em: 13-2-2014

production between 2003-04 and 2009-10, Paraná contributed with approximately 19% (CONAB, 2012).

Soybean production is not uniform among Brazilian states or even among the municipalities within a state. Thus, it becomes important to study spatial distribution to improve classical methods including the use of new technologies such as remote sensing, geographic information systems (GISs) and spatial statistics, which have been used to solve problems and provide detailed information on agriculture and environment (DORIGO et al., 2007).

Exploratory Spatial Data Analysis (ESDA) use is justified since it aims to describe spatial distribution, patterns of spatial association (spatial clustering), varied spatial regimes or other spatial instabilities (non-stationary status) and identifying outliers (non-standard observations). So, the analysis consists of observing events to see if they have some kind of systematic pattern or are randomly distributed in space (MARANDUBA JÚNIOR & ALMEIDA, 2009).

Information technology advance is one of the factors responsible for spatial statistic expansion, owing to GIS data analysis expediency. There are several methods to conduct spatial analysis; however, the Global Index or Moran's Index (I) and the Local Index of Spatial Association ($LISA$) have stood out, once they use weight matrix (W) to define neighbours, which may determine whether there is autocorrelation within data (ANSELIN, 1995).

The aim of this study was to develop a method to characterize soybean production peaks throughout the crop-years of 2003-04 and 2009-10 analysing its spatial distribution for 399 municipalities within Paraná State; as a result, showing if there is some spatial correlation among them.

MATERIAL AND METHODS

The soybean production municipal data obtained from the Paraná Department of Agriculture and Supply (SEAB, 2010) for seven crop-years between 2003/2004 and 2009/2010 were used in this study. Figure 1 shows Paraná State location within Brazil and the studied cities.

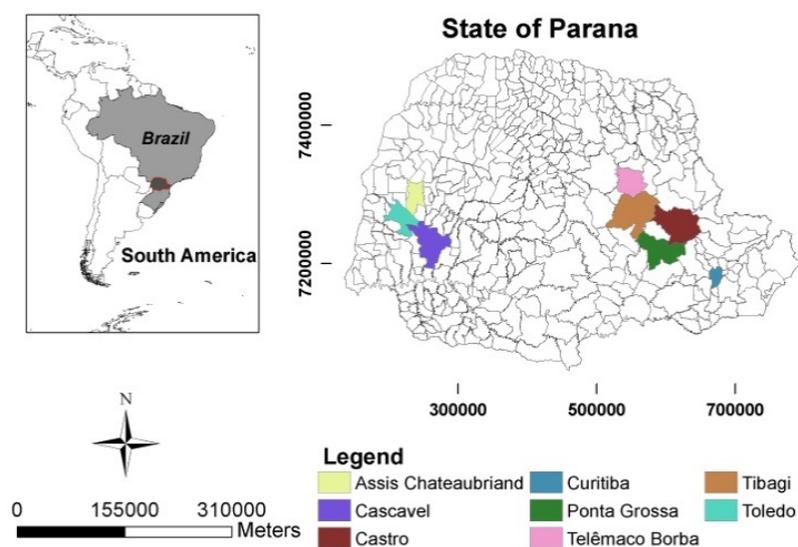


FIGURE 1. Location of Paraná state and studied cities.

The GeoDa 0.9.5-i. software was used to perform spatial data exploratory analysis and the data viewing. According to LOPES (2005), data viewing consists of attribute presentation per area by using thematic or choroplethic maps, through which is possible to check influence and behaviour that each event has on others. This is the simplest form of identifying outliers.

To identify whether the data was homogeneous, a so-called *Boxmap* was built for each crop year. The purpose of these maps, which are based on the *Boxplot* graphic, is to identify possible overall outliers, that is, municipalities with low or high soybean production outlines. Non-standard values (outliers) are those outside the interval defined by the Upper Limit ($Q_1 + [(Q_3 - Q_1) * 1.5]$) and the Lower Limit ($Q_3 - [(Q_3 - Q_1) * 1.5]$).

NOVAES et al. (2011) argue that in an exploratory spatial analysis, it is essential to check the features of spatial dependence, identifying these as attributes of a studied phenomenon correlated in space. LOPES (2005) states that three basic elements are found throughout ESDA techniques: spatial proximity matrix (W), vector of deviations (Z) and means vector (W_z).

A spatial proximity matrix is the basic element for spatial analyses, and is a useful tool for describing the spatial arrangement of municipalities. Given a set of n areas ($A_1, A_2, A_3... A_n$), a $W_{n \times n}$ weight matrix is built, where each w_{ij} element is the closeness between A_i and A_j . PIMENTEL & HADDAD (2004) state that each w_{ij} cell in the matrix $W_{n \times n}$ indicates a relation between region i and j (in a system of n regions). The cell w_{ij} is zero when the regions are not neighbours; otherwise, it is 1, which means that the element $w_{ij} = 1$ if the region A_i is a A_j neighbour, otherwise $w_{ij} = 0$, for $i \neq j = 1, 2... n$.

By using spatial proximity matrices to consider contiguity, it is important to agree the forms of neighbourhood. For this purpose, used criteria are based on queen, rook and bishop movements of a chess game (DALPOSSO, 2010)

PIMENTEL & HADDAD (2004) define that a matrix constructed based on Queen criterion considers as neighbours regions with a common edge or border, as well as a common node, which means that according to this definition microregions are neighbours when they are tangent in space. According to the same authors, a matrix built from the Rook criterion considers as neighbour exclusively regions that are border areas. ANSELIN (2002) defines neighbour regions according to Bishop criterion, that is, whether they have any node in common.

CARVALHO YWATA & ALBUQUERQUE (2011) define that besides first order neighbourhood, higher orders can also be used. Second-order neighbourhood definition, for example, the polygons i and j will be neighbours if there is another polygon k , for which i and k are first order neighbours, and j and k are also first order neighbours.

According to neighbourhood simulation presented in Figure 2, CHIARINI (2008) reports that B cells are first order with respect to A , by "Rook criterion"; with this in mind, C and D cells are second order neighbours in relation to A and also first-order with respect to B .

F	E	D	E	F
E	C	B	C	E
D	B	A	B	D
E	C	B	C	E
F	E	D	E	F

FIGURE 2. Neighbourhood criteria for first and second order neighbours. Adapted from CHIARINI (2008)

According to the same author, considering the "Queen Criterion", C and B cells are first order neighbours of A ; E , F and D cells are second order of A ; E , D and F first order of C ; and E and D are first order of B . For the 'Bishop Criterion', C cells are first order neighbours of A ; thus, D and F are considered second order of A , and also first order of C .

In the current study, we used Rook and Queen Criteria, considering first and second order neighbourhood for both aforementioned criteria.

In spatial statistics calculation, elements of the matrix W' row are multiplied by one scale data in order to the sum being normalized to 1.

A basic element used in ESDA techniques is the deviation vector Z , which represents a vector of n observations within the average. According to MARQUES (2009), in order to calculate that, we firstly have to calculate the average (μ) of the attribute under study ($y_1 \dots y_n$) considering the n objects. Each element i of Z (z_i) is obtained by subtracting the average for the corresponding attribute value ($z_i = y_i - \mu$).

Another basic element is the vector of weighted averages (W_z), which results from multiplying the transpose of the deviation vector (Z^T) by the normalized proximity matrix (W') (MARCONATO et al, 2012), obtaining a matrix W_z where each W_{zi} element contains a corresponding value of deviation average of neighbour areas to I , featuring a spatial moving average (OZON, 2011).

ESDA techniques are presented as indices, which measure spatial association (Moran's Index), scatter plot (Moran Scatterplot) and maps (*Boxmap*). According to NOVAES et al. (2011) Moran's Index is an index that expresses spatial autocorrelation, indicating grouped areas (clusters), whose attributes are spatially similar.

TEIXEIRA et al. (2008) stated that spatial weight matrix when normalized in line (elements sum is 1) provides Moran's I value as in equation 1, where Z^T is the transposed vector of deviations. Equation 2 shows deviation vector calculation and equation 3 how to calculate the weighted average.

$$I = \frac{Z^T \cdot W_z}{Z^T \cdot Z}; \quad (1)$$

$$Z = (z_i - \mu)1^T; \quad (2)$$

$$W_z = Z^T \cdot W'. \quad (3)$$

ANSELIN & LOZANO-GRACIA (2008) defines Moran's I index as a measurement whose result indicates whether data are distributed randomly in space. Moran's coefficient (I) has an expected value $E(I)$, which is defined by the following equation $E(I) = -[1/(n-1)]$, where n indicates the number of polygons within the studied area. It should be noted that unlike an ordinary correlation coefficient, this statistic is not centred at zero, and therefore, such statistics vary between -1 and +1 (MONTENEGRO & BETARELLI JUNIOR, 2008).

BETARELLI JUNIOR & ALMEIDA (2009) states that Moran's I value would be “ $-[1/(n-1)]$ ”, if there were random data distribution. ALMEIDA et al. (2005) defines that Moran I values higher or lower than the expected ($E(I) = -[1/(n-1)]$) show positive or negative spatial autocorrelation, respectively. In other words, values ranging from $-[1/(n-1)]$ to +1, indicate positive spatial correlation ($I > -[1/(n-1)]$), areas with high values (low) of a certain attribute, surrounded by areas with high values (low) of the same attribute. Secondly, values ranging from $-[1/(n-1)]$ to -1, represent negative correlation ($I < -[1/(n-1)]$), areas with high values (low) of a certain attribute surrounded with areas with low (high) values of that attribute.

With regard to testing Moran's Index significance, a reference distribution was generated using random production values and calculating a new output value for each permutation, enabling the following test hypotheses: $H_0: I = 0$ which is the null hypothesis of no spatial correlation versus an alternative hypothesis $H_1: I \neq 0$. If the descriptive level (p-value) is less than the significance level set (usually a probability of 0.05), the null hypothesis of no spatial correlation is rejected.

Otherwise, if the descriptive level is greater than or equal to the significance level, the null hypothesis cannot be rejected (DALPOSSO, 2010).

Moran's I index only returns a numeric value, which can easily lead to interpretation mistakes. An alternative to help index value interpretation is to build *Moran Scatter Plots*. MONTENEGRO & BETARELLI JUNIOR (2008) define the Moran Scatter Plot as a way to visualise overall indicators of spatial autocorrelation, revealing spatial lag of the aimed variable (i.e. average attribute of neighbourhood) on vertical and horizontal axes. Besides global measure of spatial linear association, this diagram is divided into four quadrants, namely: High-High (HH), Low-Low (LL), High-Low (HL) and Low-High (LH). Moran Scatter Plots are four-quadrant graphics representing relationship between vector Z values and weighted means (W_z).

The HH quadrant represents positive values for Z and W_z , i.e. areas with high attribute values surrounded by high value areas. The LL quadrant represents negative values for Z and W_z , that is, areas with low attribute value next to low ones. The LH quadrant represents areas with negative values for Z and positive for W_z ; these low value areas are near high value areas. The HL quadrant, which represents positive values for deviation vector Z and negative for weighted means, express areas with high values for a given attribute close to low value neighbours.

In this graphic, Moran's index (I) represents a regression line slope, that means this value indicates line inclination from W_z to Z . MONASTÉRIO et al. (2008) emphasize that Moran Scatter Plot represents a variable standardised value for each unit within the abscissa and ordinate axes.

PEROBELLI et al. (2007) indicate that the fact of Moran's index providing a single value of association as a measure for whole data may conceal spatial autocorrelation local patterns, with three different possible situations. Firstly, it involves the indication of insignificant Moran's I values, not always showing significant spatial autocorrelation. Secondly, when the Moran's I concealing local negative values and insignificant spatial autocorrelation. At third place, the evidence of negative global spatial autocorrelation can accommodate positive local spatial autocorrelation for certain data groups.

One way to avoid a wrong conclusion about the existing spatial autocorrelation, based only on Moran Index (I), is to break down the overall index due to its local constituents. According to PREDEBON et al. (2010), the local Moran's index (LISA), which is an overall measure break-up, produces specific values for each geographical area indicating statistical significance of cluster formation in similar geographical areas or isolated points for a particular attribute.

ANSELIN (1995) shows that a *LISA* aims to meet two goals: enabling significant patterns of spatial association and be a significant break-up of overall index.

The *LISA* index is represented by equation (4),

$$I_i = (Z_i \cdot W_{z_i}) / \sigma^2 \quad (4)$$

where, σ^2 is the distribution variance of deviation values.

Once statistical significance of local Moran's index has been determined, it would be useful to generate a map showing regions with local correlation that are significantly different from the other data. These regions are seen as "pockets" of non-stationary status, being areas with their own spatial dynamics and deserving detailed analysis (DALPOSSO, 2010).

RESULTS AND DISCUSSION

Figure 3 shows seven *Boxmaps* related to soybean production by municipality throughout the crop-years of: 2003/2004, 2004/2005, 2005/2006, 2006/2007, 2007/2008, 2008/2009 and 2009/2010.

According to Figure 3, there have not been major changes in production quantity for municipalities during studied period. There are large production outliers for all crop-years. The

cities Cascavel, Toledo, Assis Chateaubriand (Western Paraná) and Tibagi, Castro, Ponta Grossa (Mideastern Paraná) are among outliers with the highest productions for the seven crop-years. In many of Northwestern Paraná cities and almost the entire Eastern region (including the metropolitan region of the Greater Curitiba) null or negligible production dominate, although not configuring an *outlier* on low type in the *Boxmap*. This is justified because in Northwest most agricultural areas are intended for sugar cane production and livestock, while in East, the Atlantic forest dominates.

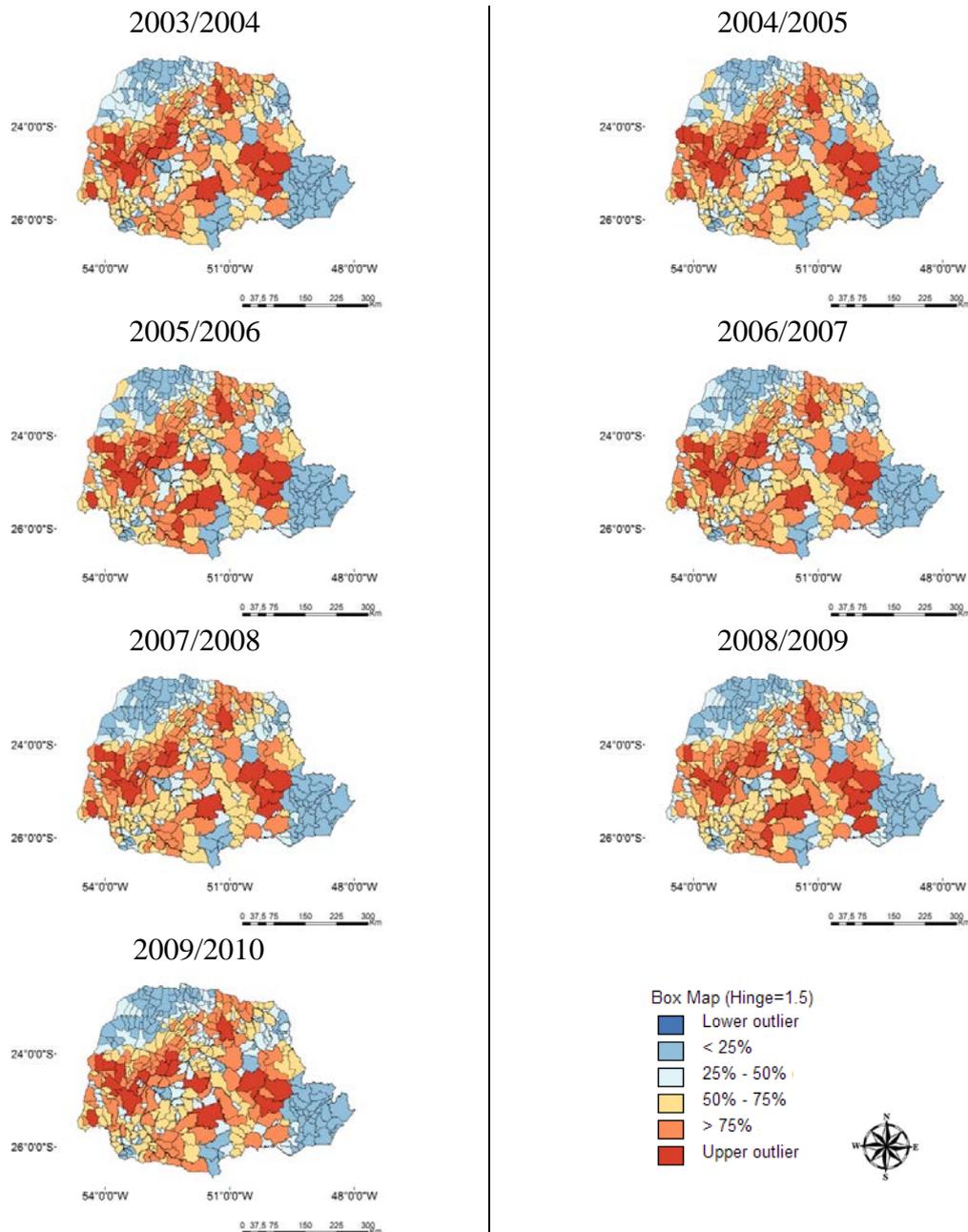


FIGURE 3. Production maps in tons (t) per municipality during all studied crop-years.

Table 1 shows the values presented in Figura 3 captions, which presents the *Boxmap* maps. These values describe intervals within the *BoxPlot*, quartile (percentile) division and upper and lower limit values. Lower Limit values are negative, being why cities that presents null production have no low-type *outlier*.

Tables 2 and 3 show Moran Index (I) values, which were obtained using different spatial proximity matrices (Queen and Rook) of first and second order respectively. The expected value (E

($I = (-[I/(n-1)])$) for Moran's Index, in both cases, considering 399 municipalities is -0.0025. Considering the first-order spatial proximity matrix, the Moran's Indexes (I) are shown in Table 2.

TABLE 1. Soybean production *Boxmap* values for each crop-year in Paraná State.

Statistics / Crop-Years	2003/04	2004/05	2005/06	2006/07	2007/08	2008/09	2009/10
Q_1	3354	3750	3000	2953	2718	2775	3916
Q_3	35347	32490	31360	41022	40678	32750	50367
Median	14300	12750	12600	17050	16727	13695	21000
Mean	25363	23858	23505	29617	29274	23358	34861
Lower limit	-44636	-39360	-39540	-54149	-54223	-42186	-65761
Upper limit	83338	75600	73900	98125	97619	77712	120044

Q1: first quartile, Q3: third quartile; Lower Limit: the lowest value for not becoming a Lower outlier; Upper Limit: the highest value for not becoming an Upper outlier; Production: Production values in tons.

All studied Moran's Index (I) values were greater than $E(I)$ (Table 2). Through this, a positive spatial correlation can be characterized, i.e. cities with high (low) production have neighbours with high (low) production. Both criteria (rook and queen) have similar Moran's I values.

TABLE 2. Moran's Index values (I) obtained by the queen and rook criteria for the matrix $W^{(1)}$.

Crop-Years	I (Queen)	I (Rook)
2003/2004	0.4146	0.4190
2004/2005	0.4293	0.4340
2005/2006	0.3815	0.3880
2006/2007	0.4088	0.4132
2007/2008	0.4183	0.4228
2008/2009	0.3947	0.3989
2009/2010	0.4252	0.4294

In case of analysing second-order neighbour matrix, it can be seen higher Moran's I values than expected ($E(I)$), which characterises a positive spatial autocorrelation among municipalities for soybean production levels.

TABLE 3. Moran's Index: Values (I) obtained by the queen and rook criteria for the matrix $W^{(2)}$.

Crop-Years	I (Queen)	I (Rook)
2003/2004	0.1866	0.1871
2004/2005	0.2233	0.2249
2005/2006	0.1667	0.1667
2006/2007	0.1908	0.1916
2007/2008	0.2103	0.2111
2008/2009	0.1520	0.1524
2009/2010	0.2055	0.2063

When analysing Moran's index for both criteria (Queen and Rook) and both neighbourhood matrices ($W^{(1)}$ and $W^{(2)}$), it is observed greater difference when we use different neighbourhood orders. Spatial autocorrelation is lower by using $W^{(2)}$ matrix (Table 3), when compared with $W^{(1)}$ (Table 2).

Moran Scatter Plots are shown in Figures 4 and 5 to assist in the interpretation of overall index (I) that are presented in Tables 2 and 3. As Moran's index represents the slope of regression line, a scatter plot allows evaluating the production value in a given municipality and the neighbour average.

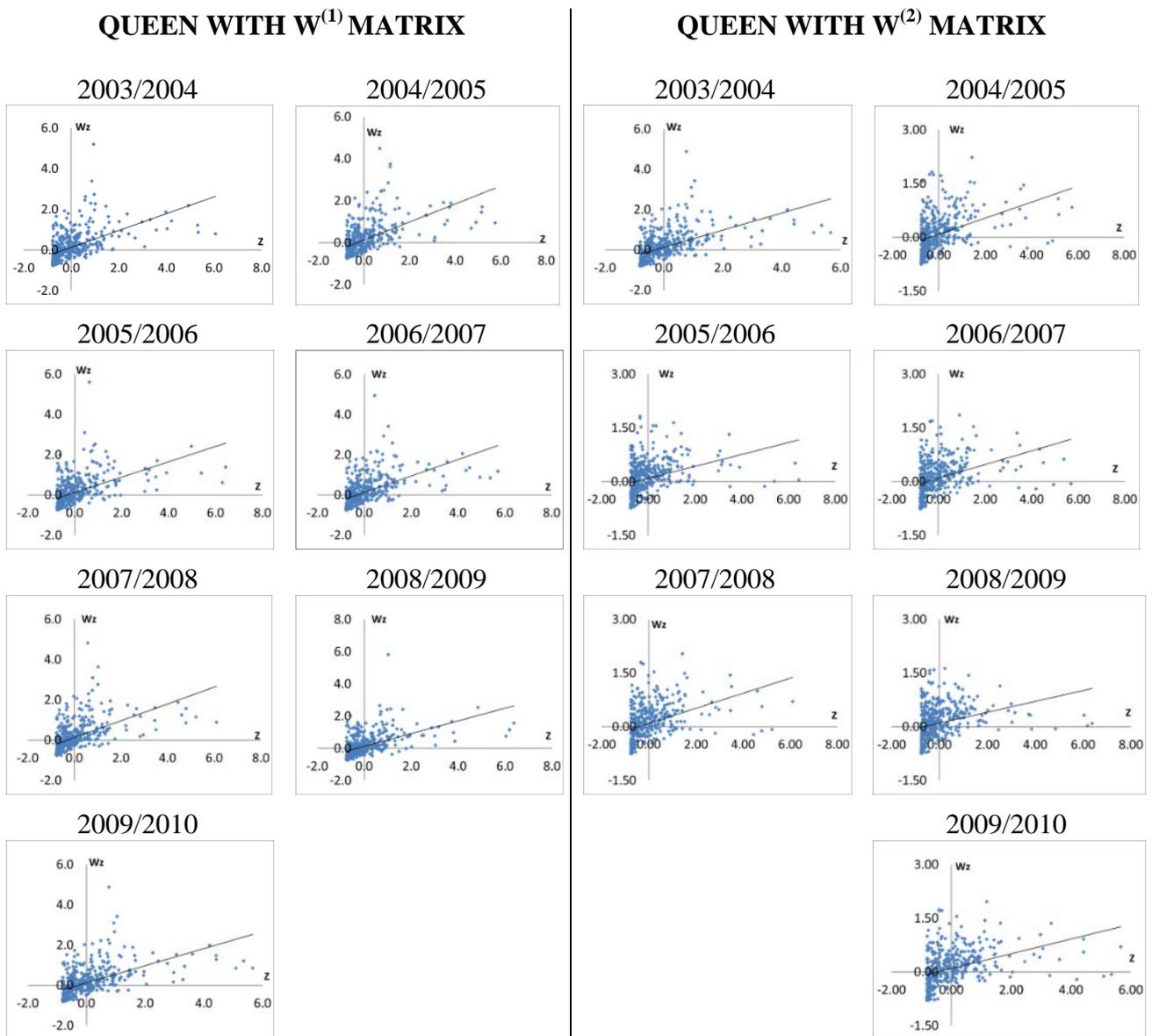


FIGURE 4. Moran scatter plot considering Queen Criterion and first and second order neighbours.

Figures 4 and 5 exemplify the Moran's index values obtained by Queen and Rook criteria with $W^{(1)}$ and $W^{(2)}$ matrices (shown in Tables 2 and 3) by means of Moran Scatter Plots. There is higher concentration of municipalities (points) in the AA (high-high) and BB (low-low) quadrants, which explains Moran's index positive value. There are also, in both figures, cities in quadrants AB and BA, which represent negative spatial association, being not only represented in numerical value of overall index (Tables 2 and 3).

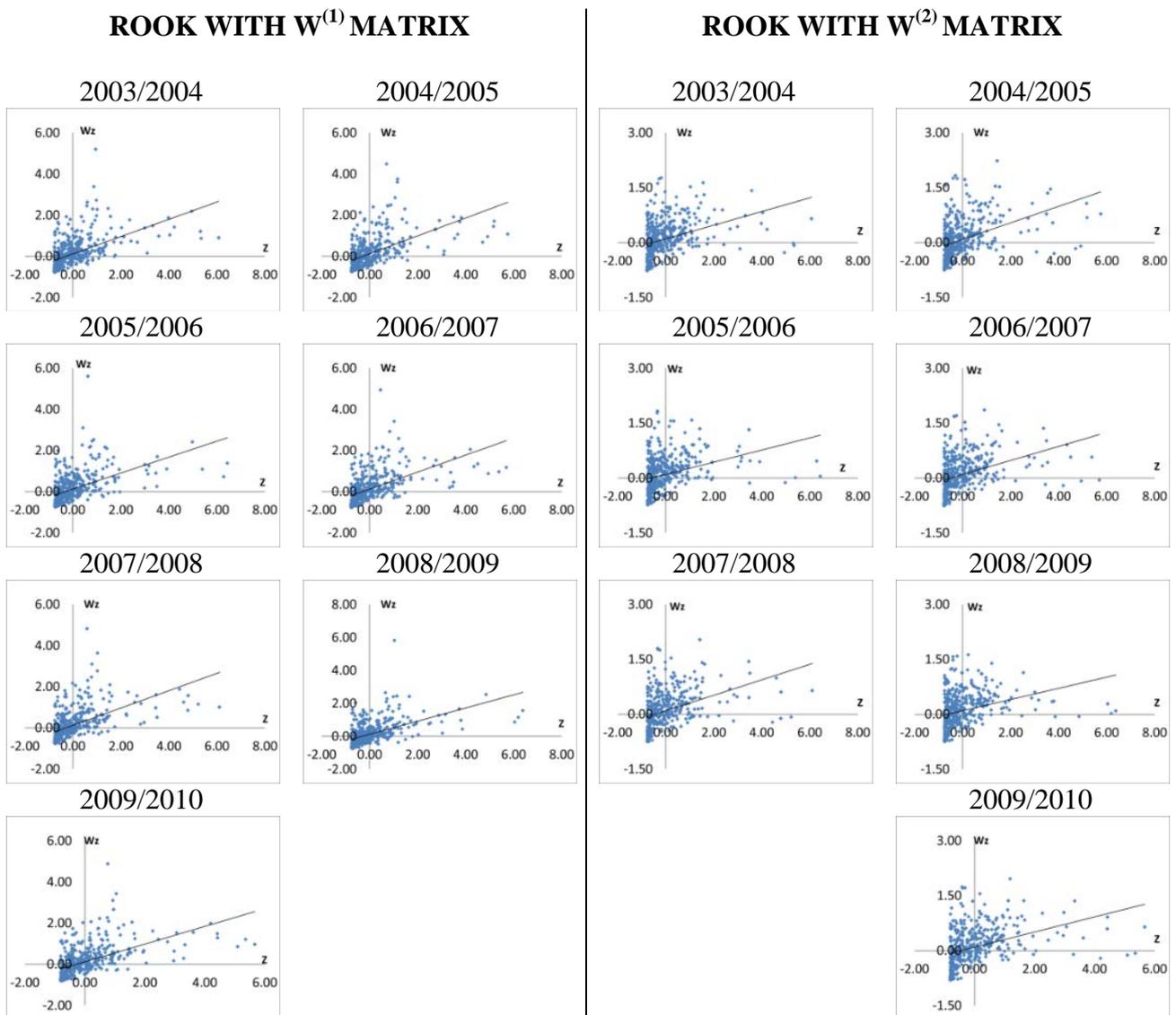


FIGURE 5. Moran scatter plot considering Rook Criterion and first and second order neighbours.

There are larger values for neighbour mean deviations (W_{zi}) using the matrix $W^{(1)}$, compared with $W^{(2)}$. That represents lower line slopes presented in matrix $W^{(1)}$ plots contrasting with matrix $W^{(2)}$. This fact makes Moran's index values (I) lower for second-order matrices, i.e. defining a neighbourhood of higher order for both criteria one can present decreasing spatial autocorrelation values.

Figure 4 and 5 shows that criterion used to determine neighbours with Moran Scatter Plots, either Queen or Rook, promoted no large variations in obtained results. According to Tables 2 and 3, Moran's index values (I) for Rook Criterion are higher compared with those from Queen. As a result, we conclude that Rook Criterion represents higher spatial correlation concerning soybean production in each city.

Evaluating values over the crop-years, there was little change in Moran Scatter Plots for each criterion and neighbourhood order. It can also be highlighted little change in spatial association orders, concluding that high-production municipalities maintained the large-scale production, while those with little or no production in same way kept the low or absent production level.

We used LISA index to analyse production spatial correlation value for each municipality. *LISA CLUSTERS MAP (LCM)* were prepared taking into account both criteria, as well as neighbourhood orders for each crop-year. By the end, clusters with low and high production were determined with only 5% significance indexes for studied crop-years.

Figure 6 shows *LCM* for Queen Criterion throughout studied years and spatial proximity matrixes of first and second order. Considering matrix $W^{(1)}$ use, the formation of large agglomerates (clusters) AA-type must be detached, i.e. high-production municipalities surrounded by high-production cities. Thus, we might highlight two major centres: the West of Paraná State, especially municipalities located around Toledo and Cascavel cities, and the East Central Paraná State, in municipalities near Castro and Tibagi cities.

Differently from the aforementioned, when analysing *LCM* for second order matrix, there was no formation of AA clusters in locations close to Castro and Tibagi cities. It also noted that for 2008/2009 crop-year, using matrix $W^{(2)}$, there was a minor AA cluster formation. As observed in able 1, this crop-year had the lowest soybean production average out of all studied crop-years.

The lower production related to 2008/2009 crop year may be related to droughts during November and December in Paraná State. This weather event had more negative effects due to much of planted soybean is from early cycle and early planting (late September and early October), therefore crop was at flowering and fruit production stage when drought happened; causing yield loss (CONAB, 2009).

Average yield for Paraná in 2008/2009 was approximately 14% lower than the other studied crop-years (CONAB, 2012). This yield reduction brought decreased grain production, even for larger areas of 91,900 ha according to CONAB (2009). Such disturbances did not restrict clusters formation for first-order neighbours as can be seen in Figure 6.

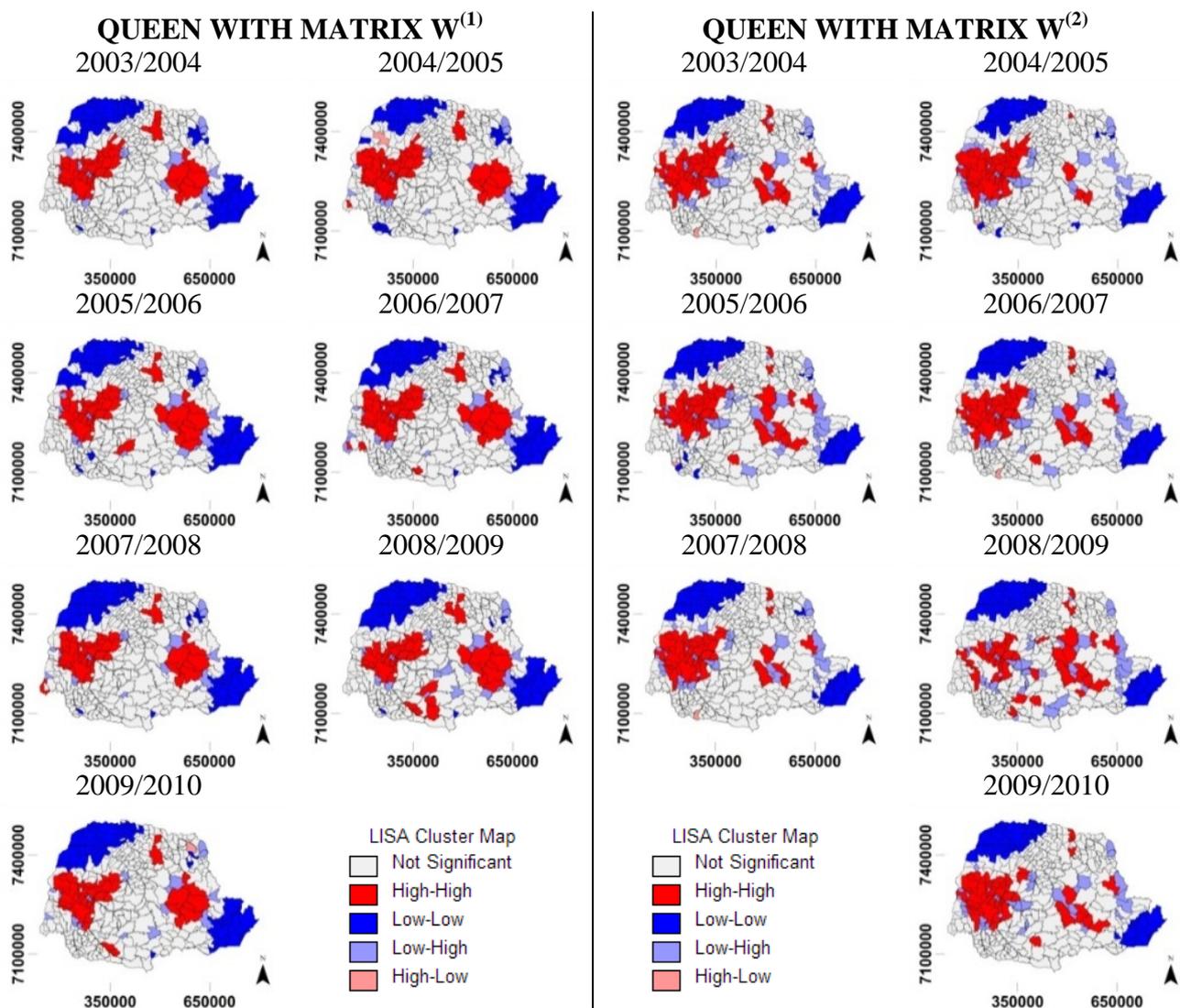


FIGURE 6. LCM considering the Queen Criterion and first and second order neighbours.

There was also two BB-type large cluster formations, where similar production level municipalities surround municipalities with little or no production, and the two clusters are located in coastal and Northwest regions.

Figure 7 represents LCM obtained when using Rook Criterion. When compared to Figure 6, concerning first order neighbours, we can observe clusters in same regions, indicating that neighbourhood criterion did not interfere in their formation, even for AA as BB cluster types.

Considering the matrix $W^{(2)}$, we can observe in Figure 7 that there is no high production cluster formation for Castro and Tibagi regions, which was already noted in Figure 6 for same neighbourhood order, with smaller AA cluster formations in 2008/2009 crop-year, what can be explained by disturbances previously mentioned.

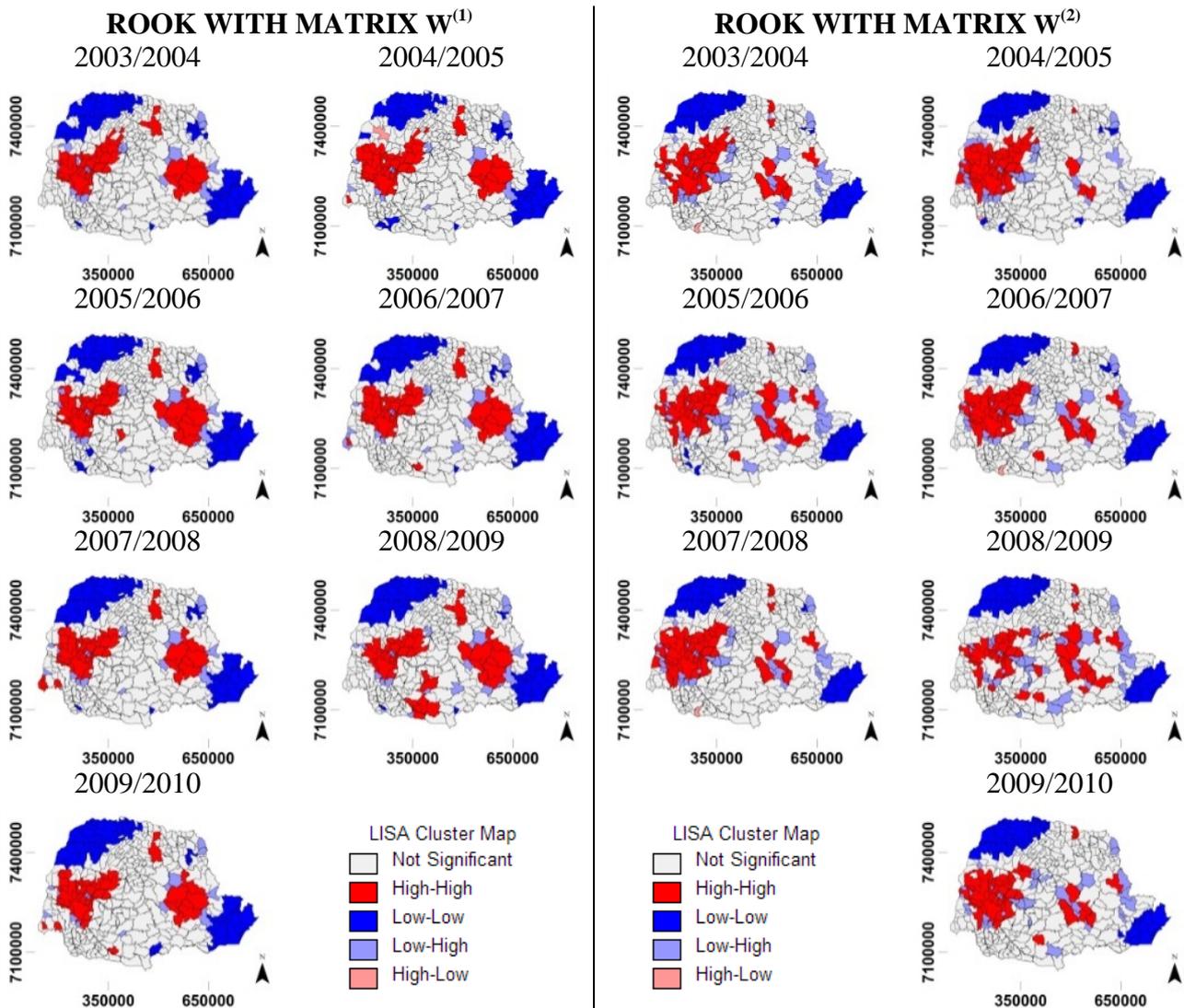


FIGURE 7. LCM considering the Rook Criterion and first and second order neighbours.

Toledo and Cascavel region formed high-production clusters in both criteria and orders matching with the findings of MERCANTE et al. (2010) who studied soybean production in some municipalities of same region once it is a great soybean area in terms of production. In the case of Castro and Tibagi regions, there was a cluster of high production when considering first order for both criteria.

One of the main Low-High type outliers is observed in Telêmaco Borba municipality. This probably occurs due to its little aptitude for grain production, since the city is a main pulp and paper

manufacturer, and according to SILVA et al. (2011) the city establishment was grounded to this sector. Therefore, there was no formation of High-Low outlier type in the region.

Regardless of criterion, there is a greater spatial autocorrelation when first order neighbourhood is adopted by observing Global Index or formed clusters. This assertive corroborates with the first law of geography: "in the world, all things are related; however, closest things relate more intensely than those more distant" (TOBLER, 1970).

There are several factors affecting spatial distribution of soybean production such as soil type, management systems, weather etc. In compliance with weather, MARIANO (2010) observed that even in the light of important technologies, climate is still largely responsible for production and yield losses.

Overall DERAL (2003) states that Paraná agriculture is performed mostly on good quality soils, which are about 90%, of Oxisol group or well-structured *Terra Roxa*, having high level of natural fertility.

Soybean cropping has been responsible for mechanised plantations growth in Paraná State (DERAL, 2003). This means that topography acts as hindrance for plantation development, as according to PEREIRA (2002), soils with more than 20% slopes are inappropriate for mechanisation at any time of year, and even for animal traction implement use.

These data enabled to understand how soybean production is spatially distributed along Paraná State, differentiating high-yield and without production regions. This kind of information is important for production flow to plot strategies that might serve not only for grain storage system, but also for production outlet (exportation).

Major soybean production centres has great importance to provide the best places to install soybean transformation industries such as animal feed, oils, milk and meat.

It is noteworthy that there are several factors affecting spatial distribution of agricultural production, such as soil types, management system, and weather among others. In this way, we highlight the great importance of undergoing studies in this line to contemplate more information layers (factors and aspects) that might influence, at anyway, spatial distribution of agricultural production.

CONCLUSIONS

A spatial profile for soybean production in Paraná State was built by using ESDA techniques, specifically with Moran's Index (I) use as global and local, varying neighbourhood matrix order and criterion.

The spatial autocorrelation values were positive among municipalities in Paraná State, i.e. soybean production values tend to be similar within neighbours regardless of adopted criteria (Queen or Rook) or even neighbourhood order ($W^{(1)}$ and $W^{(2)}$);

The neighbourhood order has interfered with obtained autocorrelation values, and first order neighbours had higher values than second ones, that is, production values tend to be more similar to in first order neighbouring than for second;

The clusters formed in Paraná West and Central Eastern regions described municipal regions with high soybean productions, featuring two great soybean production centres. Northwest and East regions (comprising the metropolitan region of Curitiba) form clusters of low soybean production over studied period.

ACKNOWLEDGEMENTS

The authors would like to thank the Paraná Department of Agriculture and Supply - SEAB-PARANÁ, the Araucaria Foundation, CNPq and CAPES for their financial support.

REFERENCES

- ALMEIDA, E.S.; HADDAD, E.A.; HEWINGS, G.J.D. The spatial pattern of crime in Minas Gerais: an exploratory analysis. *Economia Aplicada*, São Paulo, v. 9, n. 1, 2005. Disponível em: <http://www.usp.br/nereus/wp-content/uploads/TDNereus_22_03.pdf>. Acesso em: 14 dez 2012.
- ANSELIN, L. Local Indicators of Spatial Association – LISA. *Geographical Analysis*, Columbus, v.27, n.2, p.93-115, 1995. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1111/j.1538-4632.1995.tb00338.x/pdf>> Acesso em: 20 dez 2012.
- ANSELIN, L. Under the hood: issues in the specification and interpretation of spatial regression models. *Agricultural Economics*, Amsterdam, v. 27, p. 247-267, 2002.
- ANSELIN, L.; LOZANO-GRACIA, N. Errors in variables and spatial effects in hedonic house price models of ambient air quality. *Empirical Economics*, Heidelberg, v. 34, n. 1, p. 5-34, 2008
- BETARELLI JUNIOR, A. A.; ALMEIDA, E.. Os principais fatores internos e as exportações microrregionais brasileiras. *Revista de Economia Contemporânea*, Rio de Janeiro, v.13, n.2, 2009. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-98482009000200002&lng=en&nrm=iso>. Acesso em: 15 jan 2013.
- CARVALHO YWATA, A. X.; ALBUQUERQUE, P. H. M. Métodos e modelos em econometria espacial. Uma revisão. *Revista Brasileira de Biometria*, v.29, n.2, 2011. Disponível em: <http://jaguar.fcav.unesp.br/RME/fasciculos/v29/v29_n2/Alexandre.pdf>. Acesso em: 15 jan 2013.
- CHIARINI, T. *A Pobreza no espaço: uma aplicação para o Rio Grande do Sul, 2000.*, 2008. 125 f. Dissertação (Mestrado em Economia) - Faculdade de Ciências Econômicas, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2008.
- CONAB - Companhia Nacional de Abastecimento. *Acompanhamento de safra brasileira: grãos, décimo primeiro levantamento, agosto/2009*. Brasília: Conab, 2009. Disponível em: <<http://www.conab.gov.br/OlalaCMS/uploads/arquivos/ac0bd3728876ca1e9ab343877421622d.pdf>> Acesso: 30 set 2012.
- CONAB - Companhia Nacional de Abastecimento. *Soja – Brasil - Série histórica de produtividade (1976/77 – 2009/10)*. Disponível em: <<http://www.conab.gov.br/conabweb/download/safra/SojaSerieHist.xls>> Acesso em: 30 set 2012.
- DALPOSSO, G.H. *Estatística Espacial Aplicada à Agricultura de Precisão*. 2010. 80 fl. Dissertação (Mestrado em Engenharia Agrícola) - Universidade Estadual do Oeste do Paraná, Cascavel, 2010.
- DERAL – Departamento de Economia Rural. *Perfil da agropecuária paranaense. Curitiba. Novembro/2003* Disponível em: <<http://www.agricultura.pr.gov.br/arquivos/File/PDF/revista.pdf>> Acesso em: 28 de abr. 2012
- DORIGO, W.A., ZURITA-MILLA, R., WIT, A.J.W., BRAZILE, J. A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling. *International Journal of Applied Earth Observation and Geoinformation*, Enschede, v. 9, p. 165–193, 2007.
- FAOSTAT. Food and Agriculture Organization of the United Nations. *ProdSTAT – Crops. 2012*. Disponível em: <http://faostat3.fao.org/home/index_es.html?locale=es#DOWNLOAD> Acesso em: 20 set 2012.
- LOPES, S.B. *Estudo da Dependência Espacial em Modelos de Previsão de Demanda por Transportes*. 2005. 153 fl. Dissertação (Mestrado em Transportes). Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2005.

- MARANDUBA JUNIOR, N.G.; ALMEIDA, E.S. de. Análise de convergência espacial dos repasses da Lei Robin Hood. *Economia e Sociedade*, Campinas, v.18, n.3, 2009. Disponível em: < <http://www.scielo.br/pdf/ecos/v18n3/v18n3a07.pdf>>. Acesso em: 20 jan 2013.
- MARCONATO, R.; LAROCCA, A. P.; QUINTANILHA, J. A. Análise do uso de tecnologias em estabelecimentos agropecuários por meio dos índices de Moran global e local. *Revista de Política Agrícola*, Brasília, n.1, 2012. Disponível em: < <http://ainfo.cnptia.embrapa.br/digital/bitstream/item/63248/1/Analise-do-uso-de-tecnologias.pdf>>. Acesso em: 22 jan 2013.
- MARIANO, Z. de F. Precipitações pluviais e a cultura da soja em Goiás. *Revista Mercator*, Fortaleza, n.1, v.9, p. 121-134, 2010. Numero especial.
- MARQUES, A. J. Aplicação de Técnicas de Autocorrelação espacial para avaliação da organização das habitações nas mesorregiões do Estado do Paraná. *Revista Eletrônica de Geografia*, Uberlândia, v.1, n.3, p. 89-105, 2009.
- MERCANTE, E.; LAMPARELLI, R. A. C.; URIBE-OPAZO, M. A. e ROCHA, J. V.. Modelos de regressão lineares para estimativa de produtividade da soja no oeste do Paraná, utilizando dados espectrais. *Engenharia Agrícola*, Jaboticabal, v.30, n.3, 2010, Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S010069162010000300014&lng=pt&nrm=iso>. Acesso em: 30 de Janeiro de 2013.
- MONASTÉRIO, L.M.; SALVO, S.; DAMÉ, O.M.. Estrutura espacial das aglomerações e determinação dos salários industriais no Rio Grande do Sul. *Ensaio Fundação de Economia e Estatística*. v. 28, Número Especial, 2008. Disponível em: < <http://revistas.fee.tche.br/index.php/ensaios/article/viewFile/2156/2540>>. Acesso em: 18 jan 2013.
- MONTENEGRO, R.L.G.; BETARELLI JUNIOR, A.A. Análise e investigação dos fatores determinantes da inovação nos municípios de São Paulo. *Estudos Regionais e Urbanos*, Rio de Janeiro, v. 2, n.2, p. 7-31, 2008.
- NOVAES, M. R. ; RUDORFF, B. F. T ; ALMEIDA, C. M. de ; AGUIAR, D. A. de . Análise espacial da redução da queima na colheita da cana-de-açúcar: perspectivas futuras ao cumprimento do protocolo agroambiental. *Engenharia Agrícola*, Jaboticabal, v. 31, p. 572-583, 2011.
- OZON, H. O. *Índice de desenvolvimento municipal: simulando a dinâmica dos negócios com análise fatorial*. 2011. 95 fl. Dissertação (Mestrado em Desenvolvimento Econômico, Ciências Sociais Aplicadas) - Universidade Federal do Paraná, Curitiba, 2011.
- PEREIRA, L. C. *Aptidão agrícola das terras e sensibilidade ambiental: proposta metodológica*. 2002. 122p. Tese (Doutorado em Planejamento e Desenvolvimento Rural Sustentável) – Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas, Campinas, 2002.
- PEROBELLI, F.S., ALMEIDA, E.S. de, ALVIM, M.I. da S.A., FERREIRA, P.G.C. Produtividade do setor agrícola brasileiro (1991-2003): uma análise espacial. *Nova Economia*, Belo Horizonte, v.17, n.1, p. 65-91, 2007. ISSN 0103-6351.
- PIMENTEL, E., HADDAD, E. A. Análise da distribuição espacial da renda no Estado de Minas Gerais: uma abordagem setorial. *Texto para discussão, NEREUS*. São Paulo. 2004. Disponível em: < http://www.usp.br/nereus/wp-content/uploads/TDNereus_02_04.pdf> Acesso em: 25 Jun 2012.
- PREDEBON, K. M.; MATHIAS, T. A. de F.; AIDAR, T.; RODRIGUES, A. L. Desigualdade sócio-espacial expressa por indicadores do Sistema de Informações sobre Nascidos Vivos (SINASC). *Cadernos de Saúde Pública*. Rio de Janeiro, v.26, n.8, 2010. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102311X2010000800012&lng=pt&nrm=iso>. Acesso em: 25 jan 2012.

SEAB. Secretaria da Agricultura e do Abastecimento do Paraná. *Produção agrícola paranaense por Município*. Disponível em: < <http://www.agricultura.pr.gov.br/arquivos/File/deral/bdlpanet.xls>> Acesso em: 27 jul 2012.

SILVA, H. de P.; SILVA, C. L. da.; ANDREOLI, C. V. Atividade econômica de celulose e papel e desenvolvimento local: a história da Klabin e do município de Telêmaco Borba, PR. *Revista Internacional de Desenvolvimento Local*, Campina Grande, v. 12, n. 2, p. 137-148, 2011.

TEIXEIRA, R.F.A.P., ALMEIDA, E.S.de., OLIVEIRA JÚNIOR, L.B.de., ALVIM, M.I.da.S. A. Análise Espacial da Produtividade de Óleo Vegetal para Produção de Biodiesel na Zona da Mata Mineira. *Revista Eletrônica de Gestão Organizacional*. Recife, v.6, n.3, 2008. Disponível em: <<http://www.revista.ufpe.br/gestaoorg/index.php/gestao/article/viewFile/103/92>> Acesso em: 24 ago 2012.

TOBLER, W. R. A computer movie simulating urban growth in the Detroit region. *Economic Geography*, Worcester, v.46, n.2, p. 234-240, 1970.