ANALYSIS OF SPATIAL AUTOCORRELATION OF GRAIN PRODUCTION AND AGRICULTURAL STORAGE IN PARANÁ

Elizabeth G. Cima¹*, Miguel A. Uribe-Opazo², Jerry A. Johann², Weimar F. da Rocha Jr.², Gustavo H. Dalposso³

¹*Corresponding author. Universidade Estadual do Oeste do Paraná/ Cascavel - PR, Brasil. E-mail: egcima@bol.com.br

KEYWORDS
static capacity, logistics, corn, soybean, wheat and total warehouses.

ABSTRACT
This work aimed to study the spatial autocorrelation of the total static capacity storage, the total number of warehouses in 2013/2014 (CONAB) and the average of the total grain production (soybean, corn 1st and 2nd crops and wheat) in the harvest years 2008/2009 to 2013/2014 (SEAB) in Paraná State, Brazil. The study was based on Moran's global autocorrelation index, Moran’s local and Moran's bivariate correlation. It was possible to identify regions with low and high total grain production. There was positive spatial autocorrelation for the Total Static Storage Capacity (TSSC) and Total Quantity of Warehouses (TQW). For the total grain production, significant spatial autocorrelation were found. The total static storage capacity showed similarity between the studied regions. When evaluating the bivariate spatial correlation between Total Production of Harvested Grains (TPHG) in relation to the total static storage capacity and total quantity of warehouses, the presence of positive spatial correlation was observed. The results indicated that Moran's global autocorrelation and local indexes showed significant patterns of spatial autocorrelation, as well as the bivariate spatial correlation indexes in the studied variables.

INTRODUCTION
The agribusiness has been one of the pillars that sustain the Brazilian economy, contributing with 25.72% in the growth of Brazil’s Gross Domestic Product (GDP) in 2017, while the average growth was 1%, with industry and services contributing with 1.5% and 1%, respectively (CEPEA, 2017). This shows how the activities related to the field are relevant for Brazil, especially in the regions and municipalities where agricultural activities are consolidated, as in the state of Paraná (CONAB, 2017).

In the 2016/2017 harvest, according to CONAB (2016), 232.02 million tons of grains (soybean, corn, wheat and among others) were harvested in Brazil in a total planted area of 58.5 million hectares. The 2016/2017 Paraná agricultural harvest was at 40.8 million tons of grain, with a planted area of 6 million hectares in the first harvest and 1.5 million hectares in the second harvest (CONAB, 2016 and SEAB, 2016). Paraná is the second largest national producer of soybeans and corn, only behind Mato Grosso. In the case of wheat, Paraná is the largest producer, producing, in the 2016/2017 harvest, 3.41 million tons, or 62% of the national production (5.50 million tons) (CONAB, 2017).

With data from the three main products (soybean, corn and wheat) grown in the state, it is possible to carry out a better strategic planning of the production, consumption of inputs and types of management used. However, the constant production records in the agricultural harvests over the years, either by increase of productivity or by expansion of the agricultural area, have increased the demand for storage capacity, becoming more and more chronic the grain storage deficits in Brazil (Oliveira & Cicolin, 2016). A second factor that may increase the pressure for storage is associated with falling commodity prices, which means that farmers do not sell their production waiting for better prices in the future. These two factors contribute to the need for spaces to store the grains of the Brazilian harvests.

Therefore, it is necessary to evaluate grain production data and storage capacity through statistical techniques in order to understand this annual dynamic. Thus, statistical techniques that focus on the spatial analysis between regions allow the identification of

2 Universidade Estadual do Oeste do Paraná/ Cascavel - PR, Brasil.
3 Universidade Tecnológica Federal do Paraná/ Toledo - PR, Brasil.
Received in: 8-21-2017
Accepted in: 4-12-2018
patterns, behaviors of association and spatial autocorrelation that serve as reference for decision making.

The spatial association measures have been used in different areas of knowledge, among them, agrarian sciences, dealing with geographically referenced data (Gaetan & Xavier, 2010). Sass et al. (2016) reported that the exploratory analysis of spatial data allows us to discover spatial patterns in the data and to propose hypotheses that seek to describe the spatial autocorrelation, verifying if the variables are autocorrelated in the space.

Grzegorzewski et al. (2017) analyzed the spatial variability of soybean production and agro-meteorological variables in the state of Paraná through the Moran (I) global autocorrelation index and verified different sowing periods between the regions and the great climatic variability in the state. The soybean producing regions were spatially associated in the production interval in the 2003/2004 to 2009/2010 harvests in the state of Paraná, where Prudente et al. (2014) studied them from the global spatial autocorrelation. Araújo et al. (2013) applied the bivariate spatial autocorrelation analysis in spatial groupings of soybean production in the state of Paraná and identified the formation of municipalities groups, through the similarity of the variables under analysis.

The aim of this study was to analyze the spatial autocorrelation of the Total Static Storage Capacity (TSSC), the Total Quantity of Warehouses (TQW) in 2013/2014 and the average total grain production (soybean, corn 1st and 2nd harvests and wheat) from the 2008/2009 to 2013/2014 harvests. We also evaluated the bivariate spatial correlation of the total grain production with the TSSC and the TQW in 2013/2014 in the state of Paraná.

MATERIAL AND METHODS

Study Area and Data

The study area comprises 399 municipalities in the state of Paraná (Figure 1). The regions belonging to the state of Paraná stand out for the agricultural production of soybean, corn 1st and 2nd harvest and wheat. As a result of this, average data from 2008/2009 to 2013/2014 harvests were used to obtain the Total Production of Harvested Grains (TPHG) obtained from the database of the Secretariat of Agriculture Livestock and Supply (SEAB, 2015). To facilitate the visualization of these data, thematic maps of the average production of each crop were constructed.

In order to know the grain storage capacity of the state, the Total Storage Static Capacity (TSSC) and the Total Quantity of Warehouses (TQW) in 2013/2014, obtained from the National Supply Company (CONAB, 2015) and Brazilian Institute of Geography and Statistics (IBGE, 2015), were used.

Spatial Autocorrelation (Moran Index I and LISA)

For the analysis of the spatial autocorrelation, the global Moran index (I) (Equation 1) was used to evaluate the global autocorrelation and the Moran local autocorrelation index (LISA) (Equation 2), which measures the degree of spatial correlation at each specific site (Anselin & Arribas-Bel, 2013).

\[
I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (x_i - \bar{x}) (x_j - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2},
\]

that,

\[
L = \frac{1}{n} \sum_{i=1}^{n} w_{ii} (x_i - \bar{x})^2 - \frac{1}{2} \left( \frac{\sum_{i=1}^{n} w_{ii} (x_i - \bar{x})^2}{\sum_{i=1}^{n} w_{ii}} \right)^2,
\]

where:

- \( n \): number of spatial units (municipalities);
- \( x_i \) and \( x_j \): values of the attribute X considered in areas i and j;
- \( \bar{x} \): average value of attribute X in the study region;
- \( w_{ij} \): element of the normalized neighborhood matrix, corresponding to the spatial weights 0 and 1, being 0 for areas i and j that do not border each other and 1 for areas i and j that border each other.
In this study, the criterion of Queen contiguity was used (Anselin & Arribas-Bel, 2013),

\[ s_{ij} = \sum_{k=1}^{n} w_{ik} \sum_{j=1}^{n} w_{kj}, \]

that,

\[ \sigma_0^2 = \frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n}, \]

\[ x_i \]: observation of the variable of interest in the municipality \( i \) for \( i = 1, \ldots, n \),

\[ \mu \]: average of \( n \) municipalities.

The LISA serves to identify patterns of local spatial fragmentation, extreme spatial values, and to capture patterns of local association. Fu et al. (2011) argued that it is a statistic that should have, for each observation, an indication of significant spatial groups of similar values around the respective observation. According to Harries (2006), the LISA, based on the global Moran index \((I)\), can be specified for a given variable \(X_{ik}\), \( i = 1,\ldots,n \) (municipalities) \( k = 1,\ldots,p \) (variables).

The spatial exploratory analysis of the data was carried out with the help of GeoDa free software (Anselin et al., 2006). The results were presented in the form of tables and thematic maps (map of significance and LISA cluster map).

**Spatial Correlation (Moran Bivariate Index)**

According to Anselin et al. (2002), in the study of two spatially georeferenced variables, the Moran bivariate index \((I_{XY})\) is a spatial correlation index between two variables \((X, Y)\) that are obtained in the \( n \) municipalities. The bivariate Moran index \((I_{XY})\) is obtained as shown in eq. (3):

\[ I_{XY} = \frac{\sum_{j=1}^{n} z_{ij} z_{kj} w_{ij}}{\sum_{k,h=1}^{n} z_{kh}^2}, \]

that,

\[ n \]: number of municipalities;

\[ z_i \] and \( u_i \): values centered on the means of the study variables \( X \) and \( Y \) in study, respectively,

\[ z_i = (x_i - \bar{x}) \] and \( u_i = (y_i - \bar{y}) \);

\[ w_{ij} \]: element of the normalized neighborhood matrix, corresponding to the spatial weights 0 and 1, being 0 for the areas \( i \) and \( j \) that do not border each other and 1 for the areas \( i \) and \( j \) that border each other;

\[ s_{ij} \]: sum of the elements \( w_{ij} \) of the symmetric space-weight matrix \( W \), that is \( \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \).

\[ S_1^2 \text{ and } S_2^2 \text{ correspond respectively to the variances of } X \text{ and } Y, \text{ being:} \]

\[ S_1^2 = \frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n} \]

\[ S_2^2 = \frac{\sum_{i=1}^{n} (y_i - \mu)^2}{n} \]

**Box and Cox Transformation**

For the variables under study that do not present normal probability distribution, the Box and Cox transformation (Box & Cox, 1964) was carried out, as shown in eq. (4):

\[ Y_i = \begin{cases} \frac{\ln(X_i)}{\lambda}, & \text{if } \lambda = 0, \\ \frac{X_i^{\lambda}-1}{\lambda}, & \text{if } \lambda \neq 0, \end{cases} \]

that,

\( \lambda \): is a parameter that defines the transformation to be estimated from the data;

\( X_i \): corresponds to the original data, \( i = 1,\ldots,n \);

\( Y_i \): corresponds to the processed data, \( i = 1,\ldots,n \).

**RESULTS AND DISCUSSION**

**Exploratory Analysis of Total Production of Harvested Grains**

The historical average of soybeans, corn 1st and 2nd harvest and wheat production are shown in Figure 2 (between the harvest years from 2008/2009 to 2013/2014). The darker colors (Figures 2A, 2B and 2C) represent the municipalities with the highest production in tons, which predominantly are in the North East, North Central and West mesoregions. The wheat production was more centralized in the Central Eastern mesoregion, with a large presence in the municipalities of Castro, Tibagi and Guarapuava (Figure 2D). Among the crops studied, the predominance of soybean, followed by corn 2nd harvest, is evident in this analysis.

For the harvest years from 2008/2009 to 2013/2014, the total production of soybean, corn 1st and 2nd harvests and wheat was high in the municipalities of Tibagi, Ponta Grossa, Castro, Guarapuava, Campo Mourão, Cascavel, Toledo and Londrina (Figures 3A, 3B, 3C, 3D and 3F). For the other municipalities, these productions were not expressive, and this is due to the cultivation of other crops types, for example, in the Northwest mesoregion, which predominantly cultivates sugarcane, as reported by Cechim Jr. et al. (2017).
FIGURE 2. Descriptive map of the historical average (harvest years from 2008/2009 to 2013/2014) of soybean (A), corn 1st harvest (B), corn 2nd harvest (C) and wheat (D) in the state of Paraná.

In this descriptive exploratory analysis, it is evident that the regions with larger productions form a defined spatial pattern are close to each other, presenting spatial agglomerations. Among the main municipalities that produce grains and oilseeds in Paraná are Castro, Tibagi, Ponta Grossa, Assis Chateaubriand, Toledo, Cascavel, Terra Roxa and Guarapuava, which together account for 12.76% of state production (IBGE, 2017).

The variables studied presented an asymmetric behavior, which was expected, since some municipalities in the state of Paraná do not produce grains and others grow other crops, such as the Northwest mesoregion, where there is the predominance of the sugar-alcohol sector (planting and processing of sugarcane) (Cechim Jr. et al., 2017).

FIGURE 3. Descriptive map of the Total Production of Harvested Grains (t) (soybean, corn 1st and 2nd harvests and wheat) for the harvest years 2008/2009 (A), 2009/2010 (B), 2010/2011 (C), 2011/2012 (D), 2012/2013 (E) and 2013/2014 (F). Source: IBGE, CONAB and SEAB.
Global Spatial Autocorrelation of Total Production of Harvested Grains

The Box and Cox transformation ($\lambda = 0.26$) was used for the Total Production of Harvested Grains (TPHG) data from harvest years 2008/2009 to 2013/2014, so that the assumption of normality was met. The univariate Moran ($I$) global index for each crop year studied (Table 1) indicated positive spatial autocorrelation at 5% significance. The positive spatial autocorrelation for all crop years analyzed, with an average of $I = 0.6396$, shows that Paraná has municipalities with high grain production, surrounded by neighbors also with high production of these crops and municipalities with low production of these grains and surrounded by neighbors also presenting these characteristics.

TABLE 1. Moran global index ($I$) and significance test Total Production of Harvested Grains (soybean, corn 1st and 2nd harvests and wheat).

<table>
<thead>
<tr>
<th>Agricultural years</th>
<th>Moran global index ($I$)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008/2009</td>
<td>0.6489</td>
<td>0.001*</td>
</tr>
<tr>
<td>2009/2010</td>
<td>0.6497</td>
<td>0.001*</td>
</tr>
<tr>
<td>2010/2011</td>
<td>0.6544</td>
<td>0.001*</td>
</tr>
<tr>
<td>2011/2012</td>
<td>0.5943</td>
<td>0.001*</td>
</tr>
<tr>
<td>2012/2013</td>
<td>0.6394</td>
<td>0.001*</td>
</tr>
<tr>
<td>2013/2014</td>
<td>0.6510</td>
<td>0.001*</td>
</tr>
</tbody>
</table>

*Statistically significant at the 5 percent probability level.

Grzegozewski et al. (2017), when analyzing the global Moran index ($I$), identified that the spatial productivity presented positive correlation in the data and that in Paraná there are municipalities with low or high agricultural productivity surrounded by neighbors with the same characteristic.

Considering the harvest years studied, the values of the global Moran index ($I$) presented few differences among them, ranging from 0.5943 to 0.6544, characterizing the existence of spatial autocorrelation among the 399 municipalities. According to Anselin & Arribas-Bel (2013), these index values are an indication of positive spatial autocorrelation. The largest spatial autocorrelations were found for the 2009/2010, 2010/2011 and 2013/2014 harvest years, with support for the 2010/2011 harvest year, which presented the highest values of spatial autocorrelation, indicating that there was a greater similarity of TPHG among the municipalities in this harvest, when compared to other years studied (Table 1). This result corroborates with Araújo et al. (2014) who found values of the global Moran ($I$) indexes that ranged from 0.2203 to 0.8359 (between the harvest years from 2005/2006 to 2007/2008) in the soybean productivity study in the West region of Paraná.

Spatial Autocorrelation of the Total Storage Static Capacity (TSSC) and the Total Quantity of Warehouses (TQW)

The data of the TSSC and TQW were transformed through the Box and Cox transformation, so that the assumption of normality was met with parameter $\lambda = 0.12$ and $\lambda = 0.11$, respectively. Significant spatial autocorrelation of the TSSC and the TQW of the crop year 2013/2014 was found (Table 2), these two data showed similarity between the studied municipalities.

It makes sense to have relations with the municipalities, since in the state of Paraná there are municipalities with high TSSC, surrounded by municipalities also with high TSSC and municipalities with low TSSC surrounded by neighbors with low static total storage capacity, which in fact occurs with most municipalities.

TABLE 2. Moran global index ($I$) and significance test for Total Static Storage Capacity (TSSC) and Total Quantity of Warehouse (TQW) in 2013/2014.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Moran global index ($I$)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSSC</td>
<td>0.2446</td>
<td>0.001*</td>
</tr>
<tr>
<td>TQW</td>
<td>0.2675</td>
<td>0.001*</td>
</tr>
</tbody>
</table>

*Statistically significant at the 5 percent probability level.

Johann et al. (2012), when estimating and mapping the areas with soybean and corn crops in Paraná, showed that the spatial distributions of the cultivated areas showed the existence of the so-called “soy belt” contemplating from the West region to the North region. Concentrations of these same crops were also observed in the Central-Eastern and Central-South mesoregions of the state of Paraná.

Moran Local Autocorrelation Index (LISA) and Significance Test

The Moran local autocorrelation index (LISA) was calculated and the municipalities were classified in relation to the level of significance (Figure 4). The municipalities that presented the Moran local index (LISA) were discretized in different shades of green (Figures 4A, 4C, 4E, 4G, 4I, 4K), both at 0.1% and 1% significance (dark shades of green) and 5% significance (clear shades of green). On the other hand, those municipalities that did not have a significant Moran local autocorrelation index (LISA) were colorless. In the total grain production in the studied years from 2008/2009 to 2013/2014 (Figure 4), spatial patterns of clusters occur, the producing regions are similar to each other and are close to each other, allowing the identification of significant clusters (0.1%, 1% and 5%).
The Northwest mesoregion and some municipalities of the North Central region presented agglomerations of municipalities with low TPHG, as well as the Metropolitan mesoregion (dark blue color in Figure 4). The West mesoregions (comprising the municipalities of Cascavel, Toledo, Assis Chateaubriand and Terra Roxa), Western Center (comprising the municipalities of Ubiratã, Campo Mourão and Goioerê) and Eastern Center (comprising the municipalities of Castro, Tibagi and Ponta Grossa) showed clusters of municipalities with high production (red color in Figure 4).

Prudente et al. (2014) used the LISA to analyze the value of the spatial autocorrelation of soybean production for each municipality of Paraná through the LISA Cluster Map and found clustering values with high and low soybean production with a level of 5% significance in the studied harvest.

The Central South and Central North mesoregions, represented by the municipalities of Guarapuava and Londrina, are regions that have significant production, but are surrounded by regions with low TPHG. The same LISA analysis and significance test (0.1%, 1% and 5%) were carried out for TQW and TSSC for the year 2013/2014 (Figure 5).
The municipalities presented greater similarities with the p-value of 5% (light green color in Figure 5), for both TSSC and TQW. This pattern was very strong in the Metropolitan, South Center, Western and West mesoregions due to the greater total number of warehouses installed and, consequently, the greater total static capacity of the municipalities located there.

The Moran local autocorrelation indexes (LISA) are presented using the LISA Cluster Map for TQW and TSSC for 2013/2014, in Figure 6.

In Paraná, the Central, Western and Central North mesoregion presented a cluster of municipalities with high TQW surrounded by municipalities also with high TQW (Figure 6A). In the Metropolitan mesoregion, the municipality of Paranaguá stands out for having a high TQW and being surrounded by municipalities that do not have a high TQW.

In relation to the TSSC, (Figure 6B) the Paraná presented patterns of spatial clustering (clusters), represented by the color red in the mesoregions Eastern Center, Western Center and West. This shows that in these regions there is grouping with high TSSC. For the Northwest mesoregion, there is the presence of municipalities that presented clusters of municipalities with low TSSC (dark blue color Figure 6B). The Metropolitan mesoregion presented municipalities with low TSSC surrounded by the municipality of Paranaguá that presents high TSSC. For the other regions, there was no clustering, there was no defined spatial pattern.

FIGURE 6. (A): LISA Map Cluster of Total Quantity of Warehouse (TQW); (B): LISA Map Cluster of Total Static Storage Capacity (TSSC).

Spatial Correlation

In order to analyze the spatial correlations of TPHG of the harvest year 2013/2014 with TSSC and TQW of the same period, the Moran bivariate indexes ($I_{xy}$) were used (Table 3).

There was a significant positive correlation (p-value ≤ 0.05) among all the variables studied for the 2013/2014 harvest year (Table 3). The highest value for TPHG and TSSC was found (0.3890), which is justified, since there are municipalities with high and/or low total grain production surrounded by neighbors with high and/or low static total storage capacity. This was verified by SEAB (2016), who reported that the state of Paraná had no grain storage problems between the harvest years 2012/2013 and 2014/2015.

TABLE 3. Moran bivariate index ($I_{xy}$) between the Total Production of Harvested Grains (TPHG) for the Total Static Storage Capacity (t) (TSSC) and the Total Quantity of Warehouses (TQW).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Moran bivariate index $I_{xy}$</th>
<th>p-valor</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPHG vs TSSC</td>
<td>0.3890</td>
<td>0.001*</td>
</tr>
<tr>
<td>TPHG vs TQW</td>
<td>0.3669</td>
<td>0.001*</td>
</tr>
</tbody>
</table>

TPHG: Total Production of Harvested Grains; TSSC: Total Static Storage Capacity; TQW: Total Quantity of Warehouses; $I_{xy}$: Moran bivariate index; *: 5% level of significance.

CONCLUSIONS

With the spatial autocorrelation, similarities between TPHG were identified in the harvest years from 2008/2009 to 2013/2014. With the global Moran index ($I$) univariate, in the global analysis, there was spatial autocorrelation between the municipalities, which was demonstrated by the clusters.

By the TSSC, equality between the municipalities of the state was identified. For the TQW, there was a positive spatial autocorrelation among the municipalities studied. In the local analysis (LISA Moran local index), the municipalities with TPHG, TSSC and TQW also presented equivalences.

There was a bivariate spatial correlation between total grain production in relation to TSSC and TPHG in relation to TQW, in the harvest year 2013/2014.

The study showed that the TSSC and TQW vary among the regions of the state of Paraná. When associated to TPHG of the harvest year 2013/2014, they are significant, due to the existence in Paraná, municipalities with high and/or low TPHG and also in TQW and TSSC.

ACKNOWLEDGEMENTS

We would like to thanks to CAPES, CNPq and Araucaria Foundation for financial support.
REFERENCES


