

Spatial analysis of soybean yield in the western mesoregion of Paraná using agrometeorological variables

Análise espacial da produtividade da soja na mesorregião Oeste do Paraná utilizando variáveis agrometeorológicas

Análisis especial del rendimiento de la soja en la mesorregión occidental de Paraná utilizando variable agrometeorológicas

Received: 01/20/2022 | Reviewed: 01/28/2022 | Accept: 02/09/2022 | Published: 02/13/2022

Caroline Cristina Engel Gabriel

ORCID: <https://orcid.org/0000-0001-5522-9708>
Universidade Estadual do Oeste do Paraná, Brasil
E-mail: Caroline.engelg@gmail.com

Miguel Angel Uribe-Opazo

ORCID: <https://orcid.org/0000-0001-8937-4815>
Universidade Estadual do Oeste do Paraná, Brasil
E-mail: Miguel.opazo@unioeste.br

Gustavo Henrique Dalposso

ORCID: <https://orcid.org/0000-0003-2097-5748>
Universidade Tecnológica Federal do Paraná, Brasil
E-mail: gustavodalposso@utfpr.edu.br

Elizabeth Giron Cima

ORCID: <https://orcid.org/0000-0003-3539-4305>
Universidade Estadual do Oeste do Paraná, Brasil
E-mail: egcima74@gmail.com

Abstract

This paper aimed to analyze the spatial autocorrelation of soybean yield and its bivariate spatial correlation with the agrometeorological variables rainfall, mean temperature, and mean global solar radiation in 2014/2015, 2015/2016, and 2016/2017 crop years in the West of the State of Paraná – Brazil. To achieve this objective, techniques of spatial statistics of areas were used, which, through autocorrelation and spatial correlation indices, sought to identify patterns of association between soybean yield and agrometeorological variables. This research is justified because in addition to the soybean crop being the main source of food protein and vegetable oil in the world and the agrometeorological variables being the factors that most influence it, the western mesoregion of Paraná stands out with the highest production values in the state. Thus, it is important to monitor its development through spatial analysis to obtain information that will support decision making. The global and local Moran's indices showed that soybean yield is self-correlated in the municipalities of Western Paraná, identifying clusters to the west and east of the mesoregion. The significance of the bivariate spatial correlation indices confirmed the influence of the variables rainfall, mean temperature, and average global solar radiation on soybean yield.

Keywords: Spatial autocorrelation; Spatial correlation; Spatial data analysis.

Resumo

Esta pesquisa teve por objetivo analisar a autocorrelação espacial da produtividade da soja e sua correlação espacial bivariada com as variáveis agrometeorológicas precipitação pluvial, temperatura média e radiação solar global média nos anos-safra 2014/2015, 2015/2016 e 2016/2017 na mesorregião Oeste do Paraná - Brasil. Para atingir este objetivo, utilizou-se de técnicas da estatística espacial de áreas, a qual por meio de índices de autocorrelação e correlação espacial buscaram identificar padrões de associação entre a produtividade da soja e as variáveis agrometeorológicas. Esta pesquisa justifica-se, pois, além da cultura da soja ser a principal fonte de proteína alimentar e óleo vegetal do mundo e as variáveis agrometeorológicas serem os fatores que mais a influenciam, a mesorregião Oeste do Paraná se destaca com os maiores valores de produção no estado. Assim, é importante o acompanhamento do seu desenvolvimento por meio das análises espaciais visando obter informações que venham subsidiar a tomada de decisão. Os índices I de Moran global e local mostraram que a produtividade da soja está autocorrelacionada nos municípios do Oeste do Paraná, identificando-se agrupamentos a oeste e leste da mesorregião. A significância dos índices de correlação espacial bivariada confirmaram a influência das variáveis precipitação pluvial, temperatura média e radiação solar global média na produtividade da soja.

Palavras-chave: Autocorrelação espacial; Correlação espacial; Análises Espacial de dados.

Resumen

Esta investigación tuvo como objetivo analizar la autocorrelación espacial del rendimiento de la soya y su correlación espacial bivariado con las variables agrometeorológicas como lluvia, temperatura media y radiación solar global media en las cosechas 2014/2015, 2015/2016 y 2016/2017 en la mesorregión Oeste de Paraná – Brasil. Para lograr este objetivo, se utilizaron técnicas de estadística espacial de áreas que, por medio de los índices de autocorrelación y la correlación espacial, buscando identificar patrones de asociación entre el rendimiento de la soya y las variables agrometeorológicas. Esta investigación se justifica porque, además de que el cultivo de la soya es la principal fuente de proteína alimentaria y aceite vegetal en el mundo, las variables agrometeorológicas son los factores que más influyen en él. Así, también se destaca la mesorregión occidental de Paraná con los valores más altos de producción en el estado. Por ello, es importante monitorear su desarrollo mediante el análisis espacial para obtener informaciones que sirvan de apoyo a la toma de decisiones. Los índices de Moran globales y locales mostraron que la productividad de la soya está autocorrelacionada en los municipios del oeste de Paraná, identificando conglomerados al oeste y al este de la mesorregión. La importancia de los índices de correlación espacial bivariados confirmó la influencia de la lluvia, la temperatura media y la radiación solar global media en el rendimiento de la soya.

Palabras clave: Autocorrelación espacial; Correlación espacial; Análisis espacial de datos.

1. Introduction

Similar to the history of GIS – Geographic information system, spatial statistics emerged independently, scattered across a diverse set of research domains (Chan-Tack, 2014). Currently, they play a fundamental role in understanding variables related to the most diverse areas of knowledge, such as public health (Silva et al., 2020), climatology (Barboza et al., 2020), and aquaculture (Francisco et al., 2020). In the Western Mesoregion of Paraná, located in the Southern region of Brazil, there are works related to precision farming, aimed at understanding the spatial distribution of soybean productivity (Dalposso et al., 2021; Richetti et al., 2021; Silva et al., 2021). Soybean (*Glycine max* (L.) Merrill) is one of the world's main agricultural crops (CONAB, 2018). Brazil is the second largest producer of the grain: In the 2016/2017 harvest year, it obtained an average yield of 3,362 t ha⁻¹. The second largest Brazilian soybean producer is the state of Paraná, which still in the 2016/2017 cycle achieved an average yield of 3,508 t ha⁻¹. In this harvest year, the mesoregion of the West Paraná State, which obtained an average yield higher than the national average, with 3,874 t ha⁻¹ (SEAB, 2018), stood out. This soybean yield potential is determined by the interaction of soil, plant and atmosphere factors. However, the agrometeorological aspects are the main limiting effects of maximum production, since they operate at different stages of the soybean production cycle (Araujo et al., 2011). Cargnelutti Filho et al. (2009) evidence the agrometeorological variables with the greatest influence on agricultural crops: rainfall, temperature, and solar radiation. Rainfall is responsible for the water replenishment in the soil and used as a solvent for the insertion of gases, minerals and other solutes to be carried by the plant structure and is also of relevance for the heat maintenance and distribution (EMBRAPA, 2013b). Temperature, when below the ideal, can directly affect the plant metabolism and result in reduced growth or induce crop damages. In the context of solar radiation, which, despite being the main source of energy for the plants development, in a high incidence may have negative effects when simultaneously with water stress and high temperature (Ferrante & Mariani, 2018). Knowing the behavior of agrometeorological variables in the soybean crop cycle allows us to understand their influence on crop estimates, and this can be done using areas spatial statistics techniques. As explained by Araujo et al. (2014), areas spatial statistics techniques aim to describe and analyze the behavior of phenomena that occur in space to understand the spatial distribution of data, patterns of association, or different spatial regimes, through aggregated data (polygons) with geographic references (Seffrin, 2017).

Among the spatial statistical analysis, autocorrelation and spatial correlation stand out (Câmara et al., 2004). Spatial autocorrelation and spatial correlation can be measured on a global scale, which allow analyzes of data spatial associations in a region as a whole and on a local scale allow statistical inference for each location in relation to the average of a previously defined neighborhood. The most used methods to identify the autocorrelation structure and spatial correlation that best describe the data are the Global Moran's index and the Moran's Bivariate index, respectively (Simões et al., 2016). The Global Moran's index is a weighted autocorrelation coefficient that defines if the nearby areas are more similar to that expected by a

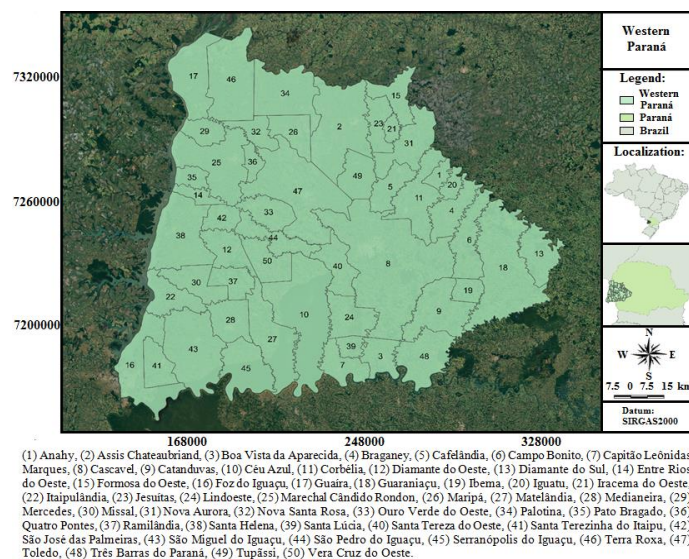
random distribution (Friche et al., 2006; Anselin, 2008; Braga et al., 2016). Seffrin (2017) used the Global Moran's Index to analyze the distribution of maize yield in the state of Paraná - Brazil and found in the harvest years 2011/2012, 2012/2013 and 2013/2014 a strong spatial autocorrelation among the municipalities of the state. To Galeano et al. (2019), the Global Moran's indices showed that the average coffee yield in the municipalities of the state of Espírito Santo - Brazil are strongly autocorrelated in the years 2011 to 2016, both on a global and local scale.

The objective of this paper is to analyze the spatial autocorrelation of soybean yield and its bivariate spatial correlation with the agrometeorological variables rainfall, mean temperature, and mean global solar radiation in the 2014/2015, 2015/2016, and 2016/2017 crop years in the West of the State of Paraná – Brazil.

2. Methodology

The study area comprises the mesoregion of the West of the State of Paraná (Figure 1), located in the Southern region of Brazil. The West of Paraná is composed of 50 municipalities that total a territorial area of 22,865 km² and a census population of 1219.558 inhabitants (IPARDES, 2018). Located in the third Paraná state Plateau, altitudes in the Western mesoregion of Paraná range from 200 to 850 m. the predominant pedological units are Red Latosols, Red Nitosol and Regolithic Neosols. The region has a semideciduous seasonal vegetation, designed by double tropical and subtropical seasons. The climate is classified according to Koeppen (1948) as a subtropical climate – Cfa, with an average temperature in the coldest month lower than 18°C, and in the hottest month higher than 22°C. The annual average temperature ranges from 21 to 23°C, being higher at lower altitudes (EMBRAPA, 2013a). The average annual rainfall ranges from 1,600 to 2,000 mm, in a way that the rainy period is concentrated between December and February, with variations from 400 to 600 mm, and the driest period from June to August, with levels from 250 to 450 mm (Zanão-Júnior et al., 2015).

Figure 1 - Western Paraná mesoregion study area.



Source: Authors (2022).

The acquisition of soybean yield variable occurred through the Department of Rural Economy (DERAL) databases and the Board of Agriculture and Supply (SEAB). The soybean (t) and area (ha) production data were obtained for the 50

municipalities that comprise the Western mesoregion of Paraná. Thus, by dividing soybean (t) production by area (ha), soybean yield (t ha⁻¹) was obtained for the 2014/2015, 2015/2016 and 2016/2017 harvests.

Agrometeorological variables were obtained from the ECMWF (European Centre for Medium-Range Weather Forecasts) model available on the Joint Research Center – JRC portal (Ecmwf, 2018). The ten-year climatic data are provided with global coverage and resized in a spatial resolution of longitude and latitude degrees, in a grid of 0.25 degrees (approximately 25 x 25 km) in raster format. The agrometeorological variables include rainfall (mm), mean temperature (°C) and average global solar radiation (kJ m⁻²) for 2014, 2015, 2016 and 2017.

For the selection of agrometeorological data, due to the same being in decline, the soybean harvest period was determined from October 1st of the initial year beginning to March 1st of the final year. Thus, for the variable rainfall, the sum of the data was obtained in the period in question, while for the variables mean temperature and average global solar radiation, the arithmetic mean was calculated for the period adopted.

Due to the agrometeorological data (rasters) covering South America, and the soybean yield data (table) at municipal level for Paraná, there is a need to standardize the data scale.

For this, the software QGIS 3.4 (Qgis Development Team, 2018) was used to structure the data. A shapefile was created that joins the data of the boundaries of the municipalities of the Western Paraná State (vector), obtained from IBGE (2016), together with the information soybean yield (table) by municipality.

The agrometeorological data, raster format, were submitted to a process of clipping at the Western mesoregion boundary, and then were transformed to shapefile data by means of the raster vectoring process.

With the data cut off at the Western level of Paraná and vectorized, the PostgreSQL extension was used to elaborate a Structured Query Language – SQL, which resulted in the selection of agrometeorological data (polygons) that intersect the municipalities boundary of the study, in order to make the sum of the polygons for the variable rainfall and arithmetic mean for the variables mean temperature and average global solar radiation. Thus, agrometeorological information was obtained for each municipality in the harvest years.

The data spatial exploratory analysis was performed in the R Project (R Core Team, 2014) and QGIS 3.4 (Cutts & Graser, 2018) software, in order to describe descriptive statistics and distributions of agrometeorological and soybean yield variables, as well as illustration in graphs and maps.

The spatial autocorrelation levels among the data from areas of the municipalities of Western Paraná State were estimated by **I** the Global Moran's Index (equation 1) and Local Indicators of Spatial Association Index (LISA) (equation 2) (Anselin et al., 2008), performed using the OpenGeoda 1.12 software (Anselin et al., 2018).

$$I = \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (z_i - \bar{z})(z_j - \bar{z})}{\sum_{i=1}^n (z_i - \bar{z})^2}, \quad (1)$$

where, **n** is the number of observations (**n** polygons); **w_{ij}** represents the elements of the spatial proximity matrix **W**; **z_i** is the value of the attribute considered in the area **i**, **z̄** is the average value of the attribute in the study region. In this case, the correlation is computed for the first-order neighbors in space, as established by the weights **w_{ij}**.

$$I_i = \frac{x_i - \mu}{\sigma_0^2} \sum_{j=1}^n w_{ij} (x_j - \bar{x}), \quad i = 1, \dots, n, \quad (2)$$

where I_i , refers to the element local index i ; x_i is the value of the attribute observed in the area i ; x_j is the value of the attribute observed in the area j ; \bar{x} it is the average of observations among the regions; w_{ij} the average value of deviations of the neighboring elements of i ; and σ_0^2 the distribution variance of the values of deviations.

The Global Moran's statistics applied as a measure of spatial dependence makes it possible to perform the hypothesis test, that is, to infer the existence or not of spatial autocorrelation. Spatial independence, called null hypothesis, is given when its value is zero. However, positive values express direct correlation and negative values an inverse correlation (Braga et al., 2016).

In the present study, the finding of positive spatial autocorrelation is interpreted as a similarity among the municipalities (polygons); therefore, municipalities that present high/low soybean yield tend to be neighbors of municipalities that also have high/low yield. On the other hand, the presence of negative autocorrelation points to a dissimilarity among the municipalities, in a way that the municipalities that reveal a high/low soybean yield may be surrounded by municipalities with low/high yield.

The identification of local spatial association patterns calculated by LISA index (I_i) follows the criteria that each municipality must have a LISA indicator and the sum of the indicators of all municipalities must match the global index (Anselin, 1995). LISA index allows to point out the significant spatial clusters of similar values around the municipality, in which positive values denote I_i the presence of spatial clusters with similar values, being high or low. For negative values of I_i , it is interpreted that spatial groupings have different values between the regions and their neighbors (Santos et al., 2016).

The spatial correlation among soybean yield, rainfall, mean temperature and average global solar radiation variables is analyzed using the Moran's Bivariate index I_{xy} (equation 3). It provided a spatial correlation index between two georeferenced variables, where each was obtained in n the municipalities (Almeida, 2012).

$$I_{xy} = \frac{\sum_{i=1}^n \sum_{j=1}^n u_i z_j w_{ij}}{S_0 \sqrt{S_u^2 S_z^2}}, \quad (3)$$

being that I_{xy} , it indicates the variable bivariate index xy in relation to y ; n it is the number of observations (polygons); $z_j = (x_j - \bar{x})$ and $u_i = (y_i - \bar{y})$ are the values observed centered on the means of the variables x ; and $y w_{ij}$ represents the m element of the spatial proximity matrix W ; $S_u^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n}$ and $S_z^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$ represent the respective variances x and y (Anselin, 2003).

For the calculation of the Global Moran's index, LISA index I_i and Moran's bivariate index I_{xy} , it is necessary to establish the spatial proximity matrix (W) beforehand, to represent a measure of proximity among the municipalities (polygons) (Leite, 2016). Thus, the W matrix was calculated from three criteria of contiguity (queen), Euclidean distance and closest neighbors.

The contiguity criterion is given in accordance with the concept that two places are neighbors when they share a common physical border. Among the most widely used conventions, Queen stands out, in which all the units that divide any kind of border with the unit in question are considered neighbors (Vidigal et al., 2018). The Euclidean distance considers the distances among the centroids (geometric center) of the areas, weighted in relation to the Euclidean distance (d), i.e., $w_{ij} = d_{ij}$, so that the distance is established by the researcher (Campos & Chagas, 2017). The closest neighbor criterion makes a comparison of the average distance among the closest neighbors with a set of points with a defined pattern (Santos et al.,

2016). To complement the spatial relationships analysis of agrometeorological variables and soybean yield, moran's map and lisa map were elaborated, which allow the indices interpretation using georeferenced illustrations.

3. Results and Discussion

The Western Paraná State mesoregion stands out for presenting one of the highest soybean yields in the state (SEAB, 2017). Table 1 presents the descriptive statistics for the study variables: Soybean yield (Prod), rainfall (Prec), mean global solar radiation (RadM) and mean temperature (TempM).

In the 2014/2015 harvest year, the average soybean yield in the Western Paraná State cities was 3.42 t ha⁻¹, in agrometeorological conditions of 3.030 mm of rainfall, 222,556 kJ m⁻² of global solar radiation incidence at an average temperature of 24.64 °C. In the southern region of Brazil, climatic conditions were adequate for soybean development, however, in the western region of Paraná, irregular precipitation had an impact on the crop (CONAB, 2015a). The highest values of standard deviation and variance are seen in the variables Prec. (1,183.43 mm) and RadM (1,778.87 kJ m⁻²): This variation is explained by the occurrence of discrepant precipitation in the municipality of Cascavel (6,197 mm).

Table 1 - Descriptive statistics of the variables in study in the 2014/2015 crop year.

Statistics	Prod	Prec	RadM	TempM
Minimum	2.231	1,013	219,925	22.92
1 st Quartile	3.131	2,335	221,194	24.11
Median	3.421	2,870	222,012	24.66
Mean	3.418	3,030	222,556	24.64
3 rd Quartile	3.744	3,530	223,575	25.32
Maximum	4.14	6,197	227,136	26.01
Standard deviation	0.427	1,183.43	1,778.87	0.833
CV(%)	12.48	39.06	0.799	3.38

Prod: soybean yield (t ha⁻¹); TempM: average temperature (°C); RadM: average global solar radiation (kJ m⁻²); Prec: rainfall (mm); CV: coefficient of variation. Source: Authors (2022).

In the 2015/2016 harvest year (Table 2), it was observed that the average soybean yield (3.43 t ha⁻¹) and the average temperature (24.24 °C) were similar to the previous harvest. However, there was a higher rainfall, with an average of 6,493 mm, which resulted in excess humidity, and a lower incidence of average solar radiation of 186.152 kJ m⁻², which reduced the luminosity. These climatic adversities are explained by the presence of the phenomenon "EL Niño", which caused the main factors of yield reduction in view of the estimated one for this harvest year (SEAB, 2016).

Table 2 - Descriptive statistics of the variables in study in the 2015/2016 crop year.

Statistics	Prod	Prec	RadM	TempM
Minimum	2.900	2,958	182,286	22.63
1 st Quartile	3.310	4,860	185,141	23.69
Median	3.410	6,464	186,530	24.27
Mean	3.434	6,493	186,152	24.24
3 rd Quartile	3.594	7,634	187,651	24.86
Maximum	4.165	13,078	189,765	25.56
Standard deviation	0.235	2,432.65	1,999.62	0.809
CV(%)	6.85	37.47	1.07	3.34

Prod: soybean yield (t ha⁻¹); TempM: average temperature (°C); RadM: average global solar radiation (kJ m⁻²); Prec: rainfall (mm); CV: coefficient of variation. Source: Authors (2022).

In 2016/2017, the soybean yield showed the highest average among the studied harvests, 3.87 t ha⁻¹ (Table 3), and this harvest year obtained values that exceeded the last record yield in Paraná, 3.42 t ha⁻¹ in 2011 (SEAB, 2017).

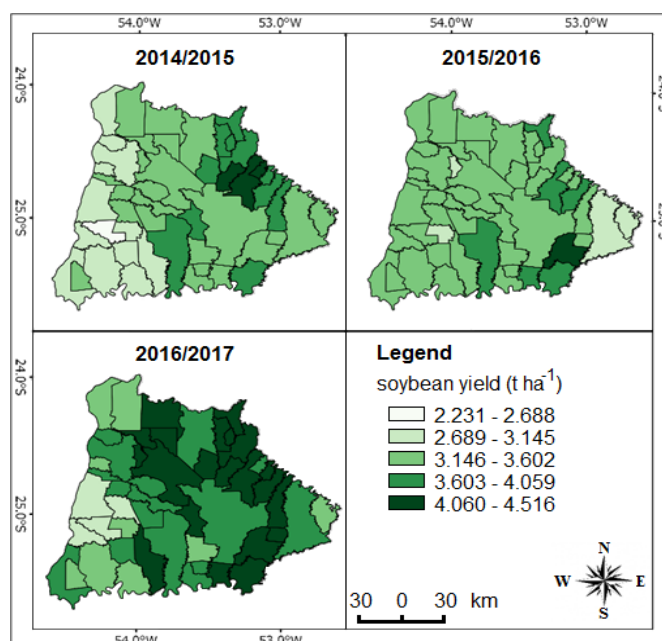
Table 3 - Descriptive statistics of the variables in study in the 2016/2017 crop year.

Statistics	Prod	Prec	RadM	TempM
Minimum	3.000	1,558	217,602	22.16
1 st Quartile	3.625	2,689	219,993	23.23
Median	3.958	3,406	222,145	23.88
Mean	3.874	3,421	222,747	23.84
3 rd Quartile	4.175	4,020	225,821	24.52
Maximum	4,515	6,575	228,828	25.27
Standard deviation	0.369	1,256.19	3,275.27	0.883
CV(%)	9.54	36.72	1.47	3.70

Prod: soybean yield (t ha⁻¹); TempM: average temperature (°C); RadM: average global solar radiation (kJ m⁻²);
 Prec: rainfall (mm); CV: coefficient of variation. Source: Authors (2022).

This year-harvest's voluminous and quality yield had virtually perfect agrometeorological conditions for the crop development, with regular rains that totaled an average of 3,421 mm of precipitation, average temperature of 23.84 °C, and 222,747 kJ m⁻² of global solar radiation, which resulted in greater luminosity in the soybean development period, which benefited production. Figure 2 illustrates the spatial distribution of soybean yield (t ha⁻¹) in the municipalities of Western Paraná State. It is observed that the highest soybean yield class, 4,060 – 4,516 t ha⁻¹, occurred in 42% of the municipalities in the 2016/2017 harvest year, while in the 2014/2015 and 2015/2016 harvest years, it occurred in 6% and 2% of the municipalities, respectively. However, most municipalities in 2014/2015 (46%) and in 2015/2016 (78%) maintained yield between 3.146 – 3.602 t ha⁻¹. It was in 2014/2015, in the city of Missal, that the lowest soybean yield was found among the harvest years under study, with a yield between 2,231 – 2,688 t ha⁻¹.

Figure 2 - Spatial distribution of soybean yield (t ha⁻¹) in the municipalities of Western Paraná in the 2014/2015, 2015/2016 and 2016/2017 harvest years.

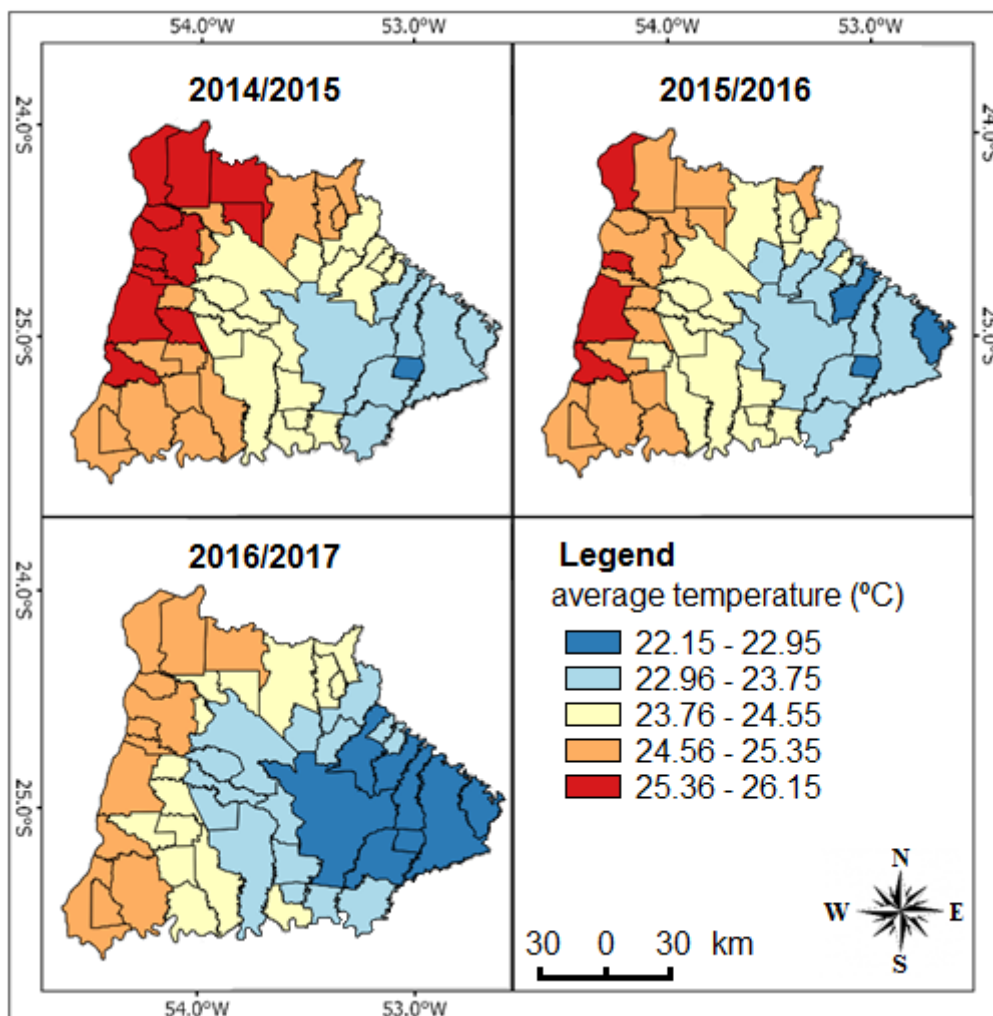


Source: Authors (2022).

Figure 3 shows the spatialization of the agrometeorological variable mean temperature in the three-year-harvest period. It is noted that the highest temperatures (25.36 – 26.15 °C) are at the west of the Western Paraná State mesoregion, and the lowest temperatures (22.15 – 22.95 °C) are at the east. Regarding the soybean yield spatial distribution, the highest yields were observed, where the temperature ranged from 22.15 to 24.55 °C.

In the 2014/2015 harvest year, the highest temperatures occurred in the study region, which resulted in lower soybean yield. The high temperatures in this harvest year reflect the occurrence of high temperatures in the Pacific Ocean, close to the Equator (Conab, 2015b), causing most of the country to have temperatures above normal (Conab, 2014). On the contrary the 2016/2017 harvest year had milder temperatures and higher soybean yield. However, a small temperature variability between 22.15 – 26.15 °C is observed in the study area and also within the ideal range from 20 to 30 °C for soybean development, as stated by Embrapa, (2013b).

Figure 3 - Spatial distribution of average temperature (°C) in the municipalities of Western Paraná in the 2014/2015, 2015/2016 and 2016/2017 harvest years.



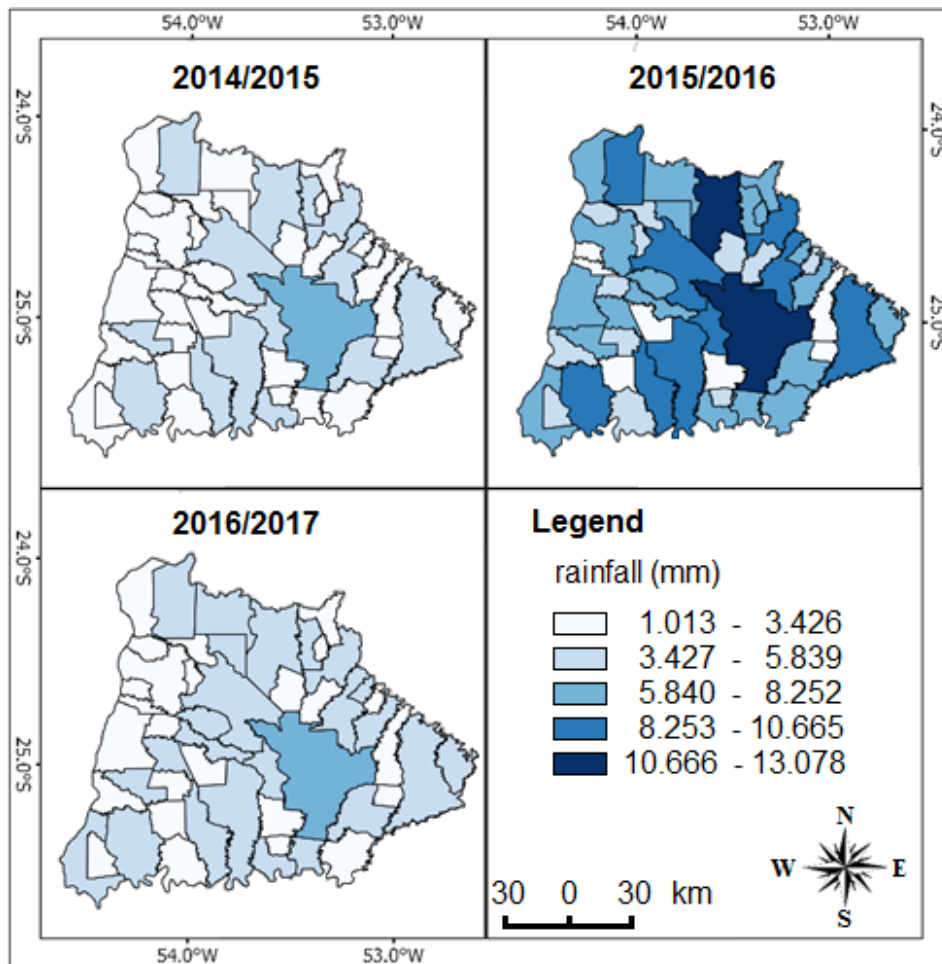
Source: Authors (2022).

Figure 4 shows the rainfall spatial analysis. It is observed that this agrometeorological variable has a greater spatial variability of distribution in the municipalities of Western Paraná State. This spatial variability is explained by occurrences of convective rainfalls (or “summer rains”), that is, high rainfall volume in small areas. An example in this scenario of

convictional rainfalls occurred in the 2015/2016 harvest, in which the municipality of Cascavel accumulated between 10,666 – 13,078 mm of rainfall; however, the neighboring municipalities Lindoeste, Campo Bonito and Ibema indicated a rainfall of 1,013 – 3,426 mm.

In 2014/2015 and 2016/2017 harvest years, the occurrence of rainfalls was between 1,013 and 8,252 mm. An analysis of the rainfall distribution with the soybean yield distribution shows that the highest yields occurred in municipalities with rainfall between 1,013 – 5,839 mm; however, the form of irregular or regular occurrences of these rainfalls should be considered. The highest amount of rain observed in the 2015/2016 harvest year was an event perceived throughout Brazil. Excess water at critical moments in the development of soybeans in this crop year harmed Brazilian production, generating losses and causing a significant number of requests for agricultural insurance (3756) when compared to requests in crop years 2014/2015 (282) and 2016/2017 (260) (Embrapa, 2019).

Figure 4 - Figure Spatial distribution of rainfall in the municipalities of Western Paraná in the 2014/2015, 2015/2016 and 2016/2017 harvest years.



Source: Authors (2022).

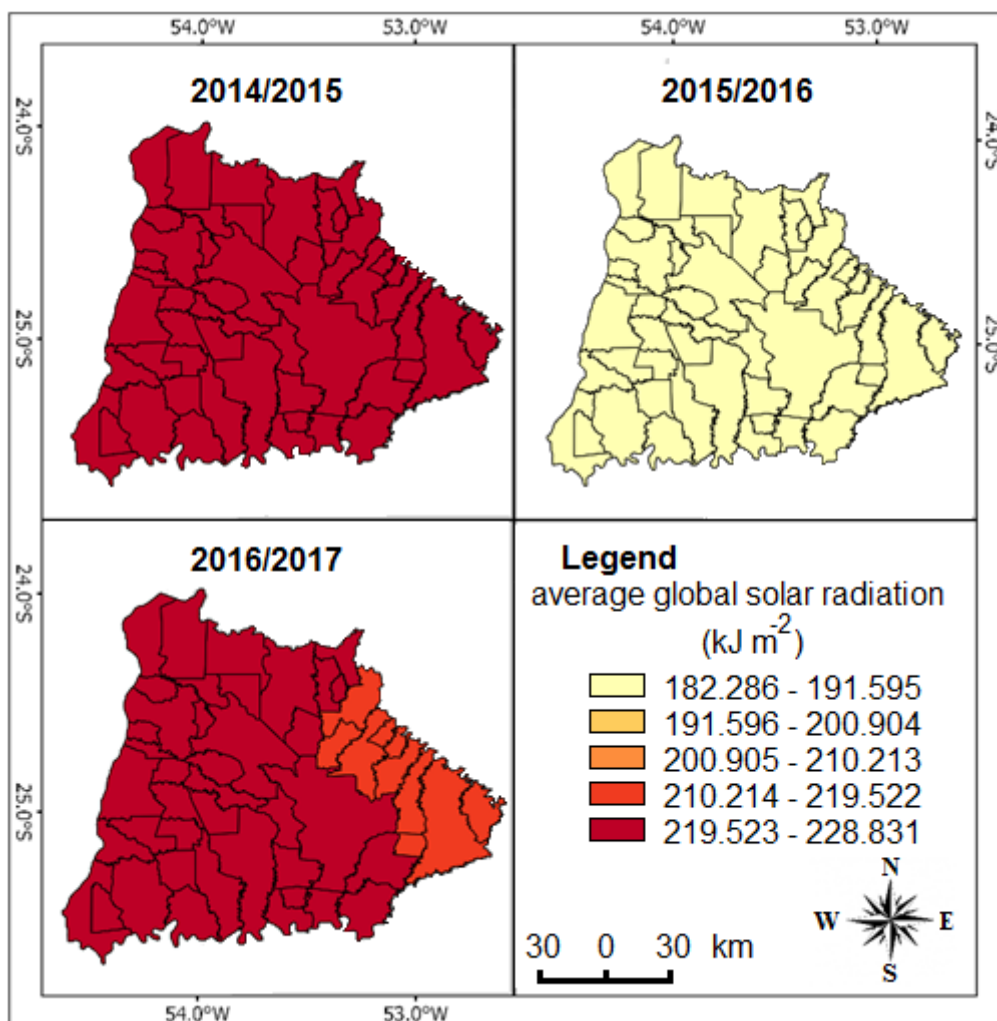
The spatial distribution of the average global solar radiation is shown in Figure 5 and it is observed that the average global solar radiation is the variable that demonstrates the least spatial variability. The 2015/2016 harvest is highlighted, because it indicates the lowest incidence of global solar radiation, between 182,286 – 191,595 KJ m⁻² in all the municipalities in the mesoregion. Contrarily, in 2014/2015, the highest incidences that exceeded 219,523 KJ m⁻² occurred in the municipalities of Western Paraná State.

For the spatial autocorrelation analysis through the calculation of I the global and local Moran's Index, the previous creation of spatial proximity matrices (W) is necessary, which dictate the neighborhood rules for polygons, that is, for each municipality of the Western Paraná state mesoregion. Thus, in the OpenGeoda software, W proximity matrices were created in the following criteria: contiguity (queen), distance between centroids and closest neighbors.

In the queen contiguity convention, polygons that divide any type of boundary are neighbors; thus, the value of 1 is assigned to neighboring polygons and the value of 0 for non-boundary-sharing polygons.

The W matrix in the criterion of distance among the centroids weights the boundaries in relation to the Euclidean distance among the centroids of the polygons (municipalities) under study, that is, the one that obtains the minimum distance established in the calculation is considered a neighbor. In the matrix elaborated, the software resulted in the minimum distance found that ensures to each place at least one neighbor was 23 km. Therefore, municipalities with a distance of less than 23 km are classified as neighbors, in contrast, municipalities that exceed this distance are not considered neighbors.

Figure 5 - Spatial distribution of average solar radiation (kJ m^{-2}) in the municipalities of Western Paraná in the 2014/2015, 2015/2016 and 2016/2017 harvest years



Source: Authors (2022).

The closest neighbor criterion elaborates a **W** matrix to ensure that all the polygons obtain the same number of neighbors. In the study, it was considered that all the municipalities have as neighbors the four closest borders.

Table 4 shows the spatial autocorrelation values by **I** the Global Moran's index and the respective significance levels of the variables soybean yield (Prod), mean temperature (TempM), average global solar radiation (RadM) and rainfall (Prec). The variables were analyzed according to the criteria contiguity (queen), distance among centroids and closest neighbors.

The descriptive levels (p-value) from tests with 9,999 random permutations presented in brackets are less than 0.1%, which allows to affirm that the variables, except for the variable rainfall (Prec), have significant spatial autocorrelation at 1% probability level. Similar spatial autocorrelation values are observed among the three neighborhood criteria. The criterion distance among the centroids that indicated the highest self-occurrence values for the variables in all years/harvest is highlighted, on the contrary, from the closest neighboring criterion, which presented the lowest values.

Soybean yield has a higher spatial autocorrelation in the 2014/2015 harvest year, both in the contiguity criterion (queen) and in the distance criteria among centroids and the closest neighbor, being that in the distance criterion among the centroids it reached the value **I = 0.6875**. The years 2015/2016 showed the lowest self-correlation values, with the closest neighbors criterion with the value **I = 0.243**. However, all the years/harvest presented a moderate spatial autocorrelation index, which allows to conclude that the municipalities of Western Paraná state are surrounded by municipalities with similar soybean yield. Grzegozewski (2017) also detected **I** a positive autocorrelation (**I** > 0.73), significant at the 5% probability level, in the soybean yield distribution in the 399 municipalities of the state of Paraná in 2011/2012, that is, it was noted that there are municipalities with high or low soybean yield surrounded by municipalities in the same scenario.

Table 4 - Global Moran I index for the soybean yield and agrometeorological variables.

Variables	Contiguity (queen)			Distance between centroids			Nearest neighbor		
	2014/2015	2015/2016	2016/2017	2014/2015	2015/2016	2016/2017	2014/2015	2015/2016	2016/2017
Prod	0.6363 (0.001) **	0.292 (0.001) **	0.5956 (0.001) **	0.6875 (0.001) **	0.3632 (0.003) **	0.6569 (0.001) **	0.624 (0.001) **	0.2437 (0.005) **	0.5669 (0.001) **
TempM	0.857 (0.001) **	0.8423 (0.001) **	0.8469 (0.001) **	0.8904 (0.001) **	0.8662 (0.001) **	0.869 (0.001) **	0.8105 (0.001) **	0.7776 (0.001) **	0.785 (0.001) **
RadM	0.7001 (0.001) **	0.7453 (0.001) **	0.8659 (0.001) **	0.7833 (0.001) **	0.8991 (0.001) **	0.9327 (0.001) **	0.7042 (0.001) **	0.8471 (0.001) **	0.8875 (0.001) **
Prec	-0.0371 (0.414) NS	-0.0603 (0.325) NS	-0.1162 (0.152) NS	0.1123 (0.183) NS	0.1109 (0.172) NS	0.0468 (0.327) NS	0.04 (0.265) NS	0.0061 (0.385) NS	-0.0442 (0.396) NS

Prod: soybean yield (t ha⁻¹); TempM: average temperature (°C); RadM: average global solar radiation (kJ m⁻²); Prec: rainfall (mm); Corresponding p-values in parentheses; ** Statistically significant at 0.01 probability. Source: Authors (2022).

Lopes et al. (2017) state that, as a result of the strong spatial correlation of soybean yield, there is an increase in susceptibility to the contagious effect of natural phenomena, such as frost, pests, droughts, among other circumstances resulting in the drop in production and yield. The highest spatial autocorrelation values are presented by the agrometeorological variables mean global solar radiation (**I = 0.9327**), followed by the mean temperature (**I = 0.8904**). The

indices of these variables indicate the lowest spatial variability, i.e., they are more homogeneous distributed among the municipalities of the study, which is an expected characteristic of these variables, due to the fact that nearby regions demonstrate similar agrometeorological characteristics. Rainfall indicated the lowest spatial autocorrelation values. This occurs due to greater spatial variability in the occurrence of rainfalls. Araujo et al. (2014) also obtained the lowest values of spatial autocorrelation in rainfall for the municipalities of Western Paraná in the 2005/2006, 2006/2007 and 2007/2008 harvest years.

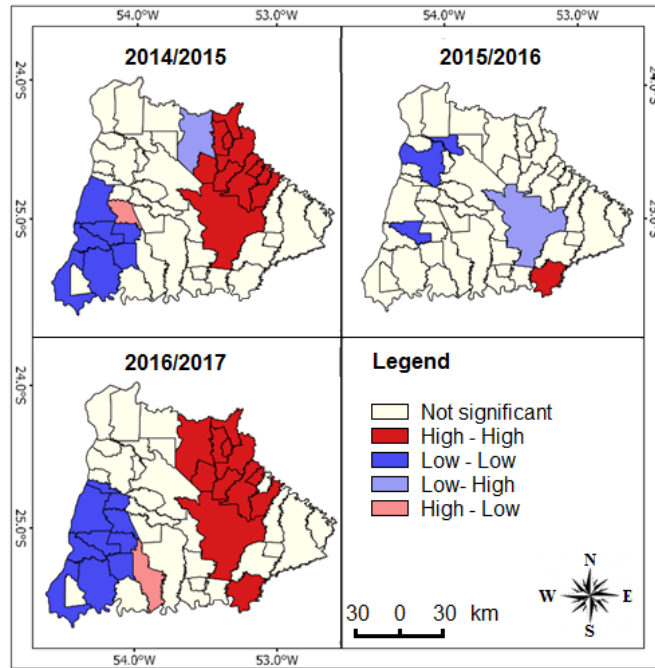
Still in rainfall, negative indices were found in the contiguous neighborhood (queen) criterion, the presence of negative autocorrelation indicates a dissimilarity, that is, municipalities of the Western Paraná state with low values are close to municipalities with high values, or vice-versa. In this case, the weighted average is higher for neighbors than the value measured at the local itself. These occurrences of rainfalls are triggered by isolated events in the municipalities.

Due to the large number of areas in the Western Paraná State, 50 municipalities, the Local Indicators of Spatial Association Index (LISA) was calculated, which allows a better groupings analysis, where the queen's contiguity was used as the neighborhood criterion. The groupings of the study region are shown in Figure 6.

In the 2014/2015 harvest, a high soybean yield cluster was observed at a level of 1% significance, composed of 11 municipalities. In 2015/2016, only Três Barras do Paraná municipality presented high yield and is also surrounded by municipalities with high yield, as a result of the intense irregular rainfall resulting from the “El Niño” phenomenon that reduced soybean yield in this harvest. In the 2016/2017 harvest, climate conditions were favorable for soybean development, such as regular rainfall and greater indecency of luminosity; consequently, high yield returned in a larger number of municipalities. In 2015/2016 there was an expressive increase in municipalities with non-significant soybean yield (45 municipalities), and low yield is seen in the municipalities Marechal Candido Rondon, Missal and Nova Santa Rosa. However, in 2016/2017 there were 10 municipalities with low yield. Still in Figure 6, dissimilarities clusters are observed, that is, municipalities with low soybean yield with high soybeans yield neighbors (low – high), and vice-versa (high – low).

In the 2014/2015 harvest, Assis Chateaubriand is a municipality with low soybean yield, but with high yield neighbors, which also occurs with Cascavel in the 2015/2016 harvest. It is noteworthy that these municipalities accumulated the most voluminous occurrences of rainfall in the region. However, in the 2014/2015 harvest, the city of Diamante d'Oeste indicated a high soybean yield and is surrounded by neighbors with low yield, the same fact that occurs in Matelândia in the 2016/2017 harvest.

Figure 6 – LISA cluster map for soybean yield.

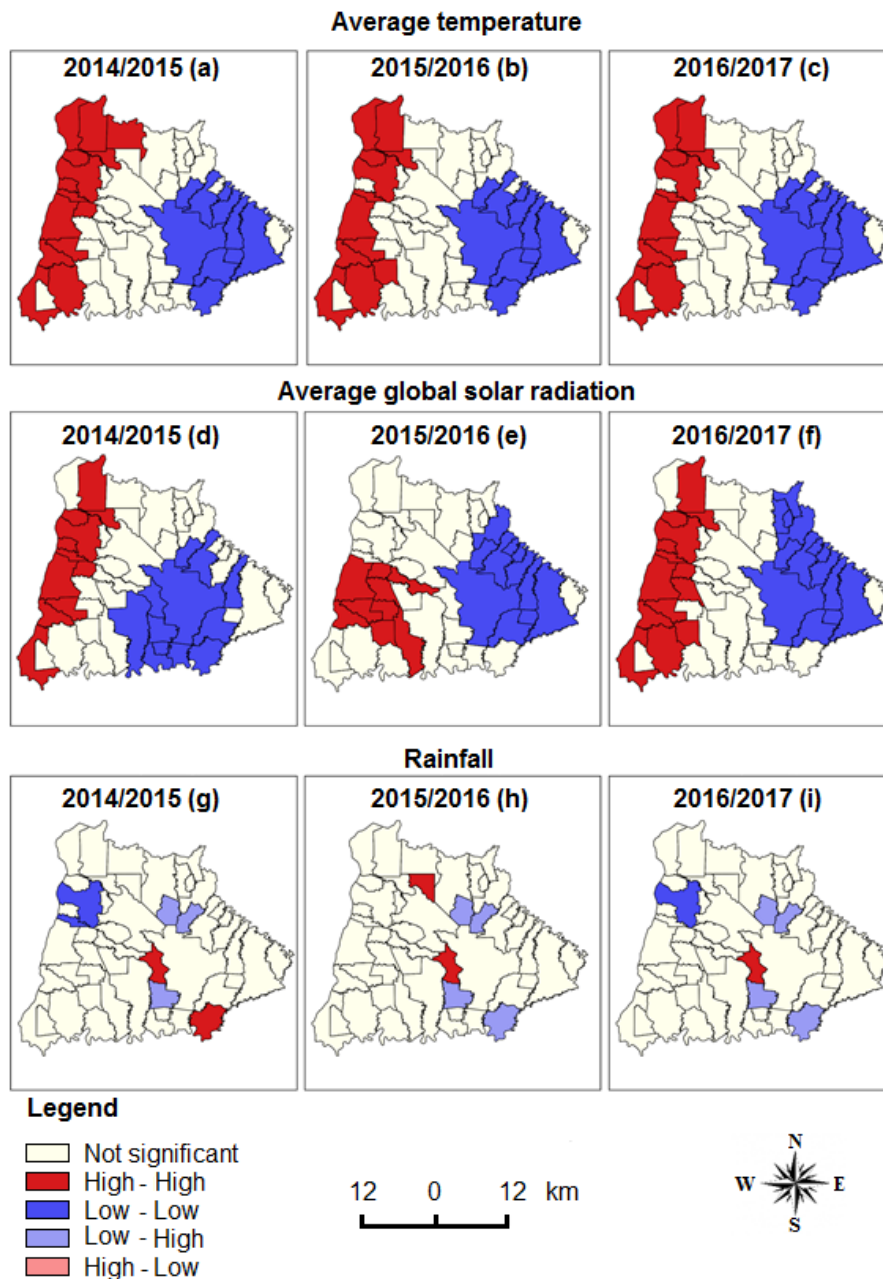


Source: Authors (2022).

The LISA cluster map for agrometeorological variables in each harvest year are shown in Figure 7. The municipalities of Western Paraná show that for the average temperature (TempM) and average global solar radiation (RadM) there are clusters of similarities, High-High concentrates in the Eastern region and Low-Low clusters in the western region of the territory under study. It is observed that the rainfall distribution (Prec) is the one that has the greatest variability and the highest number of municipalities that are not significant at the 5% probability level.

The analysis of **I** Local Moran's index (LISA) allowed to conclude that the agrometeorological variables mean temperature, average global solar radiation and rainfall have little variation between the 2014/2015, 2015/2016 and 2016/2017 harvests in the municipalities of Western Paraná state.

Figure 7 - LISA cluster map for the agrometeorological variables.



Source: Authors (2022).

The spatial correlations between the agrometeorological variables and soybean yield were obtained by the Moran's Bivariate index and their values are shown in Table 5. The variables Prod and TempM and the variables Prod and RadM indicated a negative spatial correlation, i.e. inversely proportional, and significant at a level of 1% of significance in all the harvest years under analysis, this leads to the conclusion that high temperature and solar radiation values indicate low soybean yield. However, the variable Prec presented a positive spatial correlation compared to Prod in the three years, a fact that reflects a directly proportional relationship, and therefore, low rainfall occurrences cause low soybean yield.

Despite the low correlation, the values between Prod and Prec were significant at 1 and 5% probability levels. The highest spatial correlation indices are between the variables TempM and RadM, expected values, since the temperature tends to be higher as the solar radiation increases.

The correlation of the variables TempM and RadM with Prec indicated negative and significant values in the three harvests, allowing to conclude that high rainfalls occurrences reflect in low temperatures and solar radiation. All the indices were significant, which confirms the influence of the agrometeorological variables chosen in soybean yield.

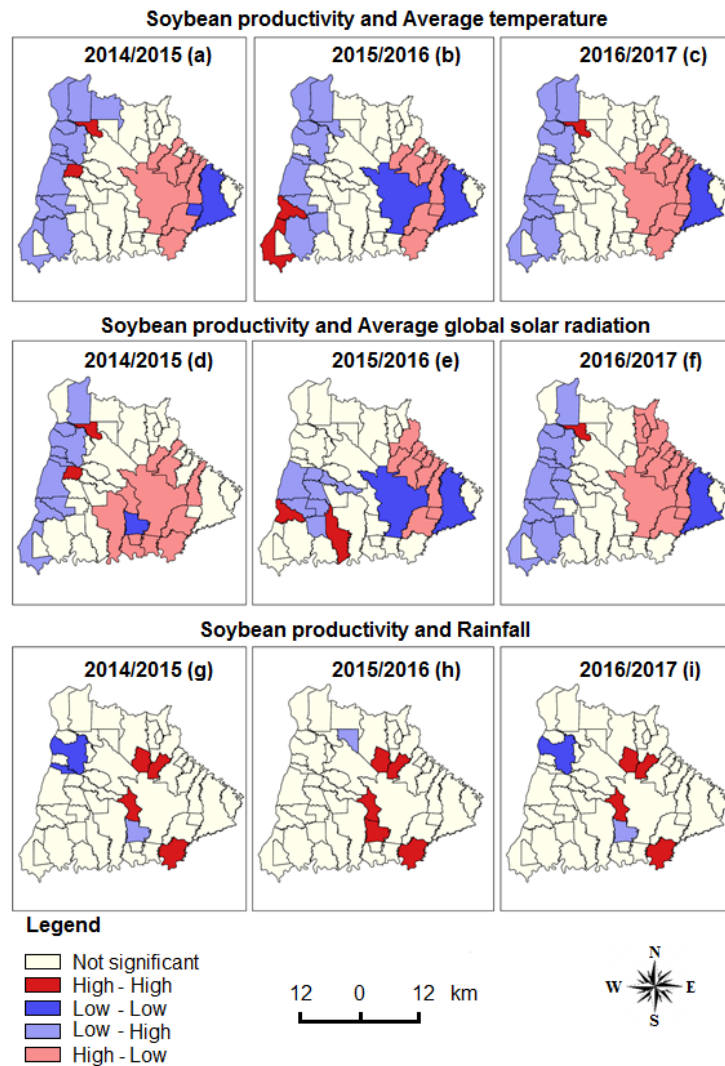
Table 5 - Bivariate Global Moran I index for the soybean yield and agrometeorological variables

Variables	Prod			TempM			RadM		
	2014 /2015	2015 /2016	2016 /2017	2014 /2015	2015 /2016	2016 /2017	2014 /2015	2015 /2016	2016 /2017
TempM	-0.4787 (0.001) **	-0.2742 (0.001) **	-0.4759 (0.001) **						
RadM	-0.4956 (0.001) **	-0.1758 (0.001) **	-0.5605 (0.001) **	0.6401 (0.001) **	0.614 (0.001) **	0.7688 (0.001) **			
Prec	0.2693 (0.001) **	0.1336 (0.037) *	0.174 (0.012) *	-0.2698 (0.001) **	-0.2164 (0.004) **	-0.1622 (0.014) *	-0.2996 (0.001) **	-0.1974 (0.005) **	-0.2022 (0.005) **

Prod: soybean yield ($t\ ha^{-1}$); TempM: average temperature ($^{\circ}C$); RadM: average global solar radiation ($kJ\ m^{-2}$); Prec: rainfall (mm); Corresponding p-values in parentheses; ** Statistically significant at 0.01 probability; *Statistically significant at 0.05 probability. Source: Authors (2022).

The relationship between the agrometeorological variables of the study and the soybean yield in the municipalities of Western Paraná state were also analyzed in local coverage through the BILISA cluster map (Figure 8).

Figure 8 – BILISA cluster map of soybean yield compared to the agrometeorological variables.



Source: Authors (2022).

The relationship between soybean yield and mean temperature (Figure 8a) shows a greater number of municipalities with dissimilar correlation, i.e., when a high/low soybean yield was obtained, low/high mean temperature occurred.

A similar behavior was observed in the 2015/2016 and 2016/2017 harvests, where most of the municipalities indicated a correlation of dissimilarity. Similarly, to the average temperature variable, the average global solar radiation indicated in greater quantity municipalities with a dissimilar correlation with soybean yield and little variation among the harvests (Figures 8d, 8e and 8f). Rainfall showed few municipalities with significant correlations with soybean yield (Figures 8g, 8h and 8i).

4. Conclusion

A cluster of municipalities with low productivity was identified in the west and a cluster with high productivity was identified in the east of the monitored region. The results allowed us to conclude that soybean yield is self-correlated in the cities of West Paraná state, globally and locally. When soybean yield was correlated with agrometeorological variables, a negative correlation was observed with mean temperature and mean global solar radiation, and a positive correlation was observed with rainfall. The significance of the indices confirmed the influence of agrometeorological variables on soybean

yield. Thus, it is evident that the analysis of agrometeorological information using methods of spatial statistics of areas can facilitate strategic decisions aimed at increasing and sustaining agricultural production.

Acknowledgments

Council for Scientific and Technological Development (CNPq); Coordination for the Improvement of Higher Education Personnel (CAPES); Post-Graduate Program in Agricultural Engineering (PGEAGRI); Federal Technological University of Paraná (UTFPR) and Spatial Statistics Laboratory (LEE-UNIOESTE).

References

- Almeida, E. S. (2012). *Econometria Espacial Aplicada*. Alínea.
- Anselin, L. (1995). Local indicators of spatial association–LISA. *Geographical Analysis*, 27(2), 93-115. doi: <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- Anselin, L. (2003). *GeoDa 1.4.6 User's Guide*. University of Illinois.
- Anselin, L.; Le Gallo, J.; Jayet, H. (2008). *Spatial panel econometrics*. In: *The econometrics of panel data*. Springer, 625-660.
- Anselin, L.; Syabri, I.; Kho, Y. (2018). GeoDa Version 1.12. <https://geodacenter.github.io>
- Araújo, E. C. de; Uribe-Opazo, M. A.; Johann, J. A. (2014). Modelo de regressão espacial para estimativa da produtividade da soja associada a variáveis agrometeorológicas na região oeste do estado do Paraná. *Engenharia Agrícola*, 34, (2), 286-299. doi:<https://doi.org/10.1590/S0100-69162014000200010>.
- Araujo, M. A.; Souza, J. L. M. de; Tsukahara, R. Y. (2011). Modelos agro-meteorológicos na estimativa da produtividade da cultura da soja na região de Ponta Grossa, Estado do Paraná. *Acta Scientiarum. Agronomy*, 33(1), 23-31. doi: 10.4025/actasciagron.v33i1.5062.
- Barboza, E.N.; Caiana, C.R.A.; Neto, F.C.B. (2020). Rainfall analysis in the Central-South Region of Ceará: A study of the period (1980-2009). *Research, Society and Development*, 9(6), e18963304. doi: <http://dx.doi.org/10.33448/rsd-v9i6.3304>.
- Braga, G. B.; Marques, M. A. F.; Braga, B. B. (2016). Um diagnóstico da distribuição espacial da atenção primária à saúde no sudeste brasileiro. *Journal of Health Sciences*, 18(1), 37-47. doi:10.17921/2447-8938.2016v18n1p41-51.
- Câmara, G.; Carvalho, M. S.; Cruz, O. G.; Correa, V. (2004). *Análise espacial de áreas. Análise espacial de dados geográficos*. Embrapa, 157-182.
- Campos, R. B. A.; Chagas, A. L. S. (2017). Identificação e caracterização das subcentralidades de emprego no município de São Paulo. *Working Paper Series*, (33), 1-20. <https://ideas.repec.org/p/spa/wpaper/2017wpecon33.html>.
- Chan-Tack, A.M. (2014). The Case for Spatially-Sensitive Data: How Data Structures Affect Spatial Measurement and Substantive Theory. *Historical Social Research*, 39(2), 315-346. <https://www.ssoar.info/ssoar/handle/document/38487>.
- Cargnelutti Filho, A.; Matzenauer, R.; Maluf, J. R. T.; Radin, B. (2009). Variabilidade temporal e espacial da precisão das estimativas de elementos meteorológicos no Rio Grande do Sul. *Ciência Rural*, 39 (4), 962-970. <https://www.redalyc.org/pdf/331/33115802039.pdf>.
- Conab - Companhia Nacional de Abastecimento. (2019). Acompanhamento da safra brasileira. Soja em números, safra 2017/2018. <https://www.conab.gov.br/info-agro/safras>.
- Conab - Companhia Nacional de Abastecimento. (2014). Boletim de monitoramento agrícola. Cultivos de verão – Safra 2014/2015. 3 (20). Retrieved October 9, 2021, from <https://www.conab.gov.br/info-agro/safras/graos/monitoramento-agricola>.
- Conab - Companhia Nacional de Abastecimento. (2015a). Boletim de monitoramento agrícola. Cultivos de verão – Safra 2014/2015. 4 (1). <https://www.conab.gov.br/info-agro/safras/graos/monitoramento-agricola>.
- Conab - Companhia Nacional de Abastecimento. (2015b). Boletim de monitoramento agrícola. Cultivos de verão – Safra 2014/2015. 4 (2). <https://www.conab.gov.br/info-agro/safras/graos/monitoramento-agricola>.
- Cutts, A., Grasen, A. (2018). *Learn QGIS: Your step-by-step guide to the fundamental of QGIS 3.4*. Packt Publishing, 1-272.
- Dalposso, G. H.; Uribe-Opazo, M. A.; De Bastiani, F. (2021). Spatial-temporal analysis of soybean productivity Using Geostatistical Methods. *Journal of Agricultural Studies* 9(2), 283-303. https://econpapers.repec.org/article/mthjas888/v_3a9_3ay_3a2021_3ai_3a2_3ap_3a283-303.htm.
- Ecmwf - European Centre for Medium-Range Weather Forecasts. (2018). ECMWF 10-daily data. 2018. <https://www.ecmwf.int/en/forecasts/datasets/>.
- Embrapa – Empresa Brasileira de Pesquisa Agropecuária. (2019). Eventos climáticos adversos e seus impactos para as culturas de soja, milho e trigo no Brasil. Londrina: Embrapa Soja, 48 p.
- Embrapa – Empresa Brasileira De Pesquisa Agropecuária. (2013a). *Sistema brasileiro de classificação de solos*. 3ª Ed. Centro Nacional de Pesquisa de Solos, 1- 412.

- Embrapa - Empresa Brasileira De Pesquisa Agropecuária. (2013b). *Tecnologias de produção de soja - Região Central do Brasil* - 2014. Embrapa Soja.
- Ferrante, A. & Mariani, L. (2018). Agronomic management for enhancing plant tolerance to abiotic stresses: High and low values of temperature, light intensity, and relative humidity. *Horticulturae*, 4, (3), 1-21. doi: <https://doi.org/10.3390/horticulturae4030021>.
- Francisco, H.R.; Coldebella, A.; Corrêa, A.F.; Feiden, A. S. (2020) Spatial analysis of point events to estimate the productive potential of the Nile tilapia (*Oreochromis niloticus*). *Research, Society and Development*, 9 (9), e855998038. doi: <http://dx.doi.org/10.33448/rsd-v9i9.8038>.
- Friche, A. A. de L.; Calaffa, W. T.; César, C. C.; Goulart, L. M. de F.; Almeida, M. C. de M. (2006). Indicadores de saúde materno infantil em Belo Horizonte, Minas Gerais, Brasil, 2001: Análise dos diferenciais intra-urbanos. *Cadernos de Saúde Pública*, 22(9), 1955-1965. doi: <https://doi.org/10.1590/S0102-311X200600900027>.
- Galeano, E. A. V.; Ferrão, R. G.; Souza, R. C.; Tanques, R. C. (2019). Mudança na Distribuição Espacial da Produtividade da Cafeicultura no Espírito Santo nos Anos 2011-2016. *Multi-Science Research (MSR)*, 2, (1), 88-112. <https://biblioteca.incaper.es.gov.br/digital/bitstream/123456789/3827/1/distribuicao-espacial-galeano.pdf>.
- Grzegozewski, D. M.; Uribe-Opazo, M. A.; Johann, J. A.; Guedes, L. P. C. (2017). Spatial correlation of soybean productivity, enhanced vegetation index (EVI) and agrometeorological variables. *Engenharia Agrícola*, 37(3), 541-555. doi: <https://doi.org/10.1590/1809-4430-eng.agric.v37n3p541-555/2017>
- Ibge - Instituto Brasileiro de Geografia e Estatística (2016). Produção Agrícola Municipal. Culturas Temporárias Permanentes. from https://biblioteca.ibge.gov.br/visualizacao/periodicos/66/pam_2016_v43_br.pdf
- Ipardes-Instituto Paranaense de Desenvolvimento Econômico e Social. (2018). Perfil da região geográfica Oeste Paranaense. from http://www.ipardes.gov.br/perfil_municipal/MontaPerfil.php?codlocal=706&btOk=ok.
- Koepfen, W. (1948). *Climatologia: con un estudio de los climas de la tierra. Fondo de Cultura Económica*.
- Leite, s. D. (2016). Análise da distribuição espacial da pobreza na região norte do brasil. *Revista economia e desenvolvimento*, 15 (2), 215-231. doi:<https://doi.org/10.22478/ufpb.1517-9354.2016v15n2.36083>.
- Lopes, P. F.; Chain, C. P.; Menezes, T. P. de; Prado, J. W. do; Carvalho, F. de M. (2017). Gestão de riscos e dependência espacial em seguros agrícolas: uma análise da produtividade da soja. *Interciência*, 42(8), 503-508. <https://www.redalyc.org/journal/339/33952871004/html/>.
- Qgis Development Team. (2018). QGIS Geographic Information System. *Open source geospatial foundation project*.
- R Core Team. (2014). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna.
- Richetti, J.; Johann, J. A.; Uribe-Opazo, M. A. (2021). Crop modeling with less data: The FAO model for soybean yield estimation. *Engenharia Agrícola*, 41 (2), 196-203. doi: <https://doi.org/10.1590/1809-4430-Eng.Agric.v41n2p196-203/2021>
- Santos, A. de P.; Rodrigues, D. D.; Santos, N. T.; Gripp Junior, J. (2016). Avaliação da acurácia posicional em dados espaciais utilizando técnicas de estatística espacial: proposta de método e exemplo utilizando a norma brasileira. *Boletim de Ciências Geodésicas*, 22 (4), 630-650. <https://revistas.ufpr.br/bcg/article/view/49633>.
- Seab - Secretaria Da Agricultura e Abastecimento. (2016). Boletins Informativos -Colheita de soja na reta final. Departamento de Economia Rural (DERAL). http://www.agricultura.pr.gov.br/sites/default/arquivos_restritos/files/qas/4912/soja_29_mar_2016.pdf.
- Seab-Secretaria Da Agricultura e Abastecimento. (2017). Boletins Informativos – Produtividade recorde. Departamento de Economia Rural (DERAL). http://www.agricultura.pr.gov.br/sites/default/arquivos_restritos/files/qas/5709/soja_28_julho_2017.pdf.
- Seab - Secretaria da Agricultura e Abastecimento. (2018). Produção agrícola por município. Departamento de Economia Rural (DERAL). <http://www.agricultura.pr.gov.br/modules/conteudo>.
- Seffrin, R. (2017). Análise exploratória de dados espaciais aplicada a produtividade de milho no estado do Paraná. 97f. Dissertação (Mestrado em Tecnologias Computacionais para o Agronegócio) - Universidade Tecnológica Federal do Paraná, Medianeira.
- Simões, R.; Guimarães, C.; Godoy, N.; Velloso, T.; Araújo, T.; Galinari, R.; Chein, F. (2016). Rede urbana da oferta de serviços de saúde: uma análise de clusters espaciais para Minas Gerais. *Anais*, 1-27.
- Silva, E.M.; Gasparin, P.P.; Paludo, A.; Becker, W.R.; Guedes, L.P.C.; Johann, J.A. (2021). Soybean productivity and agrometeorological variables assessed from the perspective of Spatial and Circular Statistics. *Journal of Agricultural Studies*, 9 (3), 303-325. https://econpapers.repec.org/article/mthjas888/v_3a9_3ay_3a2021_3ai_3a3_3ap_3a303-325.htm
- Silva, V.A.; Ribeiro, L.C.S.; Esperidião, F. (2020). Municipal development and infrastructure index and its relationship with child mortality in brazilian municipalities. *Research, Society and Development*, 9 (6), e169963491, 2020. doi: <http://dx.doi.org/10.33448/rsd-v9i6.3491>
- Vidigal, V. G.; Vidigal, C. B. R.; Parré, J. L. (2018). Distribuição espacial da produtividade da soja no Rio Grande do Sul: um estudo exploratório. *Acta Scientiarum: Human & Social Sciences*, 40 (2), 1-9. doi: <https://doi.org/10.4025/actascihumansoc.v40i2.33652>
- Zanão Júnior, L. A.; Caramori, P. H.; Faria, R. T. de. (2017). *Produtividade da soja no entorno do reservatório de Itaipu*. Iapar.