

Article

The Use of Crop Yield Autocorrelation Data as a Sustainable Approach to Adjust Agronomic Inputs

Thomas M. Koutsos , Georgios C. Menexes  and Andreas P. Mamolos 

School of Agriculture, Faculty of Agriculture, Forestry and Natural Environment, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece; gmenexes@agro.auth.gr (G.C.M.); mamolos@agro.auth.gr (A.P.M.)

* Correspondence: tkoutsos@agro.auth.gr

Abstract: Agricultural fields have natural within-field soil variations that can be extensive, are usually contiguous, and are not always traceable. As a result, in many cases, site-specific attention is required to adjust inputs and optimize crop performance. Researchers, such as agronomists, agricultural engineers, or economists and other scientists, have shown increased interest in performing yield monitor data analysis to improve farmers' decision-making concerning the better management of the agronomic inputs in the fields, while following a much more sustainable approach. In this case, spatial analysis of crop yield data with the form of spatial autocorrelation analysis can be used as a practical sustainable approach to locate statistically significant low-production areas. The resulted insights can be used as prescription maps on the tractors to reduce overall inputs and farming costs. This aim of this work is to present the benefits of conducting spatial analysis of yield crop data as a sustainable approach. Current work proves that the implementation of this process is costless, easy to perform and provides a better understanding of the current agronomic needs for better decision-making within a short time, adopting a sustainable approach.



Citation: Koutsos, T.M.; Menexes, G.C.; Mamolos, A.P. The Use of Crop Yield Autocorrelation Data as a Sustainable Approach to Adjust Agronomic Inputs. *Sustainability* **2021**, *13*, 2362. <https://doi.org/10.3390/su13042362>

Academic Editors: Marc A. Rosen and Vasileios Tzanakakis

Received: 17 December 2020

Accepted: 16 February 2021

Published: 22 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: delineation of management zones; decision making; spatial analysis

1. Introduction

Agricultural fields have natural within-field variations due to ground, climate or other related factors and their multiple interactions [1] and, as expected, crop yield functions as a sensor of the local environment, reflecting the cumulative effect of all these variations [2]. The spatial variability of yield is affected by multiple factors, such as soil, geomorphology, crop traits, and additional influencing dynamic factors, such as weather-related factors, or environmental impacts from physical or human activities, and the total extent of human intervention, including all agricultural practices used [3–7]. All these factors may affect crop yield to a different extent, even if the type of crop remains the same [8,9].

As a result, although researchers agree that information regarding the influence of these factors will support better management decisions to increase crop productivity and lower farming costs [10], the influence of these factors may vary depending on the case, and estimation of this influence is complex and difficult to achieve. Therefore, in most cases, only soil-related factors are used to determine the agronomic inputs needed in time and space [11,12]. It should also be stressed that the nutrients in the soils are spatially and temporally dynamic, and their availability to the plant at any location and time depends on many factors that may also vary from area to area [13]. Factors such as organic matter in the soil and manure applications, temperature changes and rainfall patterns, previous crop, and different leaching losses, are only some of the variables that affect the final yield at each location. The complexity of the yield response makes model specification difficult [14].

To deal with field complexity, growers and scientists have focused on obtaining more agricultural field data. Since the early 1990s, both have realized that using modern agriculture technologies (mainly precision agriculture techniques and yield monitor technology) enhances the ability to conduct on-farm trials and collect more precise yield data, with the

hope that these can provide the opportunity to improve their whole-farm decision-making process [15,16]. In terms of a more detailed mapping of crop yields and understanding of within-field variations, yield monitors mounted on combine harvesters have significantly improved, and have had widespread use for the last ten years. As a result, the management of agricultural crops is gradually becoming data-driven, with large amounts of potential information available from crops or soil sampling [13,17]. Data from high-end technology sensors, remote-sensing data for crops and soils, and weather and climate measurements enhance the available database for data-driven agricultural management [18]. Finally, the aim of the collection and analysis of these data is their transformation into meaningful information that will facilitate decision-making and improve the management of agricultural fields [13,16,17], according to the several spatial and temporal variability factors involved, which is broadly termed as precision agriculture [19].

However, although large amounts of yield data are becoming available for analysis, several issues have been reported and yield data analysis has difficulties as a decisive component of the decision-making process concerning the better management of fields [13]. Several studies [9,14,20] have shown that it is difficult to bind yields with soil or weather conditions to determine management zones (MZs) due to the many dynamic factors involved; therefore, precision agriculture technologies face issues in terms of clearly establishing a new profitability approach. To deal with uncertainty, field analysis has started to be based on spatial analysis of yield data of more than one year, hoping to obtain a better understanding of the within-field spatial variability and then use these insights to characterize within-field spatial variability and divide a field into management zones (MZs), or areas with homogeneous properties known to impact crop yield [20]. According to precision agriculture principles, all agronomic inputs, such as irrigation, fertilizers, or pesticides, can be optimized in different management zones to maximize the profit margin [12,14,16]. This technique may be less precise than the point-to-point estimation, but it can lead to minimized input and costs, and therefore be considered as a much more sustainable approach. However, there is still a debate on whether the application of precision agriculture technologies always provides tangible evidence for their performance [21,22].

The whole-field recommendations derived from generalized recommendation systems or from farmers' experience (implemented in most growing environments) can provide a reasonable average recommendation in most years but are, overall, imprecise for individual fields [23]. For example, the Stanford-type mass balance approach for calculating site-specific N recommendations depending on soil and crop N availability is potentially appealing, because of its relative simplicity, which makes it easy to implement, but it is very generalized over diverse growing conditions and can fail at predicting field-specific needs, leading to the potentially unnecessary use of inputs while adopting a less sustainable approach [24]. Contrary to this, gaining the advantage of spatial insights from the use of model-database tools can lead to a much more sophisticated sustainable approach compared with the generalized recommendation methods, helping farmers to focus only on areas with low yield data and limit fertilizers, irrigation, or any other inputs only to areas where they are needed [25]. The approach of saving resources is in line with the definition of Sustainable Development [26]: "Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs.". The main concept of this approach is that sustainable development should comprise three different aspects: economic, social, and environmental [27]. Focusing only on areas with low crop yield not only benefits from the reduced inputs but also leads to lower outputs, such as lower emissions to the atmosphere (reduced use of tractors) or reduced chemicals loss to the environment. As a result, there is a gap that spatial analysis can fill by following a more sustainable approach, as the current work suggests: spatial autocorrelation statistical indicators can be used as a tool to identify areas with low crop yield, which farmers should focus on, to lower overall inputs, save resources and minimize environmental impacts. In this case, the whole-field recommendations can be adjusted to focus only on the low-production areas.

Previous studies have successfully explored the spatial dependence of crop yield using global and local statistics [28]. However, this study provides a cluster map comparison of the performance of the local spatial statistics that leads to the conclusion that all the given cluster maps, based on spatial statistics, can be used for the delineation of management zones, and therefore for the construction of the prescription maps on tractors to reduce overall inputs and farming costs. In addition, this work also provides the percentage of the potential reduction in overall inputs which have not been estimated in previous studies, comparing the performance of the use of different spatial statistics [28]. The comparison shows that the performance of all local spatial statistics is similar, based on their cluster maps' visualization. Therefore, the current study suggests that a short spatial analysis based on the autocorrelation of the available crop-yield data, regardless of the local spatial statistics used, can help farmers to review and adjust their input management in the fields to lower their farming costs. This strategy of using any local spatial statistics as a tool can be a sustainable approach to improve the input management, replacing the common practice of using an average of the crop yield per area of interest.

Two case studies concerning corn-field experiments, conducted in Las Rosas, Argentina, are analyzed to illustrate the applicability of using the autocorrelation of crop yield data as a sustainable approach and as an effective tool to locate and focus on low-production areas based on spatial analysis [29]. Section 2.1 introduces the two on-farm databases used as case studies. Section 2.2 provides a short review of how spatial statistics have been used to date in the delineation of low-crop-yield management zones. Section 2.3 describes the mathematical background of the spatial autocorrelation statistics used to analyze the crop yield data. Section 2.4 describes how the geostatistical analyses of the field spatial variability of crop yield were conducted to locate cultivated areas with statistically significant low yields values, surrounded by low yields. We conclude with a discussion in Section 3, the limitations in Section 4, and conclusions in Section 5.

2. Materials and Methods

2.1. Site Description and Data Collection Used

The data used in this manuscript derive from strip trials conducted at a farm called "Las Rosas" in the Rio Cuarto area, Cordoba Province, in central Argentina. The farm is located at 63°50'50" W and 33°03'04" S. The sample data are referenced in the tutorials for GeoDa spatial analysis software and freely available online [30]. These files ("Las Rosas, 1999" and "Rosas, 2001") include spatial variation measurements in monitor corn yield data (quintals/ha) associated with corresponding nitrogen fertilizer amounts and other field characteristics for the "Las Rosas" experiment for two separate years: 1991 and 2001. The "Las Rosas" experiment was conducted by incorporating six nitrogen rate treatments in three replicated blocks comprising 18 strips across the field [29,30].

The percentile cluster map (Figure 1) provides a visual exploration of the corn yield variability in the field; crop yield varies from 31.23 to 90.38 quintals/ha for Las Rosas 1999 dataset, and from 12.66 up to 117.90 quintals/ha for Las Rosas 2001 dataset, respectively. However, this cluster map cannot be used for the delineation of management zones because it cannot lead to contiguous, statistically significant areas with the same traits.

2.2. Delineation of Low-Crop-Yield Management Zones Using Spatial Statistics

Spatial statistics have been widely used to analyze spatial field properties and support decision-making to improve agricultural management [17]. Several successful efforts, using different techniques, have been made to date to produce spatial clusters that can be used as different management zones [14,28,31–34]. Researchers showed that integrating spatial crop-yield variability in the decision-making process for farming management may increase yield production [13,31,32]. Studies also have shown that adding spatial variability insights to clustering management methods improved spatial clustering for practical uses [13,28,33,35].

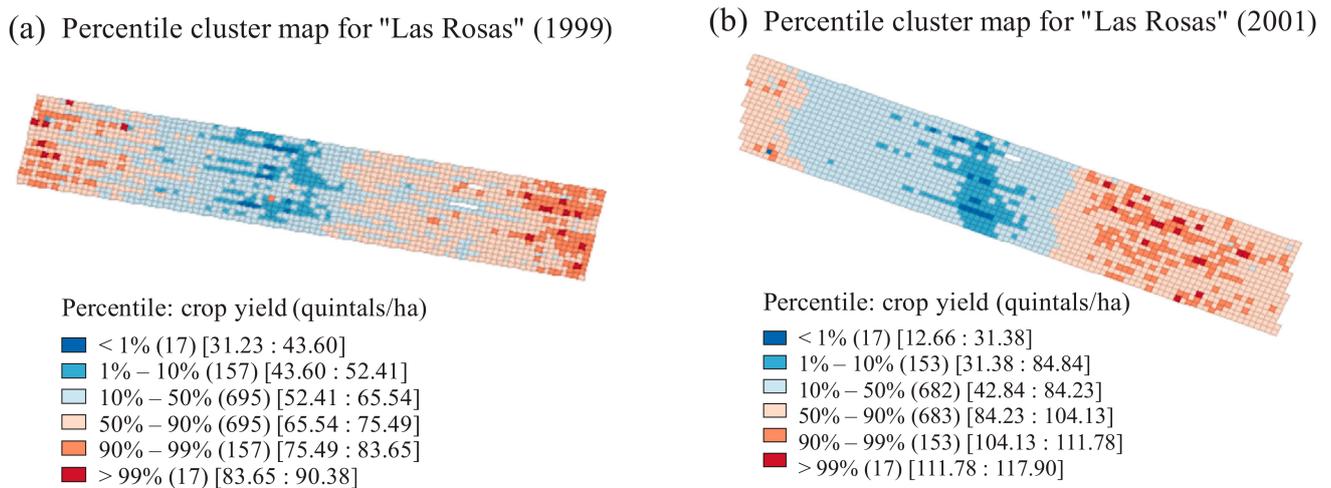


Figure 1. Percentile cluster map based on corn yield data (yield values in quintals per hectare) for: (a) "Las Rosas" farm corn yield data (1999) and (b) "Las Rosas" farm corn yield data (2001).

The delineation of MZs can be based on the characterization of soil physical variables and is achievable using regression kriging analysis and then principal component analysis and fuzzy cluster classification [36]. A multivariate spatial clustering approach has also been proposed [33] for the delineation of different MZs using spatial statistics. A regression technique (local geographically weighted regression (GWR)) has also been tested to express the spatial relationship between soil properties and in-season vegetation index, where the GWR data were finally used for the delineation of MZs [37]. Other scientists [14,17] also used a geographically weighted regression method to analyze spatially varying treatment effects in on-farm experiments. Few studies also used novel machine-learning approaches to analyze multivariable effects on crop yield [38], but they did not account for spatial variability. Other studies [34] used factorial kriging analysis based on multiple soil variables to produce spatial clusters that can represent different MZs. In sum, most of the studies focused on demonstrating the spatial relationship of soil characteristics, but they usually neglected other parameters that may affect the spatial variability of crop production and yield (climate, environmental conditions, crop, etc.).

Concerning the spatial autocorrelation of crop yield data, this has been often used to describe the degree of dependencies among neighboring observations in a field experiment, aiming to obtain an adequate sampling interval for which observations remain spatially correlated, and to design sampling protocols [12,28,39,40]. After reviewing the available literature, the authors reached the conclusion that there is no emphasis on the use of spatial autocorrelation as a sustainable approach to minimize inputs and farming costs. This research suggests the use of spatial autocorrelation of crop-yield data as a novel sustainable approach and tool for the delineation of potential low-yield management zones, aiming to limit the inputs only to the areas where they are needed.

2.3. Global and Local Spatial Autocorrelation Statistics

Spatial autocorrelation is expressed using global Moran's I and Geary's C statistics, whereas local spatial autocorrelation is described by a local indicator of spatial association, local G_i and G_i^* statistics.

The spatial autocorrelation Global Moran's I is an inferential statistic, which means that the results of the analysis are always interpreted within the context of its null hypothesis, where the values of the analyzed parameter are randomly distributed in the study area [41]. The global Moran's I can be calculated as follows [42], using Equation (1)

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

where n equals the number of observations; w_{ij} is the weight between locations i and j ; x_i and x_j are the values at locations i and j ; \bar{x} is the average over all locations of the variable.

The local indicator of spatial association (LISA) can be described [43] using Equation (2)

$$I_i = \frac{x_i - \bar{x}}{\sum_i (x_i - \bar{x})^2} \sum_j w_{ij} (x_j - \bar{x}) \quad (2)$$

where \bar{x}_i^* and \bar{x}^{2*} represent the means from the neighboring area, with x_i being excluded and included, respectively; w_{ij} represents the weight between locations i and j ; x_i and x_j are the values at locations i and j ; \bar{x} is the average over all locations of the variable. As shown above in Equations (1) and (2), a spatial weight matrix \mathbf{W} (consisting of several w_{ij} pairs) is needed for the calculation of the spatial autocorrelation. Each weight element w_{ij} , as an element of this normalized neighborhood matrix, corresponds to a pair of observations at locations i and j . Non-zero values reflect the potential spatial interaction between two observations, while zero values indicate a lack of spatial interaction [44]. The most common ways of calculating these weights are called Rook's, where w_{ij} is set to 1 if a pair shares a common edge and 0 otherwise, and Queen's weights, where w_{ij} is set if the pair shares either a common edge or a vertex and 0 otherwise [45]. By convention, w_{ii} is also defined as zero. The weight matrix \mathbf{W} can also be defined by actual distance, inverse distance with powers of 1 through 5, and k-nearest points methods [45], and it can be based in the Euclidean distance between any pair of observations, as given in Equation (3)

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (3)$$

where i and j are any two points in the given area, with respective coordinates (x_i, y_i) and (x_j, y_j) , respectively. Once d_{ij} is obtained from Equation (3), it can be used to calculate weights as inverse distance weights

$$w_{ij} = \frac{1}{d_{ij}^m} \quad (4)$$

where m is the power. In case of k-nearest weight matrices, the distances between any pair points were calculated and compared; k-nearest points were then selected and kept in the matrices. Usually, the k-closest points from 4 through 10 are selected. Row-standardization is also performed first for each matrix, to allow easier calculations of spatial autocorrelation statistics. In practice, the spatial weight matrix \mathbf{W} determines how much one observation contributes to the overall global spatial autocorrelation, because Moran's I is the summation of the product between the weight and deviation from the mean or value of the next observation [28].

The calculated variance for global Moran's I can be obtained from Equation (5) [42]

$$\text{var}_N(I) = \frac{1}{(n-1)(n+1) \left(\sum_{i=1}^n \sum_{j=1}^n w_{ij} \right)^2} \left[n^2 S_1 - n S_2 + 3 \left(\sum_{i=1}^n \sum_{j=1}^n w_{ij} \right)^2 \right] - \frac{1}{(n-1)^2} \quad (5)$$

where

$$S_1 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (w_{ij} + w_{ji})^2 \quad (6)$$

$$S_2 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (w_i + w_j)^2 \quad (7)$$

The expected I is calculated by using Equation (8)

$$E_N(I) = -(n-1)^{-1} \quad (8)$$

The significance of the global Moran's I statistic is tested based on their z -scores (simply standard deviations), using Equation (9)

$$z\text{-score} = \frac{I - E(I)}{\sqrt{\text{var}(I)}} \quad (9)$$

Geary's C statistics can be calculated by using Equation (10) [42]

$$C = \frac{n-1}{2 \sum_{i=1}^n \sum_{j=1}^n w_{ij}} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - x_j)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (10)$$

Local Geary C_i can be calculated by Equation (11) [42,43]

$$C_i = \frac{n}{\sum_{i=1}^n (x_i - x_j)^2} \sum_{j=1}^n w_{ij} (x_i - x_j)^2 \quad (11)$$

Local G_i and G_i^* statistics are described [28,42,43] in Equations (11) and (12), respectively

$$G_i = \frac{\sum_j w_{ij} x_j - \sum_j w_{ij} \bar{x}_i}{\left[\sum_j x_j^2 / (n-1) - \bar{x}_i^2 \right]^{0.5} \left\{ \left[(n-1) \sum_j w_{ij}^2 - \left(\sum_j w_{ij} \right)^2 \right] / (n-2) \right\}^{0.5}}, j \neq i \quad (12)$$

$$G_i^* = \frac{\sum_j w_{ij} x_j - \left(\sum_j w_{ij} + w_{ij} \right) \bar{x}_i^*}{\left[\sum_j x_j^2 / (n - \bar{x}^{2*}) \right]^{0.5} \left\{ \left[(n \sum_j w_{ij}^2) - \left(\sum_j w_{ij} + w_{ij} \right)^2 \right] / (n-1) \right\}^{0.5}}, \text{all } j \quad (13)$$

where \bar{x}_i^* and \bar{x}^{2*} represent the means from the neighboring area, with x_i being excluded and included, respectively. Therefore, the main difference between local G_i and G_i^* is that G_i requires x_i to be excluded from the summation, whereas G_i^* requires x_i to be included in the summation.

In the case of normal distribution of data, the threshold of 1.96 can be applied to test the significance level of z . If the z value is greater than smaller than -1.96 , this implies that the spatial autocorrelation is significant [28,46]. A p -value (observed significance level) is also calculated along with the z -score to indicate whether the difference is statistically significant and represents the probability that the observed spatial pattern was created by some random process. A very small (<0.05) p -value means that the null hypothesis can be rejected, meaning that the observed spatial pattern is not the result of a random process. Cases with very high or very low z -scores, associated with very small p -values (p -value < 0.05), indicate that it is unlikely that the observed spatial pattern reflects the theoretical random pattern represented by the null hypothesis. A statistically significant positive z -score means that similar high or similar low values cluster together, while a negative z -score means that similar values are spatially dispersed, as we expect in the case of an underlying random spatial process.

Concerning the effective sample size used in spatial modelling, it is well known that as spatial autocorrelation latent increases in geo-referenced data, the amount of duplicated information contained in these data increases too [47]. Therefore, if the n observations are (positively) spatial autocorrelated, the amount of statistical information carried by the n observations is less that it would be if the n observations were independent. It has been confirmed that the "effective sample size" is less that the actual sample size n [48]. Therefore, in the case of n datapoints as independent observations, the effective sample size is n , but if the observations are dependent then the effective sample size is less than n ,

because of the duplicated information. The reduction in this information in the context of multiple testing of local indices of spatial autocorrelation has been thoroughly examined, and it was found that the effective sample size depends on the spatial locations of the observations on the specified range of the spatial process [49–51].

2.4. Geostatistical Analyses of Field Spatial Variability of Crop Yield

Spatial autocorrelation of data can be measured either at a local or a global level. The local level represents the extent of autocorrelation within local neighborhoods, while the global level provides a total value that represents the extent of spatial autocorrelation across the entire study area [52]. Local patterns of spatial autocorrelation were found to be an appropriate perspective for understanding local instabilities, and they were expressed as local indicators of spatial association (LISA), local Getis's G_i and G_i^* , and Geary C_i statistics [28,42,43,49].

The local indicator of spatial association (LISA) as defined a statistic that satisfies two main requirements [43]: (1) LISA value gives an indication of the extent of significant spatial clustering of similar values around an observed value; (2) the sum of all LISA values for all observations is proportional to a global indicator of spatial association. The corresponding scatter plot for Moran's I values is used to provide a visual exploration of global spatial autocorrelation (Figure 2).

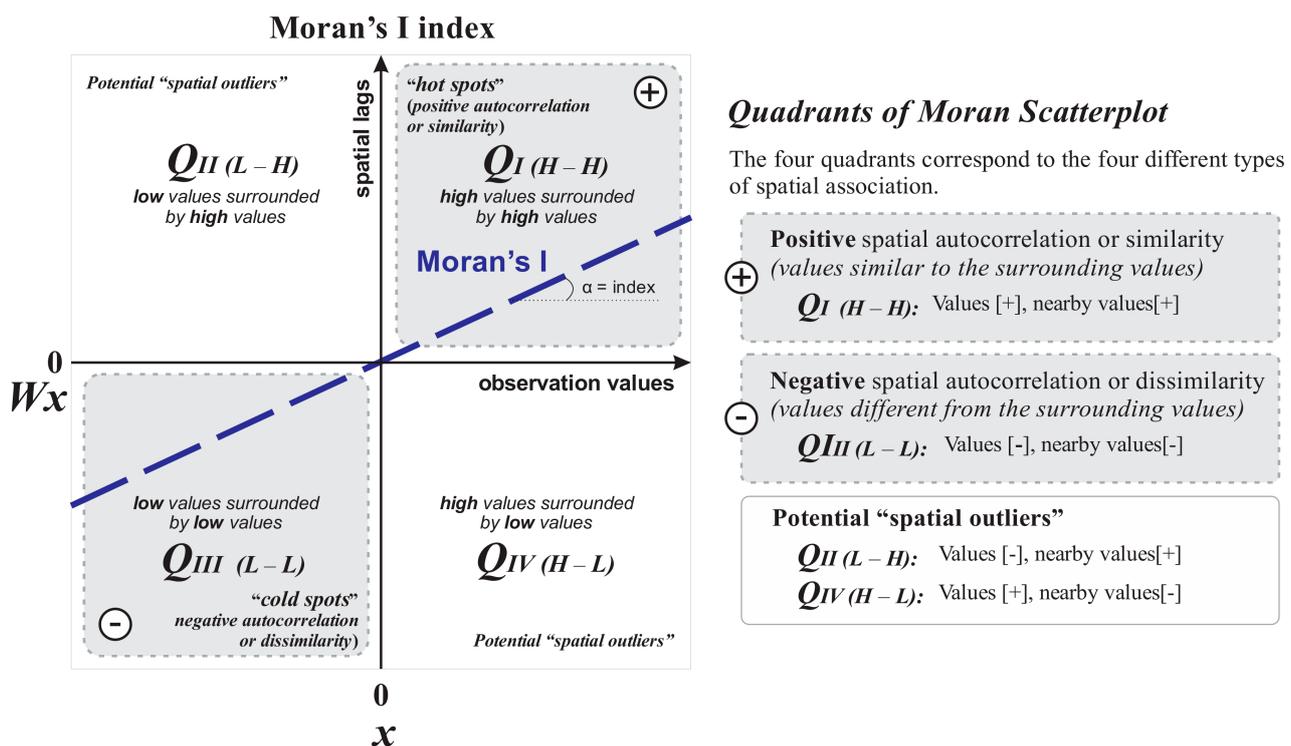


Figure 2. Quadrants of Moran Scatterplot (modified from [44]).

The Moran's I statistic is a standard measure to evaluate spatial autocorrelation and can be used as a statistical test to verify the spatial dependence of the yield crop data [44]. The null hypothesis is defined in terms of the absence of spatial autocorrelation of the examined data. In case of rejection, there is evidence that prevalent values in a specific geographical entity depend on variables in neighboring spaces. By using this statistic as a tool, the Moran's I can help the statistical spatial identification of poor production zones in every agricultural production area [42,44,53]. To check whether the null hypothesis is rejected or not (whether there is a yield-crop data spatial dependence or not), a LISA significance map for Moran's I is constructed based on the p -values calculated for each location.

LISA values allow for the computation of similarity of values with their surrounding neighbors. Therefore, five scenarios may emerge: (1) locations with high values with similar neighbors; (2) locations with low values with similar neighbors; (3) locations with high values with low-value neighbors; (4) locations with low values with high-value neighbors; (5) locations with no significant local autocorrelation. The four quadrants (Q_I to Q_{IV}) in the Moran scatter plot represent the clusters of the spatial autocorrelation [44]: Q_I : Locations with high values with similar neighbors, or *high-high (H-H)*, also known as “hot spots”, representing positive autocorrelation; Q_{III} : Locations with low values with similar neighbors, or *low-low (L-L)*, also known as “cold spots”, representing negative autocorrelation; Q_{II} : Locations with high values with low value neighbors or *high-low (H-L)*, representing potential “spatial outliers and Q_{IV} : Locations with low values with high value neighbors, or *low-high (L-H)*, representing potential “spatial outliers (Figure 2).

Moran’s I index value is usually listed at the top of the graph, showing the spatial autocorrelation of the data being examined. The slope of the regression line in Moran’s I scatterplot is an estimation of the global Moran’s I . the relative density of points in the correlation quadrants indicates how the global measure of spatial association is determined by an association between high and/or low values. The Moran’s I values range from -1 to $+1$ [48]. A higher positive Moran’s I value near $+1$ indicates high spatial autocorrelation, implying that values in neighboring positions tend to cluster together. A low negative Moran’s I value gives an indication that high and low values are interspersed. A Moran’s I value near zero means that there is no spatial autocorrelation or the data are randomly distributed. On the other hand, Geary’s C ranges from 0 to 2; whereas a zero indicates a strong positive spatial autocorrelation, a 1 shows no spatial autocorrelation, and a 2 represents a strong negative spatial autocorrelation [42,48]. The global G statistics indicate a general tendency towards the clustering of low values (negative G), high values (positive G) or none of both (non-significant). The local G_i (and G_i^*) statistics can be interpreted in the same manner and the main difference between local G_i and G_i^* is that G_i requires x_i to be excluded from the summation, whereas G_i^* requires x_i to be included in the summation. We should also stress that a positive G_i (and G_i^*) indicates a spatial clustering of high values only, but a positive LISA value is an indication of spatial clustering of either high or low values, like global Moran’s I .

When evaluating crop yield agricultural data, Moran’s I index can give an indication of the spatial autocorrelation of yield, and the Moran scatter plot provides a visual exploration of the global spatial autocorrelation of yield in the field. The quadrant Q_{III} (Figure 2) represents low values surrounded by low values (negative autocorrelation), representing low-production areas.

2.5. Cases Studies

Moran’s I index has been used to statistically measure and evaluate the spatial autocorrelation of the available corn yield data for the years 1999 and 2001 for Las Rosas farm. The result for each year is a spatial correlogram that plots the Moran’s I value for each distance for which it is measured, where the distance at which observations are no longer spatially autocorrelated is termed the spatial range, also determined by the spatial correlogram. The spatial correlogram of the available crop data was easily constructed by using GeoDa free spatial analysis software, but it can also be conducted by the `sp.correlogram` function in the `spdep` [54]-contributed package in `R`. The outcome has positive values and no negative or zero values, as expected in most site-specific data for variables at field-scales.

The result of for Univariate Moran’s I scatter plot shows a positive relationship, suggesting the existence of spatial autocorrelation in yield crop data. As expected, the slope of the regression line corresponds to statistic of Moran, meaning that the deeper the slope, the higher the degree of spatial data autocorrelation. This is also confirmed in Figure 3, where Moran’s I index was calculated for both periods (“Las Rosas” 1999, 2001). The indicator values for both years (0.701 for year 1999 and 0.957 for year 2001) prove that there is high autocorrelation between yield spatial data (Figure 3a,b) in both cases.

As a result, we expect to obtain distinct continuous areas with high crop-yield values, surrounded by high and low crop-yield values, surrounded by low values, respectively.

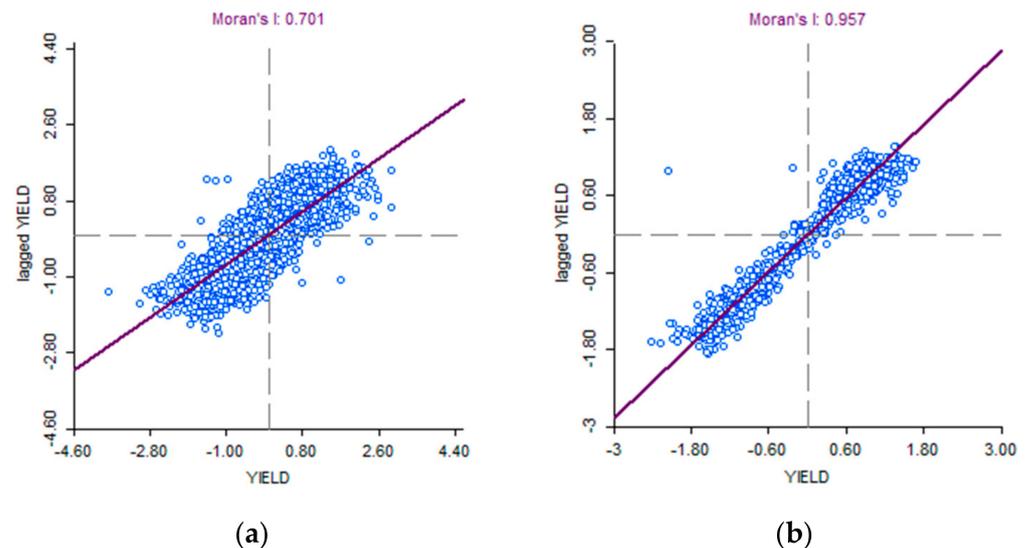


Figure 3. Univariate Moran's I scatter plot for: (a) "Las Rosas" farm corn yield data (1999) and (b) "Las Rosas" farm corn yield data (2001).

One way to check if the overall Moran's I index is statistically significant is through a simulation process in which the index is calculated with random observation samples. For both scatter diagrams, we performed a randomization with 99,999 permutations. The results show a pseudo p -value of 0.000010 (I : 0.7009; mean = -0.0006 ; sd = 0.0124; z -value = 56.5814) for "Las Rosas" farm data for the year 1999, and a pseudo p -value of 0.000010 (I : 0.9575; mean = -0.0006 ; sd = 0.0126; z -value = 76.3245) for the year 2001. In both cases, the pseudo p -value is less than 0.01 and the Moran's I index, and statistically significant at a confidence level greater than 99.9%.

The corresponding LISA cluster maps were constructed for each year (1999 and 2001) to visualize the spatial dependence of crop yield data (Figures 4a and 5a). LISA cluster map for Moran's I index offers a better visual exploration of global spatial autocorrelation of the crop yield data, where two main areas can be distinguished: (a) area with high yield values surrounded by high yield values (which corresponds to Moran's scatter plot quadrant $H-H$), and (b) an area with low yield values surrounded by low yield values (which corresponds to Moran's scatter plot quadrant $L-L$). In both LISA cluster maps (Figures 4a and 5a), two clusters ($-L$ and $H-H$ clusters) can be identified, and they can be used by farmers as two different management zones to adjust inputs in the field. These ($L-L$) areas also represent areas with low yield values surrounded by low yield values (Q_{III}). Current work suggests that input management strategy can be adjusted to focus mainly on the ($L-L$) areas, with the benefit of potentially reducing the overall inputs. In the case of the "Las Rosas" farm experiment, as expected from the high autocorrelation value of Moran's I index, significant local clusters of yield were observed within the field in the LISA cluster map for ($L-L$) areas. If overall inputs are limited to only these ($L-L$) areas, then the expected potential reduction in inputs for the "Las Rosas" farm can reach up to 74.3% (Figure 4a), and 43.2% for the year 2001 (Figure 5a).

Concerning local Getis G_i^* cluster maps (Figures 4b and 5b), discrete spatial patterns of clusters also occur, allowing the identification of (*High*) and (*Low*) clusters, like ($H-H$), and ($L-L$) clusters in LISA cluster maps, with similar areas. Therefore, we conclude that the results are almost the same and the potential reduction in inputs for the "Las Rosas" farm could be similar: up to 73.1% for the year 1999 (Figure 4b), and up to 59.2% for the year 2001 (Figure 5b).

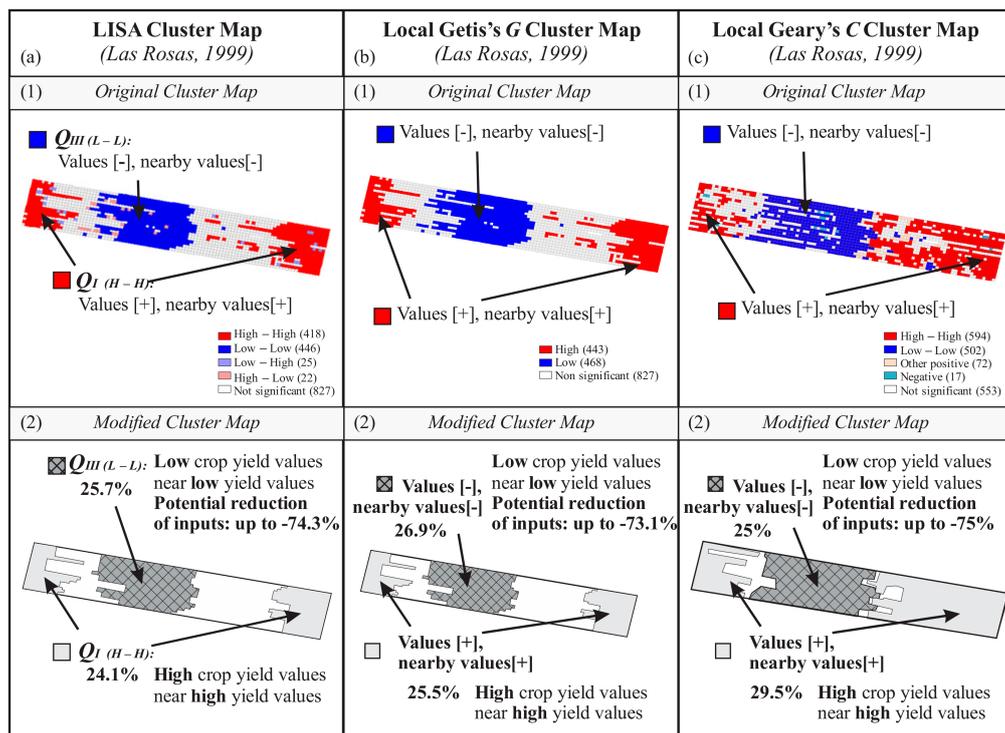


Figure 4. Cluster maps for “Las Rosas” farm corn yield data (1999) for: (a) local indicator of spatial association (LISA) and (b) Local Getis G, and (c) Local Geary C, where (1) the original cluster maps and (2) modified and simplified version of the corresponding cluster map focusing on areas with low crop-yield values. A potential reduction in inputs has been calculated in each case.

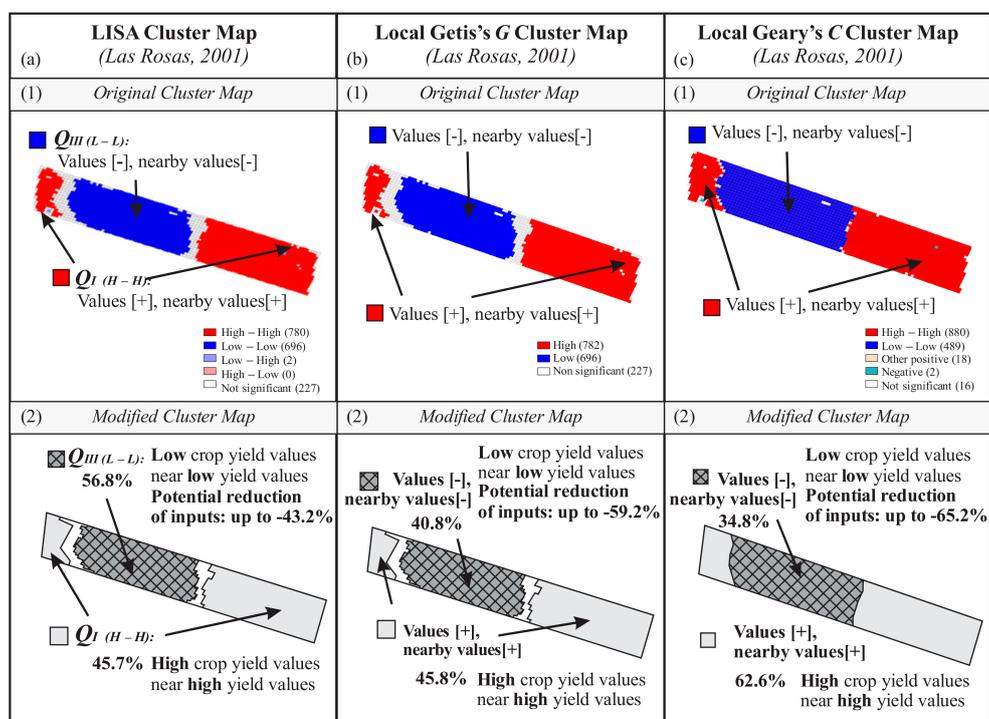


Figure 5. Cluster maps for “Las Rosas” farm corn yield data (2001) for: (a) LISA and (b) Local Getis G, and (c) Local Geary C, where (1) the original cluster maps and (2) modified and simplified version of the corresponding cluster map focusing on areas with low crop-yield values. A potential reduction in inputs has been calculated in each case.

Local Geary C cluster map can also be used for the identification of statistically significant low-production areas. It was found that the local Geary C cluster map has a similar performance, with a potential reduction in inputs of up to 75% for the “Las Rosas” farm for the year 1999 (Figure 4b), and up to 65.2% for the year 2001, (Figure 5b), respectively.

The current work suggests that an input management strategy can be based on these cluster maps (Figures 4 and 5), focusing mainly on these low-yield areas, with the benefit of potentially reducing the overall inputs. In the case of the “Las Rosas” farm experiment and the two years 1999 and 2001, if overall inputs were limited to cover the needs of these low-yield areas only, then the expected potential reduction in inputs could be significantly high compared to a uniform fertilizer application.

The corresponding significance maps for “Las Rosas” farm corn-yield datasets for the years 1999 and 2001 for (a) LISA, (b) Local Getis’s G_i , and (c) Local Geary C are given in Figures 6 and 7 (generated using the GeoDa software). In both cases, the local clusters that presented the Moran local index (LISA) were discretized in different shades, with p -values equal to or less than 0.05 (Figure 5(a1,b1)). However, those local clusters that did not have a significant Moran local autocorrelation index (LISA) were colorless.

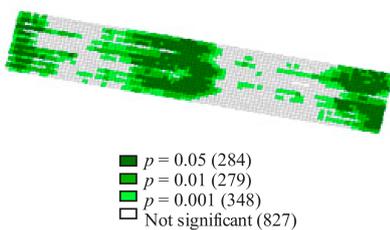
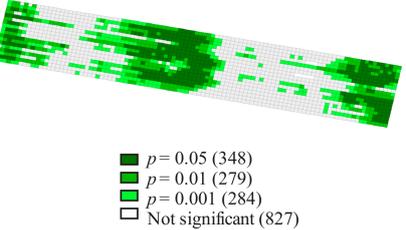
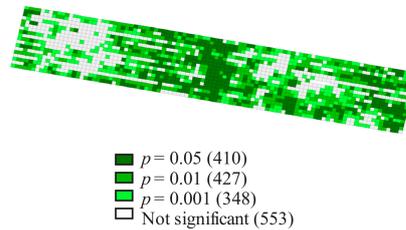
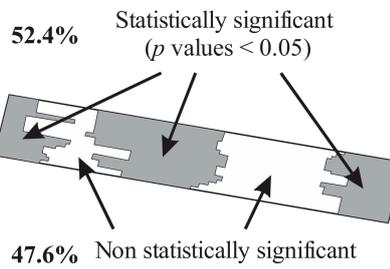
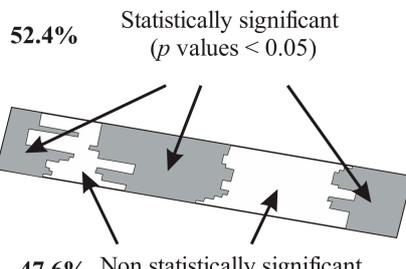
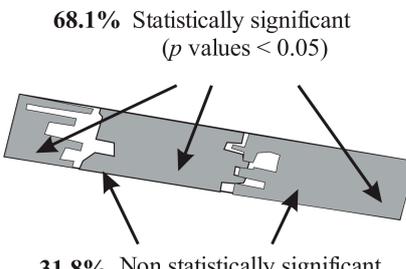
(a) LISA Significance Map (Las Rosas, 1999)	(b) Local Getis’s G_i Significance Map (Las Rosas, 1999)	(c) Local Geary’s C_i Significance Map (Las Rosas, 1999)
(1) <i>Original Significance Map</i>	(1) <i>Original Significance Map</i>	(1) <i>Original Significance Map</i>
		
(2) <i>Modified Significance Map</i>	(2) <i>Modified Significance Map</i>	(2) <i>Modified Significance Map</i>
		

Figure 6. Significance maps for “Las Rosas” farm corn-yield data (1999) for: (a) LISA, (b) Local Getis G_i , and (c) Local Geary C_i , where (1) the original cluster maps and (2) modified and simplified version, grouping statistically (p values < 0.05) and non-statistically significant areas.

The results for the local Getis’s G_i significance map are identical to LISA, while the Local Geary’s C_i significance map shows a higher percentage of statistically significant areas compared to the other two significance maps. In all cases, discrete spatial patterns of clusters occur, allowing the identification of significant clusters (Figures 6 and 7).

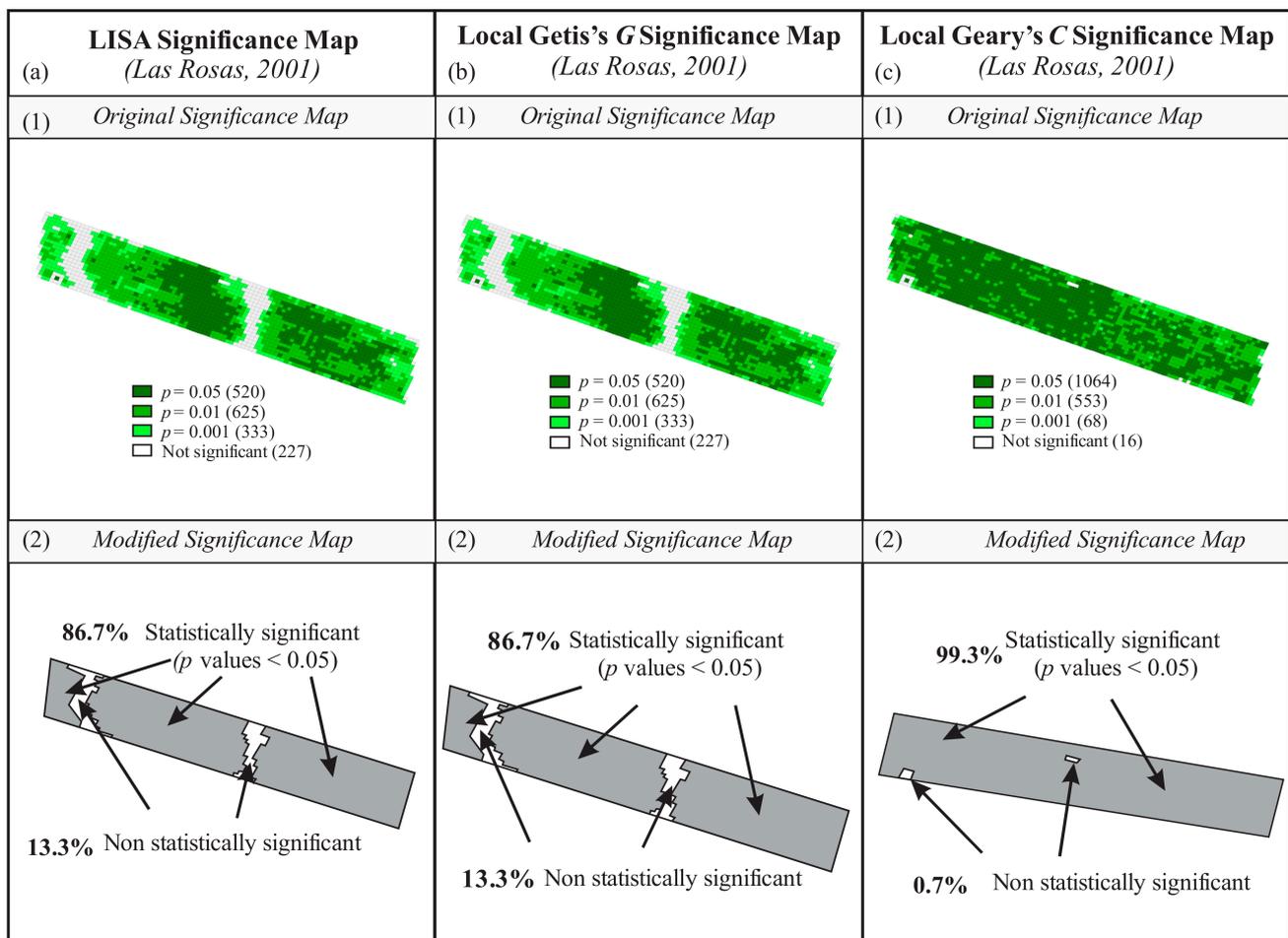


Figure 7. Significance maps for “Las Rosas” farm corn-yield data (2001) for: (a) LISA, (b) Local Getis G , and (c) Local Geary C , where (1) the original cluster maps and (2) modified and simplified version, grouping statistically (p values < 0.05) and non-statistically significant areas

The performance of the three examined spatial statistics was similar. Therefore, we conclude that the use of the LISA cluster/significance map for Moran's I is adequate to identify statistically significant low-production management zones to reduce the overall inputs.

3. Discussion

Sustainability has set the framework to diminish the environmental footprint of farming, while ensuring the food security and economic viability of agriculture, resulting in the development of precision agriculture and the use of spatial statistics to sustainably optimize the management of cultivated fields by addressing the spatial variability of several field parameters [14,15,19,28,31].

Although spatial autocorrelation was defined years ago, most of the studies to date with spatial autocorrelation for spatial dependence at the global or local scale focused on spatial econometrics [44,53]. Newer studies have explored the application of these statistics to the understanding of spatial dependence of crop yield in site-specific crop management, to evaluate the application of global and local autocorrelations by exploring the spatial variability of cotton lint yield and yield pattern changes under different weather scenarios and comparing the effects of weight selection on spatial autocorrelation [28]. However, there is no report on comparison of these local spatial statistics regarding their performance in terms of the percentage of the potential reduction in overall inputs based on their cluster maps. Therefore, in the current study, we constructed cluster maps for the three most-used spatial local statistics (LISA, Local Geary's C_i , and local Getis's G_i) for two different years using a different type of crop (corn instead of cotton) from that used

in the previous studies [28]. The results show that the performance of these three local statistics concerning the expected benefits of reduced inputs was almost similar, and any of the above can be used for the delineation of management zones.

Current study shows that a short and costless (given that crop yield monitor data are available) spatial analysis using any local spatial statistic on the available crop yield data can help the analyst, whether the farmer or a third party, to obtain an indication of the autocorrelation of crop yield values and have a better understanding of the within-field distributions. The various local spatial statistics (LISA, Local Geary's C_i , and local Getis's G_i) are equally effective methods to identify local spatial patterns and, therefore, statistically significant areas such as low- or high-production areas.

As suggested in the current study, this evidence can be provided by using a spatial analysis index, such as the univariate Moran's I followed by the univariate Moran's I scatter plot and can be visualized by constructing the LISA cluster map to present statistically significant areas with low crop yield. This cluster map that presents the autocorrelation of yield values in the sampled locations in the field is not just a map; it can be used as a recommendation to determine areas with a similar production performance and, in this case, to identify areas the inputs in the field should be focused on. As a result, this cluster map can help growers to quickly identify field patches or statistically significant areas with low yield values ($L-L$) and obtain the detailed information needed for the construction of potential prescription maps based on the delineation of different management zones. Alternatives to LISA include the local Geary's C_i or local Getis G_i cluster map, which can also lead to the identification of statistically significant low-production areas and, therefore, can help the adjustment of inputs, the same as the LISA cluster map. In our case studies, if we could focus only on statistically significant low-production zones, these cluster maps could lead to a potential reduction in agronomic inputs ranging from 43.2% to 74.3% based on LISA, 65.2% to 75% for local Geary's C_i , and 69.8% to 73.1% based on the local Getis's G_i cluster map.

The case studies presented in this manuscript show that the spatial analysis of the available crop yield data (with no other complex regression tests) can easily result in a cluster autocorrelation map that the farm manager can feasibly implement in a timely manner. The delineation of low-yield management zones can be based on statistically significant local clusters and can be used to review and adjust the overall inputs. Therefore, the result of this spatial analysis should be treated as a sustainable approach that can provide growers with a production/input recommendation of how to improve crop performance using minimized inputs and farming costs.

The expected benefits from conducting a spatial analysis of the available crop yield data can be summarized to the following:

- Quick focus on areas with poor crop-yield performance;
- Less tillage needed for the areas with high crop performance (clusters: $H-H$, *High*);
- Reduced seed for the areas with high crop performance (clusters: $H-H$, *High*);
- Less energy is needed due to reduced fuels since fewer hours are needed for machinery use (clusters: $H-H$, *High*);
- Reduced inputs (adjusted fertilizer amounts and irrigation) for the areas with high crop performance (clusters: $H-H$, *High*);
- Economic benefits for the farmer due to the reduced input amounts;
- Environmental benefits (from reduced input amounts, less chemical leaching to the environment, reduced emissions due to reduced machinery use);
- Better planning for the input needs and better future crop management.

Previous studies have successfully explored the usefulness of global and local spatial analysis in helping to delineate practical management zones, but how the insights of the autocorrelation of crop yield data could be transformed into the delineation of management zones was vague, and there was no estimation of the potential reduction in inputs [13,28]. Compared to these previous studies, the current work supports the novel idea of transforming the original cluster maps of local spatial statistics into modified prescription maps that

will be used to change the input management and, therefore, lower inputs and farming costs. Compared to previous works, the current study shows that a potential reduction in overall inputs can be estimated regardless of which local statistic is used.

In addition, stored cluster maps constructed at different times can be used for further statistical and economic analysis, regression with other parameters or to perform comparisons between years and periods. Therefore, using spatial analysis on available crop yield data based on the cluster map of a local statistic can provide insights to perform a yield history evaluation of the available historical crop yield data. This technique can also be used as an assessment tool of the efficiency of the strategy and the agronomy practices used, and therefore “trigger” input adjustments and help to improve the overall management of the field.

4. Limitations

It should be stressed that focusing only on low-yield areas may not always be safe; spatial autocorrelation gives a strong statistical indication of the low-production areas, or else for the areas that should be the focus of attention. However, depending on the high-crop yield values, even for high-production areas, a customized input management may be needed to cover overall field needs. Moreover, focusing only on low-production areas and adjusting inputs does not imply that if these needs are covered, the production would present higher yield values the next period, because low crop yields may be due to reasons (crop protection, agricultural practices, and environmental conditions prevailing in the area) other than a lack of nutrients in the soil.

Spatial autocorrelation of data can be measured at the global level to provide a total value that represents the extent of spatial autocorrelation across the entire study area or at a local level to show the extent of autocorrelation within local neighborhoods. Local patterns of spatial autocorrelation were expressed as local indicator of spatial association (LISA), local Getis G_i and G_i^* , and Geary C_i statistics. In the presented case studies, these spatial statistics provide almost the same information, which can be used for the delineation of management zones. However, results may differ depending on the parameters used (different parameters defining the weight matrix for a given dataset).

Therefore, the proposed sustainable approach aims to diminish the environmental footprint of farming by replacing the whole-field input recommendations with more precise recommendations derived from the autocorrelation of crop-yield data, focusing only on how to limit the agricultural inputs (mainly water and fertilizers) to areas where they are needed based on the yield spatial variability. It should also be stressed that the application of the proposed methodology does not depend on the size of the crop area and can be applied at a local, regional, or global area.

5. Conclusions

The current work should be considered as an extension of previous studies that attempted to explore the application of local spatial statistics to understanding the spatial dependence of crop yield in site-specific crop management. In this study, three of the most-used local statistics were tested for their performance (potential reduction in overall agronomic inputs) using a different crop to previous studies [13,45] and cluster maps as a tool to delineate different management zones (crop yield data gathered in two different years). It was found that the use of spatial autocorrelation of crop yield data, regardless of the local spatial statistic used, can provide solutions to farmers to minimize overall inputs by providing a safe and clear statistical method that can identify low-production areas based on the autocorrelation of the spatial variability in crop output at the regional level, ultimately leading to a sustainable increase in farm productivity. Crop yield data, collected by yield mapping systems on tractors with the aid of GPS technology, self-calibrating yield monitors and sensors can be used for this kind of spatial analysis. This work shows that conducting a basic spatial analysis on available yield monitor data is a relatively easy process, regardless of the spatial statistic used, as well as feasible and costless, and, in a

short time, it provides a better understanding of the within-field variations, aiming to improve decision-making concerning input management while supporting sustainability.

The advantages of performing a base spatial analysis on available crop-yield data can be summarized as follows:

- Autocorrelation of yield data to reveal areas with low yield values;
- Spatial distribution and mapping of the crop-yield data;
- Yield history evaluation by performing yield comparisons between years;
- Identification of areas with very low yield values that require additional attention;
- Insights for the delineation of the management zones for the field, aiming to improve inputs and reduce costs.

Conducting a spatial analysis can lead to a cluster map that can easily be used for the delineation of different management zones and then for the construction of possible prescription maps on the tractors, or they can simply be used to review and adjust the input strategy adopted by farmers. Our suggestion is that if yield monitor data are available, then a short spatial analysis can easily be conducted to reveal areas with low crop performance, where attention is required. This process can help farmers to make better decisions on the input management of the fields to reduce farming costs, instead of using an average yield estimation for calculating the needed inputs. Especially in the case of large, cultivated areas, current spatial analysis is imperative as it can help to significantly improve the partial budgeting and lower farming costs.

The novelty of the presented approach can be summarized as follows:

- Promotes sustainability by providing a clear and easy geostatistical way to reduce overall inputs and focus only on cultivated areas with low yield;
- Adapts spatial autocorrelation of crop yield data to on-farm experimentation;
- Allows assessment of spatially varying treatment effects;
- Outlines a statistically principled approach which enables the delineation of management zones based on spatially varying crop yield data;
- Demonstrates statistical analyses on two example datasets using free spatial analysis software (such as GeoDa spatial analysis software [30]);
- Compares the performance of the three most used spatial statistics in the potential reduction in overall agronomic inputs;
- Supports the idea of transforming the cluster maps of local statistics into prescription maps for the delineation of management zones;
- Provides an estimation of the potential reduction in inputs based on the cluster maps of local spatial statistics.

This study suggests the use of autocorrelation of crop yield data as a sustainable approach that can easily reveal statistically significant low-yield areas where farmers should focus on providing the nutrients, or any other inputs needed, regardless of the local spatial statistic being used. Instead of using an average of crop yield to calculate the input amounts needed, the proposed spatial autocorrelation approach supports sustainability and offers more accuracy, leading to minimized inputs and lower farming costs. The current work provides a safe way to quantify the potential reduction in overall inputs based on the cluster maps of local statistics.

Quantifying the crop yield spatial variability in the process of determining low-yield MZs can lead to a better understanding of the field needs and to a better management of inputs and costs. Incorporating the crop yield spatial autocorrelation can also contribute to the further development of multivariable spatial analysis to improve agricultural practices. Keeping records of the spatial analysis of the field yield variability will also provide insights for data-mining, decision-guiding and more precise agricultural modeling.

Author Contributions: Conceptualization, T.M.K.; methodology, T.M.K. and G.C.M.; formal analysis, G.C.M.; investigation, T.M.K., G.C.M., A.P.M.; data curation, G.C.M.; writing—original draft preparation, T.M.K.; writing—review and editing, T.M.K., G.C.M., A.P.M.; visualization, T.M.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yang, C.; Peterson, C.L.; Shropshire, G.J.; Ottawa, T. Spatial Variability of Field Topography and Wheat Yield in the Palouse Region of the Pacific Northwest. *Trans. ASAE* **1998**, *41*, 17–27. [[CrossRef](#)]
2. Legg, B.J.; Stafford, J.V. Precision Agriculture—New Technologies. In *Proceedings of the Brighton Crop Protection Conference—Pests & Diseases: Proceedings of an International Conference: Held at the Brighton Centre & the Stakis Brighton Metropole Hotel, Brighton, UK, 16–19 November 1998*; British Crop Protection Council: Alton, UK, 1998.
3. French, R.J. Soil factors influencing growth and yield of narrow-leaved lupin and field pea in Western Australia. *Aust. J. Agric. Res.* **2002**, *53*, 217. [[CrossRef](#)]
4. Vieira, S.R.; González, A.P. Analysis of the spatial variability of crop yield and soil properties in small agricultural plots. *Bragantia* **2003**, *62*, 127–138. [[CrossRef](#)]
5. Gu, S.; Du, G.; Zoldoske, D.; Hakim, A.; Cochran, R.; Fugelsang, K.; Jorgensen, G. Effects of irrigation amount on water relations, vegetative growth, yield and fruit composition of Sauvignon blanc grapevines under partial rootzone drying and conventional irrigation in the San Joaquin Valley of California, USA. *J. Hortic. Sci. Biotechnol.* **2004**, *79*, 26–33. [[CrossRef](#)]
6. Van Leeuwen, C. Terroir: The effect of the physical environment on vine growth, grape ripening and wine sensory attributes. *Manag. Wine Qual.* **2010**, 273–315. [[CrossRef](#)]
7. Scudiero, E.; Teatini, P.; Corwin, D.L.; Deiana, R.; Berti, A.; Morari, F. Delineation of site-specific management units in a saline region at the Venice Lagoon margin, Italy, using soil reflectance and apparent electrical conductivity. *Comput. Electron. Agric.* **2013**, *99*, 54–64. [[CrossRef](#)]
8. Plant, R.E. Site-specific management: The application of information technology to crop production. *Comput. Electron. Agric.* **2001**, *30*, 9–29. [[CrossRef](#)]
9. Basnet, B.; Kelly, R.; Jensen, T.; Strong, W.; Apan, A.; Butler, D. Delineation of management zones using multiple crop yield data. In *Proceedings of the 16th Triennial Congress of the International Soil Tillage Research Organisation (ISTRO 2003)*, Brisbane, Australia, 13–18 July 2003.
10. Masseroni, D.; Ortuani, B.; Corti, M.; Gallina, P.M.; Cocetta, G.; Ferrante, A.; Facchi, A. Assessing the Reliability of Thermal and Optical Imaging Techniques for Detecting Crop Water Status under Different Nitrogen Levels. *Sustainability* **2017**, *9*, 1548. [[CrossRef](#)]
11. Masseroni, D.; Facchi, A.; Gandolfi, C. Is Soil Water Potential a Reliable Variable for Irrigation Scheduling in the Case of Peach Orchards? *Soil Sci.* **2016**, *181*, 232–240. [[CrossRef](#)]
12. Haghverdi, A.; Leib, B.; Washington-Allen, R.; Wright, W.C.; Ghodsi, S.; Grant, T.; Zheng, M.; Vanchiasong, P. Studying Crop Yield Response to Supplemental Irrigation and the Spatial Heterogeneity of Soil Physical Attributes in a Humid Region. *Agriculture* **2019**, *9*, 43. [[CrossRef](#)]
13. Chopin, P.; Blazy, K.-M. Assessment of regional variability in crop yields with spatial autocorrelation: Banana farms and policy implications in Martinique. *Agric. Ecosyst. Environ.* **2013**, *181*, 12–21. [[CrossRef](#)]
14. Ohana-Levi, N.; Bahat, I.; Peeters, A.; Shtein, A.; Netzer, Y.; Cohen, Y.; Ben-Gal, A. A weighted multivariate spatial clustering model to determine irrigation management zones. *Comput. Electron. Agric.* **2019**, *162*, 719–731. [[CrossRef](#)]
15. Griffin, T.; Brown, J.; Lowenberg-Deboer, J. Yield Monitor Data Analysis Protocol: A Primer in the Management and Analysis of Precision Agriculture Data. *SSRN Electron. J.* **2007**. [[CrossRef](#)]
16. Perea, R.G.; Daccache, A.; Díaz, J.A.R.; Poyato, E.C.; Knox, J.W. Modelling impacts of precision irrigation on crop yield and in-field water management. *Precis. Agric.* **2018**, *19*, 497–512. [[CrossRef](#)]
17. Rakshit, S.; Baddeley, A.; Stefanova, K.; Reeves, K.; Chen, K.; Cao, Z.; Evans, F.; Gibberd, M. Novel approach to the analysis of spatially-varying treatment effects in on-farm experiments. *Field Crop. Res.* **2020**, *255*, 107783. [[CrossRef](#)]
18. Wolfert, S.; Ge, L.; Verdouw, C.; Bogaardt, M.-J. Big Data in Smart Farming—A review. *Agric. Syst.* **2017**, *153*, 69–80. [[CrossRef](#)]
19. Zhang, N.; Wang, M.; Wang, N. Precision agriculture—A worldwide overview. *Comput. Electron. Agric.* **2002**, *36*, 113–132. [[CrossRef](#)]
20. Zhang, X.; Shi, L.; Jia, X.; Seielstad, G.; Helgason, C. Zone mapping application for precision-farming: A decision support tool for variable rate application. *Precis. Agric.* **2010**, *11*, 103–114. [[CrossRef](#)]
21. McBratney, A.; Whelan, B.; Ancev, T.; Bouma, J. Future Directions of Precision Agriculture. *Precis. Agric.* **2005**, *6*, 7–23. [[CrossRef](#)]
22. Balafoutis, A.; Beck, B.; Fountas, S.; Vangeyete, J.; van der Wal, T.; Soto, I.; Gómez-Barbero, M.; Barnes, A.; Eory, V. Precision Agriculture Technologies Positively Contributing to GHG Emissions Mitigation, Farm Productivity and Economics. *Sustainability* **2017**, *9*, 1339. [[CrossRef](#)]

23. Morris, T.F.; Murrell, T.S.; Beegle, D.B.; Camberato, J.J.; Ferguson, R.B.; Grove, J.; Ketterings, Q.; Kyveryga, P.M.; Laboski, C.A.; McGrath, J.M.; et al. Strengths and Limitations of Nitrogen Rate Recommendations for Corn and Opportunities for Improvement. *Agron. J.* **2018**, *110*, 1–37. [[CrossRef](#)]
24. Sela, S.; van Es, H.M.; Moebius-Clune, B.N.; Marjerson, R.; Moebius-Clune, D.; Schindelbeck, R.; Severson, K.; Young, E. Dynamic Model Improves Agronomic and Environmental Outcomes for Maize Nitrogen Management over Static Approach. *J. Environ. Qual.* **2017**, *46*, 311–319. [[CrossRef](#)] [[PubMed](#)]
25. Scharf, P.C.; Shannon, D.K.; Palm, H.L.; Sudduth, K.A.; Drummond, S.T.; Kitchen, N.R.; Mueller, L.J.; Hubbard, V.C.; Oliveira, L.F. Sensor-Based Nitrogen Applications Out-Performed Producer-Chosen Rates for Corn in On-Farm Demonstrations. *Agron. J.* **2011**, *103*, 1683–1691. [[CrossRef](#)]
26. World Commission on Environment and Development (W.C.E.D.). *From One Earth to One World: An Overview*; Oxford University Press: Oxford, UK, 1987.
27. Tanguay, G.A.; Rajaonson, J.; Lefebvre, J.-F.; Lanoie, P. Measuring the Sustainability of Cities: A Survey-Based Analysis of the Use of Local Indicators. *SSRN Electron. J.* **2009**. [[CrossRef](#)]
28. Ping, J.L.; Green, C.J.; Zartman, R.E.; Bronson, K.F. Exploring spatial dependence of cotton yield using global and local autocorrelation statistics. *Field Crops Res.* **2004**, *89*, 219–236. [[CrossRef](#)]
29. Lambert, D.M.; Lowenberg-Deboer, J.; Bongiovanni, R. A Comparison of Four Spatial Regression Models for Yield Monitor Data: A Case Study from Argentina. *Precis. Agric.* **2004**, *5*, 579–600. [[CrossRef](#)]
30. GeoDa Data and Lab. Available online: <https://geodacenter.github.io/data-and-lab/> (accessed on 17 December 2020).
31. Batchelor, W.D.; Basso, B.; Paz, J.O. Examples of strategies to analyze spatial and temporal yield variability using crop models. *Eur. J. Agron.* **2002**, *18*, 141–158. [[CrossRef](#)]
32. Mueller, N.D.; Gerber, J.S.; Johnston, M.; Ray, D.K.; Ramankutty, N.; Foley, J.A. Closing yield gaps through nutrient and water management. *Nat. Cell Biol.* **2012**, *490*, 254–257. [[CrossRef](#)]
33. Peeters, A.; Zude, M.; Käthner, J.; Ünlü, M.; Kanber, R.; Hetzroni, A.; Gebbers, R.; Ben-Gal, A. Getis–Ord’s hot- and cold-spot statistics as a basis for multivariate spatial clustering of orchard tree data. *Comput. Electron. Agric.* **2015**, *111*, 140–150. [[CrossRef](#)]
34. Bevington, J.; Scudiero, E.; Teatini, P.; Vellidis, G.; Morari, F. Factorial kriging analysis leverages soil physical properties and exhaustive data to predict distinguished zones of hydraulic properties. *Comput. Electron. Agric.* **2019**, *156*, 426–438. [[CrossRef](#)]
35. Nie, K.; Wang, Z.; Du, Q.; Ren, F.; Tian, Q. A Network-Constrained Integrated Method for Detecting Spatial Cluster and Risk Location of Traffic Crash: A Case Study from Wuhan, China. *Sustainability* **2015**, *7*, 2662–2677. [[CrossRef](#)]
36. Moral, F.J. Comparison of different geostatistical approaches to map climate variables: Application to precipitation. *Int. J. Clim.* **2009**, *30*, 620–631. [[CrossRef](#)]
37. Scudiero, E.; Teatini, P.; Manoli, G.; Braga, F.; Skaggs, T.H.; Morari, F. Workflow to Establish Time-Specific Zones in Precision Agriculture by Spatiotemporal Integration of Plant and Soil Sensing Data. *Agronomy* **2018**, *8*, 253. [[CrossRef](#)]
38. Goldstein, A.; Fink, L.; Meitin, A.; Bohadana, S.; Lutenberg, O.; Ravid, G. Applying machine learning on sensor data for irrigation recommendations: Revealing the agronomist’s tacit knowledge. *Precis. Agric.* **2018**, *19*, 421–444. [[CrossRef](#)]
39. Cressie, N. *Statistics for Spatial Data*; John Wiley & Sons: New York, NY, USA, 1993.
40. Yin, X. Geostatistical Analyses of Field Spatial Variability of Cotton Yield. *J. Geosci. Environ. Prot.* **2016**, *4*, 75–87. [[CrossRef](#)]
41. Mathur, M. Spatial autocorrelation analysis in plant population: An overview. *J. Appl. Nat. Sci.* **2015**, *7*, 501–513. [[CrossRef](#)]
42. Cliff, A.D.; Ord, J.K. *Spatial Processes*; Pion: London, UK, 1981; p. 266.
43. Anselin, L. Local Indicators of Spatial Association-LISA. *Geogr. Anal.* **1995**, *27*, 93–115. [[CrossRef](#)]
44. Anselin, L. *Spatial Econometrics: Methods and Models*; Kluwer Academic Publishers: Boston, MA, USA, 1988.
45. Berry, B.J.L.; Marble, D.F. *Spatial Analysis: A Reader in Statistical Geography*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1968; p. 512.
46. Zhang, C.S.; McGrath, D. Geostatistical and GIS analyses on soil organic carbon concentrations in grassland of south-eastern Ireland from two different periods. *Geoderma* **2004**, *119*, 261–275. [[CrossRef](#)]
47. Ronny, V.; Consuelo, M. Effective Sample Size in Spatial Modeling. In Proceedings of the 58th World Statistical Congress, Dublin, Ireland, 21–26 August 2011.
48. Clifford, P.; Richardson, S. Testing the association between two spatial processes. *Stat. Decis.* **1985**, *2*, 155–160.
49. Getis, A.; Ord, J. Seemingly independent tests: Addressing the problem of multiple simultaneous and dependent tests. In Proceedings of the 39th Annual Meeting of the Western Regional Science Association, Kanuui, HI, USA, 28 February 2000.
50. Getis, A.; Ord, J.K. The Analysis of Spatial Association by Use of Distance Statistics. *Geogr. Anal.* **2010**, *24*, 189–206. [[CrossRef](#)]
51. Ord, J.K.; Getis, A. Local Spatial Autocorrelation Statistics: Distribution issues and an application. *Geogr. Anal.* **1995**, *27*, 286–306. [[CrossRef](#)]
52. Mueller-Warrant, G.W.; Whittaker, G.W.; Young, W.C. GIS Analysis of Spatial Clustering and Temporal Change in Weeds of Grass Seed Crops. *Weed Sci.* **2008**, *56*, 647–669. [[CrossRef](#)]
53. Anselin, L.; Bongiovanni, R.; Lowenberg-de Boer, J. A Spatial Econometric Approach to the Economics of Site-Specific Nitrogen Management in Corn Production. *Am. J. Agric. Econ.* **2004**, *86*, 675–687. [[CrossRef](#)]
54. Bivand, R. Spdep: Spatial Dependence: Weighting Schemes, Statistics and Models. R Package Version 0.4-21. 2008. Available online: <http://cran.rproject.org/src/contrib/Descriptions/spdep.html> (accessed on 12 February 2021).