Recent Spatial Patterns of Grains and Fruits Crop Production in the Brazilian Northeast

Danielle A.P. Torres[1]

Marcelo de O. Torres[2]

Carlos Wagner A. Oliveira[1]

[1] Department of Regional and Urban Studies, Institute for Applied Economic Research. SBS-Quadra1-BlocoJ-Ed.BNDES-3rd floor. 70076-900 - Brasília - DF - Brazil

[2] Department of Economics. Catholic University of Brasília. Campus Avançado Asa Norte SGAN 916 Norte Av. W5. 70790-160 Brasília - DF- Brazil.

English Abstract

This paper uses *município*-level data in the Northeast region in Brazil to identify and analyze spatial patterns on value of production and yield of fruits and grains. Moran's I statistics are generated and used to test for spatial autocorrelation, and to prepare cluster maps that locate value of production and yield "hot spots" and "cold spots." Research results demonstrate that both variables are spatially auto correlated in some parts of the Northeast, and where correlated, high (low) value of production and yield *municípios* tend to be located next to high (low) *municípios*. Maybe more importantly the results set the stage for the use of spatial econometrics for a future multivariate analysis of agricultural value of production and yield of fruits and grains in the Northeast, and particularly in the region of Juazeiro/Petrolina a well-know area of fruit production in Brazil.

Key Words: cluster analysis, spatial autocorrelation, value of production, yield.

Résumé

Cet article utilise des données de la région nordeste du Brésil au niveau municipal pour identifier et analyser l´existence d´ autocorrélation spatiale pour les variables valeur de la production et rendement agricole des fruits et céréales. Le *I* de Moran a été utilisé pour évaluer l'autocorrélation spatiale global et le *I* de Moran locale (*LIM*) a été utilisé pour identifier dans des cartes les *municípios* avec valeur de la production et rendement

agricole plus élevés dans un voisinage qui lui ressemble (plus-plus) et *municípios* avec valeur de la production et rendement agricole moins élevés dans un voisinage qui lui ressemble (moins-moins). Les résultats indiquent l'existence d'une autocorrélation spatiale et suggèrent l'utilisation de l'économétrie spatiale pour une analyse multivariée du rendement et valeur de la production des fruits et céréales dans la région de Juazeiro/Petrolina, région connue au Brésil pour sa production des fruits.

Mots-clés: analyse de *cluster*, spatial autocorrelation spatiale, valeur de production, rendement.

1. Introduction

During the last decade, Brazil has experienced a boom in agricultural production. Helped by favorable international markets, relatively lower costs of production, and domestic household income rises, the Brazilian agricultural production and area has increased considerably from 1997-2007. During the same period, total production increased faster than production indicating a rise in land productivity measured by the partial factor productivity ratio (total agricultural production over total area under plow.) Brazilian Northeast productivity has also increased during this period, but still lags behind when compared to the rest Brazil.

More interesting, the pattern in the Northeast has not been homogenous neither in terms of its spatial distribution nor in terms of the product mix. While total harvested area has increased by 12% in the same period, land area in grains has increased by 21% and in fruits by 20%. And within these groups, the crops that showed larger increases were soybeans in grains, and grapes, guava and mangoes in fruits. In terms of location, grains production has increasingly been concentrated in the western portion of Bahia state, while fruits in northern Bahia and west of Pernambuco states.

While this agricultural production pattern and product mix shift in these areas may be of no surprise, recently released data indicate that some municipalities or micro regions within high value fruit production areas have remained low value and specialized in grains, and some municipalities in regions specialized in grains have shifted to higher value fruit production. Visual inspection suggests some patterns in the spatial distribution of fruits and grains production in the Brazilian northeast, but this must be confirmed by statistical analysis.

This paper lays the groundwork for examining the spatial patterns of grains and fruit productivity in the Brazilian northeast. We examine whether or not location matters in determining the agricultural productivity and the value of production in a given municipality. Cluster maps are developed and used to identify grains and fruit productivity and value of production 'hot and cold spots.'

In addition, we examine employment, average income in the agricultural sector of the Northeast region and an index of human development (IHD), as a proxy for a social indicator of the region, to see whether the existence of the fruit cluster has been beneficial for the region. In other words we investigate whether employment, average income, and the IHD are higher in the clustered areas compared to the entire Northeast Region.

2. Background - The Origins of Fruit Clustering in Petrolina/Juazeiro

Originally the main economic activity in the region of Petrolina/Juazeiro was livestock (Andrade, 1980) and agriculture was only for subsistence. In 1948, the federal government created the San Francisco River Valley Commission (CVSF) that later became the San Francisco River Valley Development Agency (CODEVASF), a government institution that targeted the promotion of navigation, irrigation, agricultural and industrial development in the San Francisco river valley. CODEVASF implemented public irrigation projects; enlisted different size growers and agricultural processing firms in each project; provided incentives for agricultural industries to establish in the region; and in 1988 supported the creation of a growers association (VALEXPORT) that was key to promoting exports.

By providing different sized lots, CODEVASF explicitly established a structure of production consisting of both large and small growers. Hundreds of small growers received irrigation-ready lots (with on-farm pumps, canals, and drainage system installed and ready for use), guidance on what to produce and technical assistance, and also facilitated access to credit and buyers. Initially, growers paid a nominal fee for the lots and received most of the support at no cost. Later, in the late 1980s, CODEVASF began charging water fees and drastically reduced its technical assistance. There was also a noteworthy technological dimension to this strategy. CODEVASF promoted a sequence of crops that facilitated the learning process of small growers, most of whom had never previously worked with irrigated agriculture. Thus, growers first produced a combination of annual crops, including beans, corn, and melon, followed by the wide-spread adoption of industrial tomato crops, and subsequently higher-value fruit crops,

including mangoes and grapes. The transition from phase to phase involved a combination of conventional and more innovative support policies to help growers in each, consecutively more difficult, phase.

Nowadays in the San Francisco Valley Region growers produce fresh fruit for both domestic and exports markets and have learned different strategies for upgrading, including to establish their relationship with each other and local support agencies. Large, medium, and small growers, coexist allowing the evaluation of the diffusion/reach of upgrading efforts developed by grower organizations and public sector agencies. Finally the region has agglomerated several firms and has access to a skilled labor pool, local input suppliers and support services (agricultural consultants, transportation, repair shops).

The public sector has had an important role in promoting the economic activity through irrigation, specialized education (through technical schools and two public colleges), agricultural research through the Brazilian Agricultural Research Corporation (EMBRAPA) and technical Assistance through Enterprise Technical Assistance and Rural Extension (EMATER). It has had also an active role in developing the infra-structure in urban and rural areas such as communication, transportation, structure for commercialization and dam construction (Sobradinho dam in particular) and in providing incentives to private investments for irrigation through subsidies from Northeast Development Superintendence (SUDENE). .

3. The Production Chain

The value chains in which growers participate have changed in recent years, together with the coordination of activities between growers and buyers (“governance”). The global market for fresh fruit has changed in i) its consolidation of global retailers which are gradually substituting small, national retailers and local vendors, and ii) its sourcing strategies, where the need to define and control for product and process standards has forced buyers to develop tighter relationships with their suppliers (importers and/or growers).

The increased participation of large food retailers has changed the relationship between growers and buyers. Whereas previously this relationship was generally segmented with an intermediary, retailers are increasingly shifting away from middlemen and wholesalers to alternative, more direct forms of procurement. These alternative strategies include formal and informal contracts directly with growers and the establishment of their own distribution centers, practices which allow the

supermarkets greater leverage in enforcing their quality and safety standards. The restructuring of food retail has therefore given more power to retailers and their importers.

The greater power of importers and buyers in these chains has meant mounting pressures for growers to make the necessary changes in their products and production processes to meet the demands of these buyers. That is, growers are under greater pressures to upgrade because they now have fewer buyers and these buyers are more demanding than ever.

4. Data

The data on agricultural production value and average yield by crop and municipality were drawn from the Municipal Agricultural Production (PAM) surveys performed annually by the Brazilian Geography and Statistic Institute (IBGE). Information on the number of employees in the agricultural sector and an average wage comes from the Ministry of Labor and an index of human development, from the Federation of Industries of the State of Rio de Janeiro (FIRJAN).

5. Methodology

The main question to be addressed in this paper is whether the observed pattern of fruit and grain production across the Northeast of Brazil is equally likely as any other spatial pattern rather than clustered. We evaluate this pattern by focusing at the *município* level and primarily on the value of production and yield for mangoes, melons and grapes, in the case of fruits, and for an aggregated measure of grains (corn, beans and soybeans). These are either the main crops in terms of value or volume in the region. We also extend the analysis by looking at employment, average income and at an index of social economic development. If the values of one of these variables in a given *município* are associated with values, of the same variable, in its surrounding *municípios*, the variable is then spatially auto correlated, which indicates that the pattern of this specific variable across the space in the Northeast is clustered.

To measure spatial autocorrelation, we use the global Moran´s *I* statistic. Originally developed by Moran 1948, autocorrelation statistics have been extended and applied in several different contexts such as in Amarasinghe *et al*., 2000; Anselin 1996, Cliff and Ord 1981, Griffith 2003; and Palmer-Jones and Sen, 2006; Pinkse ,2003, to name a few. In the Brazilian context it has been applied for the analysis of clusters of agricultural productivity, Perobelli et al 2007.

The measure of spatial autocorrelation is given by the global Moran´s *I* statistic, defined as

$$I = \frac{\sum_{i,j} w_{ij}(y_i - \bar{y})(y_j - \bar{y}) / \sum_{i,j} w_{ij}}{\sum_i (y_i - \bar{y})^2 / N}.^{1} \quad (1)$$

Where (y_i) is the value of a given variable in *município i*, $\bar{y}$ is the average value of this same variable over the entire Northeast region, and (y_j) is its value in *município j*. The term $(y_i - \bar{y})(y_j - \bar{y})$ is an element of a spatial weighting matrix, with the average value of the variable standardized around the sample mean, and w_{ij} is an element of a spatial weighting matrix. If *município i* shares a common boundary with *município j*, then w_{ij} =1, otherwise w_{ij} =0. This definition of neighboring areas is based on *rook contiguity*.

In this application the weighting matrix is row standardized, in which the weights are defined as $w_{ij}^s = \frac{w_{ij}}{\sum_j w_{ij}}$, such that $\sum_j w_{ij}^s = 1$.[2] The row standardization has two implications: 1) it implies equal weights across neighbors of a same *município*, and 2) it implies that the sum over all elements of the row-standardized weight matrix (w_{ij}^s) is equal to the total number of observations (N); that is, in $\sum_{i,j} w_{ij} = N$.[3] Therefore, (1) can be re-written as

[1]In the numerator of *I*, $\sum_{i,j} w_{ij}(y_i - \bar{y})(y_j - \bar{y})$ is a gamma statistic with (y_i) and (y_j) as the random variables, and as such, it is scale-dependent. In order to make it scale-independent, we divide it by $\sum_{i,j} w_{ij}$ and by a consistent estimator of the variance of the average yield (y_i) $\sum_i (y_i - \bar{y})^2 / N$.

[2] For example, if a *município i* has 4 neighbors $w_{ij}^s = 1/4$.

[3] With row standardization, the sum of weights in each row becomes 1. Since there is one row for each *município* in the sample, there are *N* rows. Therefore, the sum over all weights in the matrix, $\sum_{i,j} w_{ij}$, is *N*.

$$I^s = \frac{\sum_{i,j} w_{ij}^s (y_i - \bar{y})(y_j - \bar{y})}{\sum_i (y_i - \bar{y})^2} \qquad (1')$$

If *município*s with an above-average (below-average) value of a given variable are surrounded by neighboring *município*s with above-average (below-average) value, the cross product term $(y_i - \bar{y}) * (y_j - \bar{y})$ becomes positive, making $I^s > 0$, and implies that there is positive spatial autocorrelation. On the other hand, if *município*s with above-average (below-average) values are surrounded by neighboring *município*s with below-average (above-average) values, the cross product term $(y_i - \bar{y}) * (y_j - \bar{y})$ is negative, making $I^s < 0$, and implying that there is negative spatial autocorrelation. The closer I^s gets to zero, the weaker the evidence to support spatial autocorrelation and of a clustered pattern.

The *Moran´s I* result is however a global result statistic and does not tell us where the clusters might be, but rather suggests only that the spatial pattern is not random. To locate the clusters we turn to local indicators of spatial association or *LISA* (Anselin, 1995). *LISA* is a class of statistics that provides location-specific information (by *município*, in this case) and estimates the extent of spatial autocorrelation between the value of a given variable in a particular location and the values of that same variable in locations around it. Through inference analysis we are able to identify spatial clusters. ‘Hot-spots’ in the case of high-value *municípios* surrounded by high-value *municípios*) and ‘cold-spots’ in the case of low-value *municípios* surrounded by low-value municípios. These clusters might be comprised of a single *município* and its contiguous neighbors, or a larger set of contiguous *municípios* for which the *LISA* is statistically significant. We use the Local Moran’s *I* statistic or *LMI*, one of several statistics that falls within the *LISA* definition and it is defined as follows: [4]

$$LMI_i = \frac{x_i}{\sum_i x_i^2} \sum_j w_{ij} x_j \ . \qquad (2)$$

Where, $x_i = y_i - \bar{y}$ and $x_j = y_j - \bar{y}$ and y_i and y_j are, respectively, values of a specific variable for *municípios* *i* and *j*, and $\bar{y}$ is the sample mean. Spatial weights, w_{ij}, are

[4] See Anselin (1995) for examples of other *LISA* statistics, such as the *Local Geary*.

defined as before: $w_{ij} = 1$ if the *i* and *j* *municípios* are contiguous neighbors, $w_{ij} = 0$ otherwise, based on *rook contiguity*.

Analogous to the Global Moran's *I*, positive values of *LMI* indicate positive spatial autocorrelation, i.e., that a given *município* is surrounded by similar *municípios*, either above or below the Northeast region average. On the other hand, negative values of *LMI* indicate negative spatial autocorrelation, i.e., that a given *município* is surrounded by *municípios* with dissimilar measures.

6. Results

Table 1 below shows the *Moran´s I* statistics for yield and value of production for fruits and grains. Further Table 2 shows the *Moran´s I* for employment, average wage for 2007 and the IHD for the years available (2000 and 2005). The *município*-level global *Moran´s I* statistics for the value of production in mangoes, grapes, melons and grains in the northeast were 0.19, 0.12, 0.12, 0.35, respectively, and they changed to 0.18, 0.30, 0.28 and 0.41 in 2007. Except for mangoes, the *Moran's I* increased from 1997 to 2007. For yield, results are positive and higher, and also increased from 1997 to 2007.

Table 1. *Moran´s I* Values

Variable	***Morans´I* 1997**	***Morans´I* 2007**
Value of Production		
Mangoes	0.19	0.18
Grapes	0.12	0.30
Melons	0.12	0.28
Grains	0.35	0.41
Yield		
Mangoes	0.36	0.41
Grapes	0.20	0.33
Melons	0.45	0.49
Grains	0.60	0.63

We can conclude that there is a positive relationship or, in other words, that municipalities with high(low) yield and with high(low) value of production in fruits and grains relative to the state-wide mean tended to be located near municipalities with above(below)-average yield and value of production. Even though these *Moran's I* statistics are lower than expected they are statistically significant and confirm the existence of positive spatial autocorrelation.[5]

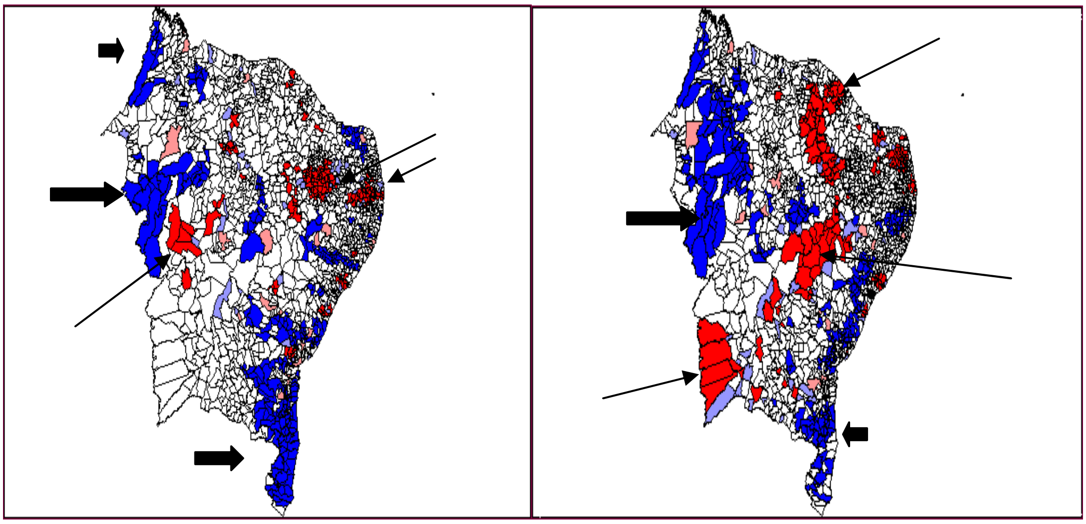
After finding that the spatial pattern is not random, we investigate where the clusters are located across the region. To do that we calculate the Local Moran's I (*LMI*) for each municipality and use these to locate the clusters. Looking at the maps, yield clusters emerge. Figure 1, below, shows in 1997 three small 'hot spots' with positive, and 'high-high' spatial correlation (thin arrow) indicating spatial clusters of *municípios*

[5] GeoDa™ is used to calculate all statistics and clusters maps in this paper. The test for significance uses the permuation approach. This software was developed by the Center for Spatially Integrated Social Science (CSISS) at the University of Illinois, Urbana-Champaign, Urbana, IL, USA.

with high yield in mangoes. There are also three 'cold spots' with positive, and 'low-low' spatial correlation (thick arrow) indicating spatial clusters of *municípios* with small yield in mangoes. In 2007 the map changed and three different and larger 'hot spots' appeared while one of the 'cold spots' increased (in the west) the other decreased (in the south).

Figure 1. *LMI* – Mangoes Yield

1997 **2007**



As a conclusion, there was an increase in the area formed by the clusters with high yield in mangoes across the whole region, while the area with low yield clusters decreased in the south and increased in the west of the Northeast region. Visually, this suggests a contagion process of high yield in mangoes across the region. Next we observe the yield of two other fruits and grains to investigate whether there are also clusters of high or low yield and whether they are located in the same areas formed by the mangoes cluster.

Figure 2, below, shows one 'hot spot' with high yield in grapes in the years of 1997 and 2007 in the same region as the mangoes cluster. That is, the region that includes the *municípios* of Juazeiro in the state of Bahia and of Petrolina in the state of Pernambuco. Other *municípios* that make part of the mangoes and grapes production cluster are: Lagoa Grande and Santa Maria da Boa Vista in the state of Pernambuco and Casa Nova, Curaca, Sento Se, Sobradinho in the state of Bahia.

Figure 2. LMI – Grapes Yield

1997 **2007**

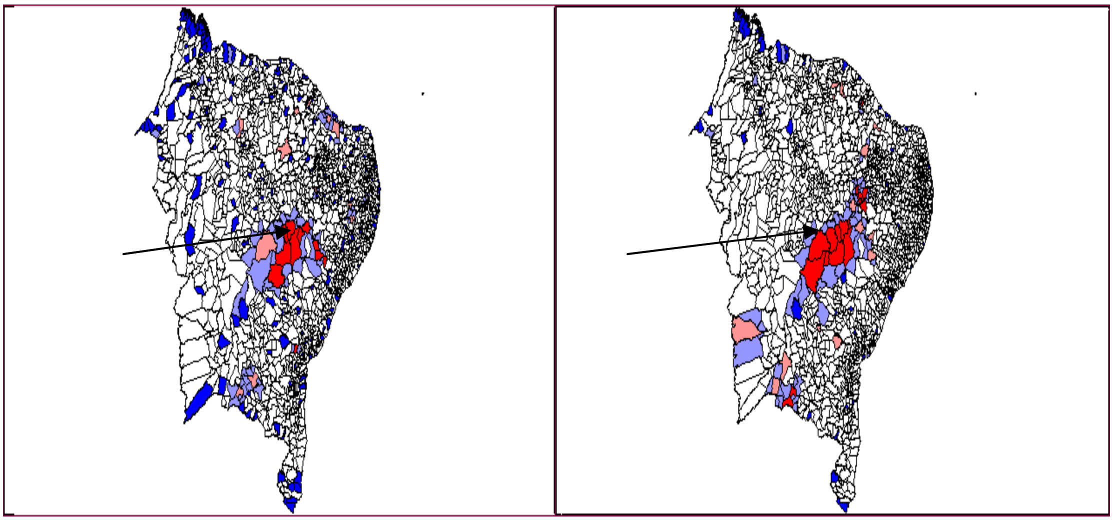


Figure 3, below, shows the cluster maps for melon yield and once more we found a cluster in the area Juazeiro/Petrolina, which visually did not change in the years 1997 and 2007. The figure also shows another cluster in the states of Rio Grande do Norte and Ceará that also remains the same through the years studied.

Figure 3. *LMI* – Melons Yield

1997 **2007**

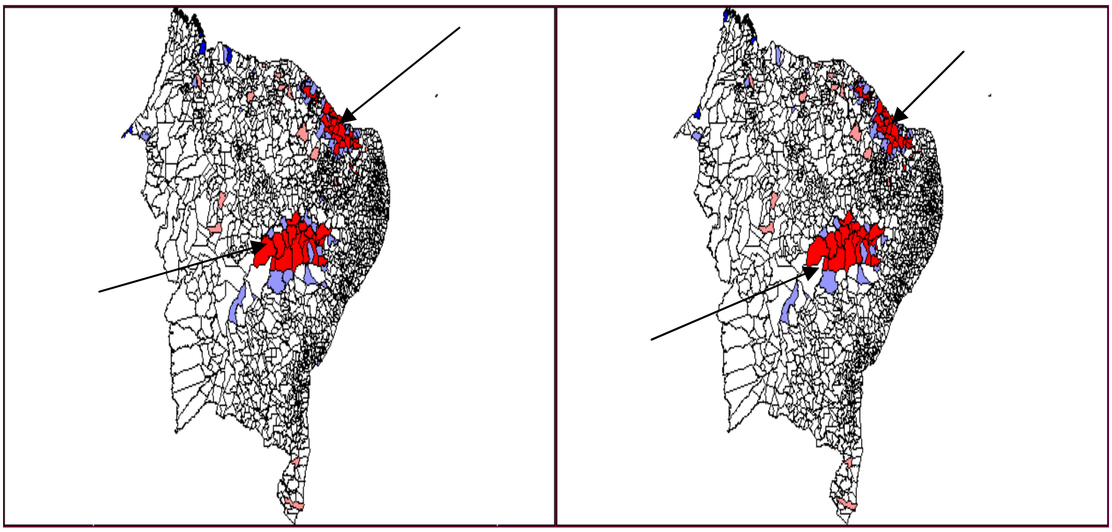
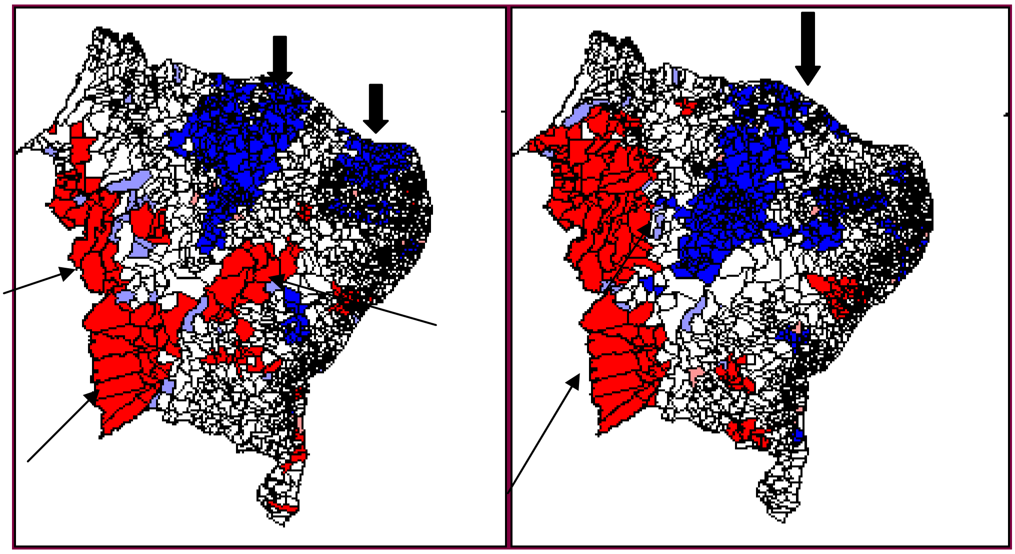


Figure 4, below, shows in 1997 two 'hot spot' for grains yield, located in the west of Bahia and the region of Juazeiro/Petrolina. There are also two 'cold spots' located in the north and northeast of the region.

Figure 4. *LMI* – Grains Yield

1997 **2007**



In 2007 the map has changed. The Juazeiro/Petrolina region was no longer a 'hot spot' of grain yield whereas the area in the west increased as a 'hot spot' and the area in the northeast coast was no longer a 'cold spot'. Through the map we observe the area with low yield of grains moving from the coast towards the inside of the region and the high yield clusters increasing all the way through the west of the Northeast region.

In conclusion, the *LMI* maps, presented in Figures 1 to 4, suggest that there are two patterns of spatial correlation of high fruits and grains yield in the Brazilian northeast. Fruits are located in the middle of the Northeast in the region of Juazeiro/Petrolina and grains are located in the west of the state. Among the fruits, we observed that the high yield clusters of mangoes are somewhat spreading while the cluster of grapes and melons stayed basically the same size in the two periods studied. The grains cluster also increased and is larger than the fruit cluster.

We now look at the three other variables at the *município* level: employment, average income and an index of human development. The main goal is to observe whether the Juazeiro/Petrolina area is also a 'hot spot' with respect to these variables. At this stage we do not intend to make any causality analysis between variables, but to explore whether more productive and high-value areas would have employees with higher incomes and higher quality of life than areas with low yield and low value of production.

Table 2. *Moran´s I*

Variables	***Moran´s I***	***Moran´s I***
Employment in the Agricultural Sector (2007)		0.23
Average Wage in the Agricultural Sector (2007)		0.21
IHD (2000 and 2005, respectively)	0.17	0.27

The results on Table 2 show a positive spatial correlation of all three variables, which, albeit low in magnitude, are statistically significant and have the expected positive relationship. In Figure 5 we notice that 'hot spots' for the variables are located in the same clusters of high fruit and grain yield, but they are smaller.

Figure 5. *LMI* - Employment and Wages in the Agricultural Sector – 2007

Employment　　　　Average Wage

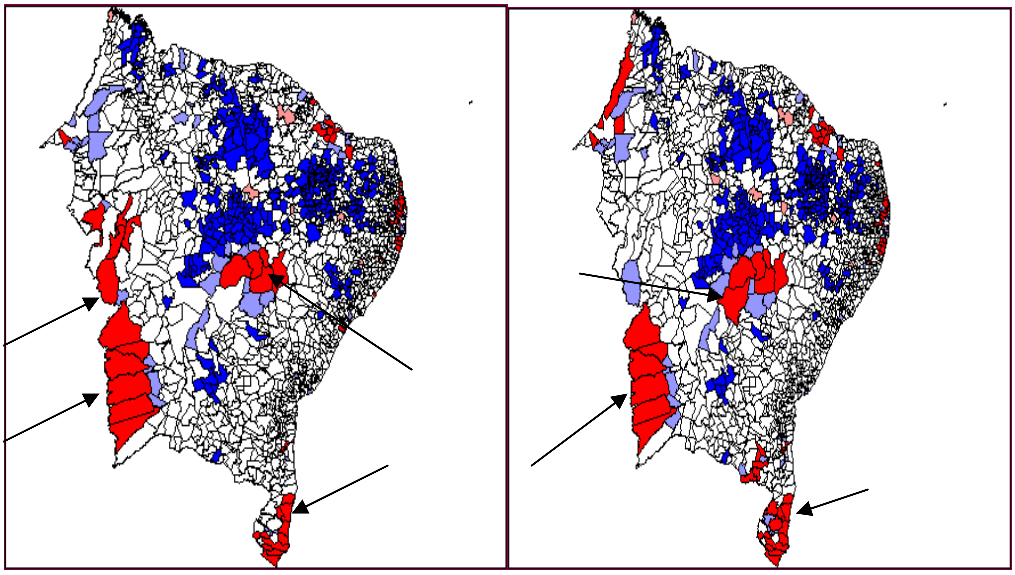


Figure 6. *LMI* - IHD

2000 **2005**

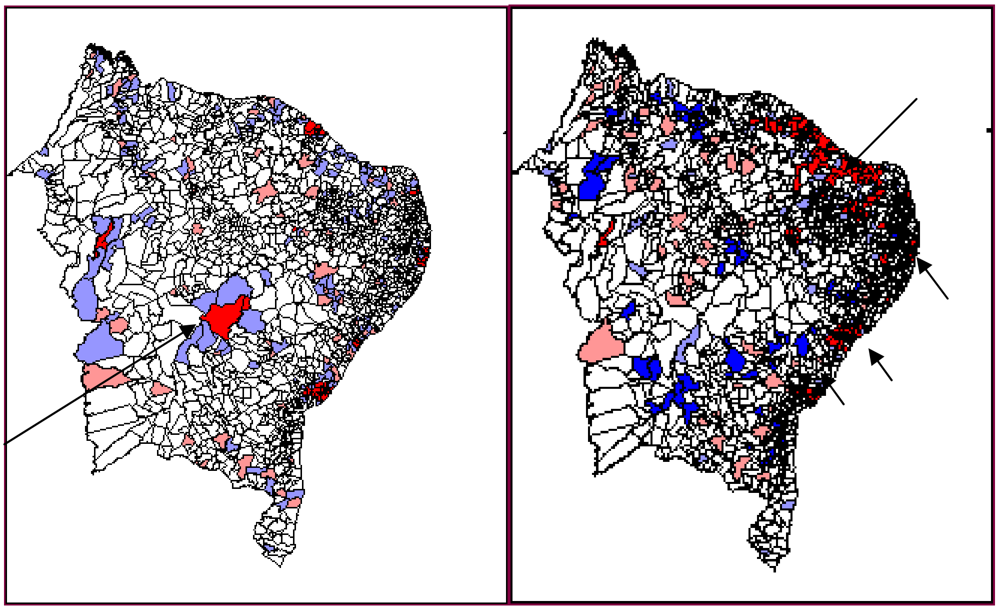


Figure 6 above shows the clusters with 'hot spots' of human development. Surprisingly only two *municípios* in the Juazeiro/Petrolina region are clustered in 2000 and none in 2005. In fact, a cluster emerges only in the very urbanized coast in the east and in some other metropolitan areas, in 2005 in particular.

7. Conclusion

This paper uses the case of the Brazilian northeast to examine spatial patterns of agricultural productivity and to identify clusters of municipalities that have been unable to increase their value of production and yield and those that have been successful in doing so. Special techniques are used to control for the potential links between area and production in identifying statistically significant agricultural productivity 'hot spots.' In this paper, we use *município*-level data to identify and analyze spatial patterns of value of production, yield and employment and average wage in the agricultural sector of the Brazilian Northeast region.

First, we found that value of production and yield is spatially auto correlated in the Northeast – i.e., observed spatial patterns of both variables are not likely to be random. More specifically, our results indicate a positive spatial autocorrelation of value of production and yield in the Northeast; with *municípios* high(low) yield and with high(low) value of production in fruits and grains relative to the state-wide mean

tended to be located near municipalities with above(below) average yield and value of production.

Looking into the local patterns of the spatial distribution of yield we discovered that *municípios* located in the region of Juazeiro/Petrolina, in the middle of the Northeast, have higher fruits yield. We also observed that among the fruits mangoes are somewhat spreading while the cluster of grapes and melons stayed basically the same size in the two periods studies. Also, whereas grapes and melons have only ‘hot spots’ clusters mangoes yield have ‘cold spots’ areas.

Looking at the *LMI*’s of grains yields in 1997 and 2007, we found an area with low grains yield moving from the coast towards the inside of the region and the high grains yield clusters increasing all the way through the west of the Northeast region. We also found that in 1997 the Juazeiro/Petrolina region was also a ‘hot spot’ of grains, but later in 2007 the ‘hot spot’ was no longer there. These results indicate a specialization of fruit production in the Juazeiro/Petrolina area.

Next, we examined the *LMI*’s for employment, average wages and an index of human development, to explore whether an area with high yields and high value of production, such as the Juazeiro/Petrolina region would also present ‘hot spots’ for the variables mentioned. We found ‘hot spots’ for employment and average wage located in the same clusters of high fruit and grain yield, but these ‘hot spots’ were smaller. In the case of the index of human development, we found only one ‘hot spot’ located in the coast.

Our results indicate that there are some specialization of fruits and grains production in the Northeast, the clusters of *municípios* that comprised high (low) yields of fruits and grains ‘hot spots’ (‘cold spots’) did not correspond to state-level boundaries and policies to improve agriculture production and productivity should take into account these spatial distribution.

Another finding was that in the areas that have clusters with high yield, not necessarily they have clusters with high employment and high average wage. In addition, when calculating the *LMI* of the index of human development, we did not find clusters in the Juazeiro/Petrolina and in the west of Northeast, areas with high yields of fruits and grains and that we expected to have IHD ‘hot spots’. We need to investigate further, but it might be the case that in these regions there is a distribution problem.

In conclusion, policy action to increase agricultural productivity has the potential to geographically target low productivity areas, but little scientific information is available to suggest whether or not such a strategy should be pursued, or to guide such targeting efforts if indeed they were determined to be useful. Future work will help determine the causal relationships between agricultural production, on the one hand, and an array of factors that could influence the productivity of agriculture (e.g., climate, access to water, access to markets, etc.), on the other; such information will be essential for developing proper policy instruments and for geographically targeting them correctly.

8. References

Amarasinghe, U., Madar S., & Markandu. A. 2005. Spatial clustering of rural poverty and food insecurity in Sri Lanka. ***Food Policy***, 30: 493–509

Andrade, M. 1980. ***Terra e o homem no Nordeste***. São Paulo: Ciências Humanas.

Anselin, L. 1996. The Moran scatterplot as an ESDA tool to assess local instability in spatial association. In M. Fischer, H. Scholten, and D. Unwin (Eds.), ***Spatial analytical perspectives on GIS in environmental and socioeconomic sciences***: 111–125. London: Taylor and Francis.

Anselin, L. 1995. Local indicators of spatial association – *LISA*. ***Geographical Analysis***, 27: 93-115.

Griffith, D. 2003. ***Spatial autocorrelation and spatial filtering: Gaining understanding through theory and scientific visualization***. Berlin: Springer.

Moran, P. 1948. The Interpretation of statistical maps. ***Journal of The Royal Statistical Society***, 10:243-251.

Palmer-Jones, R. and K. Sen. 2006. It is where you are that matters: the spatial determinants of rural poverty in India. ***Agricultural Economics***, 34: 229–242.

Perobelli, F. S., Almeida, E. S., Silva, M. & Ferreira P. 2007. Produtividade do setor agrícola brasileiro (1991-2003): uma análise espacial. ***Nova Economia***, 17: 65-91.

Pinkse, J. 2003. Moran-Flavored tests with nuisance parameters: examples. In Anselin, L., R. Florax, S. Rey (Eds.), ***Advances in spatial econometrics: Methodology, tools and applications***. Berlin: Springer.