ICAS VI
SIXTH INTERNATIONAL CONFERENCE ON AGRICULTURAL STATISTICS
2013

Typology of sugarcane production in Brazil: the use of multivariate statistics on municipal data

Maria do Carmo Fasiaben
Embrapa Informática Agropecuária
Campinas, Brasil
maria.fasiaben@embrapa.br

Aryeverton de Oliveira
Embrapa Informática Agropecuária
Campinas, Brasil
ary.fortes@embrapa.br

Alexandre Maia
Universidade Estadual de Campinas, Instituto de Economia
Campinas, Brasil
gori@eco.unicamp.br

Fábio Marim
Embrapa Informática Agropecuária
Campinas, Brasil
fabio.marim@embrapa.br

Maxwell Almeida
Instituto Brasileiro de Geografia e Estatística
Rio de Janeiro, Brasil
maxwell.almeida@ibge.gov.br

Octávio de Oliveira
Instituto Brasileiro de Geografia e Estatística
Rio de Janeiro, Brasil
octavio.oliveira@ibge.gov.br

ABSTRACT

Given the large territorial extension and the high social and economic diversity, Brazil has a remarkable variability in agricultural cropping systems. The description and the understanding of this variability is fundamental for proposing research gaps, technology transfer and appropriate public policies for the sector. Sugarcane is used for several purposes on farms, such as household

consumption, energy and sugar production, and forage production. Data collected during the 2006 agricultural census, accomplished by the Brazilian Institute of Geography and Statistics (IBGE), shows that 192,931 farms (3.7% of Brazilian farms) reported having grown sugarcane in 2006. This paper addresses the classification and characterization of the sugarcane producing municipalities in Brazil, using techniques of multivariate statistical analysis (factor and cluster analysis). The 41 variables used were created from the data collected by the 2006 agricultural census, covering 3,576 municipalities. Data went through a sugarcane filter, and was then regrouped by municipality. Those variables gather socioeconomic and technological information on the farms, such as land usage, harvested area, production goal, productivity, input usage, use of industrial wastes, irrigation, source of producer's income, percentage of the income that comes from sugarcane, family or conventional farming, size of herds, distance from the farms to sugar mills, among the most important. Analyses identified 9 different groups of sugarcane production in the municipalities, remarking large variability of sugarcane sector in Brazil, and the clear spatial differences of production and technology use in the territory. The results of the statistical analysis and the characteristics of the groups were discussed among scholars specialized in sugarcane research and were considered coherent with Brazilian reality.
Keywords: Sugarcane production, Typology of municipalities, Factor analysis, Cluster analysis

1. Introduction

Sugarcane (Saccharum spp.) is of major social and economic importance in Brazil. It is one of the most important commodities in Brazilian agribusiness, contributing to the country's energy and food security. Products as sugar, ethanol and biomass for energy are produced from sugarcane (Goldemberg, 2007). Brazilian production of sugarcane has been expanding since early-2000's, now reaching areas where it had never been planted before, mainly driven by the rise of ethanol consumption in the internal market.

For upscaling from the field to regions involves transport of knowledge between regimes of increasing spatial scale (Anderson et al., 2003). This is important because an useful application of biophysical and economic analysis is the possibility of extrapolation of field experiments for large areas, enhancing the understanding on the cropping systems and crop responses to different management strategies, for instance. To do so it is required to select the most typical cropping systems in each region based on socioeconomic and biophysical drivers, mainly for simulation and extrapolation procedures. The objective of this paper is to perform an analysis of the main biophysical and socioeconomic drivers which would be applicable for identifying typical cropping systems in each region of Brazil.

2. Methodology

2.1. Data source

This study used data from special tabulations of the Brazilian Agricultural Census 2006, conducted by the Brazilian Institute of Geography and Statistics (IBGE). In this census 5,175,636 agricultural establishments were investigated, of which 192,931 (3.7% of total) reported having grown sugarcane in 2006.

In this study, we used tabulation with aggregate data per municipality that had farms with sugarcane in 2006. Given the principle of non-identification of farms, database included 3,576 municipalities that produced sugarcane in 2006.

2.2. Typology of municipalities

In this study, factor analysis was initially applied to 41 variables derived from census data of IBGE, in order to identify factors that represent the diversity of municipalities which are producers of sugarcane. After, the factors with the most significant contributions to explain the total variability of the data were used as criteria for classification by cluster analysis. To better understand the results, we present below a brief description of the statistical techniques used and the selected variables.

2.3. Factor analysis

Factor analysis allows exploring the unknown dimensionality of observable quantitative variables. Suppose, first, a set of *n* observable variables *X*. The technique assumes that they can be expressed by linear combinations of *m* unobservable factors *F* (where m ≤ n) no self-correlated (KIM & MUELLER, 1978). In other words:

$$X_i = a_{i1}F_1 + ... + a_{im}F_m + d_i U_i \qquad (1)$$

Where the coefficients *a* inform the relationship between observable variables and the new hypothetical factors (unobservable) *F*. These factors *F* are also called common factors, and help to explain the variability of the *n* observable variables. The variables *U* refer to behavior not explained by common factors.

The central aim of the technique is to get *m* common factors *F*, that reasonably explain the total variability of the *n* observable variables *X*. Some important indicators to understand the results are (CUADRAS, 1981):

• Communality (h^2): represents the portion of the total variability of the *i-th* observable variable *Xi* explained by the *m* common factors *F*;

• Total variability explained by each factor (λ): represents the discriminatory power of the *j-th* factor in relation to all observable variables. It also can be expressed in relative terms, i.e. as a percentage of the total variability of the observable variables.

Among the techniques employed to achieve the common factors, the principal components one is often preferred due to its operational simplicity and the consistency of its results with the analytical reality (CUADRAS, 1981).

The interpretation of factors involves the analysis of their linear correlation coefficients *a*, considering their importance in predicting each observable variable. The interpretation process can be facilitated by the rotation of the factors, a linear transformation which sometimes is able to make the relationship between the factor and the observable variables more clear and objective, however without altering the explanatory power of the factors. Among the most commonly used techniques of rotation, the varimax rotation maximizes the variance of the squared coefficients (SAS, 2009). In this study, we chose the technique of principal components - from PROC FACTOR in SAS - with varimax rotation - option ROTATE = VARIMAX - which provided the results that showed more consistency with the analytical reality of the study.

2.4. Cluster Analysis

Cluster analysis aims to define hierarchical groups of observations, so that the differences between members of the same group are minimal. Early in the process, each observation represents a cluster. The two closest clusters are merged to form a new cluster that can replace it and so on, until only one remains (CRIVISQUI, 1999).

There are a number of methods that can be applied in this process and the main difference between them is how the distance (or dissimilarity) between the clusters is calculated (SAS, 2009). One of the most widely adopted methods is the *Ward* method, which employs a strategy of aggregation based on analysis of variance between and within the formed groups. This method sets hierarchical groups such that the variances within the groups are minimal and the variances between the groups are maximized. These variances can also be viewed as Euclidean distances from the center of gravity (average population) of the standardized values of *P* quantitative variables of interest.

The criterion of aggregation of each stage is to find the next class that minimizes the variability within the new group. Early in the process, we have a zero degree of generalization (all observations are distinct from each other) and at the end of the process we have 100% of generalization (all observations are similar to each other). The definition of the number of groups depends on the degree of generalization intended to adopt as well as the feasibility of analytical groups formed.

2.5. Variables employed

The observable variables used in this study are: 1. Total area of sugarcane; 2. Average area of farms with sugarcane; 3. Average area of sugarcane in farms; 4. Total area of sugarcane / explored area; 5. Total area of grassland / explored area; 6. Pasture planted area / total area of grassland; 7. Total cropland area / explored area; 8. Total area of sugarcane / cropland area; 9. Percentage of farms with sugarcane; 10. Percentage of farms specialized on sugarcane; 11. Average area of farms specialized on sugarcane; 12. Percentage of sugarcane produced that is sold; 13. Percentage of farms that fertilized sugarcane; 14. Percentage of farms that employed vinasse; 15. Percentage of farms that made irrigation; 16. Percentage of farms that made crop rotation; 17. Percentage of farms that made biological control of diseases and / or parasites; 18. Percentage of farms that performed mechanical harvesting; 19. Percentage of farms that perform manual harvesting; 20. Percentage of farms that performed mechanical and manual harvesting; 21. Percentage of sugarcane area in farms that operating mechanical harvesting; 22. Percentage of sugarcane area in farms operating manual harvesting; 23. Percentage of sugarcane area in farms operating manual and mechanical harvesting; 24. Productivity of sugarcane; 25. Amount of sugarcane produced; 26. Amount of sugarcane sold; 27. Production value of the sugarcane / total production value; 28. Value of sugarcane sold / total value of sugarcane produced; 29. Percentage of revenue that comes from agriculture; 30. Percentage of revenue that comes from vegetable production; 31. Percentage of revenue that comes from sugarcane; 32. Percentage of revenue that comes from non-agricultural activities undertaken in the farms; 33. Percentage of revenue that comes from external revenue of the farms; 34. Area of land owners / total area of farms with sugarcane; 35. Area of land tenant / total area of farms with sugarcane; 36. Average number of tractors per farm; 37. Average number of harvesters per farm; 38. Number of farms with tractor / total farms with sugarcane; 39. Number of farms with harvester / total farms with sugarcane; 40. Distance from the farms to ethanol and sugar plants; 41. Total number of cattle in farms that produced sugarcane.

3. Results

3.1. Common Factors of sugarcane production

Factor analysis was used to reduce the 41 variables described earlier to only 10 indicators (common factors). The choice of the number of factors was based on the expressiveness of the marginal contribution of each indicator to explain total variability of observable variables. The 10 common factors explained 69% of the total variability of the original variables.

The matrix of correlations between variables and factors and the variability of each variable explained by the factors is showed in Appendix 1. This correlation structure provides arguments to the data interpretation, as briefly described in Table 1.

Table 1: Common factors description.

Factors	Description
Factor 1: Farms specialized in the production of sugarcane	Factor 1 has the highest discriminatory power among the factors identified, representing 28.7% of the total variability of 41 observable variables. It has a strong positive correlation with the percentage of farms specialized on sugarcane (variable 10); percentage of value of production arising from sugarcane (variable 27); percentage of sugarcane sold (variable 28); percentage of revenue from crop production (variable 30); percentage of revenue from sugarcane (variable 31) and sugarcane yield (variable 24). It also shows medium and positive correlation for the variable crop area / area explored (variable 6) and medium and negative correlation for distance to the plant (variable 40).
Factor 2: Importance of sugarcane for the municipality	This factor accounts for 7.2% of the total variability and shows strong positive correlation with total area of sugarcane (variable 1), with the total amount of sugar cane produced and sold (variables 25 and 26, respectively) and the percentage of establishments producing sugarcane in the municipality (variable 9).
Factor 3: Prevalence of manual and mechanical harvesting	Represents 5.8% of the total variability and has strong positive correlation with the percentage of establishments that combine mechanical and manual harvesting (variable 20) and the percentage of area harvested mechanically associated with manual harvesting (variable 23). We observe strong and negative correlation for both percentages of establishments operating exclusively manual harvest (variable 19) and the percentage of area with manual harvesting (variable 22).
Factor 4 - Size of farms producing sugarcane	It accounts for 5.2% of the variability showing a strong positive correlation with the average area of farms with sugarcane, average area of sugarcane in farms and average area of farms specialized on sugarcane (variables 2, 3 and 11, respectively). These variables reflect the dimensionality of the activity of sugarcane production in the farms.
Factor 5 - Use of technology	Accounts for 4.5% of the variability of the observable variables. Presents medium to strong correlation (positive) with variables 36 to 39, which represent, respectively, the number of tractors per farm, the number of harvesters per farm, the percentage of farms with tractor and percentage of farms with harvesters, and the percentage of farms that fertilize sugarcane (variable 13). Positively correlated with farms performing crop rotation (variable 16).
Factor 6 - Areas in lease and use of vinasse	Represents 4.0% of the total variability and is strongly and positively correlated with the variables area under lease / total area of farms (variable 35) and strongly and negatively correlated with the variable area conducted by owners / total area of farms (variable 34). It also presents medium and positive correlation for percentage of farms that employed vinasse (variable 14).
Factor 7 - Importance of pasture and cattle	Accounts for 3.7% of the variability. Shows a strong positive correlation with the variables grazing area / area explored (variable 5); area with planted pasture / grazing area (variable 6) and size of herds (variable 41).
Factor 8 - Sources of revenue	Represents 3.4% of the total variability and is strongly and positively correlated to the percentage of revenues from agriculture (variable 29) and strong, but negatively correlated to the percentage of revenues from sources other than agriculture (variable 32). In this case, the smaller the value of the factor, the greater is the proportion of non-agricultural revenue for the farm.
Factor 9 - Mechanical harvest	Accounts for 3.1% of the total variability and shows a strong positive correlation with the percentage of farms that perform mechanical harvesting and the percentage of the area that had mechanical harvesting (variables 18 and 21, respectively).
Factor 10 - Proportion of the area occupied by sugarcane	Accounts for 2.9% of the variability of the observable variables. Shows a strong positive correlation with variables such as sugarcane area / area explored (variable 4) and sugarcane area / crop area (variable 8).

Source: Authors

3.2. Typology of sugarcane producer municipalities

The 10 common factors were used as criteria for grouping in the cluster analysis using the Ward’s minimum variance method. The selection of the number of groups in the analysis was based on the discriminatory power of common factors for explaining the variability of the combinations and the coherence of the clusters found.

The high heterogeneity observed in farming systems demanded to take into account large number of groups to represent a major part of the factors variability. Thus, we selected 10 clusters to represent the variability of all municipalities producing sugarcane in Brazil. In the final analysis, one cluster (cluster 10, with 5 municipalities) was disregarded because of its inconsistent data on agricultural production, probably due to errors in measurement of variables.

Table 2 shows mean values of the factors for each identified cluster. Moreover, the clusters obtained in the analysis were characterized in more detail through census database provided by IBGE, filtered for farms growing sugarcane in 2006.

The results were discussed with sugarcane experts and pointed out as representative of Brazilian reality. The types of municipalities growing sugarcane are described below.

Table 2- Mean values of the common factors

CLUSTER	1	2	3	4	5	6	7	8	9	10
Factor 1	-0,444	0,223	-0,549	-0,232	0,605	**1,114**	**1,194**	0,777	0,13	-1,315
Factor 2	-0,04	-0,139	-0,192	-0,054	-0,204	-0,558	**3,774**	0,22	1,829	-0,75
Factor 3	-0,06	-0,374	-0,069	-0,164	**2,315**	-0,213	0,27	0,083	0,304	-1,614
Factor 4	-0,135	0,036	-0,173	-0,04	-0,07	0,399	-0,46	-0,37	**8,711**	1,245
Factor 5	-0,08	-0,372	**1,282**	-0,156	-0,091	0,1	-0,084	0,13	**1,557**	-0,362
Factor 6	-0,217	-0,244	0,129	-0,079	-0,125	**3,309**	0,509	0,682	0,687	-0,671
Factor 7	**0,994**	-0,307	-0,931	0,056	-0,27	0,283	-0,19	-0,124	0,246	-0,996
Factor 8	0,357	0,328	0,28	**-2,042**	0,241	0,146	0,265	0,173	0,117	0,415
Factor 9	-0,021	-0,113	0,066	-0,054	-0,419	-0,281	-0,167	**5,803**	-0,422	-0,063
Factor 10	0,051	-0,11	-0,116	-0,053	0,02	0,259	0,181	0,195	-0,145	**21,684**
Number of occurrences	845	1237	453	469	266	119	103	56	23	5

Source: Authors

Cluster 1: Municipalities with non-specialized sugarcane farms, related to pasture and cattle production

Cluster 1 (Figure 1, Appendix 2) is composed of 845 municipalities (23.6% of occurrences). On average, farms have 122.8 ha of total area and 3.9 ha growing sugarcane. They are non-specialized sugarcane farms, where the cane is used for feedstock with yield of 25.1 ton/ha. It represents 7% of the explored area. They are primarily dedicated to livestock (76% of the area growing pastures, and 68% of these being planted pastures), and 16% is used to grow crops. In this cluster, 18% of the total area is occupied by natural vegetation. Most of the revenue comes from agriculture.

Cluster 2: Municipalities with diversified farms, lower use of technology, non-specialized in sugarcane

Cluster 2 (Figure 2, Appendix 2) is composed of 1237 municipalities (34.6% of occurrences). On average, farms have 69.7 ha of total area and 17.4 ha growing sugarcane. They are non-specialized sugarcane farms, where the cane is used for feedstock with yield of 23.1 ton/ha. These are farms with diversified production and, on average, 40% of the area is maintained for pastures (more than half is natural pasture) and 44% for crops. In this cluster, 22% of total area is occupied by natural vegetation. Most part of the revenue comes from agriculture, but other incomes play important role.

Cluster 3: Municipalities with diversified farms, greater use of technology, non-specialized in sugarcane

Cluster 3 (Figure 3, Appendix 2) is composed of 453 municipalities (12.7% of occurrences). On average farms have 140.7 ha of total area and 50.7 ha growing sugarcane, where mean yield is 29.4 ton/ha. Crops use the major part of area (59%) followed by pastures (28%). 36% of pasture area are occupied by sowed grass. About 18% of the total area is occupied by natural vegetation. More than half of the farms in this cluster reported as using crop rotation, as well as tractors and harvesters. About 90% of the revenue of the farms in this cluster come from agriculture.

Cluster 4: Municipalities with farms where external revenues are relevant

Cluster 4 (Figure 4, Appendix 2) is composed of 469 municipalities (13.1% of occurrences). On average, farms have 61.6 ha of total area and 3.9 ha growing sugarcane. They are non-specialized sugarcane farms, where mean yield is 20 ton/ha. Area is primarily dedicated to pastures (52% of the area growing pastures, and 50% of these being planted pastures), and 28% is used for grown crops. About 27% of the total area is occupied by natural vegetation. 52% of the revenue of the farms come from agriculture, and 39% come from non-agricultural activities.

Cluster 5: Municipalities with farms specialized in the production of sugarcane, manual harvesting and manual and mechanical harvesting

Cluster 5 (Figure 5, Appendix 2) is composed of 266 municipalities (7.4% of occurrences). Farms producing sugarcane in this cluster have 249 ha of total area. On average, farms have 98 ha growing sugarcane, with mean yield of 48 ton/ha. Area is primarily dedicated to crops (62%) and 31% for pastures (in which 45% are sowed grass). 13% of the total area occupied by natural vegetation. In average, 97% of the revenue of the farms comes from agriculture, and 60% from sugarcane. They are sugarcane dedicated farms managed by the owners. In 2006, harvest was mainly done using fire and manpower.

Cluster 6: Municipalities with farms specialized in the production of sugarcane, with a predominance of leases

Cluster 6 (Figure 6, Appendix 2) is composed of 119 municipalities (3.3% of occurrences). Farms in this cluster have 536 ha of total area, and in average they have 362 ha growing sugarcane, with mean yield of 72 ton/ha. Area is primarily dedicated to crops (62%), and 15% to pastures (in which 60% are sowed grass). Almost all of the revenue comes from agriculture, being 82% from sugarcane. Farms are specialized in sugarcane, but land was taken on lease (56% of the area). In 2006 manual harvesting were also predominant. 15% of farms employed vinasse.

Cluster 7: Municipalities great sugarcane producers

Cluster 7 (Figure 7, Appendix 2) represents 2.9% of total cases and has 103 municipalities. Sugarcane is clearly important for these municipalities in such way that nearly 80% of the explored area is used for sugarcane. In this cluster, 77% of farms declared being specialized in sugarcane. Farms have a mean average area of 342 ha, from which 270 ha grow sugarcane. The area occupied by natural vegetation equivalent to 7% of the total area. On average, 97% of revenue comes from agriculture, from which 89% come from sugarcane. In 2006, manual harvesting was prevalent. In this cluster, 65% of sugarcane area are managed by owners and mean yield is 74 ton / ha.

Cluster 8: Municipalities with farms specialized in the production of sugarcane, mechanical harvesting

Cluster 8 (Figure 8, Appendix 2) has 56 municipalities, which represent 1.6% of total occurrences. Sugarcane producers in this cluster have an average area of 369 ha whose average area of sugarcane has 221 ha. The crop areas represent 73% of the explored area of farms and areas of pasture, 25%. The pasture area, 48% are natural pastures. The natural vegetation occupies an area near 8% of the total area of farms. In this cluster, the agricultural income also represents virtually all of the revenue from business, 79% from sugarcane. In this cluster the areas cultivated by owners predominate (67% of the area), compared to holdings. The main feature of this cluster refers to the importance of mechanical harvesting, which responded to 61% of total area in 2006. The average yield of sugarcane is 69 ton/ha.

Cluster 9: Municipalities with large farms producing sugarcane

The Cluster 9 (Figure 9, Appendix 2) represents 0.6% of cases, with 23 municipalities. What best characterizes this cluster is the size of the farms cane producers, with 4542 hafrom which3278 ha is for growing sugarcane. Here, 92% of the explored area correspond to crops, and of these, sugarcane occupies on average 93%. The area occupied by natural vegetation equivalent to 8% of the total area. All revenue comes from agriculture, and 95% from sugarcane. In this cluster there is manual and mechanical harvesting. In this cluster, 47% of the cultivated area of sugarcane is managed by owners. The average yield of sugarcane is 75 ton/ha.

The set of all types of municipalities producing sugarcane in Brazil can be seen in Figure 10, Appendix 2.

3.3. Integrated discussion

A sugarcane production typology should consider the evolution of production systems within the context of modern agribusiness supply chains. Local productive arrangements and a systemic view of the agribusiness should be analyzed, and the aim of this section is to interpret possible relations between clusters and factors that drive the sugarcane production. Clusters present a spatial dispersion, which, in a broader sense, indicates that different production systems are under development. Some drivers and facts should be identified and emphasized.

Logistics and integrated production

One important variable in the analysis was the association of the production areas with biorefineries locations, captured by the Factor 1. Logistics costs are directly influenced by the distances from sugarcane mills, both for the application of inputs and industrial residues, like vinasse, associated with Cluster 6 and 7, and for the harvesting and transportation of sugarcane production to biorefineries. The competitiveness of the facility is also associated with locational specificities of the assets, establishing that the relationship between industry and agricultural production system should be introduced as a crucial issue in the definition of the typology. Industrial and rural production in farms follows a clear trend of integrated planning and operations. Biorefineries organize industrial and agricultural activities with an integrated analysis at strategic, tactical and operational levels. The reform of areas, the organization of harvesters and scheduling of processing on window of harvest are examples of decisions integrated in those levels. It should be considered that the capability of a supply chains coordination of agricultural and industrial operations are in evolution, and different governance structures varies through different regions.

Flexibility of sugarcane production and usage in association with other agriculture products

Those are areas of recent expansion where sugarcane occupies traditional crop fields. Factor 2 explains part of the behavior of clusters 7 and 9. It is explained by issues related with comparative advantage of alternative crops, historically established. The organization of other supply chains, as grains and oranges, shows that the proximity implies the relative importance of sugarcane over the production system but, including the association with crop rotation during reforms of the fields.

Factor 3, associated with cluster 5, approaches to some typical situations sugarcane agribusiness, where manual harvest is being applied on important production areas, as the Northeast region. The production system also coexists with other agricultural or industrial activities. This type of operation is not dominant, but for families agriculture can be an important element.

Factor 4 is also a representative association between income and other agricultural inputs also available to other activities inside farms. The cluster 9 associated with this factor probably represents the organization of production in territories occupied by other crops. There is use of crop rotation, which may also occur in the reforms of sugarcane, due to technical advantages and economic conditions to establish partnerships in the regions.

But the factor 5, associated with cluster 3, also incorporates areas where sugarcane could be produced to supply the dairy cattle, in the South Region in Brazil, as showed by figure 3. The dominance of sugarcane is not absolute, including the usage for sugar and ethanol production.

Substitution of other agribusiness systems

There was a great expansion of sugarcane over pasture areas. The national livestock, especially beef cattle, is in a development process that incorporates technologies more slowly than other sectors. The activity is in large extent based on the exploitation of pastures on natural conditions, with low investment in the maintenance of productive capacity. The activity occupies soils with bigger percentage of sand, with relative low capacity of annual crops. There is a consequent degradation of the production areas. It seems that sugarcane occupies grasslands in marginal areas within farms and, within some of the biggest farms, can occupy soils with higher capacity. Livestock can be easily moved to areas with lower aptitude to agriculture, where mechanization is impractical and soils have lower fertility, or within areas of agricultural frontier with a recent history of deforestation. It is therefore an activity that can be naturally a donor of area for sugarcane production. In expansion areas of sugarcane production, this competition should take effect. Also in this case the cane can find a way to coexist with other crops, in reform periods of plantations. Competition cane with regionally organized activities, as the orange production, should be a favorable element to regional diversification. The two crops engender long-term land occupation cycles, which limits the possibilities of coexistence with other forms of cultivation. The sugarcane may also face competition with annual crops, such as soybeans, must submit large cultivated areas to enable the production of high investments.

Search for new areas over larger farms

Factor 4 suggests also an association between farms size and areas of specialization in the production of sugarcane. There is a trend for large size farms to be associated with activities with low intensity in use of capital and labor over soils with lower capacity inside the farms. Using better soils to support crop systems, with double annual crop cycles, the conversion of production to sugarcane should encounter limits. Larger farms in the regions are interesting for the installation of new ventures where land ownership has a third part, not the plants, which can create a relationship of power not so favorable.

Trends to monoculture and their impacts

This could be one key to link sugarcane production with economic, social and environmental implications. It is an important element in the analysis of the impacts of production in national level. Due to economic advantages, a broader view of this topic associated with sustainable intensification and usage of capital in substitution of land is needed. The confluence of gains in the scale of cultivations in favorable environmental conditions also stimulates the organization of production chains around specific production systems. Logistics and vertical integration, discussed previously are linked topics.

One advantage of large scale production is the dissemination of conservation practices in terms of soil management. The biorefineries should preserve the ability of soils and adopt an effective system to ensure return on investment. Issues related to mechanized harvesting, soil conservation and best practices, investments in biorefineries, the use of straws for energy cogeneration and the application industrial residues over cultivated land, are examples of issues linked with the intensification.

Specialization in production can promotes specific technologies that save natural resources. The imposition of environmental restrictions on areas, by legal requirement, for example, can create a more diverse landscape in the areas of monoculture, and can be a crucial factor for ensuring the provision of ecosystem services. Of great environmental importance is the reduction of sugarcane burning, in terms of reducing emissions of pollutants and greenhouse gases. The introduction of plants in a region should impact positively on these elements to increase the pressure for technical and economic efficiency of production, creating an efficient life cycle.

4. Conclusion

The study highlighted the dissemination of sugarcane in the Brazilian territory. Its cultivation reflects, however, different farms' objectives and, consequently, huge differentiations in the use of production factors, technology employed and yields obtained. Production systems oriented to household consumption and forage production coexist, even in nearby regions, with integrated systems to agroindustry of sugar and alcohol.

This work, based on 2006 Agricultural Census data, shows the diversity that characterizes the production of sugarcane in the country and analyzes the spatial distribution of producing municipalities, according to the importance of the venture, the characteristics of producers, predominant production technologies, and economic outcomes. Multivariate analysis allowed classifying municipalities producers consistently with the national reality.

The typologies of farming systems help to orientate actions of research and technology transfer and the formulation of public policies related to the sector.

REFERENCES

Anderson, M.C., Kustas, W.P., Norman, J.M., 2003. Upscaling and downscaling—A regional view of the soil–plant–atmosphere continuum. *Agronomy Journal* 95, 1408–1423.

Cuadras, C. M. *Métodos de análisis multivariante*. Barcelona: EUNIBAR – Editorial Universitária de Barcelona S. A., 1981

Crivisqui, E. Presentación de los métodos de clasificación. Programa Presta, ULB, 1999.

Goldemberg, J. 2007. Ethanol for a Sustainable Energy Future. Science 315, no. 5813 (February 9): 808-810.

IBGE. Censo Agropecuário 2006. Disponível em: <http://www.sidra.ibge.gov.br/bda/pesquisas/ca/default.asp?o=2&i=P>. Diversos acessos.

Kim, J.; Mueller, C. W. **Factor Analysis**: statistical methods and practical issues. Iowa: University of Iowa, 1978.

SAS. Base SAS® **Procedures guide**. Cary, NC: SAS Institute Inc, 2009.

APPENDIX 1

Matrix of correlations between variables and factors, communalities and percentage of variability explained

Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8	Factor9	Factor10	Communalities
1	0,207	**0,842**	0,144	0,314	0,069	0,126	-0,029	0,047	0,078	0,073	0,906
2	0,093	0,133	0,094	**0,813**	0,198	0,111	0,103	0,040	0,090	-0,015	0,769
3	0,099	0,240	0,098	**0,830**	0,144	0,170	-0,015	0,023	0,063	0,091	0,829
4	0,388	0,206	0,145	0,134	0,042	0,191	-0,113	0,037	0,055	**0,748**	0,848
5	-0,422	-0,113	-0,042	-0,133	-0,077	-0,293	**0,672**	0,011	-0,039	0,122	0,770
6	-0,003	0,005	0,022	0,003	0,019	0,062	**0,617**	0,006	-0,012	-0,009	0,386
7	**0,497**	0,150	0,120	0,095	0,199	0,353	-0,547	0,098	0,063	-0,031	0,771
8	0,047	0,003	-0,024	0,056	-0,012	-0,035	0,015	0,031	0,016	**0,907**	0,831
9	0,100	**0,578**	0,119	-0,326	0,139	0,009	-0,121	-0,006	-0,012	0,073	0,503
10	**0,705**	0,256	0,294	0,066	0,110	0,158	-0,065	-0,037	0,125	0,178	0,743
11	0,098	0,279	0,071	**0,802**	0,148	0,127	0,020	0,022	0,040	0,058	0,778
12	0,869	0,068	0,131	0,112	0,055	0,059	-0,069	0,110	0,059	-0,019	0,817
13	0,035	0,114	0,141	-0,147	**0,549**	0,159	-0,397	0,085	-0,033	0,134	0,566
14	0,217	0,049	0,148	0,298	0,176	**0,454**	-0,018	-0,009	0,023	0,098	0,408
15	0,128	-0,131	0,088	0,294	-0,163	-0,098	-0,110	-0,029	-0,016	0,067	0,182
16	-0,289	-0,020	-0,037	-0,025	**0,427**	0,173	-0,414	0,081	-0,050	0,027	0,479
17	0,158	-0,044	0,005	0,169	0,307	0,252	-0,020	0,027	-0,068	0,110	0,231
18	0,153	0,057	0,260	0,103	0,163	0,073	-0,003	0,044	**0,816**	0,051	0,807
19	-0,253	-0,087	**-0,795**	-0,154	-0,161	-0,113	0,030	-0,041	-0,397	-0,045	0,929
20	0,232	0,078	**0,879**	0,135	0,107	0,100	-0,037	0,031	-0,010	0,045	0,876
21	0,127	0,114	0,046	0,033	0,085	0,047	-0,017	0,028	**0,847**	0,016	0,760
22	-0,275	-0,084	**-0,714**	-0,145	-0,150	-0,154	-0,047	-0,062	-0,342	-0,015	0,783
23	0,246	0,325	**0,721**	0,004	0,077	0,130	-0,038	0,056	-0,070	-0,001	0,719
24	**0,550**	0,204	0,230	0,030	0,321	0,292	0,114	0,027	0,128	0,240	0,673
25	0,200	**0,842**	0,133	0,332	0,105	0,146	-0,016	0,040	0,098	0,062	0,925
26	0,212	**0,831**	0,144	0,324	0,122	0,140	-0,010	0,042	0,104	0,022	0,909
27	**0,747**	0,192	0,240	0,107	0,089	0,192	0,030	-0,037	0,121	0,259	0,793
28	**0,872**	0,067	0,125	0,114	0,050	0,059	-0,086	0,103	0,060	-0,024	0,822
29	0,187	0,029	0,060	0,035	0,028	0,034	0,090	**0,919**	0,048	0,025	0,898
30	**0,683**	0,071	0,066	0,111	0,089	0,155	-0,360	0,320	0,035	-0,029	0,755
31	**0,839**	0,185	0,225	0,131	0,088	0,188	-0,043	0,103	0,111	0,133	0,892
32	-0,104	-0,044	-0,054	-0,001	-0,108	-0,054	-0,026	**-0,929**	-0,024	-0,036	0,896
33	-0,215	-0,021	0,002	0,081	-0,281	0,010	-0,195	0,068	0,088	0,003	0,183
34	-0,153	-0,136	-0,090	-0,103	-0,030	**-0,831**	0,037	-0,046	-0,079	0,029	0,763
35	0,218	0,137	0,181	0,023	0,101	**0,807**	0,003	0,060	0,062	0,049	0,771
36	0,205	0,334	0,101	0,374	**0,547**	0,036	0,026	0,002	0,149	-0,037	0,628
37	0,054	0,319	-0,005	0,091	**0,538**	-0,028	-0,019	0,000	0,226	-0,105	0,466
38	0,244	0,102	0,249	0,109	**0,704**	0,085	0,032	0,081	0,140	0,049	0,676
39	0,028	-0,006	0,120	0,193	**0,688**	0,055	-0,049	0,067	0,116	-0,022	0,549
40	**-0,406**	-0,089	-0,224	0,008	-0,130	-0,072	-0,361	-0,068	-0,084	-0,199	0,427
41	-0,081	-0,037	-0,038	0,087	-0,015	0,066	**0,588**	0,125	0,006	-0,083	0,390
Variância total (%)	28,7	7,2	5,8	5,2	4,5	4,0	3,7	3,4	3,1	2,9	
Variância acumulada (%)	29	36	42	47	51	55	59	63	66	69	

Source: Authors.

APPENDIX 2

Figure 1 – Geographic location of Cluster 1 (845 municipalities)
Source: Authors. Map prepared with Philcarto (http://philcarto.free.fr)

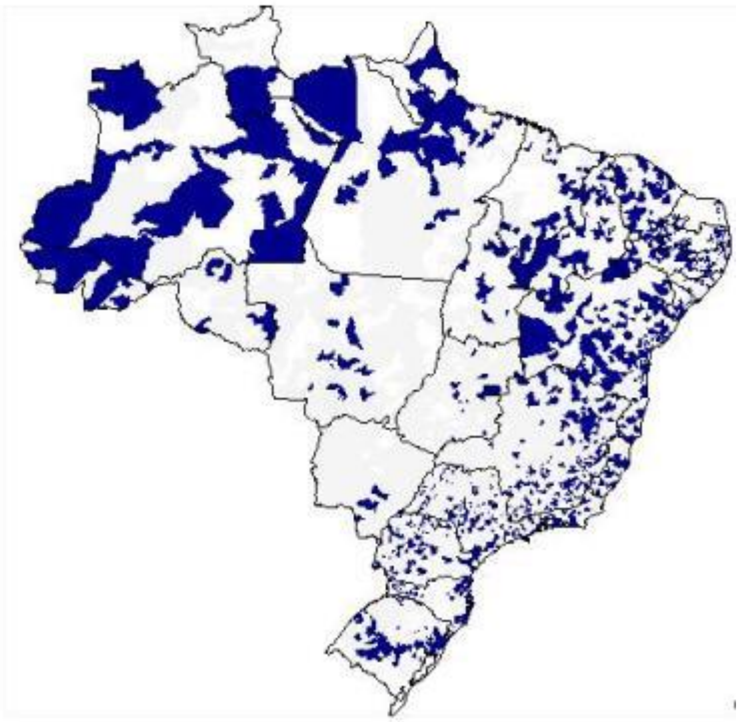

Figure 2 – Geographic location of Cluster 2 (1237 municipalities)
Source: Authors. Map prepared with Philcarto (http://philcarto.free.fr)

Figure 3 – Geographic location of Cluster 3 (453 municipalities)
Source: Authors. Map prepared with Philcarto (http://philcarto.free.fr)

Figure 4 – Geographic location of Cluster 4 (469 municipalities)
Source: Authors. Map prepared with Philcarto (http://philcarto.free.fr)

Figure 5 – Geographic location of Cluster 5 (266 municipalities)
Source: Authors. Map prepared with Philcarto (http://philcarto.free.fr)

Figure 6 – Geographic location of Cluster 6 (119 municipalities)
Source: Authors. Map prepared with Philcarto (http://philcarto.free.fr)

Figure 7 – Geographic location of Cluster 7 (103 municipalities)
Source: Authors. Map prepared with Philcarto (http://philcarto.free.fr)

Figure 8 – Geographic location of Cluster 8 (56 municipalities)
Source: Authors. Map prepared with Philcarto (http://philcarto.free.fr)

Figure 9 – Geographic location of Cluster 9 (23 municipalities)
Source: Authors. Map prepared with Philcarto (http://philcarto.free.fr)

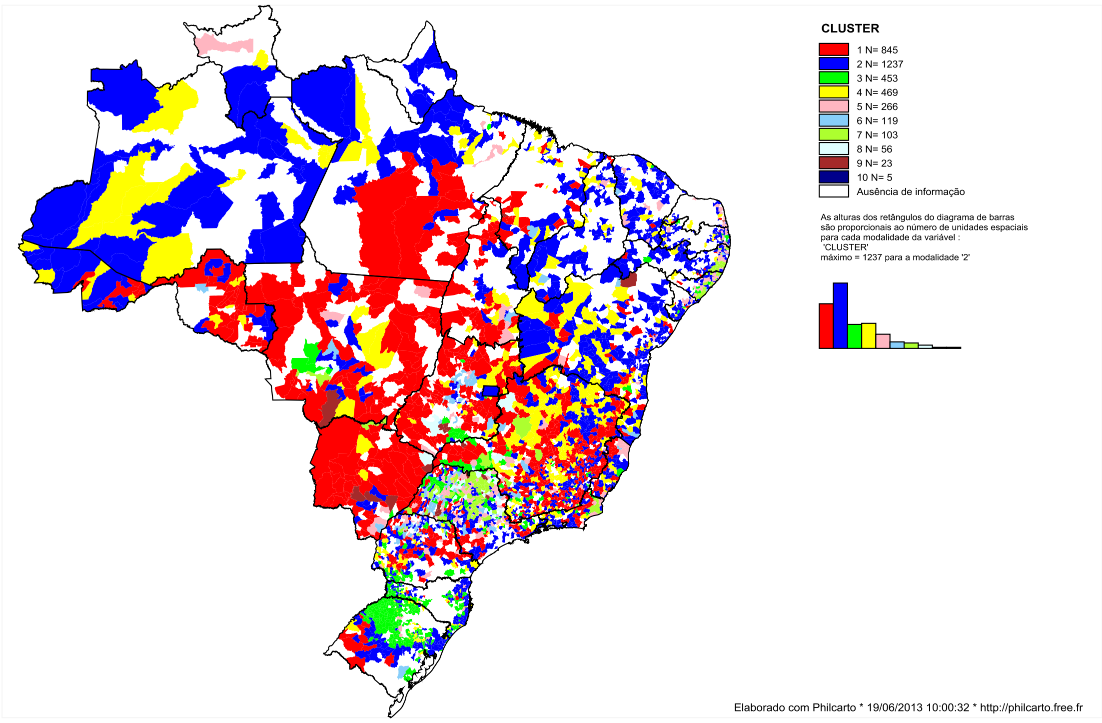

Figure 10 – Typology of Brazilian Sugarcane Producer Municipalities.
Source: Authors. Map prepared with Philcarto (http://philcarto.free.fr)