DIVERSIFICATION, MID-TERM RAINFALL AND AGRICULTURAL PRODUCTION. AN ANALYSIS OF THE BRAZILIAN NORTHEAST REGION

Pietro Andre Telatin Paschoalino[1]
José Luiz Parré[2]
André Luiz Squarize Chagas[3]

RESUMO

Os estudos sobre diversificação agrícola no território brasileiro têm ganhado maior notoriedade nos últimos anos, e se apresenta como um tema de pesquisa que precisa avançar em direção ao entendimento de como esta variável se relaciona com outros indicadores agrícolas, em particular, com a produção, principalmente nas regiões que apresentam elevadas particularidades climáticas, como a região do Nordeste brasileiro. Desta forma, o objetivo do presente estudo é avaliar a relação da diversificação agrícola e da precipitação de médio prazo, nas microrregiões nordestinas com a produção agrícola desta região, levando ainda em consideração possíveis efeitos espaciais. A estratégia metodológica incluiu estimação de modelos em dados em painel e painel espacial para as microrregiões do Nordeste brasileiro, considerando os anos 2006 e 2017. O banco de dados foi composto das informações disponibilizadas pelo Censo Agropecuário, Pesquisa Agrícola Municipal e de dados da University of East Anglia. De forma geral, a diversificação agrícola apresentou relação inversa com a produção agrícola, além disso, a precipitação de médio prazo se mostrou uma variável altamente relacionada com a produção agrícola.
Palavras-Chave: Diversificação agrícola; ruralidade; Precipitação.
Códigos Jel: Q10; Q50; Q57.

ABSTRACT

Studies on agricultural diversification in Brazil have gained greater notoriety in recent years, and it is a research topic that needs to advance towards understanding how this variable relates to other agricultural indicators, in particular production, especially in regions with high climatic particularities, such as the Brazilian Northeast. This study evaluates the relationship between agricultural diversification and medium-term rainfall in the Northeastern micro-regions and agricultural production, considering possible spatial effects. The methodological strategy included estimating panel data and spatial panel models for the micro-regions of the Brazilian Northeast, considering the years 2006 and 2017. The database consisted of information from the Agricultural Census, the Municipal Agricultural Survey, and data from the University of East Anglia. In general, agricultural diversification showed an inverse relationship with agricultural production, and medium-term rainfall proved to be a variable highly related to agricultural production.

Keywords: Agricultural diversification; Rurality; Rainfall.
JEL CODES: Q10; Q50; Q57.
Área Temática 17: Desenvolvimento rural e local

[1] Professor do colegiado de economia da Universidade Estadual do Norte do Paraná (UENP); contato: pietro.paschoalino@uenp.edu.br
[2] Professor do Departamento de Economia da Universidade Estadual de Maringá (UEM); contato: jlparre@uem.br
[3] Professor do Departamento de Economia da Faculdade de Economia, Administração, Contabilidade e Atuária da Universidade de São Paulo (FEA-USP); contato: achagas@usp.br

1. INTRODUCTION

Research in Brazil has consistently shown a decline in agricultural diversification, even at different geographic levels (PIEDRA-BONILLA et al., 2020a; PARRÉ; CHAGAS, 2022). This is a significant finding, particularly in the context of climate change, which disproportionately affects developing countries reliant on agriculture (TOL, 2018).

From a theoretical point of view, the relationship between production and diversification is multifaceted. For producers, diversification serves as a risk management tool, particularly in dry environments. However, it also comes with the trade-off of reduced expected returns due to risk reduction, despite other uncertainties such as pests, diseases, and prices (DI FALCO; CHAVAS, 2009; CULAS; MAHENDRARAJAH, 2005).

However, diversification has been gaining greater evidence in recent times as this may present economies of scope in production since diversified farms with cost complementarities can achieve greater efficiency than specialized farms. (DE ROEST; FERRARI; KNICKEL, 2018).

The urgency of studying agricultural diversification in Brazil becomes apparent when considering the need for a conclusive understanding of its impacts on agricultural production. The region under analysis, the Northeast, spans 1.561.177,8 km², with 969.589,4 km² of that falling within the semi-arid region. Many locations in this region experience rainfall between 280 and 800 mm in annual averages (DE ARAÚJO, 2011). Furthermore, the average crop production in the northeast region (analyzed by microregion) was lower in 2017 than in 2006, as revealed in the results of this study, underscoring the need for immediate attention to this issue.

Those facts demand attention since, according to (DI FALCO; CHAVAS, 2008; DONFOUET *et al.,* 2017), in the lack of rainfall, it is expected that the productivity of the agroecosystem can be increased through a higher level of crop biodiversity.

Some of the studies that attempted to verify the relationship between agricultural production or productivity and agricultural diversification in Brazil (and in the northeast region) are the studies by Paschoalino and Parré (2023), Paschoalino and Parré (2022), Parré, Chagas and Arends-Kuenning (2022), and Parré and Chagas (2022).

In general terms, Paschoalino and Parré (2023) find a negative (and significant) relationship between diversification and the value of agricultural production. However, Paschoalino and Parré (2022) found a positive (and significant) relationship between land productivity and diversification in the northeast microregions. Parré, Chagas, and Arends-Kuenning (2022) found a negative and significant relationship between Productivity (Gross Value of Agricultural Production (GVP)/Planted Area) and diversification.

Therefore, the results found for analyses in Brazil are not conclusive since one of them showed a positive relationship between the variables. The results may depend on whether the production variable is measured as total or weighted (productivity). Furthermore, there are issues yet to be analyzed in previous works; the production and diversification variable has not been elaborated in physical terms (tons), as they were measured in monetary values, nor were climate variables included as explanatory (or control) variables.

Thus, the objective of the study is to evaluate how agricultural diversification, measured through the Shannon index, based on the quantity produced per crop in tons, is related to agricultural production measured in tons in the northeast region, using panel data and spatial panel data for the years 2006 and 2017, with data at the level of Brazilian geographic microregions.

Furthermore, we seek to include the climate variable of precipitation among the explanatory variables measured in millimeters per month, computed as a 5-year average, following Piedra-Bonilla *et al*. (2020b).

The results can serve as a significant step in understanding the relationship between production and agricultural diversification in Brazil, as few studies that had Brazil as an object of study include a climate variable as a regressor using Spatial Econometrics. It is, therefore, essential for the literature, serving as a basis for subsequent studies and providing necessary answers to the public sector in Brazil since the uncertainty generated by climate change.

2. LITERATURE REVIEW

The following two subsections aim to analyze only some articles on the topic but to demonstrate the results of some relevant and actual published articles. The first subsection focuses on articles that analyze countries other than Brazil. The subsequent section focuses on results related to the topic, considering Brazil as the region of analysis, as there is a scarcity of articles on the Brazilian Northeast region.

2.1 International Studies

The studies by Di Falco and Zoupanidou (2016), Bellon et al. (2020), Ndip et al. (2023), and Kumar et al. (2024) are some of the relevant articles on the topic for different countries.

Di Falco and Zoupanidou (2016) focused on the value of gross production (revenues for all products) through unbalanced panel data at the farm level data for Italy from 1981 to 2003. In addition to crop diversification (number of crops cultivated on the farm and/or number of livestock activities), the paper also includes soil fertility quality through a soil fertility index (categorical variable) to analyze the gross production per farm. (DI FALCO; ZOUPANIDOU, 2016).

Estimations via the Arellano–Bond two-step dynamic panel data GMM estimator demonstrate that diversification (considered endogenous) is positively related to gross product, and the level of fertility of land used (if it is high) also has a positive impact. Furthermore, due to the interaction of the variables, it was possible to note that diversity becomes more important if poor fertility levels are considered (degradation, following the authors) (DI FALCO; ZOUPANIDOU, 2016).

Bellon et al. (2020) examined the relationship between crop diversity and self-consumption of food crops and cash income from crops sold by smallholder farmers in northern Ghana. The authors were interested in checking whether smallholder farmers may benefit more from a diversification or a specialization agricultural development strategy for improving their livelihoods. The study results suggest that crop diversification benefits these farmers more than specialization. Crop diversity is positively associated with the self-consumption of food crops and monetary income from crops sold. This finding suggests that increasing crop diversity opens up household market opportunities and contributes to self-consumption.

Ndip et al. (2023) explore the relationship between land fragmentation and crop diversification using survey data from Cameroon. The outcome variable of interest is crop diversification, measured by the number of crops the household grows on different plots (count). The measured fragmentation is the number of plots cultivated by the household using the Shannon-Weaver index. In addition to fragmentation, the authors used covariates that can affect diversification, such as socioeconomic characteristics variables (age of household head, gender of household head, household size, availability of alternative sources of income); farming characteristics, such as farming experience; and institutional factors, such as access to extension services. The results indicate that farmers with more fragmented lands are more likely to diversify than those with fewer. Given that most smallholder agricultural households are the primary source of home-consumed food, fragmentation ensures that they cultivate diverse crops to provide a heterogeneous food basket for the household.

Finally, Kumar et al. (2024) utilized a Panel Autoregressive Distributed Lag model approach to determine the factors affecting crop diversification in the 28 Indian states. They calculated the crop diversification index using Theil's entropy index for Indian states and found that it has risen in most of them. The authors used variables representing agricultural infrastructure (electricity, rural road density), agronomic factors (fertilizers, irrigation), agricultural landholding (cropping intensity, operational holding), economic factors (gross state agricultural domestic product), agricultural finance (credit) to test its impact on crop diversification. The analysis finds that cropping intensity, gross state domestic product, rural road density, and operational holding have led to crop diversification. In contrast, credit, fertilizer, irrigation intensity, and electricity have led to crop concentration.

2.2 Brazilian studies

As pointed out in the introduction, Piedra-Bonilla et al. (2020a) verified the diversification of agricultural production in Brazilian municipalities through the Shannon index of production value, using data from the Municipal Agricultural Survey between 1987 and 2017, it can be stated that agricultural production in Brazil presents low diversification. Besides that, over time, it has become more and more specialized, as there was a drop in the Shannon index during that period.

Paschoalino and Parré (2023), through spatial regressions at the level of microregions for the year 2017, found a negative relationship between diversification and agricultural production when using as a dependent variable the sum of the value of agricultural production of permanent and temporary crops (R$ thousand) and as a measure of diversification the Shannon index based on the area planted or intended for harvesting (ha) or based on the value of production (R$ thousand).

Paschoalino and Parré (2022) verified the relationship between agricultural diversification and land productivity in the Northeast microregions using regression via panel data for the years 2006 and 2017, in which land productivity was used as the dependent variable (value of the production of temporary and permanent crops divided by the area harvested from such crops, in each microregion). Diversification was measured using the Shannon index based on the area planted or destined for harvesting the 64 crops used in the research. In general, they found a positive relationship between the variables.

Parré, Chagas and Arends-Kuenning (2022) used the Simpson index as a measure of diversification based on the Gross Value of Production (GPV) of temporary and permanent crops, horticulture, forestry, and livestock (Gross Value Sold of heads of cattle, pigs, and poultry), and evaluated how the variables farm size and farmland use are related to this measure, through regressions by spatial panel data using data at the Minimum Comparable Areas (MCA) level for the years 1996, 2006 and 2017. Among the explanatory variables used in the study, the authors used productivity – Gross Value of Agricultural Production (GVP)/Planted Area. Through the regressions, it was found that there was a negative and significant relationship between diversification and

Piedra-Bonilla et al. (2020b), with an ordered probit for the year 2006, verified the influence of climate variability on the probability of a municipality being classified with higher categories of diversification. The diversification was measured by the Simpson Index constructed with the Gross Value of Production (GVP) of each crop in the municipality, dividing this diversification into four categories and using values of the 5-year moving average of temperature and precipitation (summer and winter seasons and their variability) as explanatory variables. The authors find that the effects of increased temperature and precipitation presented ambiguous results on diversification. However, the greater the variability of precipitation and temperature, the greater the probability of the municipality being classified as very diversified.

3. METHODOLOGY

3.1 Variables and Data

The study aimed to estimate the following agricultural production function:

$$Q_{it} = f(L_{it}, K_{it}, A_{it}, D_{it}, P_{it}, C_{it}) \quad (1)$$

Where Q_{it} is the sum of the quantities produced in each micro-region[4] i, in each year t, expressed in tons of 63 agricultural products from temporary and permanent crops set out in the Municipal Agricultural Survey (PAM)[5].

Where L_{it} means the total number of personnel occupied in agricultural establishments for each year according to the agricultural census (December 31, 2006, and September 30, 2017). K_{it} is the total number of existing tractors in agricultural establishments in each year. A_{it} is the total area of establishments in the microregion (in hectares). D_{it} is the Shannon index used to measure the diversification of the crops in each microregion and year.

The precipitation is measured in millimeters per month,[6] and to compute the mid-term impact on the production of crops, as pointed out in Piedra-Bonilla *et al.* (2020b), it is necessary to calculate an average for a more extended period, specifically in this article, P_{it} represents a 5-year average for each year (2006 or 2017) in each microregion. Thus, the precipitation for a given microregion in 2006 was obtained through the average precipitation from January 2002 to December 2006 (monthly values), while for 2017, it was obtained by the average of the values from January 2013 to December 2017. So, how the values were measured in mm/month, and the final value is an average of 60 months; the results are also in mm/month. Finally, C_{it} is the vector of the establishment control variables (socioeconomic variables). In this case, two variables were used, and they were the proportion of establishments that received some technical guidance about the total number of establishments in the micro-region and the proportion of establishments in the

---

[4] This type of variable, instead the variable in monetary values was used in Di Falco et al. (2010). In Brazilian studies in the agricultural economics field, it also stands out in Perobelli et al. (2007) and Antunes and Stege (2020), despite being counted as productivity (dividing them by the area).

[5] As the pineapple crop is counted as a thousand fruits instead of tons, the conversion described in Perobelli et al. (2007) was applied, and the conversion factor are equal to 1,81, thus, to find the value in tons, the units in thousand fruits were multiplied by 1.81. Coconut (coco-da-baia) was disregarded due to the absence of a conversion factor.

[6] The precipitation data were obtained by the gridded time-series dataset (CRU DATA) version 4.06 from the University of East Anglia with a resolution of 0.5º x 0.5º. To extract the precipitation values in each microregion polygon, the software R was used, by the exact_extract function with the "mean" argument through the exactextract package.

micro-region under analysis with managers (producer or administrator) aged 55 years or more, about the total number of establishments.

About the D_{it} variable, the diversification index, it was used the Shannon index, which, according to Magurran (1988):

$$S = -\sum_{i=1}^{s} p_i \ln p_i \tag{2}$$

Where p_i is the proportion of the produced crop i in tons in the microregion, the index was not calculated by the planted area because, following Piedra-Bonilla *et al.* (2020b), there are successive or/and simultaneous crops in the same year and place. This fact can generate planted areas that exceed the geographical area of the micro-region, in addition to the fact that the variable planted area may contain measurement errors.

It is important to highlight that, unlike other indexes, the Shannon index was chosen because it is sensitive both to the increase in the number of crops as well as to the uniformity of the different crops planted (DI FALCO; CHAVAS, 2008), in addition to be used extensively in the literature (DI FALCO E CHAVAS, 2008; DONFOUET ET AL., 2017).

Following Di Falco et al. (2010), the chosen variables follow the literature that considers production as a function of inputs, where socioeconomic and physical characteristics are also generally included.

3.2 Method and empirical model

Panel data were used to estimate the production function indicated in (1) for the microregions of the Brazilian Northeast region since the data are from individuals in more than one year (2006 and 2017).

The model was estimated using 187 microregions (Northeast region of Brazil) as individuals in two years, generating 374 observations and a balanced panel. In that text, the estimated model can be described as:

$$lnQ_{it} = \beta_L lnL_{it} + \beta_K lnK_{it} + \beta_A lnA_{it} + \beta_D D_{it} + \beta_P lnP_{it} + \beta_Y Year_t + \beta_O lnOri_{it} + \beta_{Ag} lnAge55_{it} + \alpha_i + u_{it} \tag{3}$$

The variables were discussed in the previous section but summarized in Table 1. Furthermore, spatial panel estimation was also used for equation 3, given the possibility of spatial autocorrelation.

Table 1. Variables used in the empirical model.

Variable	Resumed Meaning	Source
lnQ_{it}	Natural logarithm of the total produced quantity of 63 crops (temporary and permanent) in each microregion, expressed in tons.	Municipal Agricultural Survey (PAM).
lnL_{it}	Natural logarithm of total personnel occupied in agricultural establishments in each microregion for each year according to the agricultural census.	Agricultural Census (2006) and Agricultural Census (2017).
lnK_{it}	Natural logarithm of number of existing tractors in agricultural establishments in each microregion in each year[7].	Agricultural Census (2006) and Agricultural Census (2017).
lnA_{it}	Natural logarithm of the total area of establishments in the microregion (in hectares) in each year.	Agricultural Census (2006) and Agricultural Census (2017).
D_{it}	Shannon index measured by quantity produced (tons).	Municipal Agricultural Survey (PAM).
lnP_{it}	Natural logarithm of precipitation for each microregion measured by five-year average (monthly data expressed in mm/month) of 2002-2006 or 2013-2017.	CRU DATA – TS4.06 – Data from the University of East Anglia.
$Year_t$	Time fixed effect of 2017 year – Dummy variable equal one if the year is 2017 and 0 if year is 2006.	-
$lnOri_{it}$	Natural logarithm of the proportion of establishments that received some technical guidance about the total number of establishments in the microregion in each year.	Agricultural Census (2006) and Agricultural Census (2017).
$lnAge55_{it}$	Natural logarithm of the proportion of establishments in the microregion under analysis with managers (producer or administrator) aged 55 years or more, about the total number of establishments in each year.	Agricultural Census (2006) and Agricultural Census (2017).

Source: The authors.

This section will highlight the equation for the fixed effects model because it was the most empirically appropriate for the models. Models can be the spatial autoregressive model (SAR), spatial error model (SEM), or spatial autoregressive lag and error model (SARAR) (INSEE, 2018). Equation 4 demonstrates the SAR model for panel data with fixed effects (INSEE, 2018).

$$lnQ_{it} = \lambda \sum_{i \neq j} w_{ij} y_{jt} + x_{it}\beta + \alpha_i + u_{it} \tag{4}$$

Where the explanatory variables of equation (4) are represented by k vectors x_{it} of dimension $(1, k)$ with the parameters to be estimated represented by vector β with dimension $(k, 1)$. Furthermore, w_{ij} is part of a spatial weighting matrix W_N of dimension $(N, N;)$ therefore, this model includes the lag of the dependent variable $(\sum_{i \neq j} w_{ij} y_{jt})$ (INSEE, 2018).

In turn, equation 5 demonstrates the SEM empirical model for equation (3) considering fixed effects (INSEE, 2018).

$$lnQ_{it} = x_{it}\beta + \alpha_i + u_{it} \tag{5}$$
$$u_{it} = \rho \sum_{i \neq j} w_{ij} u_{jt} + \varepsilon_{it}$$

[7] It was necessary to add one to the variable, after building the panel (one was added in the two years of data) since there was a value of 0 in one microregion in 2017, making it impossible to take a logarithm.

Therefore, $u_{it} \sim IID\ (0, \sigma^2)$. In this case, the spatial autoregressive error term ($\rho \sum_{i \neq j} w_{ij} u_{jt}$) captures the spatial interaction. Finally, the SARAR model applied to equation (3) using fixed effects can be represented by equation (6) (INSEE, 2018).

$$lnQ_{it} = \lambda \sum_{i \neq j} w_{ij} y_{jt} + x_{it}\beta + \alpha_i + u_{it} \tag{6}$$
$$u_{it} = \rho \sum_{i \neq j} w_{ij} u_{jt} + \varepsilon_{it}$$

With $\varepsilon_{it} \sim IID\ (0, \sigma^2)$, this model captures the spatial interaction both through the spatial autoregressive error term and through the lag of the dependent variable. All three models were estimated by maximum likelihood.

4 RESULTS AND DISCUSSION

4.1 Descriptive statistics

This section presents the descriptive statistics of the variables used in the study. Table 2 presents the standard deviation and average for 2006 and 2017 and considers the pooled data of the variables used in the empirical models (logarithm) and level.

Table 2. Descriptive statistics.

	2006		2017		Panel	
Statistic	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Q_{it}	506,843.70	1,191,738.00	435,944.90	971,447.90	471,394.30	1,086,309.00
L_{it}	41,171.44	32,919.82	34,100.18	27,937.94	37,635.81	30,694.57
A_{it}	406,814.60	444,547.70	379,111.50	471,897.30	392,963.00	458,021.60
$D_{it}(quantity)$	1.3	0.57	1.26	0.59	1.28	0.58
Ori_{it}	0.11	0.08	0.1	0.07	0.11	0.07
$Age55_{it}$	0.38	0.05	0.45	0.05	0.42	0.06
P_{it}	79.45	30.96	65.28	27.97	72.37	30.3
K_{it}[8]	334.91	478.32	449.47	673.12	392.19	585.93
lnQ_{it}	11.9	1.59	11.37	1.95	11.64	1.79
lnL_{it}	10.29	0.89	10.11	0.85	10.2	0.87
lnA_{it}	12.38	1.11	12.23	1.16	12.31	1.13
$lnOri_{it}$	-2.42	0.69	-2.5	0.66	-2.46	0.67
$lnAge55_{it}$	-0.97	0.14	-0.8	0.11	-0.89	0.16
lnP_{it}	4.31	0.37	4.1	0.4	4.2	0.4
lnK_{it}	5.27	1.06	5.48	1.17	5.37	1.12

Source: Authors based on Municipal Agricultural Survey (PAM) and Agricultural Census.

[8] Number of tractors are considering one more unit in both years, this question was necessary to transform the data into logarithm, due to the zero number of tractors in one microregion in 2017.

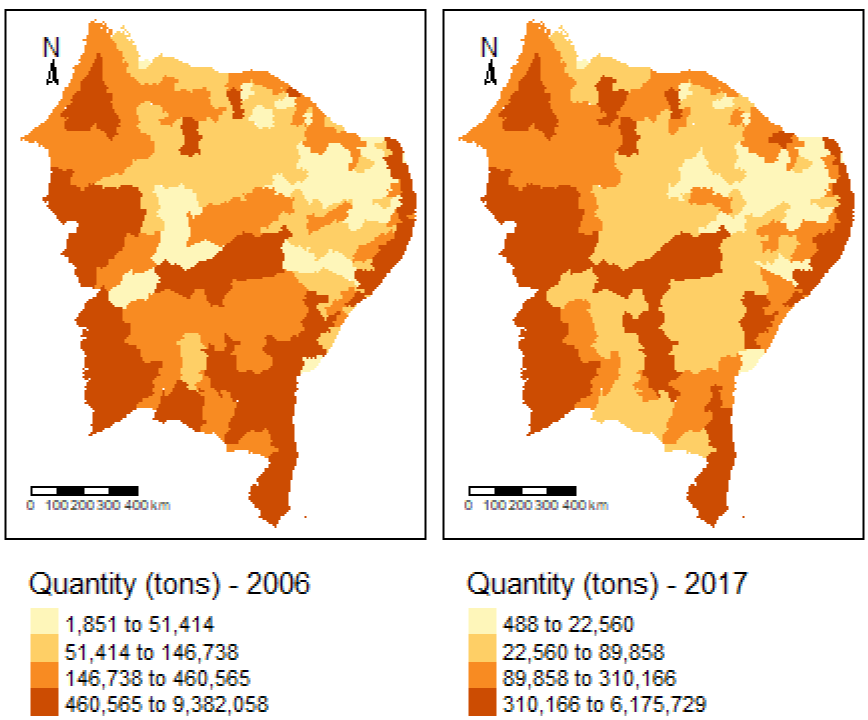
According to Table 2, it is possible to verify some interesting questions about the behavior of the data between years. Firstly, the average production of the analyzed crops decreased from one period to the next; that is, the total production of the analyzed crops decreased, which becomes highly worrying.

It is also interesting to note that the diversification index decreased from one period to another. The average rainfall in micro-regions was also reduced. Remember that each microregion's precipitation averages five years of monthly precipitation. This shows that, on average, precipitation has decreased when considering a medium term, which may be due to climate change. This result may be associated with a reduction in production.

Finally, it is noted that despite the slight increase in the average number of tractors per microregion, both the number of employed personnel occupied and the total area of establishments decreased, showing less use of inputs for agricultural production.

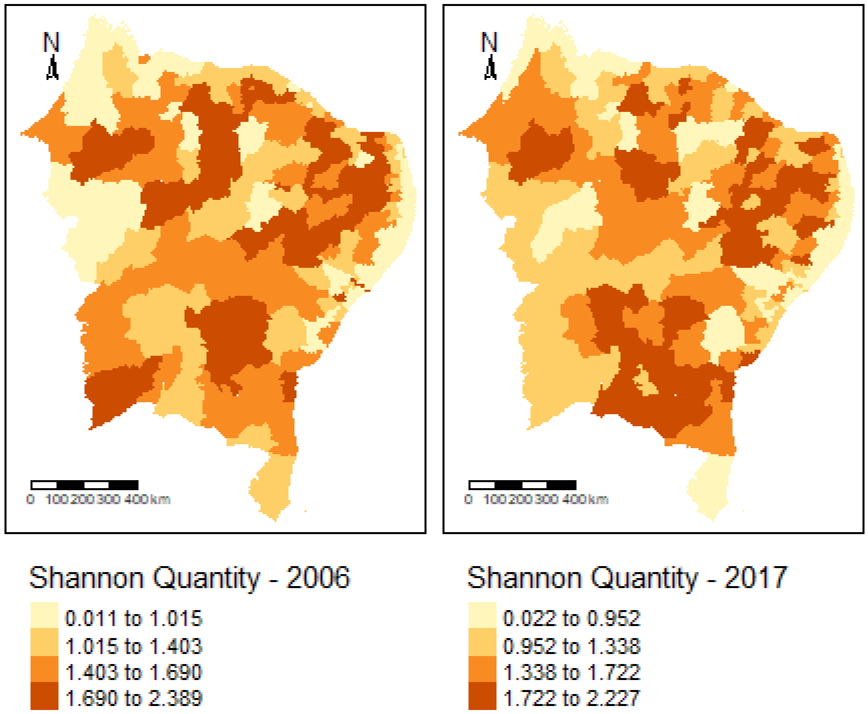
Quantile maps were generated to verify the variables' spatial distribution. Figure 1 shows the distribution of agricultural production in the microregions of Northeastern Brazil. Figure 2 shows the distribution of the agricultural diversification indicator, while Figure 3 shows the distribution of precipitation.

Figure 1. Map with quantiles of agricultural production in the northeast region of Brazil (2006 and 2017).



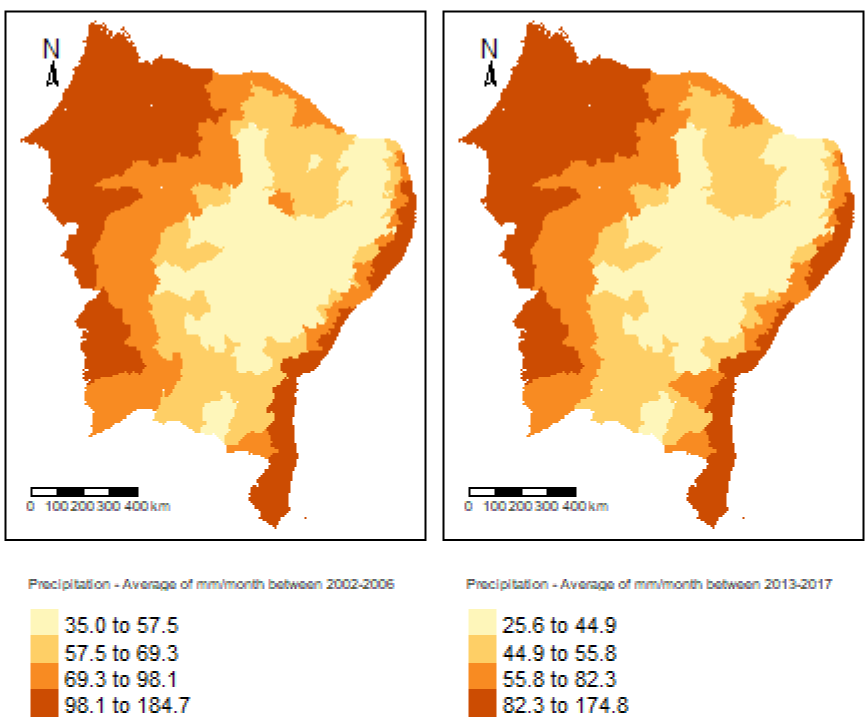
Source: Authors based on Municipal Agricultural Survey (PAM).

Figure 2. Map with Shannon index quantiles in Brazil's northeast region (2006 and 2017).



Source: Authors based on Municipal Agricultural Survey (PAM).

Figure 3. Map with precipitation quantiles based on a five-year mm/month average in the northea st region of Brazil (2006 and 2017).



Source: Authors based on CRU DATA.

These maps are best analyzed together. For example, it is possible to verify that regions with greater agricultural production can be related with various regions with a high volume of precipitation, with the opposite also being true (as in the semi-arid region).

Furthermore, it is also noted that regions with greater production (such as the coast and others) have less agricultural diversification (but only sometimes).

Once we have discussed the descriptive statistics, the following subsection shows the results obtained with the empirical estimations.

4.2 Results

Firstly, Table 3 shows the estimations of the quantity produced concerning the previously presented variables. The results include the estimation of pooled, random, and fixed effects.

It's verified that capital and labor positively correlate with the quantity produced, but only capital keeps its significance in the fixed specification. Precipitation also presents the expected sign, and an increase of one percent in monthly medium-term precipitation is related to a more than a 3% increase in the quantity produced. It is also verified that diversification presents a negative sign on production and that results are significant statistically.

The tests to define the best model were implemented using the software R (Pftest, plmtest, and Hausman), with the fixed effects model being the most recommended. However, heteroscedasticity was verified, and the covariance matrix estimation was carried out using the Arellano method; the results are displayed in Table 4.

Although the variables of interest maintain statistical significance, we verified through Table 3 an inconclusive result about cross-sectional dependence (Breusch-Pagan LM test and Peasaran CD test) and the presence of spatial correlation. Therefore, Table 5 brings the Lagrange and Spatial Hausman multiplier tests.

As can be seen, the robustness of the LM test for spatial error was not significant. Therefore, the SAR model is the most appropriate. Furthermore, the SAR model was estimated by considering both fixed and random effects, and using the Hausman test, it was stated that it is best to consider fixed effects. The results of the Sar model estimations with fixed effects are analyzed in Table 6.

Table 3. Estimations of lnQ_{it} model with Shannon-Quantity.

	Pooled	Random Effects	Fixed Effects
	(1)	(2)	(3)
lnL_{it}	0.47^{***}	0.42^{***}	0.32
	(0.10)	(0.12)	(0.25)
lnK_{it}	0.71^{***}	0.60^{***}	0.24^{**}
	(0.08)	(0.08)	(0.12)
lnA_{it}	-0.12	-0.06	0.07
	(0.08)	(0.09)	(0.18)
D_{it}	-1.34^{***}	-1.23^{***}	-1.02^{***}
	(0.12)	(0.11)	(0.15)
lnP_{it}	1.29^{***}	1.46^{***}	3.05^{***}
	(0.17)	(0.21)	(0.90)
$Year_t$	-0.52^{***}	-0.62^{***}	-0.21
	(0.13)	(0.10)	(0.26)
$lnOri_{it}$	0.06	-0.13	-0.26^{***}
	(0.10)	(0.08)	(0.10)
$lnAge55_{it}$	0.82^{*}	1.67^{***}	1.68^{***}
	(0.47)	(0.45)	(0.63)
Constant	1.98^{*}	1.78	
	(1.11)	(1.33)	
Observations	374	374	374
R^2	0.69	0.62	0.53
Adjusted R^2	0.68	0.61	0.02
F Statistic	101.30^{***} (df = 8; 365)	587.58^{***}	25.06^{***} (df = 8; 179)
Hausman			Chi² = 30.51; p-value = 0.00
BP test for heteroscedasticity			BP= 374; p-value = 0.00
BP LM test for cross-sectional dependence			Chi² =34782, p-value = 0.00
Pesaran CD test for cross-sectional dependence			z = -0.10, p-value = 0.92
Randomized W test for spatial correlation of order 1.			p-value = 0.08

Source: Authors based on research data.
Notes: *p<0.1; **p<0.05; ***p<0.01; Standard error in parentheses; RW test realized with queen contiguity matrix.

Table 4. Fixed effects with Arellano Heteroscedasticity-Consistent Covariance Matrix Estimation.

Variable	Fixed Effects
lnL_{it}	0.32
	(0.20)
lnK_{it}	0.24^{**}
	(0.10)
lnA_{it}	0.07
	(0.15)
D_{it}	-1.02^{***}
	(0.18)
lnP_{it}	3.05^{***}
	(0.91)
$Year_t$	-0.21
	(0.27)
$lnOri_{it}$	-0.26^{***}
	(0.09)
$lnAge55_{it}$	1.68^{**}
	(0.74)

Source: Authors based on research data.
Notes: *p<0.1; **p<0.05; ***p<0.01; Standard error in parentheses

Table 5. Lagrange Multiplier test and Spatial Hausman test for specification.

	Statistic	P-value
Test for Spatial lag dependence	63.30	0.00
Test for Spatial error dependence	42.79	0.00
Robust test for spatial lag dependence	22.25	0.00
Robust test for spatial error dependence	1.73	0.19
Spatial Hausman	37.30	0.00

Source: Authors based on research data.

Table 6. SAR fixed effects model.

	Estimate
lnL_{it}	0.34**
	(0.16)
lnK_{it}	0.17**
	(0.08)
lnA_{it}	-0.11
	(0.12)
D_{it}	-0.97***
	(0.09)
lnP_{it}	1.57***
	(0.58)
$Year_t$	-0.14
	(0.16)
$lnOri_{it}$	-0.18***
	(0.06)
$lnAge55_{it}$	0.96**
	(0.40)
λ	0.48***
	(0.06)

Source: Authors based on research data.
Notes: *p<0.1; **p<0.05; ***p<0.01. Standard error in parentheses.

Thus, the significant variables in the a-spatial model continue to present statistical significance (although such results are not corrected for heteroscedasticity). The marginal effects are necessary to correctly analyze the results, showing the indirect, direct, and total impacts shown in Table 7.

Table 7. Impacts of SAR FIXED Effects specification.

	Direct	p-value	Indirect	p-value	Total	p-value
lnL_{it}	0.49	0.03	0.14	0.03	0.64	0.03
lnK_{it}	0.24	0.04	0.07	0.05	0.32	0.04
lnA_{it}	-0.16	0.37	-0.05	0.37	-0.21	0.37
D_{it}	-1.43	0.00	-0.42	0.00	-1.85	0.00
lnP_{it}	2.32	0.01	0.68	0.02	3.00	0.01
$Year_t$	-0.21	0.38	-0.06	0.38	-0.28	0.38
$lnOri_{it}$	-0.26	0.00	-0.08	0.01	-0.34	0.00
$lnAge55_{it}$	1.42	0.02	0.42	0.03	1.84	0.02

Source: Authors based on research data.

We can see that the total impacts of the variables for Labor, capital, precipitation, and experience of microregions showed a positive and significant sign. At the same time, diversification and technical orientation had a total negative and statistically significant impact.

4.3 Discussion

The two main results of the previous section can be highlighted as follows: A highly positive relationship between medium-term precipitation and production and a negative relationship between diversification and production.

Regarding diversification, the results align with Paschoalino and Parré (2023) for Brazil while disagreeing with Paschoalino and Parré (2022). Even though the last of the two also considered the Northeast region, they used land productivity as the dependent variable, and this fact may be one of the explanations for the positive relationship with diversification.

Despite the dependent variable being the diversification index, Parré, Chagas and Arends-Kuenning (2022) also found a negative relation between the variables considering Brazil, and Kidane and Zegeye (2018) after verifying that diversification was endogenous, also found a negative relationship between the variables (for Ethiopia), however, not statistically significant and the possible reason for the negative signal is that diversified systems are complex to manage and requires correct skills and input compared to the specialized systems.

Despite this, this result does not mean that diversification should be something negative for farmers. Firstly, the results are aggregated at the level of geographic micro-regions, and the results may be presented differently if data were available at the Farm level.

Furthermore, the evaluated result only shows the relationship between quantity (tons) and diversification. Revenues or expenses are not considered, nor is diversification's impact on the variability of these revenues. Therefore, the result needs to be analyzed carefully, even because the diversification carried out by the farmers may need to be correctly taking advantage of the possibilities of economies of scope (the mix of crops used could have been better).

Still, the results are essential in showing that at a more aggregated geographic level, diversification does not present a different relationship in the northeast region about what was found for the rest of the country in Paschoalino and Parré (2023), even with its edaphoclimatic characteristics of that region.

Furthermore, the precipitation result is important, showing that climate limitation (precipitation) is directly related to agricultural production in the region. Donfouet et al. (2017) found a positive and significant relationship between the annual rainfall (over 30 years) and crop production in France. Furthermore, turn on a red light and such results need to be analyzed with other studies showing climate change's impact in Brazil's northeast region.

These studies include Araújo et al. (2014) and Martins et al. (2019). Martins et al. (2019) used simulation models considering different levels of Co2 emissions in maize yields in northeast Brazil. In general, they found that relative to rainfed agriculture, the drop in productivity is significant (mainly at the end of the century), and to sustain the current level of productivity, it is necessary to use irrigation that would significantly increase the amount of water needed.

Araújo et al. (2014) show through a Tobit model that temperature and precipitation levels were important in explaining the productivity of cassava, corn, and sugar cane crops in the northeast region. Furthermore, climate projections from the third IPCC report verify how temperature changes can affect the productivity of such crops, stating that, in general, the productivity of such crops will be lower than what could be achieved if the climate projections prove to be correct.

CONCLUSION

The present study started by observing that agricultural diversity is decreasing in Brazil. It raised the possibility that this fact could have even more significant impacts in the Northeast region of Brazil since it includes the Brazilian semi-arid region, with annual rainfall up to 800 mm.

Furthermore, the literature that analyzed the effect of agricultural diversification in Brazil should have considered the control for precipitation in its econometric models, nor did it even consider production in terms of quantity.

Therefore, the article uses econometrics and spatial econometrics to reduce the gaps in the literature on the topic in Brazil.

According to the results, agricultural diversification was inversely related to agricultural production, showing that the effects of specialization are essential. Furthermore, it highlighted the role of rainfall (medium-term monthly average) on agricultural production, providing information that, together with other studies that predict rainfall, can indicate how agricultural production will behave in the face of climate change.

For future studies, the importance of the availability and use of micro-data is highlighted, which would make it possible to find instruments for the diversification variable, considering it endogenous.

REFERENCES - APA

Antunes, L., Stege, A. L. (2020). Análise espacial da convergência da produtividade agrícola brasileira (1990-2018). *IN*: XXIII encontro de economia da região sul.

Araújo, P. H. C.; Silva, F. D. F.; Gomes, M. F. M.; Féres, J. G.; Braga, M. J. (2014). Uma análise do impacto das mudanças climáticas na produtividade agrícola da região nordeste do Brasil. *Revista Economica do Nordeste*, 45(3), 46-57.

Bellon, M.R.; Kotu, B.H.; Azzarri, C.; Caracciolo, F., (2020). To diversify or not to diversify, that is the question. Pursuing agricultural development for smallholder farmers in marginal areas of Ghana. *World Development*, V. 125, https://doi.org/10.1016/j.worlddev.2019.104682.

Culas, R.; Mahendrarajah, M. (2005). Causes of diversification in agriculture over time: evidence from Norwegian farming sector. In: *CONGRESS OF THE EUROPEAN ASSOCIATION OF AGRICULTURAL ECONOMIST*S, 11., 2005, Copenhagen. Proceedings. Copenhagen: EAAE.

De Araújo, S. M. S. (2011). A região semiárida do nordeste do Brasil: questões ambientais e possibilidades de uso sustentável dos recursos. *Rios Eletrônica-Revista Científica da FASETE*. ano 5 n. 5.

De Roest, K; Ferrari, P; Knickel, K. (2018). Specialisation and economies of scale or diversification and economies of scope? Assessing different agricultural development pathways. *Journal of Rural Studies*, v. 59, p. 222-231. Doi: https://doi.org/10.1016/j.jrurstud.2017.04.013.

Di Falco, S., & Chavas, J. P. (2009). On crop biodiversity, risk exposure, and food security in the highlands of Ethiopia. *American Journal of Agricultural Economics*, 91(3), 599-611.

Di Falco, S., & Zoupanidou, E. (2017). Soil fertility, crop biodiversity, and farmers' revenues: Evidence from Italy. *Ambio*, 46, 162-172.

Di Falco, S., Bezabih, M., & Yesuf, M. (2010). Seeds for livelihood: crop biodiversity and food production in Ethiopia. *Ecological Economics*, 69(8), 1695-1702.

Di Falco, S.; Chavas, J. P. Rainfall shocks, resilience, and the effects of crop biodiversity on agroecosystem productivity. *Land Economics*, v. 84, n. 1, p. 83-96, 2008.

Donfouet, H. P. P., Barczak, A., Détang-Dessendre, C., Maigné, E. (2017). Crop production and crop diversity in France: a spatial analysis. *Ecological Economics*, 134, 29-39.

INSEE. *Handbook of spatial analysis*: Theory and practical application with R. (2018).

Kidane, M. S., & Zegeye, E. W. (2018). Crop diversification and productivity in semiarid and sub-humid maize-legume production systems of Ethiopia. *Agroecology and Sustainable Food Systems*, 42(10), 1106-1127.

Kumar, C.R., Nayak, C. & Pradhan, A.K. (2024). Status and determinants of crop diversification: evidence from Indian States. *Lett Spat Resour Sci* 17, 1. Doi: https://doi.org/10.1007/s12076-023-00366-4

Magurran, A. E. *Ecological diversity, and its measurement*. (1988). Princeton University Press.

Martins, M. A., Tomasella, J., & Dias, C. G. (2019). Maize yield under a changing climate in the Brazilian Northeast: impacts and adaptation. *Agricultural water management*, 216, 339-350.

Ndip, F.E.; Molua, E.L.; Mvodo, M.E.S; Nkendah, R.; Choumbou , R.F.D.; Tabetando, R.; Akem, N.F. (2023). Farmland Fragmentation, crop diversification and incomes in Cameroon, a Congo Basin country. *Land Use Policy*, Vol. 130. https://doi.org/10.1016/j.landusepol.2023.106663

Parré, J.L., Chagas, A.L.S. (2022). Determinants of agricultural diversification in Brazil: a spatial econometric analysis. Lett Spat Resour Sci 2022. https://doi.org/10.1007/s12076-021-00295-0

Parré, J. L.; Chagas, A. L. S.; Arends-Kuenning, M. P. (2022). The effect of farm size and farmland use on agricultural diversification: A spatial analysis of Brazilian municipalities. In: *ENABER*. 22st. Salvador, 2022.

Paschoalino, P. A. T.; Parré, J. L. (2022). Diversificação agrícola e produtividade da terra nas microrregiões nordestinas: uma análise para o período 2006 e 2017. *Revista de Desenvolvimento Econômico.* Ano XXIV – v. 1, n. 51, p. 235 – 260.

Paschoalino, P. A. T.; Parré, J. L. (2023). Diversificação e produção agrícola no Brasil: Uma análise por modelos espaciais. *Revista de Política Agrícola*, v. 32, n. 1, p. 121-140.

Perobelli, F. S., et. al. (2007). Produtividade do setor agrícola brasileiro (1991-2003): uma análise espacial. *Nova economia*, 17, 65-91.

Piedra-Bonilla, E. B.; Braga, C. A. S.; Braga, M. J. (2020a) Diversificação agropecuária no brasil: conceitos e aplicações em nível municipal. *Revista de Economia e Agronegócio*, v. 18, n. 2, p. 1-28, 2020a.

Piedra-Bonilla, E. B., da Cunha, D. A., & Braga, M. J. (2020b). Climate variability and crop diversification in Brazil: An ordered probit analysis. *Journal of Cleaner Production*, *256*, 120252.

Tol, R. S. (2018). The economic impacts of climate change. *Review of environmental economics and policy*.