RESR Revista de Economia e Sociologia Rural
ISSN 1806-9479
ARTICLE

Diversification, mid-term rainfall and agricultural production. An analysis of the Brazilian Northeast region

Diversificação, precipitação de médio prazo e produção agrícola. Uma análise para a região do nordeste brasileiro

Pietro André Telatin Paschoalino[1]*, José Luiz Parré[2], André Luis Squarize Chagas[3]

[1]Universidade Estadual do Norte do Paraná (UENP), Cornélio Procópio (PR), Brasil. E-mail: pietro.paschoalino@uenp.edu.br
[2]Universidade Estadual de Maringá (UEM), Maringá (PR), Brasil. E-mail: jlparre@uem.br
[3]Faculdade de Economia, Administração, Contabilidade e Atuária (FEA), Universidade de São Paulo (USP), São Paulo (SP), Brasil. E-mail: achagas@usp.br

How to cite: Paschoalino, P. A. T., Parré, J. L., & Chagas, A. L. S. (2025). Diversification, mid-term rainfall and agricultural production. An analysis of the Brazilian Northeast region. Revista de Economia e Sociologia Rural, 63, e289389. https://doi.org/10.1590/1806-9479.2025.289389

Abstract: Studies on agricultural diversification in Brazil have gained greater notoriety in recent years, and it is a research topic that requires further exploration to understand how this variable relates to other agricultural indicators, particularly production, especially in regions with highly particular climates, such as the Brazilian Northeast. This study evaluates the relationship between agricultural diversification and medium-term rainfall in the northeastern microregions 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.

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 essa 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 a 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 pelas 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.

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 involves a trade-off, where reduced expected returns accompany risk reduction, even amid uncertainties such as pests, diseases, and prices (Di Falco & Chavas, 2009; Culas & Mahendrarajah, 2005).

However, diversification has been gaining greater attention in recent times, as it may present economies of scope in production, since diversified farms that benefit from cost complementarities can achieve greater efficiency compared to specialized farms (De Roest et al., 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 km^2, 969,589.4 km^2 of which being the semi-arid. Many locations in this region experience rainfall between 280 and 800 mm in annual average (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, a fact that underscores the need for immediate attention to this issue.

Those facts demand attention since, according to (Di Falco & Chavas, 2008; Donfouet et al., 2017), under conditions of reduced rainfall, agroecosystem productivity is expected to increase through higher levels of crop biodiversity.

Among the studies that attempted to verify the relationship between agricultural production or productivity and agricultural diversification in Brazil (and in the Northeast), we mention the studies by Paschoalino & Parré (2022, 2023), Parré et al. (2024), and Parré & Chagas (2022).

In general, Paschoalino & Parré (2023) found a negative and significant relationship between diversification and the value of agricultural production. However, Paschoalino & Parré (2022) found also a positive and significant relationship between land productivity and diversification in the Northeast microregions. Parré et al. (2024) 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. Theoretical Foundation

The following two subsections aim not only to analyze some few articles on the topic, but also 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, having Brazil as a sample, since there is a scarcity of articles addressing Brazilian Northeast.

2.1 International Studies

The studies by Di Falco & Zoupanidou (2017), Bellon et al. (2020), Ndip et al. (2023), and Kumar et al. (2024) are some of the most relevant international articles on the topic.

Di Falco & Zoupanidou (2017) focused on the gross production value (revenue from all products) using unbalanced panel data at the farm level for Italy (from 1981 to 2003). In addition to crop diversification (number of crops cultivated and/or number of livestock activities on a farm), the paper also includes soil fertility quality, measured with a soil fertility index (categorical variable) to analyze the gross production per farm (Di Falco & Zoupanidou, 2017).

Using the Arellano–Bond two-step dynamic panel data GMM estimator, they demonstrated that diversification (considered endogenous) is positively related to gross production, and that the level of fertility (if it is high) also has a positive impact. Furthermore, due to the interaction of the variables, it was found that diversity becomes more important if poor fertility levels are considered (“degradation”, according to the authors) (Di Falco & Zoupanidou, 2017).

Bellon et al. (2020) examined the relationship between crop diversity, food self-consumption (value), and cash income from crops sales among smallholder farmers in northern Ghana. The authors were interested in checking whether smallholder farmers may benefit more from a diversification or a specialization for their livelihoods. The study results suggest that crop diversification benefits these farmers more than specialization. Crop diversity is positively associated with food self-consumption and money from crops sales. This finding suggests that diversification increases household market opportunities and contributes to self-consumption.

Ndip et al. (2023) analyses the relationship between land fragmentation and crop diversification using survey data from Cameroon. Crop diversification, measured by the number of crops the household grown 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 variables (age of the household head, their gender, household size, alternative income sources); farming characteristics (e.g., such as farming experience); and institutional factors (e.g., access to extension services). The findings indicate that farmers with more fragmented lands are more likely to diversify. Given that most smallholder farming households use their own production as the main source of food for domestic consumption, fragmentation ensures that they grow diverse crops to provide a heterogeneous food basket for the family.

Finally, Kumar et al. (2024) employed a Panel Autoregressive Distributed Lag model to determine the factors affecting crop diversification over India’s 28 states. They measure crop diversification using Theil’s entropy index for the 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), landholding (cropping intensity, operational holding), economy (gross state agricultural domestic product), agricultural financing (credit) in order to test their impact on crop diversification. The analysis found 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 & 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 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 & 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 et al. (2024) used the Simpson index as a measure of diversification based on the Gross Production Value (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 spatial panel regressions 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 productivity.

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. 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 found 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:

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

Where Q_{it} is the sum of the quantities produced in each micro-region[1] *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)[2].

Where L_{it} means the total number of personnel occupied in agricultural establishments for each year according to the agricultural census (as of 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.

Precipitation is measured in millimeters per month,[3] 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 from 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. Therefore, since 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 relative to the total number of establishments in the microregion and the proportion of establishments in the microregion under analysis with managers (producer or administrator) aged 55 years or more relative to the total number of establishments.

With regard to the D_{it} variable, the diversification index, the Shannon index was used, 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 using the planted area because, following Piedra-Bonilla et al. (2020b), there are successive and/or simultaneous crops in the same year and place. This situation can generate planted areas that exceed the geographical area of the microregion, 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 to both the increase in the number of crops and the uniformity of the different crops planted (Di Falco & Chavas, 2008), in addition to its extensive use in the literature (Di Falco & 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 features are also generally included.

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

[2] Since 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 is 1,81. Thus, to convert to tons, the units in thousand fruits were multiplied by 1.81. Coconut (coco-da-baía) was disregarded due to the absence of a conversion factor.

[3] 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, using the exact_extract function with the "mean" argument through the exactextract package.

3.2 Method and empirical model

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

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

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

The variables were discussed in the previous section but summarized in Table 1.

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[4].	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), covering 2002-2006 or 2013-2017.	CRU DATA – TS4.06 – Data from the University of East Anglia.
$Year_t$	Time fixed effect for the year 2017– 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 relative to 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, relative to the total number of establishments in each year.	Agricultural Census (2006) and Agricultural Census (2017).

Instituto Brasileiro de Geografia e Estatística (2024).
Source: The authors.

Furthermore, spatial panel estimation was also used for Equation 3, given the possibility of spatial autocorrelation. 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) (Institut National de la Statistique et des Études Économiques, 2018). Equation 4 demonstrates the SAR model for panel data with fixed effects (Institut National de la Statistique et des Études Économiques, 2018).

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

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

[4] It was necessary to add one to the variable after the panel was built (one was added in the two years covered by the data), since there was a value of 0 in one microregion in 2017, making it useless as logarithm.

the lag of the dependent variable $\left(\sum_{i \neq j} w_{ij} y_{jt} \right)$ (Institut National de la Statistique et des Études Économiques, 2018).

In turn, Equation 5 demonstrates the SEM empirical model for Equation 3 considering fixed effects (Institut National de la Statistique et des Études Économiques, 2018).

$$lnQ_{it} = x_{it}\beta + \alpha_i + u_{it} \qquad (5)$$

$$u_{it} = \rho \sum_{i \neq j} w_{ij} u_{jt} + \varepsilon_{it}$$

Therefore, $u_{it} \sim IID\left(0, \sigma^2\right)$. 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 (Institut National de la Statistique et des Études Économiques, 2018).

$$lnQ_{it} = \lambda \sum_{i \neq j} w_{ij} y_{jt} + x_{it}\beta + \alpha_i + u_{it} \qquad (6)$$

$$u_{it} = \rho \sum_{i \neq j} w_{ij} u_{jt} + \varepsilon_{it}$$

With $\varepsilon_{it} \sim IID\left(0, \sigma^2\right)$, this model captures the spatial interaction both through the spatial autoregressive error term and also considering 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.

According to Table 2, interesting observations can be made about the behavior of the data between years. First, the average production of the analyzed crops decreased from one period to the next; that is, the total production of the analyzed crops declined, which is highly concerning.

It is also interesting to remark that the diversification index decreased from one period to another. The average rainfall in microregions was also declined. It is important to remember that each microregion's precipitation is actually five years of monthly precipitation average. 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 was observed that despite the slight increase in the average number of tractors per microregion, both the number of employed personnel and the total area of establishments decreased, showing less use of inputs for agricultural production.

Quantile maps were generated to verify the spatial distribution of the variables. 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.

These maps offer the best analysis when combined. For example, it is possible to verify that regions with greater agricultural production are related to various regions experiencing high 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 not very frequently).

Table 2. Descriptive statistics.

Variable	2006		2017		Panel	
	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
$D_{it}\left(quantity\right)$	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_5}	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.

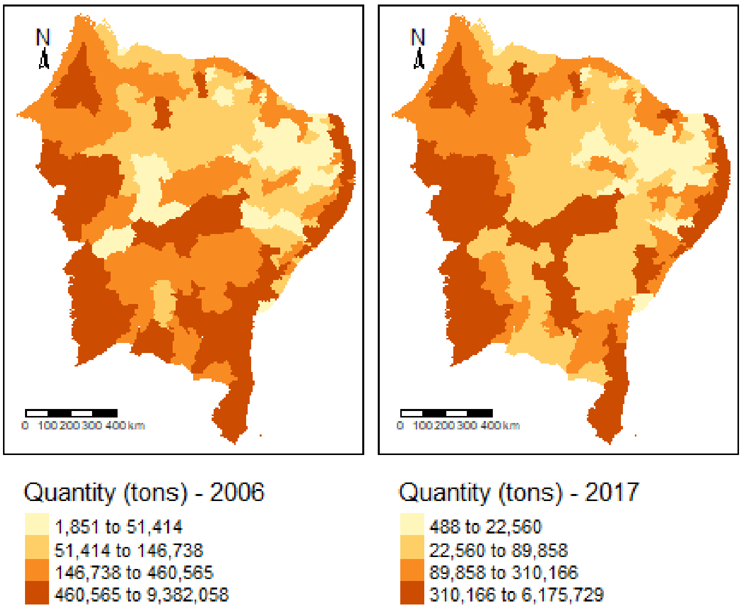

Figure 1. Map with quantiles of agricultural production in the northeast region of Brazil (2006 and 2017). **Source:** Authors based on Municipal Agricultural Survey (PAM).

[5] The number of tractors considers 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.

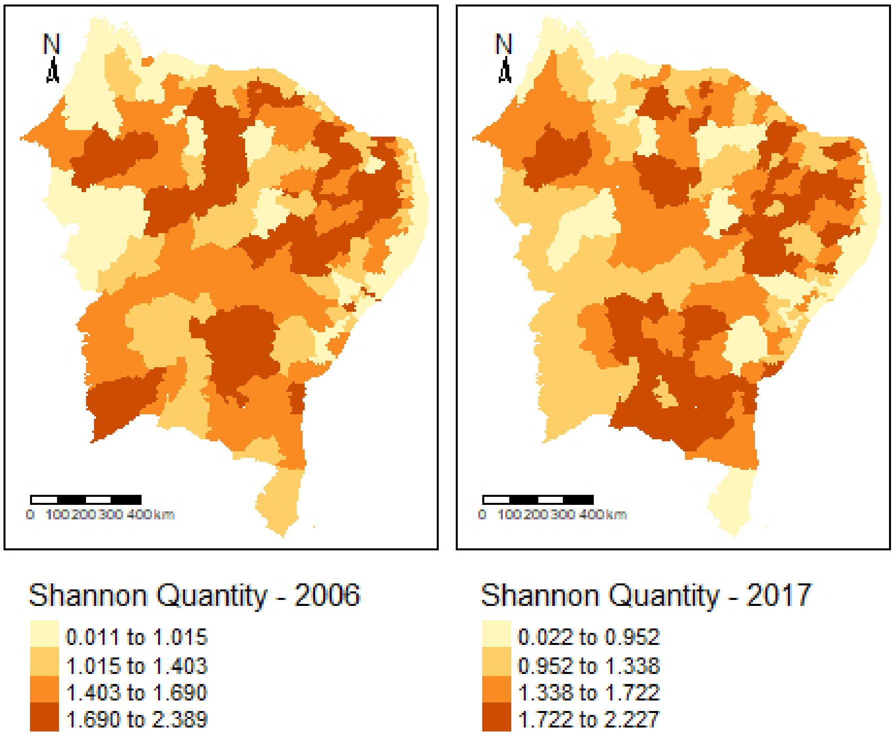

Figure 2. Map with Shannon index quantiles in Brazil's northeast region (2006 and 2017).
Source: Authors based on Municipal Agricultural Survey (PAM).

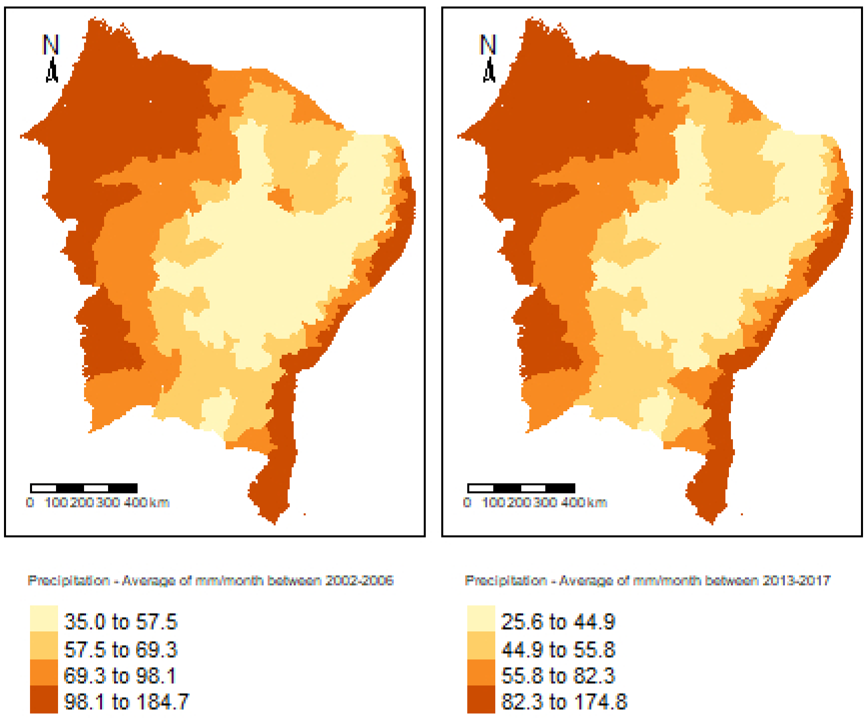

Figure 3. Map with precipitation quantiles based on a five-year mm/month average in the Northeast of Brazil (2006 and 2017).
Source: Authors based on CRU DATA.

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

4.2 Results

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

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

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

It is verified that capital and labor positively correlate with the quantity produced, but only capital retains 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 an increase of more than 3% in the quantity produced. It is also verified that diversification presents a negative sign for production and that these results are statistically significant.

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 in Table 3 an inconclusive result regarding cross-sectional dependence (Breusch-Pagan LM test and Pesaran CD test) and the presence of spatial correlation. Therefore, Table 5 show the Lagrange and Spatial Hausman multiplier tests.

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

	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)

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

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 determined that it is best to consider fixed effects. The results of the SAR model estimations with fixed effects are analyzed in Table 6.

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.

Thus, the significant variables in the non-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 as shown in Table 7.

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)

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

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 in the 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 & Parré (2023) for Brazil, while disagreeing with Paschoalino & Parré (2022). Even though the latter also considered the Northeast, 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 et al. (2024) also found a negative relationship between the variables in Brazil. Kidane & 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 sign is that diversified systems are complex to manage and require correct skills and input compared to specialized systems.

Despite this, this result does not mean that diversification should be considered negative for farmers. Firstly, the results are aggregated at the level of geographic micro-regions, and the results may differ 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 not 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 compared to what was found for the rest of the country in Paschoalino & Parré (2023), even with its edaphoclimatic characteristics.

Additionally, 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. This finding demands caution, since such results need to be analyzed alongside other studies showing climate change's impact on Brazil's Northeast.

These studies include Araújo et al. (2016) and Martins et al. (2019). Martins et al. (2019) used simulation models considering different levels of CO_2 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, which would significantly increase the amount of water needed.

Araújo et al. (2016) showed 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. 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.

5 Conclusions

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 of up to 800 mm.

Furthermore, the literature that analyzed the effect of agricultural diversification in Brazil ought to 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, the study highlighted the role of rainfall (medium-term monthly average) in 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.

Authors' contributions

PATP: Conception of the study, Data collection, Analysis and interpretation, Writing of the manuscript, and Critical review.

JLP: Writing of the manuscript, Critical review.

ALSC: Writing of the manuscript, Critical review.

Financial support:

Nothing to declare.

Conflicts of interest:

Nothing to declare.

Ethics approval:

Not applicable.

Data availability:

Research data is not available.

* Corresponding author:

Pietro André Telatin Paschoalino. pietro.paschoalino@uenp.edu.br

References

Antunes, L., & Stege, A. L. (2020). Análise espacial da convergência da produtividade agrícola brasileira (1990-2018). In *Anais do XXIII Encontro de Economia da Região Sul*. Niterói: ANPEC.

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*, *5*(5), 1-10.

Araújo, P. H. C., Silva, F. D. F., Gomes, M. F. M., Féres, J. G., & Braga, M. J. (2016). Uma análise do impacto das mudanças climáticas na produtividade agrícola da região nordeste do Brasil. *Revista Economica do Nordeste*, *45*(3), 46-57. http://doi.org/10.61673/ren.2014.118

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*, *125*, 104682. http://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 *Proceedings of the 11th Congress of the European Association of Agricultural Economists.* Copenhagen: EAAE.

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*, *59*, 222-231. http://doi.org/10.1016/j.jrurstud.2017.04.013

Di Falco, S., & Chavas, J. P. (2008). Rainfall shocks, resilience, and the effects of crop biodiversity on agroecosystem productivity. *Land Economics*, *84*(1), 83-96. http://doi.org/10.3368/le.84.1.83

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. http://doi.org/10.1111/j.1467-8276.2009.01265.x

Di Falco, S., Bezabih, M., & Yesuf, M. (2010). Seeds for livelihood: crop biodiversity and food production in Ethiopia. *Ecological Economics*, *69*(8), 1695-1702. http://doi.org/10.1016/j.ecolecon.2010.03.024

Di Falco, S., & Zoupanidou, E. (2017). Soil fertility, crop biodiversity, and farmers' revenues: evidence from Italy. *Ambio*, *46*(2), 162-172. http://doi.org/10.1007/s13280-016-0812-7

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. http://doi.org/10.1016/j.ecolecon.2016.11.016

Institut National de la Statistique et des Études Économiques – Insee. (2018). *Handbook of spatial analysis: theory and practical application with R.* Paris.

Instituto Brasileiro de Geografia e Estatística – IBGE. (2024). *Census of agriculture.* Retrieved in 2024, August 14, from https://www.ibge.gov.br/en/statistics/economic/agriculture-forestry-and-fishing/21929-2017-2017-censo-agropecuario-en.html?edicao=25774

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. http://doi.org/10.1080/21683565.2018.1505679

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), 1. http://doi.org/10.1007/s12076-023-00366-4

Magurran, A. E. (1988). *Ecological diversity, and its measurement.* Princeton: Princeton University Press. http://doi.org/10.1007/978-94-015-7358-0.

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. http://doi.org/10.1016/j.agwat.2019.02.011

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*, *130*, 106663. http://doi.org/10.1016/j.landusepol.2023.106663

Parré, J. L., Chagas, A. L. S., & Arends-Kuenning, M. P. (2024). The effect of farm size and farmland use on agricultural diversification: a spatial analysis of Brazilian municipalities. *Agricultural and Food Economics*, *12*(1), 27. http://doi.org/10.1186/s40100-024-00323-9

Parré, J. L., & Chagas, A. L. S. (2022). Determinants of agricultural diversification in Brazil: a spatial econometric analysis. *Letters in Spatial and Resource Sciences*, *15*(2), 173-195. http://doi.org/10.1007/s12076-021-00295-0

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*, *1*(51), 235-260. http://doi.org/10.36810/rde.v1i51.7485

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

Perobelli, F. S., Almeida, E. S., Alvim, M. I. S. A., & Ferreira, P. G. C. (2007). Produtividade do setor agrícola brasileiro (1991-2003): uma análise espacial. *Nova Economia*, *17*(1), 65-91. http://doi.org/10.1590/S0103-63512007000100003

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*, *18*(2), 1-28. http://doi.org/10.25070/rea.v18i1.9501

Piedra-Bonilla, E. B., Cunha, D. A., & Braga, M. J. (2020b). Climate variability and crop diversification in Brazil: an ordered probit analysis. *Journal of Cleaner Production*, *256*, 120252. http://doi.org/10.1016/j.jclepro.2020.120252

Tol, R. S. (2018). The economic impacts of climate change. *Review of Environmental Economics and Policy*, *12*(1), 4-25. http://doi.org/10.1093/reep/rex027

Received: August 14, 2024;
Accepted: November 10, 2024
JEL Classification: Q10; Q50; Q57.

Associate Editor: Gustavo Inácio de Moraes