

Article

Canonical Analysis of the Impact of Climate Predictors on Sugarcane Yield in the Eastern Region of Pernambuco, Brazil

Rodrigo Rogério da Silva ¹, Geber Barbosa de Albuquerque Moura ¹, Fabrício Marcos Oliveira Lopes ^{1,*}, Cristina Rodrigues Nascimento ¹ and Pedro Rogério Giongo ²

¹ Department of Agronomy, Federal Rural University of Pernambuco, Dom Manoel de Medeiros Avenue, SN, Dois Irmãos, Recife 52171-900, PE, Brazil; rodrigossilvaagr@gmail.com (R.R.d.S.); geber.moura@ufrpe.br (G.B.d.A.M.); cristina.nascimento@ufrpe.br (C.R.N.)

² Campus of Santa Helena of Goiás, State University of Goiás, Via Protestado Joaquim Bueno, No. 945 Urban Perimeter, Santa Helena of Goiás 75920-000, GO, Brazil; pedro.giongo@ueg.br

* Correspondence: pabricio.lopes@ufrpe.br

Abstract

Sugarcane yield plays a crucial role in food safety and biofuel production, and it is strongly influenced by climatic variations. In this context, this study applies canonical correlation analysis (CCA) to identify the climatic predictors, such as sea surface temperature, atmospheric pressure, and wind speed, that affect sugarcane yield from 1990 to 2019. Hierarchical cluster analysis applied to the performance of 58 municipalities in the eastern region of Pernambuco identified three distinct and homogeneous groups. An analysis of the CCA for the three sugarcane yield groups and climatic variables revealed that the first canonical function was significant with $R = 0.82$ and precision of 0.67 ($p \leq 0.05$ at 95% confidence level), and that the sugarcane yield groups and climatic variables were different (Wilks' lambda = 0.14), but they were associated. Climatic variables affected the three sugarcane productivity groups, with redundancy indices of 29.7%, 52.2%, and 59.9%. Climatic variables with positive canonical weights enhance performance, while those with negative weights decrease yields. The structural canonical loads and cross-loadings reveal that sea surface temperature plays a crucial role in determining sugarcane yield, potentially influencing precipitation and temperature patterns in the region. The sensitivity analysis confirms the stability of the canonical loads and the robustness of the results, demonstrating that this research can support yield forecasting, regional agricultural policy, and drought management. Identifying climate predictors, such as sea surface temperature, wind speed, and atmospheric pressure, enables the creation of accurate models to predict sugarcane productivity, assisting farmers in planning input management, irrigation during dry periods, and harvesting. Furthermore, climate data can inform policies that encourage sustainable agricultural practices and adaptation to climate conditions, strengthening food security and guiding the selection of more resilient sugarcane varieties, increasing production resilience.

Keywords: agricultural planning; wind trade; temperature sea surface; hierarchical cluster; El Niño



Academic Editor: Dengpan Xiao

Received: 5 July 2025

Revised: 5 August 2025

Accepted: 20 August 2025

Published: 18 October 2025

Citation: Silva, R.R.d.; Moura, G.B.d.A.; Lopes, P.M.O.; Nascimento, C.R.; Giongo, P.R. Canonical Analysis of the Impact of Climate Predictors on Sugarcane Yield in the Eastern Region of Pernambuco, Brazil. *Agriculture* **2025**, *15*, 2162. <https://doi.org/10.3390/agriculture15202162>

Copyright: © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the Creative Commons

Attribution (CC BY) license

(<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Brazil's economic base is agriculture, with the sugar and ethanol sectors being imperative for gross domestic product (GDP), representing approximately 2% of the total. In addition, the sugarcane production chain generates approximately BRL 100 billion per year [1]. The Northeast region (NEB) experienced increases of 0.8% in planted area and

2.8% in average productivity, reaching a production of approximately 51 million tons in the 2020/2021 harvest [2]. The state of Pernambuco is the second-largest producer of sugarcane in NEB. The eastern sector of Pernambuco may present an increase of close to 3.5% compared with the 2023/2024 harvest. The expectation is that 14.5 million tons of the plant will be processed, compared with the 14.01 million tons processed in the 2023/2024 harvest. This production should result in the manufacture of 1.1 million tons of sugar and ethanol production [3].

One of the stages that directly influence the production capacity of sugarcane is harvesting, which has changed from manual to mechanized [4]. However, some areas may have difficulties with mechanization due to terrain characteristics, associated with the burning of straw [5]. Furthermore, the eastern sector of the Pernambuco no longer has area for expansion, making adequate management of this crop essential, aiming to increase its yield [6].

The practice of burning is criticized due to its environmental impacts, such as the loss of soil nutrients [7] and the emission of pollutants such as CO₂ and N₂O responsible for intensifying the natural phenomenon of global warming [8]. The Intergovernmental Panel on Climate Change (IPCC) concluded, among the scenarios studied, that there is more than a 50% chance of global temperatures reaching or exceeding 1.5 °C between 2021 and 2040 due to the increase in greenhouse gas emissions [9]. In Brazil, the temperature of the Earth's surface could increase by approximately 4 °C, and there could be a 10 to 15% increase in rainfall during the fall and a reduction during the summer, in addition to the intensification of climate phenomena such as El Niño and La Niña in NEB [10]. If the IPCC predictions are confirmed, sugarcane, the main crop in the Zona da Mata of Northeast Brazil, could suffer drastic reductions due to possible decreases in water availability and an increase in the expected demand for water for other uses [11]. Therefore, predicting how rainfall, average ocean surface temperature (sst) anomalies, and wind speed can affect sugarcane yield in micro and mesoregions is of great importance for decision-making by governments and civil society [12].

This study uses canonical correlation analysis (CCA), which finds linear combinations, called canonical variables, between two groups of variables by maximizing their relationship. The CCA technique explores the relations between two large groups of variables, but the researcher may be interested in examining only a few linear combinations within the set [13]. According to Trugilho et al. [14], a striking feature of canonical correlation analysis is that, unlike principal component analysis, it is invariant in relation to the scale of the variables. Canonical correlations refer to the relationship between canonical pairs, that is, linear combinations between groups of variables, whose correlation between these combinations is maximum [15]. CCA effectively describes data, verifies numerical models, and builds statistical prediction models, providing knowledge of which configurations tend to occur in two or more distinct fields and the degree of connection between them [16–19].

In Brazil, studies have already been carried out on sugarcane crops, aiming to contribute to the genetic improvement of varieties with the application of canonical correlation analysis, but they do not predict sugarcane productivity based on climatic variables [4,20–23]. Other studies have used canonical correlation analysis for climate forecasts and scenarios and diagnostic studies without relating it to sugarcane productivity, for example, Rana et al. [18] for seasonal forecasting of winter precipitation in central-southwest Asia; Forootan et al. [24] to understand rainfall variability in Australia; and Moura et al. [12] to forecast rainfall (rainy season) in eastern NEB for the four wettest months, three months in advance.

Predicting sugarcane crop productivity is essential for both governmental and producers' planning, which can reduce costs and improve harvest logistics with better productivity

management [25]. However, many farms use visual estimates of sugarcane yield by technicians based on acquired experience and information from previous harvests; therefore, this type of analysis can be biased and depends exclusively on the knowledge of the technician who evaluates the plot without investigating the errors involved [8,26]. Meanwhile, previous studies have extensively investigated the impact of climatic factors on sugarcane production [27]. These researchers employed a variety of methodologies, including relevant models, to evaluate sugarcane production. They conducted regression analyses to establish correlations between climatic factors and sugarcane yield, and assessed the effects of these factors during specific planting periods on sugarcane [28–30]. Canonical correlation analysis has not previously been employed to assess the impact of climatic predictors on sugarcane yield in the eastern sector of Pernambuco, Brazil.

In this context, the present study aimed to determine the climate variables that most influence sugarcane yield, up to three months before the harvest, in the eastern sector of Pernambuco, Brazil. Furthermore, it applied cluster analysis to group municipalities with similar sugarcane yield. Canonical correlation analysis can verify the existence and intensity of the most relevant climate variables for homogeneous groups of municipalities that produce sugarcane in the eastern sector of Pernambuco, Brazil.

2. Material and Methods

2.1. Study Area and Data

The study area (Figure 1 and Table 1) included the administrative boundaries of 58 sugarcane-producing municipalities in the eastern region of Pernambuco, Brazil, along the coast and in the transition zones between forest and native forest. According to the Koppen classification, the area has a humid tropical climate, with average monthly temperatures above 18 °C and an annual average of 25 °C [31]. The northeast and southeast trade winds blow all year round, varying only in the changing seasons. The predominant meteorological systems in the region are easterly waves, upper-level cyclonic vortex, sea breezes, and the intertropical convergence zone. Remnants of cold fronts and instability lines are also present, although less frequent [32]. The region has high air humidity levels and average annual rainfall between 1500 mm and 2500 mm. The rainy season, which runs from April to July, is a period with a higher probability of rain, with the possibility of moderate to heavy rain showers, especially on the coast [33].

The flowchart (Figure 2) shows the processing steps for determining climate predictors of average sugarcane yield.

In the 2022/2023 harvest, the area cultivated with sugarcane in the region reached 233.9 thousand hectares [37]. Data for the average sugarcane yield refer to municipal agricultural production (PAM) between 1990 and 2019, available at <https://sidra.ibge.gov.br/Tabela/1612>, accessed on 26 May 2025. Missing yield data were estimated by a one-variable linear regression model for the years 1990, 1991, and 1992 for the municipalities of Lagoa do Carro and Xexéu; 2007 to 2009 for Itapissuma; 2015 for João Alfredo, Limoeiro and Machados; 2016 for Barra de Guabiraba, João Alfredo, Limoeiro, and Machados; 2017 for São Benedito do Sul; 2018 and 2019 for Chã Grande.

The varietal census of Pernambuco showed that in the 2022/2023 season, the most common cultivars (% area harvested) were SP78-4764 (21.06%), RB92-579 (18.94%), RB04-1443 (10.76%), RB86-7515 (9.43%), SP79-1011 (8.60%), VAT90-212 (4.55%), and B8008 (2.25%), according to RIDESA (Rede Interuniversitária para o Desenvolvimento do Setor Sucroalcooleiro), available at <https://www.ridesa.com.br/censo-varietal>, accessed on 5 August 2025.

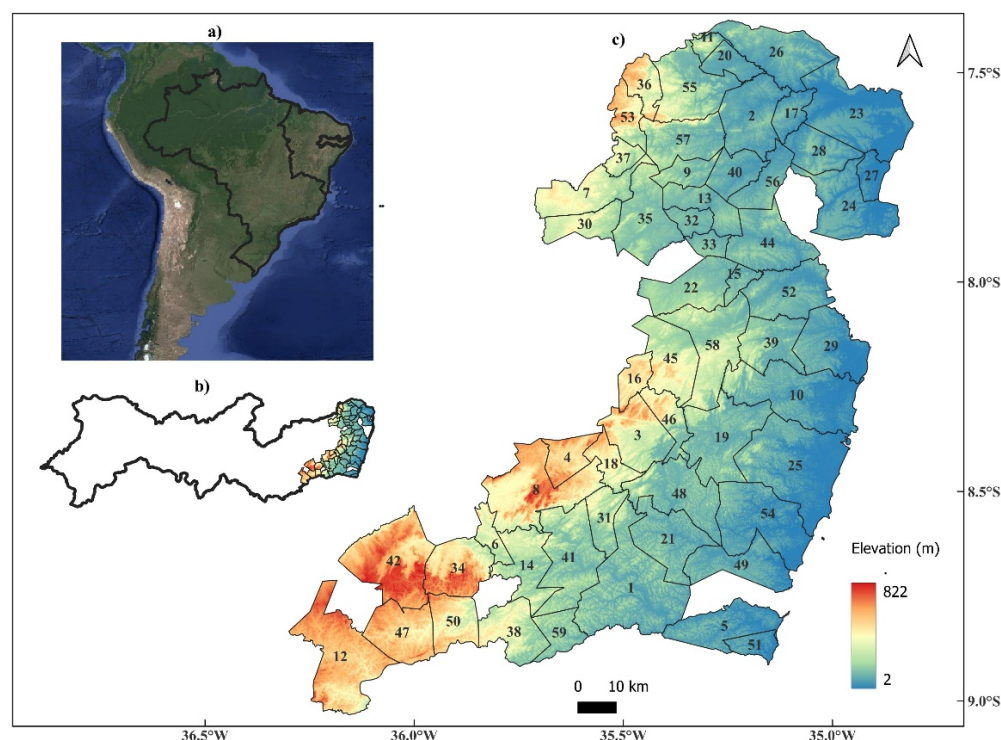


Figure 1. (a) Location of South America, Brazil, and Pernambuco; (b) State of Pernambuco and study area; (c) Elevation of the study area and identification of sugarcane producing locations. The data were obtained from Google [34], the Brazilian Institute of Geography and Statistics [35], and Farr et al. [36], respectively.

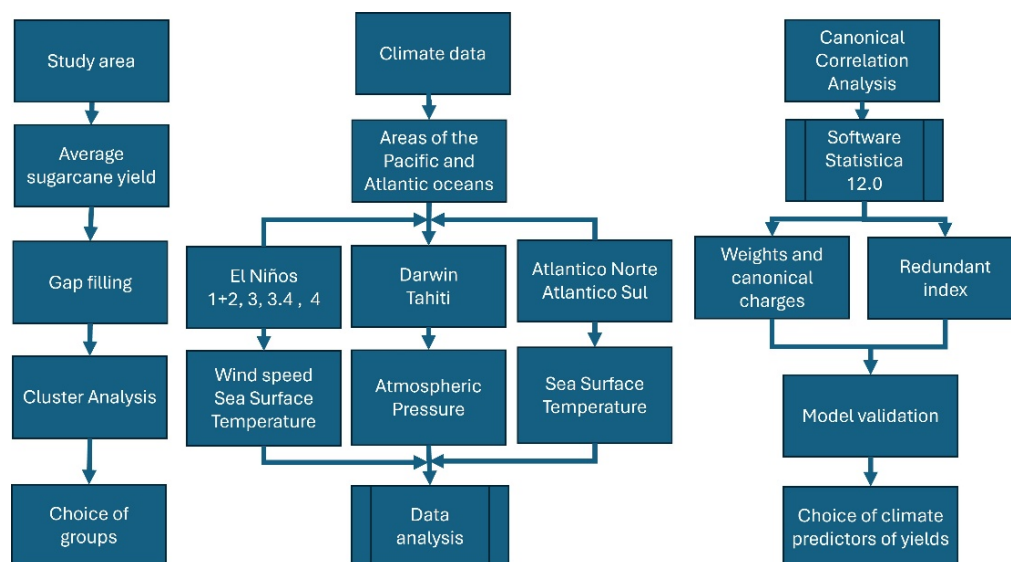


Figure 2. Flowchart with processing steps for determining climate predictors of average sugarcane yield. Source: Author.

The choice of sugarcane variety reflects niformity in cultivation practices, influenced by topography and regional conditions, as shown in Table 2. In the North Mata Zone (B), the cultivated sugarcane varieties include RB 75126, RB 763710, SP 701143, SP 784764, and SP 791011. The harvest is either manually or mechanically, in flat and plateau areas. In the Southern Mata Zone (D), the cultivated sugarcane varieties, including RB 75126, RB 763710, SP 716949, SP 784764, and SP 791011, are also planted on flat land for manual or mechanical harvesting.

MaisAgro [38] includes CTC 9001 BT (transgenic), resistant to the sugarcane borer (*Diatraea saccharalis*), along with other cultivars that are well-adapted to the region's soil and climate conditions. In general, these varieties have low soil fertility requirements and are resistant to rust. The selection of these crops depends on their productive potential and adaptability to local conditions [39–41].

The harvest of the raw sugarcane follows an average period of 15 months for the plant cane after planting and 12 months for the three successive harvests of ratoon cane [42]. The ideal period for planting sugarcane in the Zona da Mata region of Pernambuco is between January and March, during the rainy season. This period favors bud development, providing rapid sprouting and reducing the risk of disease in the stems [39,42].

Table 1. Sugarcane yields at locations in the eastern sector of Pernambuco, Brazil.

ID	Locations	ϕ (°)	λ (°)	ID	Locations	ϕ (°)	λ (°)	ID	Locations	ϕ (°)	λ (°)
1	Água Preta	−8.74	−35.53	21	Gameleira	−8.62	−35.38	41	Palmares	−8.67	−35.61
2	Aliança	−7.64	−35.17	22	Glória do Goitá	−8.02	−35.33	42	Panelas	−8.64	−36.03
3	Amaraji	−8.41	−35.46	23	Goiana	−7.61	−34.90	43	Paudalho	−7.91	−35.16
4	Barra Guabiraba	−8.45	−35.62	24	Igarassu	−7.84	−34.95	44	Pombos	−8.20	−35.38
5	Barreiros	−8.81	−35.24	25	Ipojuca	−8.44	−35.06	45	Primavera	−8.32	−35.39
6	Belém de Maria	−8.61	−35.82	26	Itambé	−7.45	−35.13	46	Quipapá	−8.85	−36.03
7	Bom Jardim	−7.81	−35.63	27	Itapissuma	−7.74	−34.89	47	Ribeirão	−8.51	−35.37
8	Bonito	−8.50	−35.74	28	Itaquitinga	−7.69	−35.04	48	Rio Formoso	−8.69	−35.17
9	Buenos Aires	−7.75	−35.35	29	Jaboatão dos Guararapes	−8.16	−35.01	49	São Benedito do Sul	−8.81	−35.91
10	Cabo S. Agostinho	−8.23	−35.20	30	João Alfredo	−7.86	−35.55	50	São José da Coroa Grande	−8.87	−35.17
11	Camutanga	−7.43	−35.29	31	Joaquim Nabuco	−8.55	−35.55	51	São Lourenço da Mata	−8.04	−35.12
12	Canhotinho	−8.92	−36.14	32	Lagoa do Carro	−7.85	−35.33	52	São Vicente Férrer	−7.62	−35.48
13	Carpina	−7.83	−35.26	33	Lagoa de Itaenga	−7.91	−35.29	53	Sirinhaém	−8.55	−35.16
14	Catende	−8.66	−35.72	34	Lagoa dos Gatos	−8.68	−35.91	54	Timbaúba	−7.56	−35.36
15	Chã de Alegria	−7.99	−35.21	35	Limoeiro	−7.88	−35.46	55	Tracunhaém	−7.77	−35.15
16	Chã Grande	−8.24	−35.48	36	Macaparana	−7.50	−35.46	56	Vicência	−7.65	−35.35
17	Condado	−7.611	−35.11	37	Machados	−7.71	−35.50	57	Vitória de Santo Antão	−8.15	−35.28
18	Cortês	−8.45	−35.52	38	Maraial	−8.84	−35.73	58	Xexéu	−8.83	−35.65
19	Escada	−8.37	−35.28	39	Moreno	−8.15	−35.14				
20	Ferreiros	−7.47	−35.25	40	Nazaré da Mata	−7.75	−35.25				

ID refers to the municipality; ϕ (°) is latitude; λ (°) is longitude.

Table 2. Varieties according to topography and region.

Varieties	Slope				Floor				Floodplain				Tableland			
	Manual		Mechanized													
	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
RB 72-454					X	X		X	X	X		X	X	X		
RB 73-2577			X				X	X			X	X	X			X
RB 75-126	X	X	X	X	X	X	X	X	X	X	X				X	X
RB 76-3710	X	X	X	X	X	X	X	X		X		X			X	X
RB 81-3804					X			X	X			X	X			
RB 83-102						X			X			X			X	
SP 70-1143	X	X			X	X				X					X	
SP 71-6949			X	X			X	X			X	X			X	
SP 77-5181													X	X		
SP 78-4764	X	X	X	X	X	X	X	X	X	X	X				X	X
SP 79-1011	X	X		X	X	X		X	X	X		X			X	X
B 8008													X		X	X

A = North Coast; B = North Mata; C = South Coast; D = South Mata Source; X = The variety is cultivated according to the topography and the region. Adapted from IPA [43], available at <http://www.ipa.br/resp20.php>, accessed on 8 August 2025.

2.2. Climate Data

This study used monthly sea surface temperature (sst) data from January 1990 to December 2019 (30 years) for the Niño 1 + 2 (0°–10° S, 90°–80° W), Niño 3 (5° N–5° S, 150°–90° W), Niño 3.4 (5° N–5° S, 170°–120° W), and Niño 4 (5° N–5° S, 160° E–150° W) areas located in the Pacific Ocean (Figure 3a) and also from the tropical North Atlantic (NA) and tropical South Atlantic (SA) (Figure 3c), regions associated with the tropical Atlantic dipole pattern [44,45], from the National Centers for Environmental Prediction (NCEP) and National Center for Atmospheric Administration (NCAA) reanalysis, available at <https://origin.cpc.ncep.noaa.gov/products/precip/CWlink/MJO/enso.shtml>, accessed on 23 July 2025. In addition, trade wind index data were collected monthly at 850 hPa in the equatorial Pacific Ocean at three different locations: the Central Pacific region spanning from 5° N to 5° S and 175° W to 140° W, the East Pacific region from 5° N to 5° S and 135° W to 150° W, and the West Pacific region from 5° N to 5° S and 135° E to 180° W.

The locations of Darwin in Australia (12°27' S; 130°50' E) and Tahiti in French Polynesia (17°40' S; 149°27' W) were used to obtain monthly sea-level atmospheric pressure (Figure 3c). Darwin and Tahiti are located in areas crucial for monitoring ENSO (El Niño–Southern Oscillation), a climate phenomenon that affects temperatures, precipitation, and winds worldwide. The atmospheric pressure at these locations can be used to calculate the Southern Oscillation (SOI), an indicator of the ENSO. All climate data were acquired from the Climate Prediction Center (CPC) of the National Centers for Environmental Prediction (NCEP), available at <https://www.cpc.ncep.noaa.gov/>, accessed on 23 May 2025.

To study the canonical correlations between the groups of average sugarcane yield and the climate group between 1990 and 2019, the average sugarcane yields and the quarterly averages of SST, wind speed, and atmospheric pressure were considered three months before the rainy season (April to July). Therefore, the months considered were November, December, and January.

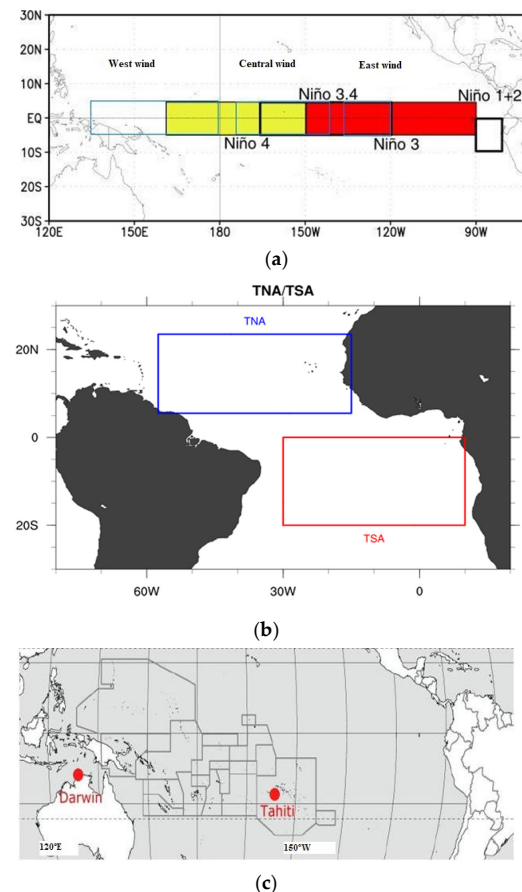


Figure 3. (a) Areas of action of El Niño 1+2, 3, 3.4, 4 and their predominant easterly, central and westerly winds used by Trenberth [46], Trenberth and Stepaniak [47], Source: Adapted from NOAA [48]; (b) Area of the North Atlantic Ocean (TNA) and South Atlantic Ocean (TSA) dipole used by Enfield [49], Source: Adapted from NOAA [48]; (c) Location of Darwin (12°27' S; 130°50' E) in Australia and Tahiti (17°40' S; 149°27' W) in French Polynesia for obtaining atmospheric pressure at mean sea level in the Equatorial Pacific Ocean used by Abril-Salcedo et al. [50], Source: Adapted from Schlanger and Heres [51].

2.3. Ward's Method for Hierarchical Clustering

In this study, we used Euclidean distance as a clustering function to estimate the similarity or dissimilarity between the average sugarcane yields of the municipalities in the cluster analysis, given by Equation (1) [52] as follows:

$$D(P_i, P_j) = \left[\sum_{q=1}^n N_i (P_{iq} - P_{jq})^2 \right]^{1/2} \quad (1)$$

where P_i and P_j represent average sugarcane yields of the municipalities in $\text{kg} \cdot \text{ha}^{-1}$, n represents the number of years, q represents the years, and N_i represents individuals (municipalities) organized into K classes.

The grouping criterion used was Ward's method [53]. This method typically produces groups with an equal number of elements, and its theoretical foundation is rooted in the principles of variance analysis. The technique considers the sum of squared deviations (SSD) of each data point in relation to the mean of its grouping and evaluates the lack of information caused by the grouping. Thus, for N_i individuals organized into K classes, X_i represents the classification of the individual within the i th class (Equation (2)).

$$\text{SSD} = \sum_{i=1}^n \left[X_i^2 - \frac{1}{n} \left(\sum X_i \right)^2 \right] \quad (2)$$

2.4. Linear Correlations

One approach used in this study was linear correlation analysis between the annual sugarcane yield of each agricultural group and climate variables. These variables included the November, December, and January averages of sstSA, sstNA, EN1+2, EN3, Central Equatorial Pacific Trade Wind Index (WC), East Pacific Equatorial Trade Wind Index (WE), West Pacific Equatorial Trade Wind Index (WW), Darwin Mean Surface Air Pressure (Darwin), and Tahiti Mean Surface Air Pressure (Tahiti). All climate variables were normalized. Specifically, the correlations focused on showing the results of the influence of climate variables three months before obtaining the annual sugarcane yield.

2.5. Canonical Variable and Canonical Correlation

Consider two groups of variables, $X^{(1)}$, and $X^{(2)}$, defined as follows: $X^{(1)} = (x_1, x_2, \dots, x_p)$ represents a vector of measures of p characteristics that belong to group I, while $X^{(2)} = (x_1, x_2, \dots, x_q)$ represents a vector of observations of q characteristics that belong to group II'. The combination of $X^{(1)}$ and $X^{(2)}$ obtains the vector $X_{(p+q) \times 1}$ and the covariance matrix Σ , as shown in Equations (3) and (4) [54].

$$X_{(p+q) \times 1} = \begin{bmatrix} X_1^{(1)} \\ X_2^{(1)} \\ \vdots \\ X_p^{(1)} \\ X_1^{(2)} \\ X_2^{(2)} \\ \vdots \\ X_q^{(2)} \end{bmatrix} = \begin{bmatrix} X^{(1)} \\ X^{(2)} \end{bmatrix} \quad (3)$$

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \vdots & \Sigma_{12} \\ p \times q & & p \times q \\ \dots & \dots & \dots \\ \Sigma_{21} & \vdots & \Sigma_{22} \\ q \times p & & q \times p \end{bmatrix} \quad (4)$$

Therefore, the p, q elements of Σ_{12} measure the correlation between the two groups. The objective is to restrict selected linear combinations of variables belonging to $X^{(1)}$ and $X^{(2)}$ instead of working with all the covariances presented in Σ_{12} . To align with Anderson's [54,55] notation, consider U as a linear combination of X^1 , and V as a linear combination of X^2 , defined by Equations (5)–(9) as follows:

$$U = a'X^1 \quad (U = a_1x^1 + a_2x^1 + \dots + a_px^1) \quad (5)$$

$$V = b'X^2 \quad (V = b_1x^2 + b_2x^2 + \dots + b_qx^2) \quad (6)$$

where $a' = [a_1, a_2, \dots, a_p]$ is the vector $1 \times p$ of weights of the characteristics of group I; and $b' = [b_1, b_2, \dots, b_q]$ is the $1 \times q$ vector of group II feature weights.

$$\text{Var}(U) = a' \cdot \text{Cov}(X^1) \cdot a = a' \cdot \sum_{11} a \quad (7)$$

$$\text{Var}(V) = b' \cdot \text{Cov}(X^2) \cdot b = b' \cdot \sum_{22} \cdot b \quad (8)$$

$$\text{Cov}(U, V) = a' \cdot \text{Cov}(X^1, X^2) \cdot b = a' \cdot \sum_{12} \cdot b \quad (9)$$

The first canonical correlation (r_1) represents the maximized relationship between U_1 and V_1 . Functions U_1 and V_1 constitute the first canonical pair associated with the canonical correlation expressed by Equation (10) [55,56] as follows:

$$r_1 = \text{Cor}(U_1, V_1) = \frac{\text{Cov}(U, V)}{\sqrt{\text{Var}(U)\text{Var}(V)}} = \sqrt{\lambda_1} \quad (10)$$

The first pair (U_1, V_1) represent linear combinations of variance equal to 1. The pair k (U_k, V_k) represents linear combinations of variance equal to 1, maximizing the correlation between all uncorrelated choices with pairs U_1 and $U_{2(k-1)}$. λ_k is the eigenvalues. The standardized variable cases use $\Sigma_{11} = R_{11}$, $\Sigma_{22} = R_{22}$, and $\Sigma_{12} = R_{12}$, where R represents a correlation matrix [55].

$$R = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \quad (11)$$

The first step is to determine the eigenvalues obtained by the equation $|R_{11}^{-1} R_{12} R_{22}^{-1} R_{21} - \lambda I| = 0$. The $R_{11}^{-1} R_{12} R_{22}^{-1} R_{21} a = \lambda a$ and $R_{11}^{-1} R_{12} R_{22}^{-1} R_{21} b = \lambda b$ eigenvalue equation calculates the eigenvectors a and b . The first canonical correlation between the linear combination of the characteristics of groups I and II is given by $r_1 = \sqrt{\lambda_1}$, where λ_1 represents the largest eigenvalue of the square matrix $R_{11}^{-1} R_{12} R_{22}^{-1} R_{21}$. The eigenvalues and eigenvectors of the p or q th expressions estimated other correlations and canonical factors. The approximate test of x^2 evaluated the significance of the hypothesis, considering all possible null canonical correlations ($H_0: r_1 = r_2 \dots r_s = 0$), with s less than $(p$ or $q)$ and n equal to the number of observations [14,55,56], expressed by Equation (12) as follows:

$$x^2 = -[n - 0.5(p + q + 3)] \log_e \left[\prod_{i=1}^s (1 - r_i^2) \right] \quad (12)$$

Equation (13) tests hypothesis H_0 where $r_k > 0$ and $r_{k+1} = r_{k+2} = \dots r_s = 0$, and the associated statistic depends on the degrees of freedom.

$$x^2 = -[n - 0.5(p + q + 3)] \log_e \left[\prod_{i=k+1}^s (1 - r_i^2) \right] \quad (13)$$

which is associated with x^2 with $(p-k)$ $(p-k)$ degrees of freedom. Only the statistically significant roots were used for subsequent studies on canonical pairs.

2.6. Interpretation of Canonical Statistical Variables

The canonical correlation coefficients (r) are the square roots of the eigenvalues $r_k = \sqrt{\lambda_k}$ interpreted as correlation coefficients. The correlations between the canonical variables are called canonical correlations. The degree of canonical correlation between two groups of variables was determined using the highest correlation coefficient as an indicator. However, the other canonical variables can also be significant and provide interpretations. The eigenvalues (λ) represent the portion of variance shared between the respective canonical variables [14,54,55].

The canonical coefficients (or canonical loading) are the coefficients of each canonical function, measuring the simple linear correlation between a variable observed in one of the groups (I or II) and the corresponding canonical pair. The canonical coefficient reflects the

importance of the variable in the respective group; the higher the coefficient, the greater the relevance of the variable. The square of the canonical correlation loads estimates the portion of variance shared by the canonical variables in each group. If the group I counts p variables and group II counts q variables, the squared canonical correlation (r^2) multiplied by the extracted variance portion, using Equations (14) and (15), obtains an estimate of redundancy [14], which measures how much the variables in the groups overlap, as follows:

$$Re_1 = [\Sigma (c_1^2)/p] \times r^2 \quad (14)$$

$$Re_2 = [\Sigma (c_2^2)/q] \times r^2 \quad (15)$$

where Re_1 and Re_2 are the redundancies of groups one and two; c_1 and c_2 represent the charges of groups one and two, respectively.

On a large dataset, canonical correlations close to $r = 0.30$ can be statistically significant. Squared, this coefficient ($r^2 = 0.09$) can calculate redundancy, which indicates how much of the variability in the variables is explained by the canonical roots [54–56]. Assessing significance is subjective by nature, but the redundancy measure is essential for determining the share of a canonical root in the variance of the variables. In this paper, three methods were employed to analyze canonical variates, using STATISTICA software, version 12.0 [57]: (1) canonical coefficients (or canonical weights), (2) canonical correlations (or canonical loadings), and (3) canonical cross-correlations.

3. Results

3.1. Cluster Analysis

The hierarchical cluster analysis resulted in three homogeneous groups of sugarcane yield, each located in a different municipality. Table 3 shows the parameters used to create the dendrogram. Group G1 had the highest average yield of 52,033.45 tons/ha, indicating that it generally performed better than Groups G2 and G3. In contrast, Group G2 had a lower average yield of 43,184.862 tons/ha, suggesting that there may be factors limiting its yield. The error pattern indicates the precise average yield of sugarcane. A minor error suggests that the average is a more confident estimate of real performance. Group G3 had a lower standard error of 675.09 tons/ha compared to Groups G1 and G2, indicating that its average yield was more consistent with the collected data. The median of group G1 was higher (53,335.86 tons/ha), which reinforces the idea that most of the yields in this group were higher. The median of group G2 was lower (43,122,036 tons/ha), indicating that most of the yields were below the median.

Table 3. Description of statistical parameters of the dendrogram for the homogeneous groups of sugarcane yield in tons per hectare (ton/ha).

Static Parameters	G1	G2	G3
Mean	52,033.45 tons/ha	43,184.862 tons/ha	49,293.44 tons/ha
Standard Error	1234.80 tons/ha	1031.22 tons/ha	675.09 tons/ha
Median	53,335.86 tons/ha	43,122.036 tons/ha	49,444.39 tons/ha
Standard Deviation	6763.28 tons/ha	5648.26 tons/ha	3697.66 tons/ha
Sample Variation	45,741,896.8 tons/ha	31,902,266.02 tons/ha	13,672,688.35 tons/ha
Kurtosis	3.85	0.596	4.23
Asymmetry	−1.64	−0.96	−1.34
Interval	32,372.80 tons/ha	20,911.76 tons/ha	20,240.74 tons/ha
Minimum	28,571.43 tons/ha	29,264.71 tons/ha	36,283.22 tons/ha
Maximum	60,944.23 tons/ha	50,176.47 tons/ha	56,523.96 tons/ha
Sum	1,561,003.52 tons/ha	1,295,545.81 tons/ha	1,478,803.23 tons/ha
Sample number	30	30	30

The standard deviation is a measure of the dispersion of values close to the mean. In this case, group G1 had the highest deviation in yields (6763.28 tons/ha), indicating greater variability in sugar cane yields. The sample variation is the measure of the dispersion of two dice. The G1 group had the most elevated variation (45,741,896.8 tons/ha), indicating that the dispersion of the two yields is not equal between the groups. In the kurtosis evaluation of the “height” and “length” in the distribution of the sugarcane yield, a kurtosis value less than 3 indicates that the distribution is more flattened than a normal distribution. The groups G1 and G3 presented values of 3.85 and 4.23, respectively, suggesting that they had extreme yields. The G2 group presented a 0.596 flattening, indicating that the performance data did not contain many extreme values. The asymmetry of the two groups of sugar cane yields was different from zero, indicating that the groups had an asymmetric distribution.

The hierarchical cluster analysis, represented by the dendrogram presented in Figure 4, resulted in three homogeneous groups of sugarcane yield, each located in a different municipality. The sugarcane yield dendrogram was determined using Ward’s hierarchical method for cluster analysis [53]. These municipalities belong to nine microregions located in the state of Pernambuco: Brejo Pernambucano, Garanhuns, Itamaracá, Mata Meridional Pernambucana, Mata Setentrional Pernambucana, Médio Capibaribe, Recife, Suape and Vitória de Santo Antão. Among them, the main producing microregions are Mata Meridional and Setentrional (Figure 4).

Table 4 presents the municipalities categorized into homogeneous groups based on sugarcane yield. Group G1 included 14 producers located in the North Mata of Pernambuco and the Middle Capibaribe. Group 1 comprises situated municipalities in the northern part of the state, specifically in the Pernambuco forest region. This area experiences less traffic and is characterized by greater facilities for mechanization, which has consequently facilitated management in other production areas of the state. The most productive commercial crop varieties, RB92579, RB867515, and RB9350 [40], have been associated with investments in irrigation, reaching percentages of more than 60% [41].

Group 2 is more heterogeneous, covering 17 municipalities in the North Mata of Pernambuco, Médio Capibaribe, Garanhuns, Recife, Brejo de Pernambuco, and Vitória de Santo Antão. In Group 2, we find the majority of municipalities located in the transition areas between the forest and the agreste, and this distribution extends across the transverse region of the state.

Group 3 contains 26 sugarcane-producing municipalities in Mata Sul de Pernambuco, Brejo de Pernambuco, Suape, Recife, Vitória de Santo Antão, and Itamaracá. The group presents some sugarcane varieties with productive potential from an agro-industry perspective, such as SP784764, SP860621, and SP86127 [39]. The localities in the southern forest represent a region characterized by the Canary Islands sector and a topography with a more pronounced decline. This area experiences greater precipitation and lower productivity, influenced not only by climatic factors but also by management practices, which are primarily manual and less intensive.

Figure 5 illustrates the spatial distribution of the three homogeneous groups of 58 sugarcane-producing municipalities in eastern Pernambuco, Brazil, from 1990 to 2019. Group 1 (represented in blue) covers the northern part of the region. Group 2 (in green) presents an irregular distribution, with concentrations in the north and southwest; however, producing municipalities 16 (Chã Grande), 29 (Jaboatão dos Guararapes), and 51 (São José da Coroa Grande) are dispersed. Finally, the producing municipalities of Group 3 represent the central area of the map (in orange), with producing municipalities 24 (Igarassu) and 27 (Itapissuma) located outside this central region [41].

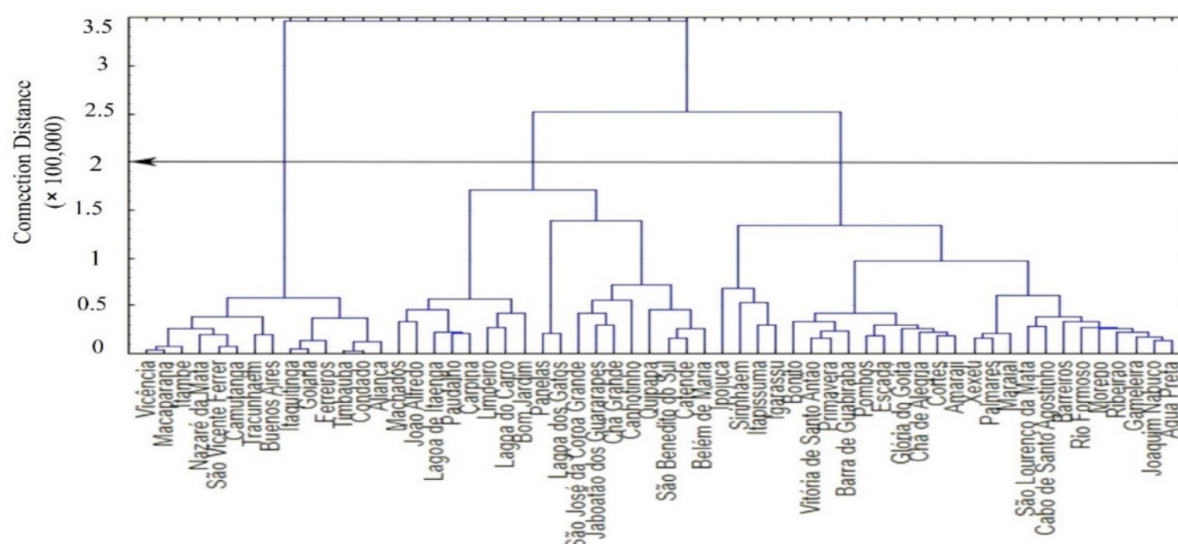


Figure 4. Dendrogram of 58 sugarcane productivity municipalities in eastern Pernambuco, Brazil, during 1990–2019. Source: Author.

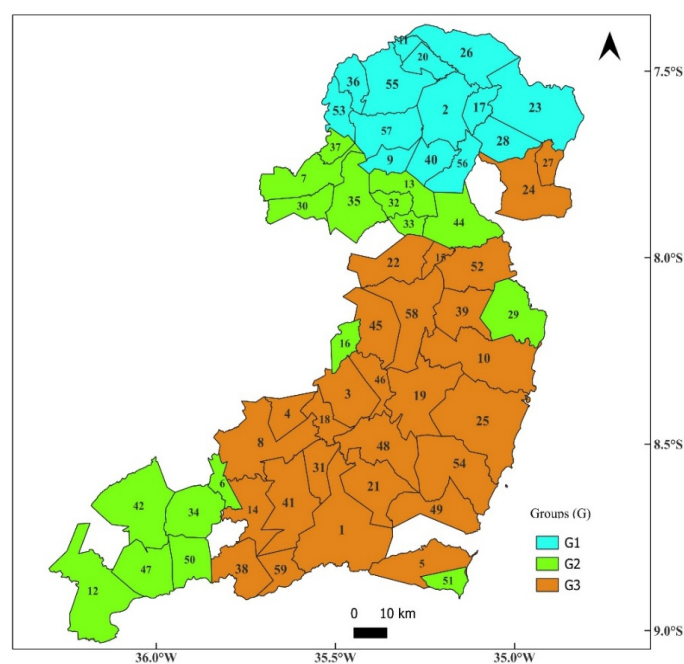


Figure 5. Spatial distribution of homogeneous groups of 58 sugarcane-producing municipalities in eastern Pernambuco, Brazil, during 1990–2019. Source: Author.

Table 4. Groups (G) of sugarcane producing municipalities in eastern Pernambuco, Brazil; id is the municipality identifier.

Groups	Municipalities (id)
G1	Aliança (2), Buenos Aires (9), Camutanga (11), Condado (17), Ferreiros (20), Goiana (23), Itambé (26), Itaquianga (28), Macaparana (36), Nazaré da Mata (40), São Vicente Férrer (53), Timbaúba (55), Tracunhaém (56), Vicência (57).
G2	Belém de Maria (6), Bom Jardim (7), Canhotinho (12), Carpina (13), Chã Grande (16), Jaboatão dos Guararapes (29), João Alfredo (30), Lagoa do Carro (32), Lagoa de Itaenga (33), Lagoa dos Gatos (34), Limoeiro (35), Machados (37), Paelas (42), Paudalho (44), Quipapá (47), São Benedito do Sul (50), São José da Coroa Grande (51).
G3	Água Preta (1), Amaraji (3), Barra de Guabiraba (4), Barreiros (5), Bonito (8), Cabo de Santo Agostinho (10), Catende (14), Chã de Alegria (15), Cortês (18), Escada (19), Gameleira (21), Glória do Goitá (22), Igarassu (24), Ipojuca (25), Itapissuma (27), Joaquim Nabuco (31), Marial (38), Moreno (39), Palmares (41), Pombos (45), Primavera (46), Ribeirão (48), Rio Formoso (49), São Lourenço da Mata (52), Sirinhaém (54), Vitória de Santo Antão (58), Xexéu (59).

3.2. Analysis Linear Correlations Three Months Before Sugarcane Yield

Determining linear correlations between meteorological variables three months before sugarcane yield is crucial to check the potential effects of multicollinearity and autocorrelation among the variables selected during the predictor selection process.

The results of the linear correlation analysis indicated significant correlations between the sugarcane yields of the G1, G2, and G3 groups and the tested predictor variables (Table 5). Although the relationship between the examined variables is positive, it is weak. However, correlation coefficients were significant (equal to or greater than 0.4) between the sugarcane yield groups and the predictor variables: sstNA, sstSA, Darwin, Tahiti, EN1+2, EN3, EN4, WE, WW, and WC. Therefore, the correlation analysis was insufficient to detect greater relationships between the groups regarding sugarcane yield and meteorological variables. This correlation did not demonstrate a high degree of linearity between the groups of variables (dependent and independent), which justifies the use of canonical correlation analysis to verify whether the correlation between these combinations is maximal. However, linear correlations equal to or greater than 0.4 between the independent variables (sst, wind speed, and atmospheric pressure) were significant with $p < 0.05$ at a confidence level of 95%, using Student's *t*-test.

Table 5. Mean, Standard Deviation (SD) and linear correlations between groups G1, G2, G3 with the variables sstNA (°C), sstSA (°C), Darwin (hPa), Tahiti (hPa), EN1+2 (°C), EN3 (°C), EN4 (°C), WE (m/s), WW (m/s), WC (m/s), three months before obtaining the total sugarcane yield; ton/ha: tons per hectare; hPa: hectopascal; m/s: meters per second.

Groups (G) and Variables	Mean	SD	G1	G2	G3	sstSA	sstNA	Darwin	Tahiti	EN1+2	EN3	EN4	WC	WE	WW
G1	52,033 ton/ha	6763 ton/ha	1.0	0.5	0.3	0.3	0.2	0.1	0.1	0.1	0.1	0.1	0.01	−0.3	0.2
G2	43,185 ton/ha	5648 ton/ha		1.0	0.8	0.3	0.4	−0.1	0.1	−0.2	0.1	0.2	0.2	0.1	0.1
G3	49,293 ton/ha	3698 ton/ha			1.0	0.3	0.2	0.1	0.2	−0.1	0.1	0.1	0.1	0.0	−0.1
sstSA (°C)	24.8 °C	0.3 °C				1.0	0.3	0.2	−0.2	0.6	0.5	0.2	−0.2	−0.4	0.1
sstNA (°C)	26.9 °C	0.3 °C					1.0	−0.2	0.1	−0.1	−0.1	−0.1	0.2	0.0	0.4
Darwin (hPa)	1007.5 hPa	0.9 hPa						1.00	−0.7	0.7	0.8	0.8	−0.9	−0.6	−0.8
Tahiti (hPa)	1011.3 hPa	0.9 hPa							1.0	−0.6	−0.7	−0.8	0.9	0.5	0.7
EN1+2 (°C)	23.0 °C	1.1 °C								1.0	0.9	0.6	−0.8	−0.9	−0.4
EN3 (°C)	25.3 °C	1.2 °C									1.0	0.8	−0.8	−0.9	−0.5
EN4 (°C)	28.5 °C	0.8 °C										1.0	−0.7	−0.6	−0.7
WC (m/s)	8.4 m/s	2.5 m/s											1.0	0.7	0.7
WE (m/s)	9.1 m/s	1.3 m/s												1.0	0.3
WW(m/s)	1.4 m/s	2.2 m/s													1.0

sstNA, sstSA, Darwin, Tahiti, EN1+2, EN3, EN4, WE, WW and WC represent the linear correlation coefficients of the average NDJ (November, December, and January) before obtaining the total sugarcane yield of the North Atlantic (NA) and South Atlantic (SA) sea surface temperatures (sst), surface atmospheric pressures in Darwin and Tahiti, sea surface temperatures in the El Nino (EN) areas (1+2, 3, 3.4, and 4), trade wind indices at 850 hPa in the Central Pacific (WC), Eastern Pacific (WE), and Western Pacific (WW), respectively.

3.3. Statistical and Practical Significance Analysis of the 3-Month Delay in Sugarcane Yields

Initially, statistical significance tests evaluated the correlations for the three canonical functions. The results presented in Table 6 indicate that only the first canonical correlation was statistically significant with *R* equal to 0.82 and with a *p*-value ≤ 0.05 at 95% confidence. In addition, multivariate tests evaluated the functions simultaneously, and Wilks' lambda significance test calculated the joint significance of the eigenvalues.

Table 6. Canonical correlation analysis and multivariate significance test.

Canonical Function	R	R ²	χ^2	df	<i>p</i>	Wilks' Lambda
1	0.82	0.67	44.18	27	0.02	0.14
2	0.62	0.38	19.37	16	0.25	0.42
3	0.56	0.32	8.62	7	0.28	0.68

R is the canonical correlation; *R*² is the canonical root, eigenvalue, and squared canonical correlation; χ^2 is the Chi-square; *df* is the degree of freedom; and *p* is the *p*-value. Wilks' lambda ranges from 0 to 1, where 0 indicates a complete difference between the groups, and 1 indicates no difference between the groups.

The Wilks' lambda test statistic indicates that the first canonical function exhibits a significant difference between the sugarcane yields and canonical sets, with a value of 0.14. Conversely, functions 2 and 3 display weaker differences between the groups, with values of 0.42 and 0.68, respectively. In addition, the calculated Chi-square value of 44.18 ($df = 27$) is greater than the critical 40.1 [58], showing an association between the sugarcane yield groups and the climate variables.

3.4. Redundancy Analysis of Independent Variables for Three Months Delay in Sugarcane Yield

The redundancy indices offer a concise summary of the relationship between variations in climatic factors and variations in sugarcane yield.

As presented in Table 7, the redundancy indices for the first canonical function between the canonical variables relate to sugarcane yields and the climatic variables. The redundancy index for sugarcane yield in group G1 is 0.297, indicating that climatic variations account for 29.7% of the variation in G1's yield. For the sugarcane yield groups G2 and G3, the redundancy indices are 0.599 and 0.522, respectively. The impact of climatic variations on yield variation was 59.9% for G2 and 52.2% for G3. An average redundancy index of 0.473 indicates that climatic variables accounted for 47.3% of the variation in sugarcane yield. The redundancy indices for each performance group provide a detailed view of how each group is affected by climatic variations, allowing the researcher to identify which groups are most sensitive to these conditions, thereby assisting in making decisions for optimizing performance.

Table 7. Calculation of the redundancy indices for the first canonical function.

Variables	Canonical Loading	Canonical Loading Squared	Average Loadings Squared	Canonical R ²	Redundancy Index
Groups of Dependent Variables					
G1	0.668	0.446			0.297
G2	0.949	0.901			0.599
G3	0.886	0.785			0.522
Sum of Square canonical loadings		2.131	0.710	0.665	0.473 *
Independent Climate Variables					
sstSA	0.427	0.182			0.121
sstNA	0.303	0.083			0.061
Darwin	0.035	0.001			0.001
Tahiti	0.222	0.034			0.033
EN1+2	−0.088	0.004			0.005
EN3	0.096	0.012			0.006
WC	0.184	0.027			0.023
WE	−0.074	0.019			0.004
WW	0.057	0.012			0.002
Sum of Square canonical loadings		0.449	0.041	0.002	0.028 *

* The redundancy index equals the average load squared multiplied by the canonical R² value.

The independent variable, represented by climatic factors, has a considerably low average redundancy index of 0.0284. However, this low value should not be considered unexpected, given the clear distinction between the dependent variable (yield) and the independent variable (climatic factors). The low redundancy of the independent statistical variation results from the relatively low shared variation in the independent statistical variation (0.0284), and not the canonical R².

The results of the redundancy analysis for the second and third canonical functions differ significantly (Table 8). Specifically, the canonical R² values are smaller, at 0.3787 and 0.3191, respectively, for the second and third functions. Furthermore, the degree of shared variance between sets of variables is notably small, at 0.1635 and 0.1811 for the dependent

statistical variables (sugarcane yield) in the second and third canonical functions and 0.1201 and 0.0401 for the independent statistical variables (climate), respectively.

Table 8. Redundancy analysis of dependent and independent variables for second and third canonical functions.

Standardized Variance of the Dependent Variables Explained by					
Their Own Canonical Variate (Shared Variance)			The Opposite Canonical Variate (Redundancy)		
Canonical Function	Percentage	Cumulative Percentage	Canonical R ²	Percentage	Cumulative Percentage
1	0.6554	0.6554	0.6655	0.4362	0.4362
2	0.1635	0.8189	0.3787	0.0619	0.4981
3	0.1811	1.000	0.3191	0.0578	0.5559

Standardized Variance of the Independent Variables Explained by					
Their Own Canonical Variate (Shared Variance)			The Opposite Canonical Variate (Redundancy)		
Canonical Function	Percentage	Cumulative Percentage	Canonical R ²	Percentage	Cumulative Percentage
1	0.0191	0.0191	0.6655	0.0127	0.0127
2	0.1201	0.1392	0.3787	0.0455	0.0582
3	0.0401	0.1794	0.3191	0.0128	0.0710

When combined with the canonical root in the redundancy index, the dependent statistical variables (sugarcane yield) produce values of 0.0619 and 0.0578, while the independent statistical variables (climate) produce values of 0.0455 and 0.0128. As a result, the second and third canonical functions have no statistical or practical significance.

3.5. Canonical Weights of the Three Canonical Functions

The magnitude of the weights represents their relative contribution to the corresponding canonical function. In this way, a positive canonical weight indicates that an increase in the variable is associated with an increase in the sugar cane yield. In contrast, a negative canonical weight suggests that an increase in the variable is associated with a decrease in the yield.

Table 9 displays the canonical weights of both the dependent and independent variables for each canonical statistical variable. The analysis of the first canonical function revealed that sugarcane yield groups G1, G2, and G3 all demonstrate positive canonical weights, indicating they contribute to increased sugarcane yield in the dependent canonical variable. For the climatic variations, the evaluation of the first canonical function revealed that the sea surface temperature (sst) of the area EN1+2, North Atlantic (sstNA) and wind in the eastern and western Pacific (OE, OW) were weighted -2.3312 , -0.0465 , -0.9910 and -0.1420 , indicating that an increase in these variables was associated with a decrease in sugarcane yield. Climatic variables with negative weights, such as EN1+2, suggested that adverse climatic conditions may negatively impact sugarcane yield.

On the other hand, the climatic variables with greater positive weights were the SST (1.0140) in the EN3 area, the atmospheric pressure in Darwin (0.9806), and the SST in the South Atlantic area (0.8243), indicating that an increase in these variables was associated with an increase in the sugarcane yield, highlighting their importance as positive predictors for yield. However, the performance groups G1, G2, and G3 exhibit positive weights, indicating a direct relationship with increased performance. In contrast, climatic variations such as the sea surface temperature (SST) of EN1+2 and wind speed have negative weights, suggesting an inverse relationship with performance. This analysis underscores the importance of monitoring specific climatic factors to optimize sugarcane yield.

Table 9. Canonical weights for the three canonical functions.

Standardized Canonical Coefficients	Function 1	Function 2	Function 3
Groups of Dependent Variables	Canonical Weights	Canonical Weights	Canonical Weights
G1	0.2544	1.0412	−0.5457
G2	0.4859	−1.5050	−0.8432
G3	0.4166	0.8277	1.3147
Independent Climate Variables			
sstSA	0.8243	0.3287	0.4054
sstNA	−0.0465	−0.3847	0.2834
Darwin	0.9806	1.0991	−0.5930
Tahiti	0.5113	0.5564	1.3655
EN1+2	−2.3312	−0.2823	0.5105
EN3	1.0140	−0.9627	0.2567
WC	0.6518	−1.2415	−1.9951
WE	−0.9910	−0.5345	1.6829
WW	−0.1420	1.2590	−0.5408

Canonical equations describe the relationship between climate variables (predictors) and sugarcane yield groups (responses) in a statistical model. The pair of equations (U1, V1) for the first canonical function of the variables generated the following Equations (16) and (17):

$$U1 = 0.2544 \times G1 + 0.4859 \times G2 + 0.4166 \times G3 \quad (16)$$

$$V1 = 0.8243 \times \text{sstSA} - 0.0465 \times \text{sstNA} + 0.9806 \times \text{Darwin} + 0.5113 \times \text{Tahiti} \\ - 2.3312 \times \text{EN12} + 1.0140 \times \text{EN3} + 0.6518 \times \text{WC} \\ - 0.9910 \times \text{WE} - 0.1420 \times \text{WC} \quad (17)$$

Equation (16) represents the canonical variable U1, which depends on sugarcane yield, expressed as a weighted combination of yield groups 1, 2, and 3. The three groups exert a direct influence on the canonical variable U1, since their weights are positive. In return, Equation (17) represents a resulting canonical variate, which is a linear combination of climate variables that affect sugarcane productivity. The variables temperature (sstSA and EN3), atmospheric pressure (Darwin and Tahiti), and wind (WC) had a positive effect, contributing to increased sugarcane yield. On the other hand, the temperature of the EN1+2 area and the wind variables (WE and WW) contributed to the decrease in sugarcane yield.

3.6. Structural Canonical Loadings for the Three Canonical Functions

Structural canonical loadings measure the simple linear correlation between an original observed variable in the dependent or independent set and the canonical statistical variable of the set. Table 10 presents the structural canonical loadings for the independent and dependent variables across the canonical functions. The analysis of the structural canonical loadings of the first canonical function indicates that the three sugarcane yield groups contributed positively to the canonical dependent variable. Specifically, sugarcane yield group G2 (0.6676) accounted for 42.26% of the variability in the dependent canonical variable, followed by group G3 with 36.82% and group G1 with 20.92%.

In evaluating the canonical loading for the first canonical function, the sea surface temperature of the South Atlantic (sstSA) has a loading of 0.4267, contributing 47.40% to the variability of the independent canonical variable. The sea surface temperature in the North Atlantic (sstNA) with 23.95% (0.3033), atmospheric pressure in Tahiti with 12.85%

(0.2222), wind in the West Pacific (WC) with 8.81% (loading = 0.1840), and the SST of EN3 with 2.42% (0.0964).

Table 10. Structural canonical loadings for the three canonical functions.

Groups of Dependent Variables	Function 1		Function 2		Function 3	
	Canonical loadings	CV (%)	Canonical loadings	CV (%)	Canonical loadings	CV (%)
G1	0.6676	20.92	0.5161	75.49	−0.5365	55.78
G2	0.9490	42.26	−0.2929	24.31	−0.1164	2.63
G3	0.8858	36.82	0.0264	0.20	0.4633	41.60
Independent Climate Variables						
sstSA	0.4267	47.40	0.2682	7.30	0.1892	13.57
sstNA	0.3033	23.95	−0.0810	0.67	−0.0837	2.66
Darwin	0.0347	0.31	0.2515	6.42	0.1696	10.91
Tahiti	0.2222	12.85	0.0628	0.40	0.1708	11.06
EN1+2	−0.0875	1.99	0.4824	23.61	0.0643	1.57
EN3	0.0964	2.42	0.3179	10.25	0.0001	0.00
WC	0.1840	8.81	−0.3654	13.54	−0.1199	5.45
WE	−0.0735	1.41	−0.5776	33.84	0.2234	18.92
WW	0.0570	0.85	0.1980	3.98	−0.3075	35.86

CV is the contribution of the variable (%) to the canonical variable, estimated as the square of the loadings divided by the sum of the squares of the loadings multiplied by 100.

It is crucial to note that while these contributions reflect the variability of the climatic variables in the independent canonical variable, they do not necessarily indicate a direct effect on sugarcane yield. The contributions of the canonical loadings of the climatic variables align with the evaluation of the canonical weights.

Therefore, the main predictors of sugarcane yield identified in this analysis are wind speed in the Central Pacific, atmospheric pressure in the Tahiti area, and SST in the El Niño 3 area, as well as in the North Atlantic and South Atlantic. These predictors represent the variables that demonstrate the high correlation and explanatory power regarding variability in sugarcane yield. The analysis highlights the importance of oceanographic and atmospheric parameters in the tropical regions of the Pacific and Atlantic Oceans.

3.7. Canonical Cross-Loadings for the Three Canonical Functions

The computation of canonical cross-loadings involves correlating each original observed dependent variable (sugarcane yield groups) directly with the independent canonical statistical variable (climatic variables) and vice versa. Table 11 presents the canonical cross-loadings for the three canonical functions.

Table 11. Canonical cross-loadings for the three canonical functions.

Groups of Dependent Variables	Function 1		Function 2		Function 3	
	Canonical Cross-Loadings	CV (%)	Canonical Cross-Loadings	CV (%)	Canonical Cross-Loadings	CV (%)
G1	0.5447	20.93	0.3176	75.50	−0.3031	55.79
G2	0.7742	42.26	−0.1802	24.30	−0.0657	2.62
G3	0.7226	36.82	0.0162	0.20	0.2617	41.59
Independent Climate Variables						
sstSA	0.3481	47.40	0.1650	7.87	0.1069	15.71
sstNA	0.2474	23.94	−0.0498	0.72	−0.0473	3.08
Darwin	0.0283	0.31	0.1548	6.92	0.0958	12.62
Tahiti	0.1813	12.86	0.0387	0.43	0.0965	12.80
EN1+2	−0.0714	1.99	0.2969	25.47	0.0363	1.81
EN3	0.0786	2.42	0.1957	11.06	0.0000	0.00
WC	0.1501	8.81	−0.2248	14.60	−0.0677	6.30
WE	−0.0600	1.41	−0.3555	36.51	0.1262	21.90
WW	0.0465	0.85	0.1218	4.29	−0.1737	41.49

CV is the contribution of the variable (%) to the canonical variable, estimated as the square of the cross-loadings divided by the sum of the squares of the cross-loadings multiplied by 100.

Analyzing the first canonical function reveals that the sugarcane yield groups G1, G2, and G3 contribute to the formation of the climatic canonical variable with 20.93%, 42.26%, and 36.82%, respectively. The climatic characteristics impacted the performance of the two groups, G2 and G3.

The analysis of the first canonical function for the climatic variables revealed that the sea surface temperature of the Southern Atlantic (sstSA) was the most significant factor in the relationship between the canonical yield and climatic variables, accounting for 47.40% of the relationship. Additionally, the variables sstNA, atmospheric pressure in Tahiti, wind speed in the Western Pacific (WC), and the SST of the EN3 area contributed positively to the canonical variation of sugarcane yields, with contributions of 23.95% (0.2474), 12.86% (0.1813), 8.81% (0.1501), and 2.42% (0.0786), respectively.

It is important to note that G1 and sstSA influenced the relationship between the canonical variables of the sugarcane yield groups and the set of climatic variables. The percentages of the structural and cross-canonical loadings are equal, indicating a direct relationship between the variables, homogeneity of the two groups, an adequate canonical analysis model, and symmetry in the canonical analysis.

3.8. Validation and Diagnosis

Table 12 presents the outcomes of the sensitivity analysis, specifically when excluding individual independent variables from the study. The canonical loads are stable and consistent in the three cases where an independent variable was removed (sstNA, WC, and EN3). A variation occurred in the canonical correlation (R) with the removal of the EN3 variable, changing only three hundredths, and in the canonical load of G1 with a difference of six-hundredths compared with the original value with all variables.

Table 12. Sensitivity analysis of the canonical correlation results to removal of an independent variable.

Complete variate		Result After Deletion of		
		sstNA	WC	EN3
Canonical correlation (R)	0.82	0.82	0.81	0.79
Canonical root (R^2)	0.67	0.67	0.65	0.63
Dependent variate				
G1	0.67	0.68	0.65	0.73
G2	0.95	0.95	0.91	0.91
G3	0.89	0.88	0.92	0.87
Shared variance	0.66	0.66	0.63	0.67
Redundancy index	0.47	0.47	0.46	0.44
Independent variate				
Canonical Loadings				
sstSA	0.43	0.43	0.46	0.46
sstNA	0.30	-	0.29	0.30
Darwin	0.04	0.04	0.06	0.06
Tahiti	0.22	0.22	0.24	0.23
EN12	−0.09	−0.08	−0.06	−0.05
EN3	0.10	0.10	0.11	-
WC	0.18	−0.18	-	0.16
WE	−0.07	−0.08	−0.09	−0.13
WW	0.06	0.06	0.05	0.08
Shared variance	0.02	0.02	0.01	0.02
Redundancy	0.03	0.02	0.03	0.03

4. Discussion

Hierarchical cluster analysis is a multivariate statistical technique that subdivides data into groups. The cluster analysis aimed to classify sugarcane yield data from the eastern sector of Pernambuco, Brazil. Variables reflecting different aspects of agricultural

production, such as local climate conditions, soil type, management practices, and varietal characteristics, were not considered. We conducted cluster analysis on sugarcane yield data using various distance measures including Manhattan (city block), Chebyshev, squared Euclidean, percent disagreement, Pearson r , and Euclidean distances, employing Ward's method. All distance measures resulted in distortions in the dendrogram of the sugarcane yield groups, except for the Euclidean distance. Since the sugarcane yield was on the same scale (tons per hectare), we chose to use the Euclidean distance for our analysis. The result of this analysis was the construction of a dendrogram, a graphical representation illustrating the structural similarity between the different groups of data.

The dendrogram revealed the existence of three homogeneous groups of sugarcane yield, with distinct characteristics attributed to specific factors that influence productivity. The first group presented high sugarcane yields associated with advanced agricultural practices and ideal growing conditions. The second group, in turn, comprised sugarcane mills in municipalities between the Zona da Mata and Agreste regions of Pernambuco. These sugarcane mills had moderate yields, reflecting a combination of appropriate management practices, but with some limitations, such as lower rainfall or suboptimal soil characteristics. However, these sugarcane mills are easy to use and mechanized, with more efficient soil and crop management compared with other productive areas of the state [59]. Finally, the third group of sugarcane mills had lower yields, possibly due to challenges such as steep topography, adverse weather conditions, and outdated agricultural practices that negatively impact production. Although this group had the highest rainfall [60,61], sugarcane yields remained below expectations due to the interaction of factors such as climate, soil characteristics, and management practices. In particular, agricultural practices, which are predominantly manual and of low intensity in sugarcane mills, have a significant impact on reducing yield. Furthermore, there has been a reduction in sugarcane production capacity since 2008, a period in which there was an intensification of mechanization, especially in harvesting [62,63]. This drop in production is due to the reduction in the quality of the harvested material, which contributes to productivity losses throughout the affected area with each harvest [64].

After analyzing the linear correlations, the independent variables SST, wind, and pressure did not show sufficient linearity with sugarcane yield. However, classic linear correlation analysis is limited in its ability to reveal complex relationships between multiple variables [65]. This approach examines the relationship between two variables at a time, which may not capture the full dynamics between sugarcane yield and meteorological variables. Furthermore, Mazouz et al. [66] add that the resulting intra-group correlation implies that individual observations are independent, which contradicts the assumption of independence required by conventional statistical models. This violation of the independence assumption may lead to biased and inaccurate results in the analysis. This limitation justifies the use of canonical correlation analysis since this technique requires fewer restrictions on the types of data with which it operates [67]. Furthermore, there is a consensus among researchers that canonical correlation is the most appropriate multivariate technique in situations where many dependent and independent variables of one or more materials [68] or events are analyzed. Canonical correlation analysis allows a more comprehensive understanding of the interactions and influences that meteorological variables exert on sugarcane yield.

Canonical correlation analysis seeks a set of prognostic equations between the sugarcane yield group and climatic variables [69]. In this way, the canonical correlation analysis and multivariate significance test results were statistically significant for the first canonical function, showing that the group of climate variables influences the three sugarcane yield groups. The growth of sugarcane has an important relationship with climate

factors [70], and these factors are required in different amounts in different periods of sugarcane growth [71]. Local meteorological factors such as rainfall, air temperature, and wind speed influence the efficiency of sugarcane yield by approximately 43% [72]. The sugarcane area of Pernambuco consists of Mata Norte, a dry region (1400 mm), and Mata Sul, where the annual rainfall regime is 1900 mm [33]. The difference in the quantity of rainfall worries many farmers who supply sugarcane. The southeast trade winds transport the humidity of the South Atlantic heat to the southern forest area, favoring rainfall and increasing the sugarcane yield. The northern forest area does not directly receive the trade winds, suffering from high temperatures and dryness [73]. Although there are no studies in the literature that use the same type of analysis applied to sugarcane yield nor established reference values for canonical correlation and its square, we can infer that the values obtained are considerably high. In linear regression analysis, the total variance of the dependent variable is a measure of the variation of the data around the mean, expressed in terms of the sum of total squares, which is not necessarily equal to 1. However, canonical correlation, which simultaneously analyzes sets of dependent variables, assumes that the total variance explained reaches 100% [67]. Therefore, the expected values in canonical correlation tend to be below this limit.

Redundancy expresses how much of the variance in one set of variables can be explained by another; it is a measure of the average proportion of variance in set Y that is accounted for by set X and is comparable to the squared multiple correlation in multiple linear regression [65]. Mazouz et al. [66] justify the use of the redundancy index due to the lack of independence caused by intragroup and temporal correlations in time series data of climate variables. Furthermore, the emphasis of the index is to explain the effect of the independent set on the dependent set rather than the correlation between the two sets, as in canonical correlation analysis. In this study, the redundancy rates were lower, representing precisely 30% in group 1, where the municipalities belong to the northern forest zone, known for having fewer climate variations with drier characteristics compared with the municipalities of the southern forest zone, which according to the cluster analysis were concentrated in group 2, presenting a redundancy rate of approximately 52%. The highest rate (60%) occurred in transition areas of the forest zone and agreste with variable topography and soil characteristics. Lower productivity averages may relate not only to the production environment but also to the lower management required. It can be inferred that the crop in this region responded better to the climate variations during the study period, providing a greater degree of prediction accuracy in the prediction due to limited influence of management.

Analysis of the results showed that the low redundancy of the independent variable (climate) resulted from the relatively low shared variance in the independent variable and not from the canonical R^2 . Hair Junior et al. [68] state that canonical loadings (or canonical structural correlations) measure the simple linear correlation between an original variable observed in the dependent or independent set and the canonical statistical variable of its respective set. The larger the loading, the more important the variable is for deriving the canonical statistical variate. The redundancy index analysis and significance tests used the first canonical function.

The canonical weights of the dependent and independent variables reflect their contributions to the canonical function [12]. For the first canonical function, the independent variables that contributed the most were EN1+2, EN3, and WE. The dependent variables associated with this function were G2, G3, and G1, whose sequence does not follow the logic of canonical functions 2 and 3, indicating complex interactions between the variables. Thus, it is crucial to consider the instability of the canonical weights in contexts of multicollinearity, which can distort their interpretation. Wildt et al. [74] suggest using

generalized jackknife statistics to provide estimated standard errors for testing the significance level of weights and canonical loadings. However, the recommendation is to use canonical and cross-loadings for a more robust analysis of the relationships between variables. Canonical loadings help to clarify the correlations between the original and canonical variables, while cross-loadings elucidate the relationships between dependent and independent variables in different contexts [75].

The canonical load is a crucial factor for understanding and interpreting the relationships between sets of variables in statistical analyses. Therefore, the canonical load of the first independent statistical function had predictors: wind speed in the Central Pacific Ocean, atmospheric pressure in Tahiti, and Sea Surface Temperature in the El Niño-3 and North and South Atlantic areas. That does not rule out the use of predictors that are known to affect sugarcane yields, such as solar radiation, rainfall, and air temperature. The predictors used have the advantage of being obtained and related with a lag of up to three months before the sugarcane harvest, allowing analyses similar to rainfall forecasts [12].

The statistical variables present low redundancy values, as evidenced by the significantly reduced canonical loadings for both variables in the second and third functions. Therefore, the negative response reflected in the lower canonical loadings and low redundancy reinforces the lack of practical significance of the second and third functions. Temperature variations in the Atlantic Ocean influence the rainfall regime in the eastern sector of Pernambuco, Brazil. Sea surface temperature influences rainfall and air temperature variations in Northeast Brazil [76]. The El Niño Southern Oscillation (ENSO) is a phenomenon observed in the Equatorial Pacific Ocean area and impacts the climate and weather of several parts of the world [77]. ENSO also influences the trade winds in the Equatorial Pacific region and the Southern Oscillation. Thus, its influence on the water regime of the NEB and, consequently, on sugarcane productivity can be inferred. The study is crucial for the management of natural resources and, consequently, for the progress of human activities in the region [78].

The validation of the canonical correlation analyses used stability analysis of the independent variables of the set. The canonical loadings demonstrated stability for the three sugarcane yield groups, with the removal of the variable EN3 in group 1. According to Moura et al. [12], the canonical loadings, like weights, vary from one sample to another. This variability suggests that the loadings, and therefore the relationships, may be sample-specific, resulting from occasional extrinsic factors.

5. Conclusions

The yield of sugarcane plays a fundamental role in food security and biofuel production, being strongly influenced by climatic variables. The ability to predict and understand the impact of these variables on cultivation is essential for optimizing agricultural practices and mitigating risks associated with climate change. In this context, a hierarchical cluster analysis conducted on 58 sugarcane-producing municipalities in the eastern region of Pernambuco identified three distinct and homogeneous groups based on their annual sugarcane production. This classification revealed clear patterns of sugarcane productivity, demonstrating how localized climatic conditions influence agricultural performance in different regions.

The application of linear correlation analysis was insufficient to detect greater relationships between groups regarding sugar cane productivity and meteorological variations. This correlation did not demonstrate a high degree of linearity between the groups of variables (productivity and climatic variables), justifying the use of canonical correlation analysis to verify that the correlation between these combinations is maximal.

The canonical correlation analysis was more significant for the first canonical function, which showed $R = 0.82$ ($p < 0.05$ with 95% confidence), indicating that the sugarcane yield groups and climatic variables are different (Wilks' $\lambda = 0.14$), but they were associated. The accuracy of sugarcane yield estimation was 0.67 for the first canonical function.

Climatic variables affect sugarcane yield, with redundancy indices for the first canonical function varying from 29.7% to 59.9% between the different groups. This difference suggests that management practices and cultivation strategies may need to be adapted according to the specific characteristics of each group to maximize productivity.

An analysis of weights for the first canonical function revealed that climatic variables such as sea surface temperature (EN1+2) and wind speed (WE) have negative weights, suggesting that adverse climatic conditions may negatively impact sugarcane yield. In contrast, the positive weights of climatic variables including sea surface temperature (SST) in the EN3 area and atmospheric pressure in Darwin indicate that these favorable conditions are associated with higher sugarcane yield.

An analysis of the structural canonical loads reveals that the sea surface temperature of the Southern Atlantic Ocean (sstSA) emerged as the most important climatic variable, contributing 47.40% to the variability of the independent canonical variable. This strong correlation indicates that oceanic conditions play a crucial role in determining sugarcane yield, potentially influencing precipitation and temperature patterns in the region. Other climatic variables, such as the sea surface temperature of the North Atlantic Ocean (sstNA), the atmospheric pressure in Tahiti, the wind speed in the Western Pacific, and the SST of the EN3 region, also showed relevant contributions, such as 23.95%, 12.85%, 8.81% and 2.42%, respectively. These discoveries underscore the complexity of interactions between climatic factors and agricultural yields.

The analysis of canonical cross-loadings correlates the groups of sugarcane yield with the climatic variables. The groups G1, G2, and G3 contribute to the formation of the canonical climatic variation by 20.93%, 42.26%, and 36.82%, respectively. The surface temperature of the South Atlantic Sea (SST) is the most significant factor, accounting for 47.40% of the relationship with climatic variations. Other variables, such as TSA, atmospheric pressure in Taiti, wind speed in the Western Pacific, and TSM in the EN3 area, also contribute positively to the canonical variation in sugar cane productivity.

The sensitivity analysis indicates that the developed canonical model is robust, showing that climatic variations have a significant, though not critical, impact on the overall structure of the analysis. The stability of the canonical loads confirms the robustness of the results, demonstrating that the research can support yield forecasting, regional agricultural policy, and drought management. Identifying climate predictors, such as sea surface temperature, wind speed, and atmospheric pressure, enables the creation of accurate models to predict sugarcane productivity, assisting farmers in planning input management, irrigation during dry periods, and harvesting. Furthermore, climate data can inform policies that encourage sustainable agricultural practices and adaptation to climate conditions, strengthening food security and guiding the selection of more resilient sugarcane varieties, increasing production resilience.

Author Contributions: Conceptualization, R.R.d.S.; Formal analysis, R.R.d.S.; Investigation, R.R.d.S.; Data curation, R.R.d.S.; Funding acquisition, R.R.d.S.; Methodology, R.R.d.S.; Resources, R.R.d.S.; Conceptualization, G.B.d.A.M.; Formal analysis, G.B.d.A.M.; Investigation, G.B.d.A.M.; Data curation, G.B.d.A.M.; Funding acquisition, G.B.d.A.M.; Methodology, G.B.d.A.M.; Resources, G.B.d.A.M.; Conceptualization, P.M.O.L.; Formal analysis, P.M.O.L.; Investigation, P.M.O.L.; Funding acquisition, P.M.O.L.; Methodology, P.M.O.L.; Resources, P.M.O.L.; Writing—original draft, P.M.O.L.; Writing—review, P.M.O.L.; Visualization, P.M.O.L.; Conceptualization, C.R.N.; Formal analysis, C.R.N.; Methodology, C.R.N.; Resources, C.R.N.; Conceptualization, C.R.N.; Conceptualization,

P.R.G.; Formal analysis, P.R.G.; Investigation, P.R.G.; Methodology, P.R.G.; Conceptualization, P.R.G.; Investigation, P.R.G. All authors have read and agreed to the published version of the manuscript.

Funding: We thank the Postgraduate Program in Agricultural Engineering at the Federal Rural University of Pernambuco (UFRPE) for the logistics necessary for this research. We would also like to express our gratitude to the National Supply Company (Conab) for granting permission for the release of the first author to pursue a Master's degree in Agricultural Engineering at UFRPE, as outlined in internal GECAP document code 040/2020. The authors are grateful for the financial support for publication through the State University of Goiás (UEG) notice, PRÓ-PROGRAMAS Nº 01/2024, SEI process nº 202400020011132.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

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

References

1. Compre Rural. Cana de Açúcar Movimenta Mais movimento mais de R\$ 100 Bi Por Ano Nobi por ano no Brasil. 2022. Available online: <https://www.comprerural.com/cana-de-acucar-movimenta-mais-de-r-100-bi-por-ano-no-brasil/> (accessed on 13 May 2025).
2. CONAB—Companhia Nacional de Abastecimento. CANA-DE-AÇÚCAR. Acompanhamento Da Safra Brasileira De Cana-De-Açúcar. Brasília. Safra 2019/20. Quarto Levantamento. CONAB: Brasília, Brazil, 2020; Volume 6, pp. 1–58. Available online: https://www.conab.gov.br/info-agro/safra/cana/boletim-da-safra-de-cana-de-acucar/item/download/31590_6cfbbc41aa04783c69113c50fa499cba (accessed on 24 February 2024).
3. NovaBio. Atual Safra de Cana-De-Açúcar Deve Crescer 3,5% em Pernambuco. 2024. Available online: <https://novabio.org/atual-safra-de-cana-de-acucar-deve-crescer-35-em-pernambuco/> (accessed on 24 February 2024).
4. Melo, P.J.D.; Lúcio, A.D.C.; Torres, E.M.J.; Lopes, S.J.; Lambrecht, D.M.; Diel, M.I. Association technological and agronomic traits sugarcane for the quantitative and qualitative production of brown sugar. *Rev. Bras. Ciênc. Agrár.* **2023**, *18*, e2980. [CrossRef]
5. Machado, S. Why Is Burning of Sugarcane Fields Still Allowed in the Country Despite the Fires? 2024. Available online: <https://www.bbc.com/portuguese/articles/cjdk41z41zno> (accessed on 13 May 2025).
6. Almeida, A.C.S.; Souza, J.L.; Teodoro, I.; Barbosa, G.V.S.; Moura Filho, G.; Ferreira Júnior, R.A. Desenvolvimento vegetativo e produção de variedades de cana-de-açúcar em relação à disponibilidade hídrica e unidades térmicas. *Ciência Agrotec.* **2008**, *32*, 1441–1448. [CrossRef]
7. Luchiari, A., Jr.; Ramos, N.P. Impacts. 2022. Available online: <https://www.embrapa.br/agencia-de-informacao-tecnologica/cultivos/cana/pre-producao/meio-ambiente/impactos#:~:text=Os%20preju%C3%ADzos%20s%C3%A3o:%20perda%20de.consumidores%20e%20dos%20pr%C3%B3prios%20agricultores> (accessed on 23 January 2025).
8. Da Silva, V.D.P.; Oliveira, S.D.D.; Dos Santos, C.A.; Silva, M.T. Risco climático da cana-de-açúcar cultivada na região nordeste do Brasil. *Rev. Bras. De Eng. Agrícola E Ambient* **2013**, *17*, 180–189. [CrossRef]
9. IPCC—Intergovernmental Panel on Climate Change. Climate Change 2021: The Physical Science Basis. Summary for Policy Makers Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Available online: <https://www.ipcc.ch/report/sixth-assessment-report-cycle/> (accessed on 24 September 2022).
10. IPCC—Intergovernmental Panel on Climate Change. Climate Change 2007: The Physical Science Basis. Summary for Policy Makers Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. 2007. Available online: <https://www.ipcc.ch/site/assets/uploads/2018/02/ar4-wg1-frontmatter-1.pdf> (accessed on 22 May 2025).
11. de Carvalho, A.L.; Menezes, R.S.C.; Nóbrega, R.S.; Pinto, A.d.S.; Ometto, J.P.H.B.; von Randow, C.; Giarolla, A. Impact of climate changes on potential sugarcane yield in Pernambuco. *Northeast. Reg. Brazil. Renew. Energy* **2015**, *78*, 26–34. [CrossRef]
12. Moura, G.B.A.; Brito, J.I.B.; Sousa, F.A.S.; Cavalcanti, E.P.; Silva, J.L.B.; Nascimento, C.R.; Lopes, P.M.O. Identificação de Preditores Para as Chuvas do Setor Leste do Nordeste do Brasil Utilizando Análise de Correlação Canônica. *Rev. Bras. Geog. Fis.* **2020**, *13*, 1463. [CrossRef]
13. Lambert, Z.V.; Durand, R.M. Some Precautions in Using Canonical Analysis. *J. Mark. Res.* **1975**, *12*, 468–475. [CrossRef]
14. Trugilho, P.F.; Lima, J.T.; Mori, F.A. Correlação canônica das características químicas e físicas da madeira de clones de *Eucalyptus grandis* e *Eucalyptus saligna*. *Rev. Cer.* **2003**, *9*, 66–80.
15. Mingoti, S.A. *Análise De Dados Através De Métodos De Estatística Multivariada: Uma Abordagem Aplicada*; UFMG: Belo Horizonte, Brazil, 2005; 297p.
16. Rossi, D.A.; da Silva Menezes, B.R.; Daher, R.E.F.; de Amaral Gravina, G.; de Lima, R.N.; da Silva Ledo, F.J.E.; Gottardo, R.D.; Campostrini, E.; de Souza, C.A.L.M. Canonical correlations in elephant grass for energy purposes. *Afr. J. Biot.* **2014**, *13*, 3666–3671.

17. L'Heureux, M.L.; Tippet, M.K.; Barnston, A.G. Characterizing ENSO coupled variability and its impact on North American seasonal precipitation and temperature. *J. Clim.* **2015**, *28*, 4231–4245. [CrossRef]
18. Rana, S.; Renwick, J.; McGregor, J.; Singh, A. Seasonal prediction of winter precipitation anomalies over Central Southwest Asia: A canonical correlation analysis approach. *J. Clim.* **2018**, *31*, 727–741. [CrossRef]
19. Nobre, D.A.C.; dos Santos Silva, F.C.; Guimarães, J.F.R.; de Resende, J.C.F.; Macedo, W.R. Análise de Trilha e Correlação Canônica nos Componentes do Desempenho de Girassol. *J. Eng. Exac. Scie.* **2018**, *4*, 364–369. [CrossRef]
20. Masri, M.I.; Shaban, S.A.; El-Hennawy, H.H.; El-Taib, A.B.A.; Abu-El-lail, F.F.B. Phenotypic and genotypic correlations and path coefficient analysis in sugarcane at first clonal selection stage. *Egypt. J. Plant Breed.* **2015**, *19*, 297–321. [CrossRef]
21. Barbosa, R.P.; Neto, F.A.; Gravina, M.L.; Gravina, G.A.; Portela, M.G.; Bezerra, A.A. Seleção antecipada de cana-de-açúcar usando análise de trilha. *Genet. Mol. Res.* **2017**, *16*, e16019038.
22. Ali, A.; Khan, S.A.; Tahir, M.; Farid, A.; Khan, A.; Khan, S.M.; Ali, N. Evaluation of various sugarcane genotypes for association of quality characters with cane yield. *Adv. Crop Sci. Technol.* **2018**, *6*, e1000371. [CrossRef]
23. Ali, A.; Khan, S.A.; Tahir, M.; Farid, A.; Khan, A.; Khan, S.M.; Ali, N. Clonal selection strategy in sugarcane (*Saccharum officinarum* L.) based on the association of quality traits and cane yield. *J. Anim. E Plant Sci.* **2019**, *29*, 889–893.
24. Forootan, E.; Khandu; Awange, J.L.; Schumacher, M.; Anyah, R.O.; Van Dijk, A.I.J.M.; Kusche, J. Quantifying the impacts of ENSO and IOD on rain gauge and remotely sensed precipitation products over Australia. *Remote Sens. Environ.* **2016**, *172*, 50–66. [CrossRef]
25. Luciano, A.C.S.; Picoli, M.C.A.; Duft, D.G.; Rocha, J.V.; Leal, M.R.L.; Maire, G.L. Empirical model for forecasting sugarcane yield on a local scale Brazil using Landsat imagery and random forest algorithm. *Comput. Electron. Agric.* **2022**, *184*, e106063. [CrossRef]
26. Castro Manrique, D.R.; Lopes, P.M.O.; Nascimento, C.R.; Ribeiro, E.P.; Silva, A.S.d. Phenological Monitoring of Irrigated Sugarcane Using Google Earth Engine, Time Series, and TIMESAT in the Brazilian Semi-Arid. *AgriEngineering* **2024**, *6*, 3799–3822. [CrossRef]
27. Mbhamali, T.; Jury, M. Climate-sensitivity of sugarcane yield in the southeastern Africa lowlands. *Int. J. Biometeorol.* **2021**, *41*, 4187–4200. [CrossRef]
28. Verma, A.K.; Garg, P.K.; Prasad, K.S.H.; Dadhwal, V.K. Variety-specific sugarcane yield simulations and climate change impacts on sugarcane yield using DSSAT-CSM-CANEGRO model. *Agric. Water Manag.* **2023**, *275*, 108034. [CrossRef]
29. Legendre, B.L. Ripening of Sugarcane: Effects of Sunlight, Temperature, and Rainfall. *Crop Sci.* **1975**, *15*, 349–352. [CrossRef]
30. Araujo, R.; Júnior, J.; Casaroli, D.; Pêgo Evangelista, A. Variation in the sugar yield in response to drying-off of sugarcane before harvest and the occurrence of low air temperatures. *Bragantia* **2016**, *75*, 118–127. [CrossRef]
31. Alvares, C.A.; Stape, J.L.; Sentelhas, P.C.; de Moraes Gonçalves, J.L.; Sparovek, G. Köppen's climate classification map for Brazil. *Meteorol. Z.* **2013**, *22*, 711–728. [CrossRef] [PubMed]
32. Nóbrega, R.S.; Vital, L.A.d.B. *Paisagens Entre Espaços Urbanos Distintos Como Elemento Para Análise Microclimática; Estudos multidisciplinares: Abordagem holística*; Ribeiro, E.P., Lima, M.S.d.R., Eds.; UFPE: Recife, Brazil, 2014.
33. APAC—Agência Pernambucana de Águas e Clima. 1º de Abril—Início da Quadra Chuvosa Do Leste de Pernambuco. 2025. Available online: <https://www.apac.pe.gov.br/noticias/23-noticias/meteorologia/632-1-de-abril-inicio-da-quadra-chuvosa-do-leste-de-pernambuco> (accessed on 23 May 2025).
34. Google, Map Data. 2015. Available online: <https://www.google.at/permissions/geoguidelines/attr-guide.html> (accessed on 7 July 2025).
35. IBGE—Instituto Brasileiro de Geografia e Estatística—Municipal Agricultural Production (PAM) Between 1990 and 2019. 2020. Available online: <https://sidra.ibge.gov.br/Tabela/1612> (accessed on 26 May 2025).
36. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L.; et al. The shuttle radar topography mission. *Rev. Geophys.* **2007**, *45*, RG2004. [CrossRef]
37. CONAB—Companhia Nacional de Abastecimento. Acompanhamento Da Safra Brasileira De Cana-De-Açúcar. CONAB: Brasília, Brazil, 2025; Volume 9. Available online: <https://www.conab.gov.br/info-agro/safras/cana/> (accessed on 25 June 2025).
38. MaisAgro. Broca-Da-Cana: 3 Descobertas Que Estão Transformando O Manejo. 2024. Available online: <https://maisagro.syngenta.com.br/inovacoes-e-tendencias/broca-da-cana-3-descobertas-que-estao-transformando-o-manejo/> (accessed on 5 August 2025).
39. Melo, L.D.; Oliveira, F.D.; Bastos, G.Q.; Anunciação Filho, C.D.; Reis, O.D. Agroindustrial performance of sugar cane cultivars in the coastal-forest zone of Pernambuco. *Ciência E Agrotecnologia* **2009**, *33*, 684–691. [CrossRef]
40. Albuquerque, A.P.d.C.; Bastos, G.Q.; Dutra Filho, J.d.A.; Souto, L.S.; Maracajá, P.B.; Simões Neto, D.E. Selection of sugarcane cultivars to the North Forest Zone of Pernambuco III: Final crop. *Rev. Verde De Agroecol. E Desenvol. Sustentável* **2016**, *11*, 39–43.
41. Simões Neto, D.E. Variedades de cana-de-açúcar no estado de Pernambuco contribuição do melhoramento clássico da RIDESA-UFRPE. *Anais da Academia Pernambucana de Ciência Agrônômica. Recife* **2008**, *5*, 125–146.
42. Rosseto, R.; Santiago, A.D. Cana. Portal Embrapa. 2022. Available online: <https://www.embrapa.br/agencia-de-informacao-tecnologica/cultivos/cana/producao/manejo/plantio> (accessed on 8 August 2025).

43. IPA—Instituto Agronomiaco de Pernamco. Cana-de-Açúcar. Folhetos Explicativos. 2025. Available online: <http://www.ipa.br/resp20.php> (accessed on 8 August 2025).
44. Moura, A.D.; Shukla, J. On the dynamics of droughts in northeast Brazil: Observations. theory and numerical experiments with a general circulation model. *J. Atm. Scie.* **1981**, *38*, 2653–2675. [CrossRef]
45. Servain, J. Simple Climatic Indices for the Tropical Atlantic Ocean and some applications. *J. Geophys. Res.* **1991**, *96*, 137–146. [CrossRef]
46. Trenberth, K.E. The Definition of El Niño. *Bull. Amer. Meteor. Soc.* **1997**, *78*, 2771–2777. [CrossRef]
47. Trenberth, K.E.; Stepaniak, D.P. Indices of El Niño evolution. *J. Climate.* **2001**, *14*, 1697–1701. [CrossRef]
48. NOAA—National Oceanic and Atmospheric Administration. 2025. Available online: <https://www.cpc.ncep.noaa.gov/> (accessed on 23 May 2025).
49. Enfield, D.B.; Mestas, A.M.; Mayer, D.A.; Cid-Serrano, L. How ubiquitous is the dipole relationship in tropical Atlantic sea surface temperatures. *J. Geogr. Res.-Oceans* **1999**, *104*, 7841–7848. [CrossRef]
50. Abril-Salcedo, D.S.; Melo-Velandia, L.F.; Parra-Amado, D. Impactos de los fenómenos climáticos sobre el precio de los alimentos en Colombia Impacts of weather anomalies on food prices in Colombia. *Ens. Sobre Política Econ.* **2016**, *34*, 146–158.
51. Schlanger, V.; Heres, J. Condições Meteorológicas: A Oscilação Do Sul E O El Nino. Environmental Science Published for Every Round the Earth. Hngarian Meteorological Service. 2006. Available online: https://klimat.czn.uj.edu.pl/enid/2_Principais_sistemas_de_vento/-_El_Ni_o_SOI_4z1.html (accessed on 7 July 2025).
52. Ultsch, A.; Lötsch, J. Euclidean distance-optimized data transformation for cluster analysis in biomedical data (EDOtrans). *BMC Bioinform.* **2022**, *23*, 233. [CrossRef] [PubMed]
53. Ward, J.R.; Joe, H. Hierarchical grouping to optimize an objective function. *J. Amer. Stat. Assoc.* **1963**, *58*, 236–244. [CrossRef]
54. Hotelling, H. The most predictable criterion. *J. Educ. Psychol.* **1935**, *26*, 139–142. [CrossRef]
55. Anderson, T.W. *An Introduction to Multivariate Statistical Analysis*, 2nd ed.; John Wiley: Hoboken, NJ, USA, 1984; pp. 439–449.
56. Anderson, T.W. *An Introduction to Multivariate Statistical Analysis*, 3rd ed.; John Wiley: Hoboken, NJ, USA, 2003; pp. 459–469.
57. STATSOFT. Statistica (Data Analysis Software System). Version 12. 2023. Available online: <https://statistica.software.informer.com/12.0/> (accessed on 7 May 2025).
58. Pearson, E.S.; Hartley, H.O. Biometrika Table for Statisticians. Table 8. Cambridge University Press: Cambridge, UK, 1966; Volume 1, pp. 137–138. Available online: <https://www.datageeks.com.br/wp-content/uploads/2024/05/tabela-qui-quadrado.jpg> (accessed on 4 July 2025).
59. de Melo, C.G.B.; Rolim, M.M.; Cavalcanti, R.Q.; da Silva, M.V.; Candeias, A.L.B.; Lopes, P.M.O.; Ortiz, P.F.S.; de Lima, R.P. RGB Composition Obtained by a UAV in the Monitoring of Sugarcane Row Gaps Using the Biophysical Index. *Agriengineering* **2025**, *7*, 17. [CrossRef]
60. Jardim, A.M.d.R.F.; Silva, M.V.d.; Silva, A.R.; Santos, A.d.; Pandorfi, H.; Oliveira-Júnior, J.F.d.; De Lima, J.L.M.P.; Souza, L.S.B.d.; Araújo Júnior, G.d.N.; Lopes, P.M.O.; et al. Spatiotemporal climatic analysis in Pernambuco State, Northeast Brazil. *J. Atmos. Sol.-Terr. Phys.* **2021**, *223*, 105733. [CrossRef]
61. Bezerra, A.C.; Silva, J.L.B.d.; Moura, G.B.d.A.; Lopes, P.M.O.; Nascimento, C.R.; Ribeiro, E.P.; Galvêncio, J.D.; Silva, M.V.d. Dynamics of land cover and land use in Pernambuco (Brazil): Spatio-temporal variability and temporal trends of biophysical parameters. *Remote Sens. Appl.-Soc. Environ.* **2022**, *25*, 100677. [CrossRef]
62. Bordonal, R.D.O.; Carvalho, J.L.N.; Lal, R.; de Figueiredo, E.B.; de Oliveira, B.G.; La Scala, N. Sustainability of sugarcane production in Brazil. *A review. Agron. Sustain. Dev.* **2018**, *38*, 13. [CrossRef]
63. Du, C.; Kulay, L.; Cavalett, O.; Dias, L.; Freire, F. Life cycle assessment addressing health effects of particulate matter of mechanical versus manual sugarcane harvesting in Brazil. *Int. J. Life Cycle Assess.* **2018**, *23*, 787–799. [CrossRef]
64. Silva, M.J.; Neves, L.O.; Correa, M.H.F.; de Souza, C.H.W. Quality indexes and performance in mechanized harvesting of sugarcane at a burnt cane and green cane. *Sugar Tech.* **2021**, *23*, 499–507. [CrossRef]
65. Afifi, A.; Clark, V.A. *Computer-Aided Multivariate Analysis*, 3rd ed.; Chapman and Hall: London, UK, 1996; p. 505.
66. Mazouz, R.; Assani, A.A.; Rodríguez, M.A. Application of redundancy analysis to hydroclimatology: A case study of spring heavy floods in southern Québec (Canada). *J. Hydrol.* **2013**, *496*, 187–194. [CrossRef]
67. Protásio, T.d.P.; Tonolili, G.H.D.; JúniorIII, M.G.; Bufalino, L.; Couto, A.M.; Trugilho, P.F. Correlações canônicas entre as características químicas e energéticas de resíduos lignocelulósicos. *Cerne* **2012**, *18*, 433–439. [CrossRef]
68. Hair Junior, J.F.; Anderson, R.E.; Tatham, R.L.; Black, W.C. *Análise Multivariada de Dados*; Bookman: Porto Alegre, Brazil, 2005; 593p.
69. Barnett, T.P.; Preisendorfer, R. Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. *Mon. Weather. Rev.* **1987**, *115*, 1825–1850. [CrossRef]
70. Zhao, Y.; Yu, L.-X.; Ai, J.; Zhang, Z.-F.; Deng, J.; Zhang, Y.-B. Climate Variations in the Low-Latitude Plateau Contribute to Different Sugarcane (*Saccharum* spp.) Yields and Sugar Contents in China. *Plants* **2023**, *12*, 2712. [CrossRef]

71. lack-Prain, S.; Shi, L.; Zhu, P.; da Rocha, H.R.; Cabral, O.; Hu, S.; Williams, M. The impact of climate change and climate extremes on sugarcane production. *GCB Bioenergy* **2021**, *13*, 408–424. [[CrossRef](#)]
72. Marin, F.R.; Lopes-Assad, M.L.; Assad, E.D.; Vian, C.E.; Santos, M.C. Sugarcane crop efficiency in two growing seasons in São Paulo State, Brazil. *Pesqui. Agropecu. Bras.* **2008**, *43*, 1449–1455. [[CrossRef](#)]
73. NOVACANA. Climas Diferentes Na Zona Canavieira de Pernambuco Afetam Produção. 2021. Available online: <https://www.novacana.com/noticias/climas-diferentes-zona-canavieira-pernambuco-afetam-producao-150321> (accessed on 1 August 2025).
74. Wildt, A.R.; Lambert, Z.V.; Durand, R.M. Applying the Jackknife Statistic in Testing and Interpreting Canonical Weights, Loadings, and Cross-Loadings. *J. Mark. Res.* **1982**, *19*, 99–107. [[CrossRef](#)]
75. Ribeiro, M.D.d.S.B. Basic Wood Density of *Eucalyptus* spp. Forest Plantations: Associations with Site AND Plantation Variables and Estimates with Artificial Neural Networks. Ph.D. Thesis, Unesp, São Paulo, Brazil, 2018. Available online: <https://repositorio.unesp.br/server/api/core/bitstreams/db46edc3-122e-47af-a20b-812f69d8986c/content> (accessed on 26 June 2025).
76. Fischer, G.R.; Diniz, G.B.; Marques, J.R.Q. Previsão do número de dias de chuva para a metade sul do Rio Grande do Sul utilizando a temperatura da superfície do mar (TSM). *Ver. Bras. Met.* **2008**, *23*, 146–151. [[CrossRef](#)]
77. Stockdale, T.N.; Alves, O.; Boer, G.; Deque, M.; Ding, Y.; Kumar, A.; Kumar, K.; Landman, W.; Mason, S.; Nobre, P.; et al. Understanding and predicting seasonal-to-interannual climate variability-the producer perspective. *Procedia Environ. Sci.* **2010**, *1*, 55–80. [[CrossRef](#)]
78. Oliveira, N.L.; Marcuzzo, F.F.N. Correlação da temperatura na parte sul do oceano Atlântico com o NDC em Mato Grosso. *GEOgrap* **2013**, *18*, 183–206. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.