

Article

# Rainfall Prediction in the State of Paraíba, Northeastern Brazil Using Generalized Additive Models

Leydson G. Dantas <sup>1</sup>, Carlos A. C. dos Santos <sup>2,\*</sup> , Ricardo A. de Olinda <sup>3</sup>, José I. B. de Brito <sup>2</sup> , Celso A. G. Santos <sup>4</sup> , Eduardo S. P. R. Martins <sup>5</sup>, Gabriel de Oliveira <sup>6</sup>  and Nathaniel A. Brunzell <sup>7</sup> 

<sup>1</sup> Center for Weather Forecasting and Climate Studies, National Institute for Space Research, Cachoeira Paulista, São Paulo 12630-000, Brazil; leydson.dantas@inpe.br

<sup>2</sup> Academic Unit of Atmospheric Sciences, Center for Technology and Natural Resources, Federal University of Campina Grande, Campina Grande, Paraíba 58109-970, Brazil; jose.ivaldo@ufcg.edu.br

<sup>3</sup> Department of Statistics, Paraíba State University, Campina Grande, Paraíba 58429-500, Brazil; prof\_ricardo@servidor.uepb.edu.br

<sup>4</sup> Department of Civil and Environmental Engineering, Federal University of Paraíba, João Pessoa, Paraíba 58051-900, Brazil; celso@ct.ufpb.br

<sup>5</sup> Research Institute for Meteorology and Water Resources, Fortaleza, Ceará 60115-221, Brazil; presidencia@funceme.br

<sup>6</sup> Department of Geography and Planning, University of Toronto, Toronto, ON M5S 3G3, Canada; gabriel.deoliveira@utoronto.ca

<sup>7</sup> Department of Geography and Atmospheric Science, University of Kansas, Lawrence, KS 66045, USA; brunzell@ku.edu

\* Correspondence: carlos.santos@ufcg.edu.br

Received: 29 June 2020; Accepted: 1 September 2020; Published: 4 September 2020



**Abstract:** The state of Paraíba is part of the semi-arid region of Brazil, where severe droughts have occurred in recent years, resulting in significant socio-economic losses associated with climate variability. Thus, understanding to what extent precipitation can be influenced by sea surface temperature (SST) patterns in the tropical region can help, along with a monitoring system, to set up an early warning system, the first pillar in drought management. In this study, Generalized Additive Models for Location, Scale and Shape (GAMLSS) were used to filter climatic indices with higher predictive efficiency and, as a result, to perform rainfall predictions. The results show the persistent influence of tropical SST patterns in Paraíba rainfall, the tropical Atlantic Ocean impacting the rainfall distribution more effectively than the tropical Pacific Ocean. The GAMLSS model showed predictive capability during summer and southern autumn in Paraíba, highlighting the JFM (January, February and March), FMA (February, March and April), MAM (March, April and May), and AMJ (April, May and June) trimesters as those with the highest predictive potential. The methodology demonstrates the ability to be integrated with regional forecasting models (ensemble). Such information has the potential to inform decisions in multiple sectors, such as agriculture and water resources, aiming at the sustainable management of water resources and resilience to climate risk.

**Keywords:** non-stationary; water resources; SST indices; Northeast of Brazil; zero adjusted Gamma distribution (ZAGA)

## 1. Introduction

The high spatial-temporal variability of rainfall in the Northeast of Brazil (NEB) is highly influenced by global teleconnection patterns associated with the El Niño-Southern Oscillation (ENSO), as well as

the sea surface temperature (SST) in the Atlantic Ocean region [1–4]. It is necessary then to investigate the influence of those phenomena on the climate of different areas of the NEB. The variability of seasonal and interannual precipitation, associated with tropical SST anomalies and their interaction with the atmosphere, is associated with how extreme and prolonged droughts will be [5–11].

The tropical Atlantic Ocean's influence on the NEB rainfall regime is very intense, mainly due to its proximity. Nevertheless, ENSO events often influence the rainfall regime by either increasing (La Niña) or decreasing (El Niño) its total amount in this region [3]. Several studies concerning the climatic variability in the NEB have been developed, such as [12–15]. These studies focus on the NEB rainfall dependence on the El Niño and La Niña events.

In El Niño years, there is a decrease in the frequency of convection in the north part of the NEB due to the subsidence of the Walker cell circulation, while in La Niña years, there is an increase in convection [16]. The presence of El Niño, coupled with negative SST anomalies over the South Atlantic Ocean, leads to more severe droughts. A decrease in daily rainfall is observed during El Niño periods, especially during the rainiest period from March to April, increasing the number of consecutive dry days during this period. In La Niña years, there is an increase in precipitation, especially during the rainy season, which can cause flooding in major urban centers [13,16,17].

Many meteorological systems act in the region, influencing the spatio-temporal rainfall regime in the state of Paraíba. The main systems are the Intertropical Convergence Zone (ITCZ) [18], Upper-Tropospheric Cyclonic Vortices (UTCV) [19], Instability Lines (IL) [20], Mesoscale Convective Complexes (MCC) [21], Easterly Wave Disturbances (EWD) [22], Frontal Systems, and Land–Sea breeze [23]. The ITCZ is the most important of these systems to be studied and monitored by the weather forecast and climate centers. It is easily identified as a band of convective clouds along the equatorial strip [20] and formed by a confluence of trade winds between the northeast and southeast [22].

The ITCZ migrates seasonally from its northernmost position (around 14° N), during August–September, to its southernmost location (about 2° S), during March–April [20]. In rainy years, when the SST anomalies in the South Atlantic Ocean region are positive, the subtropical high of the South Atlantic is weakened, causing the ITCZ to move further south from its climatic position. As a consequence, there is an increase in rainfall associated with it during the northern rainy season of the NEB [8,17,24]. This is the primary atmospheric system that determines the concentration of the short rainy season from March to April in the northern and eastern parts of the NEB [9,25].

Drought-related impacts in the NEB have been studied for decades [24,26,27], and these studies have demonstrated the vulnerability of the region due to lack of significant rainfall [16], shallow soils, crystalline basement, high evaporations, and also: lack of financial resources, a structured drought policy, organizing the monitoring/forecasting pillars, vulnerability/resilience assessment and response. The country's losses associated with natural disasters exceeded BRL 31.62 billion (BRL = Real, Brazilian currency) between 2012 and 2015 [28]. Most of these are related to drought in the NEB, representing about BRL 19.24 billion, equivalent to 60.86% of the total losses for the entire country [28]. Based on annual decrees of the Government of Paraíba, it is the area most impacted by natural disasters, most of which are related to water shortages [28].

As part of mitigating the effects of a drought, we must have a forecasting system that allows us to anticipate actions to mitigate the effects of the temporal variability. In this context, the literature has shown that it is possible to observe, through dynamic and statistical models, the relationship between ocean-atmosphere and tropical Pacific and Atlantic Oceans conditions with precipitation in NEB [17]. Time series models may be used to predict climate and hydrological data [29–32]. The Generalized Additive Model for Location, Scale and Shape (GAMLSS) is capable of observing the sensitivity of the behavior of the variable to the influence of external parameters over time [33]. Studies conducted in different regions highlight the efficiency of introducing different climate variables into the model structure as the most significant advantage of using GAMLSS [34–38].

The authors of [39] emphasize that it is possible to choose different probability distributions using this model, thus providing significant flexibility in the adjustment to different data series with GAMLSS.

The flexibility of probabilistic modeling can be observed even under non-stationary conditions, making it easier to understand how distribution parameters depend on external covariates [40–45]. Consequently, due to these characteristics, GAMLSS can be used in the analysis of the climate risk associated with hydroclimatic variables [46].

This study aims to identify the influence of SST anomalies on rainfall in the state of Paraíba using GAMLSS, to quantify the contribution of climatic indices to the distribution of rainfall and, consequently, to obtain seasonal rainfall forecasts for the state. In this way, the results may serve as support for decision-makers, since these models with local characteristics are not systematically used by the agencies responsible for the implementation of Brazil’s climate forecasts, particularly in the state of Paraíba.

## 2. Data and Methods

### 2.1. Study Area

Figure 1 shows the spatial distribution of the rain gauges in the state of Paraíba used in this study. Paraíba is located in the eastern portion of NEB, between the parallels of 06° S–09° S and the meridians of 34° W–39° W, bounded to the north by the state of Rio Grande do Norte, to the south by the state of Pernambuco, to the west by the states of Ceará and the east by the Atlantic Ocean. The population is estimated to be 4,025,558 inhabitants in 223 municipalities, with a demographic density of 6670 inhabitants/km<sup>2</sup> [47].



Figure 1. Location of the used rain gauges in the state of Paraíba.

## 2.2. Data

Rainfall data from 45 locations were used (Figure 1). Monthly time series information was provided by Development Superintendence of Northeast (SUDENE) [48] from 1948 to 1990, and State Water Management Executive Agency [49] from 1994 to 2017. The filling of missing data was performed using the Expectation-Maximization with Bootstrapping (EMB) algorithm. Following the completion of the missing data, the average seasonal corresponding to the period 1986–2015 was used for this study.

The analysis of SST behavior was performed from anomaly indices in the regions of Niño 1 + 2, Niño 3, Niño 3.4, Niño 4, and SOI (Southern Oscillation Index) in the Pacific Ocean obtained from the Climate Prediction Center/National Centers for Environmental Prediction/National Oceanic and Atmospheric Administration [50]. Anomaly indices in the TNA (Tropical North Atlantic) and TSA (Tropical South Atlantic) regions, and the decadal scales AMO (Atlantic Multidecadal Oscillation) and PDO (Pacific Decadal Oscillation), were provided by the Earth System Research Laboratory/National Oceanic & Atmospheric Administration [51].

## 2.3. Methods

### 2.3.1. Descriptive Analysis

Time series analysis describes the data structure, acquiring its characteristics, and enabling predictions of its future behavior. In this study, all calculations were performed using the R programming language [52], including its library of described functions.

Missing observations are often found in time-series studies [53] as the methodologies associated with this type of analysis often must be applied in the presence of a complete historical series. Therefore, the resampling of these data using the Expectation-Maximization with Bootstrapping (EMB) algorithm, implemented in the *Amelia II* R package [54], was carried out to fill missing data found in the time series.

This method seeks to simulate up to four possible precipitation scenarios to replace the missing or inconsistent information found in the historical series [55]. These four imputations are the combinations generated by the algorithm, in which the user can subjectively select the value to be used in the imputation of the missing data. *Amelia II* is a complete R package for multiple imputations of missing data, which works faster and is easier to use than various Monte Carlo approaches via the Markov Chain, and provides essentially the same answer [54].

It is desirable to use the series of rain gauges closest to the location requiring data resampling so that the method identifies the characteristics and patterns of variability between the series and improves the method's ability to simulate missing data. It is also necessary to verify the reliability of the resampling of these data by simulating the observed data. In order to subjectively select the most suitable data for resampling, it is imperative to understand the influence of the meteorological systems regulating the region's rainfall regime and their dependence on the SST regions. Previous studies have emphasized the evolution and effectiveness of this method in information reconstruction [56–58].

The use of hierarchical cluster analysis [59–61], using the average monthly total for each location, is proposed to improve interpretation when interpreting results. This technique quantifies the similarity between sites by recognizing patterns that divide the area into groups of homogeneous precipitation regions [61,62]. The hierarchical cluster analysis is a multivariate technique that aims to group data according to the degree of similarity between them. In this way, its application is useful to generate precipitation groups with homogeneous regions within the cluster and heterogeneity between them. It is crucial to instruct that the robustness of this technique will depend on the amount of data available in the time series, where the sample must be quite representative. As in the context of this work, 30 years of data were applied, it is admissible that the selection of groups of homogeneous regions (mesoregions) was used safely.

The *factoextra* package [63] was used to facilitate the measurement of similarity as it provides accessible functions for extracting and visualizing data output and simplifies hierarchical cluster analysis using the *Ward.D2* method [64]. The *Ward.D2* algorithm uses square Euclidean distances to

minimize errors while selecting dendrogram nodes, showing equivalence between locations. Each node indicates that there is a set (mesoregions) of cities, in which the rainfall regime has similar characteristics.

One of the difficulties in analyzing the association between the SST anomaly in the Atlantic and Pacific Oceans with precipitation of a given region is to have prior knowledge of the impact of temporal lags on this relationship. In order to obtain this information, it is possible to carry out climatological analyzes applying statistical predictions in precipitation. For this, it is necessary to use some method that can observe the sensitivity between the changes in the SST and precipitation in Paraíba and the lag associated with this relationship, before applying the main methodology.

Lag correlations between time series ( $Y_t$ ) and external variables ( $X_t$ ) are often verified using the Cross-Correlation Function (CCF). However, this technique is limited to the characteristics of a stationary series, since the autocorrelation of  $Y_t$  and  $X_t$  influences the CCF. To account for this, we use the prewhitening method [65] to identify the monthly lag between the SST climatic indices and the rainfall time series.

Prewhitening generates a visual identification of the signal in which there is a higher probability of having a temporal dependence, thus highlighting the possible value of the lag between the series. For this purpose, it is desired the vertical line of the graph (not shown), which varies between the zero of the CCF, exceeds the horizontal line that indicates the confidence level at  $p < 0.05$ . This procedure has been used and adapted by several authors to remove the influence of serial correlation, as well as to eliminate the trend structure of the series [66–69]. In this study, we used the prewhitening function found in the TSA package [70].

### 2.3.2. GAMLSS Model

A fundamental criterion for studying time series is to analyze the presence of stationarity. When applying the stationarity criteria, it is possible to choose the best adjustment of the distribution of potential models to the variables to be explored. However, when working with environmental data, such as meteorological data that follow non-stationary characteristics, mathematical transformations are needed to make the series stationary. GAMLSS [33] can adjust to a non-stationary time series, highlighting its practicality.

The GAMLSS model class was used because the data are complex, given the existence of considerable asymmetry in the data and many zeros. With this class of models, it was possible to filter the climatic indexes with greater predictive efficiency and, consequently, make a climatic forecast of precipitation. The specific method was selected due to the small amount of research using statistical modeling in the regional context, which associates the rainy season in Paraíba with the SST. There is a recently published study using GAMLSS modeling in Paraíba, but using latitude, longitude, and topography as predictor variables [35]. The innovation of this study is due to the use of oceanic indices that influence and regulate the region's rainfall regime, mainly between summer and southern autumn. Therefore, it is expected that the result found will be better because it contains the main variables that predict precipitation in the localities.

As already described, in this work, the average between three consecutive months was used. The justification for this analysis is given by the historical context when working with climatic variables. For example, the Climate Prediction Center (CPC) classifies events from El Niño or La Niña from reading the averages of the anomalies of three consecutive months in the Niño 3.4 region. Another advantage of this approach is the smoothing of the high monthly climatic variability that exists in the region. When applying this form, it facilitates to measure by simulation the volume of water to be distributed in the regions of Paraíba. This performance is verified during the application of climate forecasts in the years 2016 and 2017.

In GAMLSS models, the distribution of the response variable  $Y$  does not need to be part of the exponential distribution family. In this sense, for each distribution parameter, different additive terms can be included in the predictor. This is due to the ability to model not only the mean but also the higher-order moments, including the variance, skewness, and kurtosis parameters of the

predictor variable distribution, which is ideal for the analysis of highly variable time series. Therefore, there is a possibility that the model has the potential to capture the influence of climatic indices (Niño 1 + 2, Niño 3, Niño 3.4, Niño 4, SOI, AMO, PDO, TNA, and TSA) on rainfall behavior in the state. The model adjustment can be applied in two general ways, the first is by selecting the predictor variables with the greatest influence (statistical significance) on rainfall in Paraíba state, and the second is by differentiating (smoothing) the predictor variables with low statistical significance in the attempt to improve them. In the development of this work, only the predictor variables with the best statistical significance in the modeling were selected.

This model was selected for the development of the seasonal forecast with a horizon of three months. The use of this, in the first half of the year, was chosen to be the period with the highest water recharge in cities located in the interior of the state of Paraíba. As these localities live with frequent droughts and a few months of rain, it is essential to use this periodicity to analyze the behavior of the model's predictions with the observed values of precipitation to provide a tool that can improve the management of water resources by decision-makers.

A better understanding of how GAMLSS works can be found in [33,40,41]. In the GAMLSS model framework, it is assumed that the  $y_i$  data series are independent, for  $i = 1, \dots, n$ , with  $F_Y(y_i|\theta^i)$  being the cumulative probability density function (p.d.f.) where  $\theta^{iT} = (\theta_{i1}, \theta_{i2}, \dots, \theta_{ip})$  is a vector of  $p$  parameters related to explanatory variables and random effects. This parameter vector includes the location, scale, and shape parameters of the p.d.f. Usually, in the context of hydro-meteorological data,  $p \leq 4$ , since one to four parameter families have sufficient flexibility for characterizing continuous and discrete distributions [39].

In a GAMLSS model [33], the additive terms represent the smoothing terms that allow the flexibility of the parameter dependence of p.d.f. on climate indices. For a vector of observations of the response variable given by  $y^T = (y_1, y_2, \dots, y_n)$ , considering  $k = 1, 2, \dots, p$ , a monotonous function of  $g_k(\cdot)$  relating the  $k$ th parameter  $\theta_k$  explanatory variables random effects through an additive model written form

$$g_k(\theta_k) = \eta_k = X_k\beta_k + \sum_{j=1}^{J_k} Z_{jk}\gamma_{jk}, \tag{1}$$

where  $\theta_k$  and  $\eta_k$  are vectors of size  $n$ , for example,  $\theta_k^T = (\theta_{1k}, \theta_{2k}, \dots, \theta_{nk})$ ,  $\beta_k^T = (\beta_{1k}, \beta_{2k}, \dots, \beta_{J_k k})$  is a parameter vector with size  $J_k$ ,  $X_k$  is a known structure matrix of the order  $n \times J_k$ ,  $Z_{jk}$  is a fixed matrix of known structure  $n \times q_{jk}$ , and  $\gamma_{jk}$  is a  $q_{jk}$  dimensional random variable. When choosing  $Z_{jk} = I_n$  where  $I_n$  is an identity matrix  $n \times n$ ,  $\gamma_{jk} = h_{jk} = h_{jk}(x_{jk})$  for all combinations of  $j$  and  $k$ , thus obtaining the semi-parametric additive equation of the GAMLSS,

$$g_k(\theta_k) = \eta_k = X_k\beta_k + \sum_{j=1}^{J_k} h_{jk}(x_{jk}), \tag{2}$$

where the  $h_{jk}$  is an unknown function of the explanatory variable  $x_{jk}$  and  $h_{jk} = h_{jk}(x_{jk})$  is a vector representing the  $h_{jk}$  function of  $x_{jk}$ .

The attributes of the four parameters ( $p = 4$ ) are usually characterized by location ( $\mu$ ), scale ( $\sigma$ ), skewness ( $\nu$ ), and kurtosis ( $\tau$ ). The first two parameters  $\theta_1$  and  $\theta_2$  (Equation (1)), denoted by  $\mu$  and  $\sigma$  are defined in the literature as the location (or mean) and scale (or variance) parameters, respectively; the last two  $\nu = \theta_3$  and  $\tau = \theta_4$  are called shape parameters.

We used seven probability distributions (Table 1) that have been widely used in the climate analysis of precipitation time series, such as [71–76], with the support of the *gamlss* R package [33,77].

**Table 1.** Summary of the seven distributions considered in the study.

Distribution Function	Abbreviation	Probability Density Function
Gamma	GA	$f_Y(y \mu, \sigma) = \frac{1}{(\sigma^2\mu)^{\frac{1}{\sigma^2}} \Gamma(\frac{1}{\sigma^2})} y^{\frac{1}{\sigma^2}-1} e^{-\frac{y}{\sigma^2\mu}}$ to $y > 0, \mu > 0$ and $\sigma > 0$
Generalized Gamma	GG	$f_Y(y \mu, \sigma, \nu) = \frac{ \nu \theta^\nu z^\theta \exp[-\theta z]}{\Gamma(\theta)y}$ to $y > 0, \mu > 0, \sigma > 0$ and $-\infty < \nu < \infty$
Zero Adjusted Gamma	ZAGA	$f_Y(y \mu, \sigma, \nu) = \begin{cases} \nu \\ (1-\nu) \left[ \frac{1}{(\sigma^2\mu)^{\frac{1}{\sigma^2}} \Gamma(\frac{1}{\sigma^2})} y^{\frac{1}{\sigma^2}-1} e^{-\frac{y}{\sigma^2\mu}} \right] \end{cases}$ to $0 \leq y < \infty, \mu > 0, \sigma > 0$ and $0 < \nu < 1$
Gumbel	GU	$f_Y(y \mu, \sigma) = \frac{1}{\sigma} \exp\left[-\left(\frac{y-\mu}{\sigma}\right)\right] - \exp\left(\frac{y-\mu}{\sigma}\right)$ to $-\infty < y < \infty, -\infty < \mu < \infty$ and $\sigma > 0$
Logistic	LO	$f_Y(y \mu, \sigma) = \frac{1}{\sigma} \left\{ \exp\left[-\left(\frac{y-\mu}{\sigma}\right)\right] \right\} \left\{ 1 + \exp\left[-\left(\frac{y-\mu}{\sigma}\right)\right] \right\}^{-2}$ to $-\infty < y < \infty, -\infty < \mu < \infty$ and $\sigma > 0$
Log-Normal	LOGNO	$f_Y(y \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{y} \exp\left\{-\frac{[\log(y)-\mu]^2}{2\sigma^2}\right\}$ to $y > 0, \mu > 0$ and $\sigma > 0$
Weibull	WEI	$f_Y(y \mu, \sigma) = \frac{\sigma y^{\sigma-1}}{\mu^\sigma} \exp\left[-\left(\frac{y}{\mu}\right)^\sigma\right]$ to $y > 0, \mu > 0$ and $\sigma > 0$

In the absence of smoothing functions, the model is reduced to the parametric GAMLSS model,  $g_k(\theta_k) = X_k\beta_k, k = 1, 2, 3, 4$  estimated by the maximum likelihood (ML), with the maximum likelihood function given by [77],

$$\ell = \sum_{i=1}^n \log f(y_i|\mu_i, \sigma_i, \nu_i, \tau_i), \tag{3}$$

where  $f$  represents the p.d.f. of the response variable. However, for non-parametric models, it is necessary to resort to the penalized ML method, through the logarithm of the likelihood function given by [77],

$$\ell_p = \ell - \frac{1}{2} \sum_{k=1}^4 \sum_{j=1}^{J_k} \gamma_{kj}^T G_{kj}(\lambda_{kj}) \gamma_{kj}. \tag{4}$$

### 2.3.3. Model Evaluation Criteria

In order to compare and select the appropriate probability distribution for the rainfall time series in the different GAMLSS models, we chose the model with the lowest Akaike Information Criterion (AIC) [78] and the Bayesian Information Criterion (BIC) [79]. The independence and normality of the residues are verified by the worm plot graph, which is a qq-plot of the residues. These graphs visually help to decide if the selected model fits the behavior of the studied series when verifying whether the residuals of the simulated versus the observed value are similar [80,81]. For optimal adjustment, 95% of the points are expected to be between the two elliptical curves (unit normal quantile). Otherwise, when finding points outside the range of high-frequency elliptical curves means that the model distribution adjustments (selection model) are insufficient to explain precipitation behavior (response variable), the outliers can also be found distant from the two elliptical curves on the graph [77,82] reiterate that this step is essential to enable model optimization and computational processing and to finally identify the most significant predictors to be used in the determination of predictions, as stated by [83].

### 2.3.4. Efficiency for Seasonal Prediction

Initially, the quantile technique was discussed by [84] and addressed by other authors [85–87], describing that this method can be applied to precipitation series, where the precipitation  $Y$  of a locality over time is considered a continuous variable capable of characterizing the dry, normal and rainy periods, based on the quantile categories. A quantile of order  $p$  is a numerical value that classifies a data set between intervals, the choice of which may vary according to the desired percentage of the research.

We decided to use the quantile technique to characterize the high annual variability of rainfall found in the region due to the dependence on SST anomalies. For each month, we classified the forecasted precipitation by applying the “Type = 6” algorithm because it is used in continuous data, where observations can assume any value in an interval. The algorithm is present in the quantile function of the *stats* package, the associated equation in the interactions of this algorithm is  $m = p \cdot p[k] = k / (n + 1)$ . Thus  $p[k] = E[F(x[k])]$  [88].

Thus, it is possible not only to examine whether the response of the model corresponds to the classification of the observed value but also to determine whether the model at least captures the general precipitation behavior towards normality, in the sense of achieving the proper distribution for that period. In this case, the technique was adapted for the locations studied in Paraíba to classify the precipitation predictions as extremely dry (ED), very dry (VD), dry (D), normal (N), rainy (R), very rainy (VR) and extremely rainy (ER), as shown in Table 2.

**Table 2.** Corresponding range to climate classifications represented by their quantiles.

Quantiles	Rating
$p < Q(0.05)$	Extremely dry
$Q(0.05) < p < Q(0.15)$	Very dry
$Q(0.15) < p < Q(0.35)$	Dry
$Q(0.35) < p < Q(0.65)$	Normal
$Q(0.65) < p < Q(0.85)$	Rainy
$Q(0.85) < p < Q(0.95)$	Very rainy
$Q(0.95) < p$	Extremely rainy

Four different statistical indices were adopted to measure the model’s performance and its applicability in conducting climate studies in Paraíba, due to the need to qualify the adjustment of the simulated values to those observed. Comparisons for each homogeneous locality and regions were made by Mean Absolute Error (MAE), Percent Bias (PBIAS), Root Mean Square Error (RMSE), and Coefficient of Determination ( $R^2$ ). The MAE and the RMSE represent the average deviation between the observed and predicted data. The PBIAS assesses the trend that the mean of the predicted values has with the observed ones. Therefore, positive and negative values indicate overestimation and underestimation of precipitation by the model, respectively. Finally,  $R^2$  indicates the proportion in which the variable can be explained by the model [29].

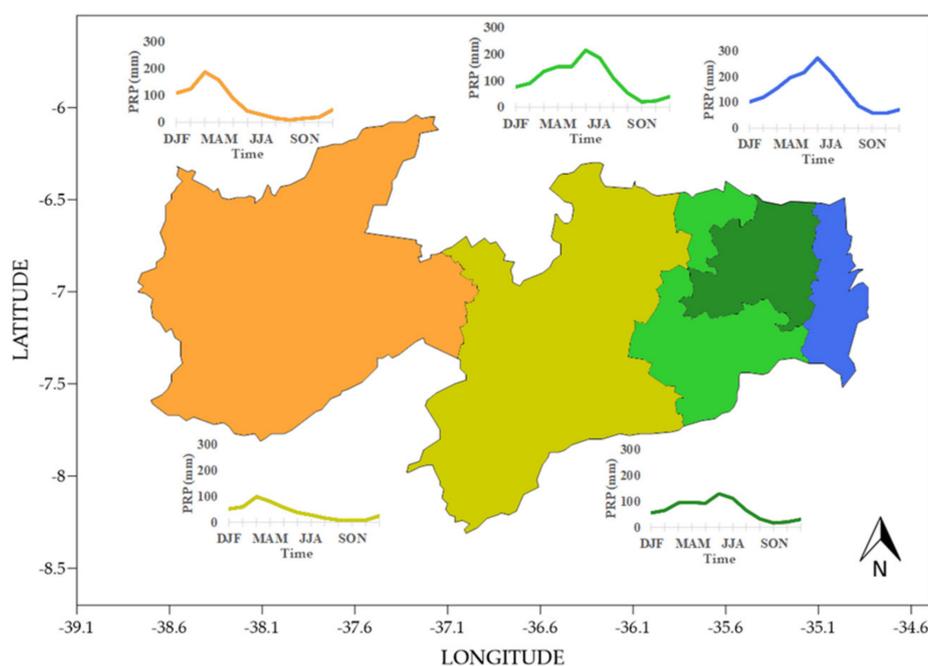
This enabled the quantification of the underestimation, overestimation, and the best adjustments of the model. Optimal fit estimation occurs when the estimation error indicated by MAE, PBIAS, and RMSE is nearby to zero, and the  $R^2$  is next to one [29]. As noted, the study was designed to follow multiple statistical criteria to achieve efficiency and enhance time series analysis and its ideal procedures, significantly improving the product generated by the GAMLSS model.

## 3. Results

### 3.1. Series Investigation

The quantile cluster analysis facilitates the identification of five homogeneous precipitation regions in the state of Paraíba (Figure 2), which were R1 (Litoral), R2 (Brejo), R3 (Agreste), R4 (Cariri and

Carimataú) and R5 (Sertão). After obtaining the locations belonging to each homogeneous region (mesoregion), three representative cities were selected from each of the mesoregions, and pre-whitening was applied to calculate the lag between the precipitation time series and the climatic indices obtained from the regions of Niño1 + 2, Niño 3, Niño 3.4, Niño 4, SOI, TNA, and TSA (Table 3), except AMO and PDO. The lag value of the abnormal levels of AMO and PDO due to the characteristics of the series being multidecadal and decadal, respectively, is not precisely indicated by the pre-whitening method. Since the calculation obtains signals with strong multicollinearity, the technique would indicate the existence of several lags (lag = 0, 1, ..., 12, ..., n months), which could lead to an error in the interpretation of the results. As a result, the highest correlation level for these two indices is lag 0 [89], so we applied a lag equal to zero.



**Figure 2.** Distribution of the five homogeneous regions by cluster analysis in the state of Paraíba. (PRP = monthly precipitation, DJF = December, January and February; MAM = March, April and May; JJA = June, July and August; SON = September, October and November).

**Table 3.** The time lag between Sea Surface Temperature (SST) anomalies and precipitation in the state of Paraíba (SOI = Southern Oscillation Index; AMO = Atlantic Multidecadal Oscillation; PDO = Pacific Decadal Oscillation; TNA = Tropical North Atlantic and TSA = Tropical South Atlantic).

Region	Niño 1 + 2	Niño 3	Niño 3.4	Niño 4	SOI	AMO	PDO	TNA	TSA
R1	0	0	0	7	3	0	0	3	1
R2	0	0	0	7	3	0	0	3	1
R3	0	0	0	7	3	0	0	3	1
R4	2	2	2	6	3	0	0	3	2
R5	2	2	2	6	3	0	0	3	2

It is necessary to justify that there are situations in which two or three “ideal” lag values are indicated by pre-whitening due to the different seasonal dependencies. Therefore, we generated Table 3 and compared this result with those already described in the existing rainy season literature in the eastern and northern parts of the NEB [8,90–95], which is the target of this research.

### 3.2. Application of the GAMLSS Model

The performance of the GAMLSS model class was verified by adjusting the non-linear characteristics (shape) of the precipitation time series (similar to [45]), and by assessing the dependence of the series at 45 locations in Paraíba based on SST anomaly indices. SST anomalies, classified as predictive variables, have been used in the distributions to obtain the best fit and combination.

The distributions that best fit the Paraíba precipitation time series can be seen in Table 4. The GG distribution resulted in the best model of rainfall in the regions R1, R2, and R3, which are localities in the state where there are no long periods with zero precipitation and are close to the east coast. The application of the GG distribution at the other locations of the study (i.e., R4 and R5) would not be adequate due to the presence of zeros in the observed data. It should be noted that the total annual rainfall of R3 is less than or equal to R5. However, R3 has more months of zero precipitation than R5. The precipitation deficit in R3 for the June to December period is often affected by negative TSA anomalies [95].

**Table 4.** Selection of Generalized Additive Models for Location, Scale and Shape (GAMLSS) by Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) criteria corresponding to each study location.

Location	Region	Distribution	AIC	BIC
Alhandra	R1	GG	4268.73	4338.68
Mataraca	R1	GG	4042.68	4081.54
Santa Rita	R1	GG	3978.83	4033.24
Areia	R1	GG	3954.90	4013.20
Bananeiras	R2	ZAGA	3946.78	3989.53
Mamanguape	R2	GG	4000.79	4055.20
Sapé	R2	GG	3838.31	3900.49
Araruna	R3	GG	3638.50	3696.79
Campina Grande	R3	GG	3618.09	3656.95
Ingá	R3	ZAGA	3571.23	3617.87
Itabaiana	R3	GG	3571.47	3629.76
Umbuzeiro	R3	ZAGA	3701.84	3752.36
Barra de Santa Rosa	R4	ZAGA	3313.38	3360.02
Caraúbas	R4	ZAGA	3183.71	3242.00
Pedra Lavrada	R4	ZAGA	3148.88	3207.17
Picuí	R4	ZAGA	3239.29	3285.92
Pocinhos	R4	ZAGA	3189.53	3240.05
Salgadinho	R4	ZAGA	3292.98	3335.73
Santa Luzia	R4	ZAGA	3344.64	3406.82
São João do Cariri	R4	ZAGA	3356.13	3410.54
São João do Tigre	R4	ZAGA	3276.69	3334.98
Soledade	R4	ZAGA	3297.82	3332.80
Sumé	R4	ZAGA	3376.54	3419.28
Taperoá	R4	ZAGA	3501.20	3551.72
Água Branca	R5	ZAGA	3675.94	3730.34
Aguiar	R5	ZAGA	3670.64	3721.16
Belém do Brejo do Cruz	R5	ZAGA	3553.75	3592.61
Bonito de Santa Fé	R5	ZAGA	3751.72	3802.24
Brejo do Cruz	R5	ZAGA	3674.20	3724.72
Cajazeiras	R5	ZAGA	3853.81	3896.55
Catingueira	R5	ZAGA	3587.68	3630.43
Catolé do Rocha	R5	ZAGA	3678.56	3740.74
Conceição	R5	ZAGA	3641.92	3711.87
Imaculada	R5	ZAGA	3554.64	3628.47
Mãe D'Água	R5	ZAGA	3463.90	3518.31

Table 4. Cont.

Location	Region	Distribution	AIC	BIC
Malta	R5	ZAGA	3612.42	3666.82
Manaira	R5	ZAGA	3631.18	3689.47
Nova Olinda	R5	ZAGA	3714.25	3780.31
Patos	R5	ZAGA	3656.41	3718.59
Pombal	R5	ZAGA	3646.57	3708.77
Princesa Isabel	R5	ZAGA	3694.23	3748.63
São João do Rio do Peixe	R5	ZAGA	3710.70	3776.77
São José de Piranhas	R5	ZAGA	3746.43	3816.37
Sousa	R5	ZAGA	3685.23	3739.64
Triunfo	R5	ZAGA	3666.80	3728.98

The inflated zero model, i.e., the ZAGA distribution function, provided the best time-series adjustments with zero presence, according to the AIC and BIC. Consequently, given that regions R4 and R5 present frequent periods of drought, the ZAGA distribution was the best way to characterize the behavior of the series.

Once the characteristic distribution for each location was determined, we verified the residual behavior of the model and the predictions under two scenario hypotheses. In the first scenario (S1), only the location parameter ( $\mu$ ) was used in the model equation, i.e., only the mean of coefficients of the predictor variables associated with the SST was used in the framework of the GAMLSS model, both in residual analysis and in the obtaining of climatic predictions. For the second scenario (S2), the location ( $\mu$ ), scale ( $\sigma$ ) and asymmetry ( $\nu$ ) parameters are adopted to not only using the mean of the distribution of the GAMLSS model but also inserting information related to the behavior of the variability and asymmetry associated with the SST indices.

Understanding the pattern of climate variability increases the accuracy found in climate modeling methodologies [96–98]. Therefore, the estimation of climatic parameters (Figures 3 and 4) allows the identification of probable variables that regulate the state rainfall regime. The GAMLSS model structure penalizes  $p$ -values above 5%, avoids multi-collinear problems, and increases the model’s predictive capacity to select which climate indices have the highest correlation with rainfall at each location.

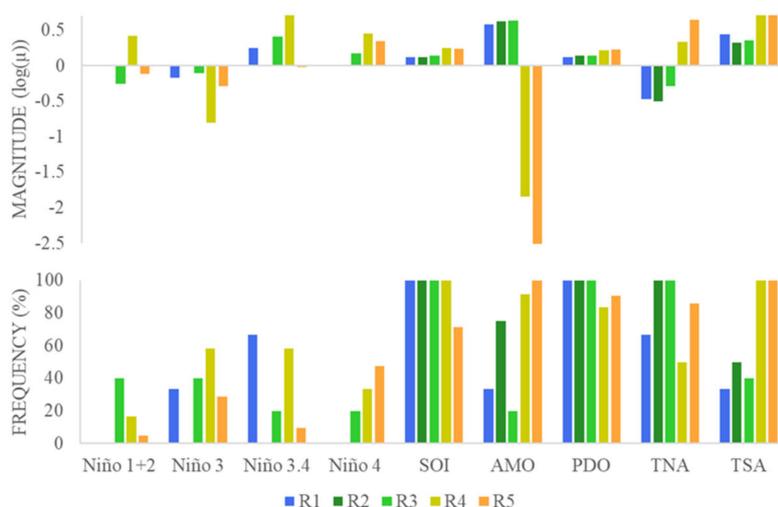
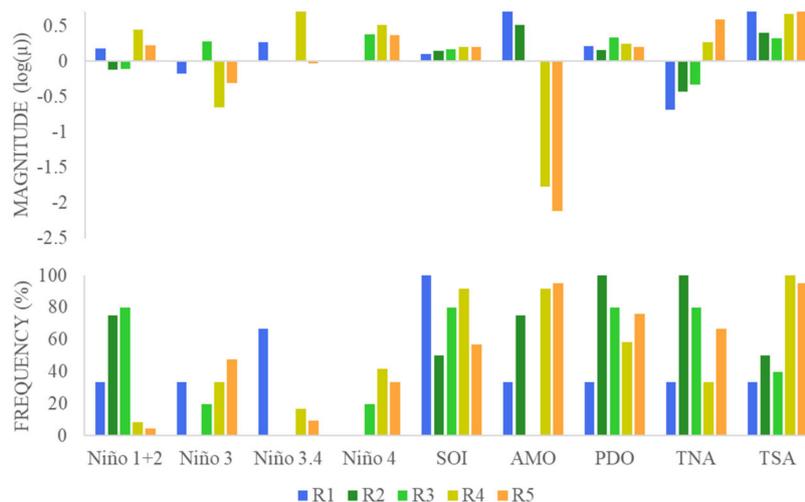


Figure 3. Magnitude of the parameter  $\mu$  for each of the five homogeneous study regions (R1–R5), associated with the climatic indices and their absolute frequency under S1 at the 5% significance level in the state of Paraíba.



**Figure 4.** Magnitude of the parameter  $\mu$  for each of the five homogeneous study regions (R1–R5), associated with the climatic indices and their absolute frequency under S2 at the 5% significance level in the state of Paraíba.

Figures 3 and 4 illustrate the frequency and magnitude of the parameters obtained for the climatic indices for each region. Figure 3 represents S1, where only the mean behavior of the climatic indices is used. When using the other parameters of variance and asymmetry, as in the equations expressed in S2, it is possible for the model to remove some of the climatic indices from the equation of the mean parameter, to reallocate the equations associated with the parameters of variance and asymmetry.

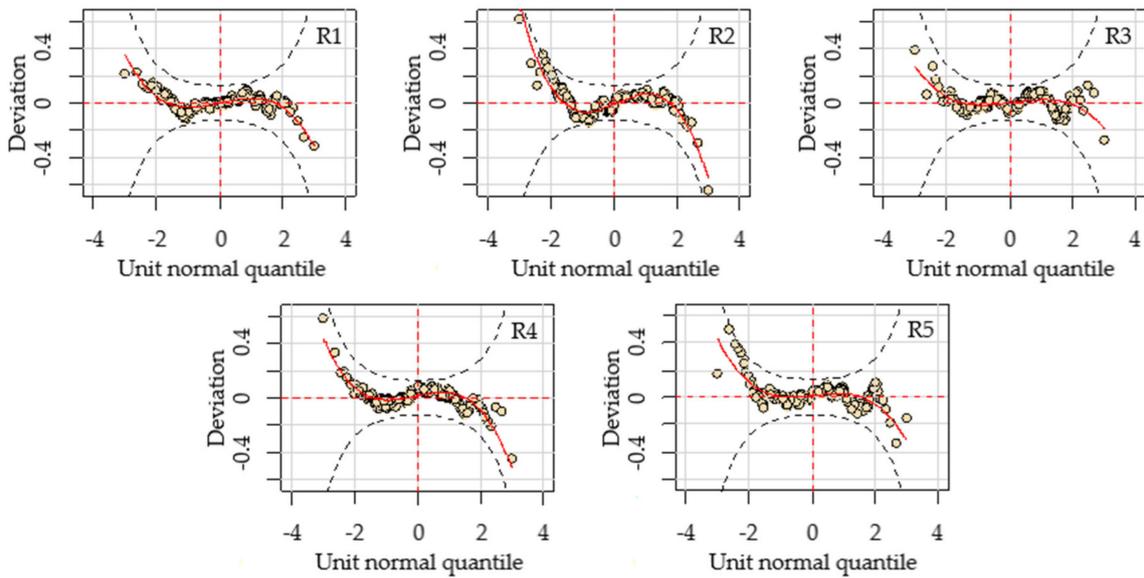
The parameters estimated by the GAMLSS model (Figures 3 and 4) indicated that the indices associated with the Niño regions in the Tropical Pacific Ocean showed a low frequency of contribution to the state's rainfall regime when compared to the other analyzed indices. This does not mean that the Tropical Pacific Ocean region plays a weak role in the state since the large-scale contribution associated with the Niño region influences the behavior of smaller systems, such as those associated with SST anomalies in the tropical Atlantic Ocean. It is worth noting that [20] also observed this contribution.

Therefore, it is essential to highlight the influence of Niño 1 + 2 in the R3, Niño 3 in the R4, Niño 3.4 in the R1 and R4, and Niño 4 in the R5. The SOI, AMO, PDO, TNA, and TSA indices are found to be more prevalent in rainfall behavior, as shown in Figures 3 and 4, with SOI and PDO being the most present. For this reason, the potential of these indices can be highlighted as good predictors, as already stated by some authors [36,99,100]. It is also noted that during the positive phase of SST in the North Atlantic Ocean, the behavior of TNA in the R1, R2, and R3 regions, and, in particular, the magnitude of AMO in the R4 and R5, makes it difficult for precipitation to occur in these regions.

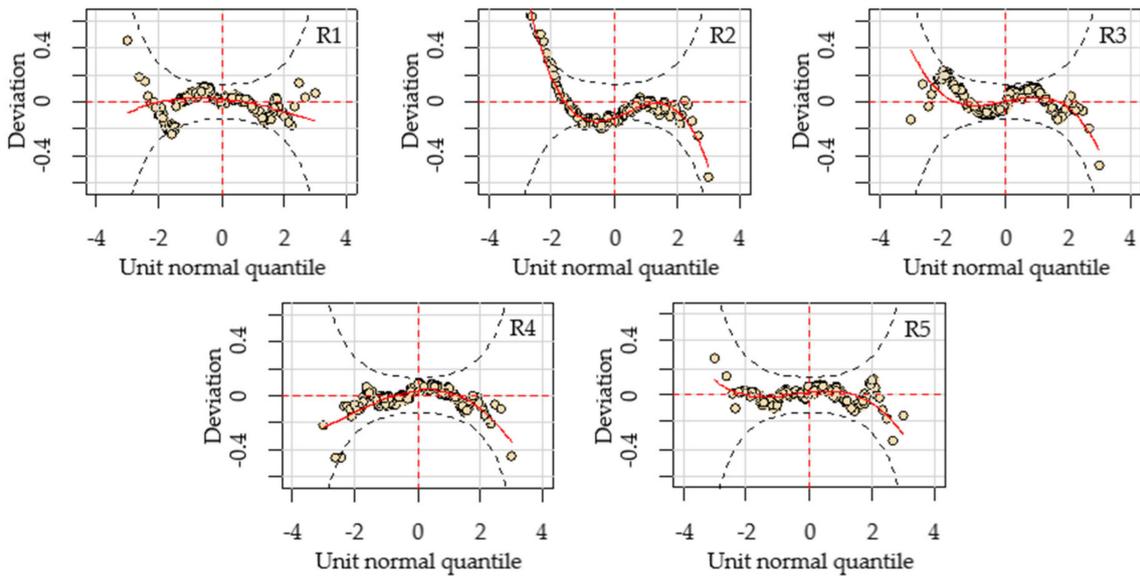
The quality of the linear relationship between precipitation and climatic indices by the GAMLSS model can be verified by using only the practicality found in the information from the worm-plot graph observed in Figures 5 and 6, as confirmed by [101].

Characteristics of the quality of the model can be verified in most of the analyzed locations because the points of the worm-plot graphs (Figures 5 and 6) are close to the solid red line, with little oscillation, close to the zero-deviation line, and within the 95% confidence intervals (dotted elliptical curves). This indicates that the model fits the time series better during the rainy season of the regions. However, it is necessary to be careful, since the large deviations found in the upper and lower normalized quantiles can be interpreted as a limitation of the model in capturing the occurrence of maximum and minimum accumulated precipitation because the curve behavior is asymmetric and leptokurtic at the ends of the worm-plot graphic. This limitation of time series adjustment associated with observed environmental data has already been verified in recent studies [34,77]. Therefore, caution is needed

when using this methodology in periods of irregular rainfall, such as those that occur during the dry season.



**Figure 5.** Residual diagnosis of the GAMLSS model characteristic of each homogeneous region in scenario S1, by structural worm-plot analysis.



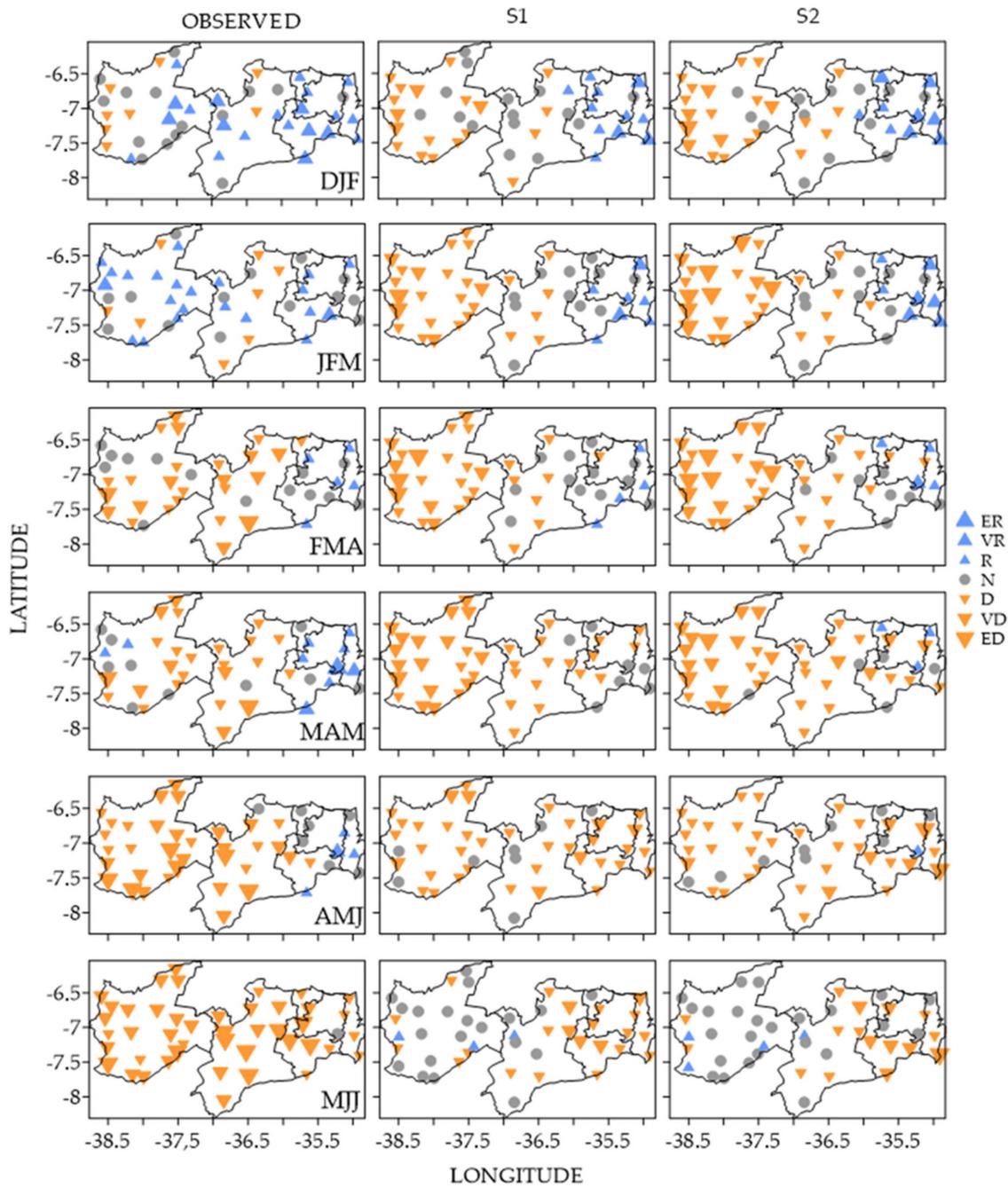
**Figure 6.** Residual diagnosis of the GAMLSS model characteristic of each homogeneous region in scenario S2, by structural worm-plot analysis.

### 3.3. Quantile Prediction

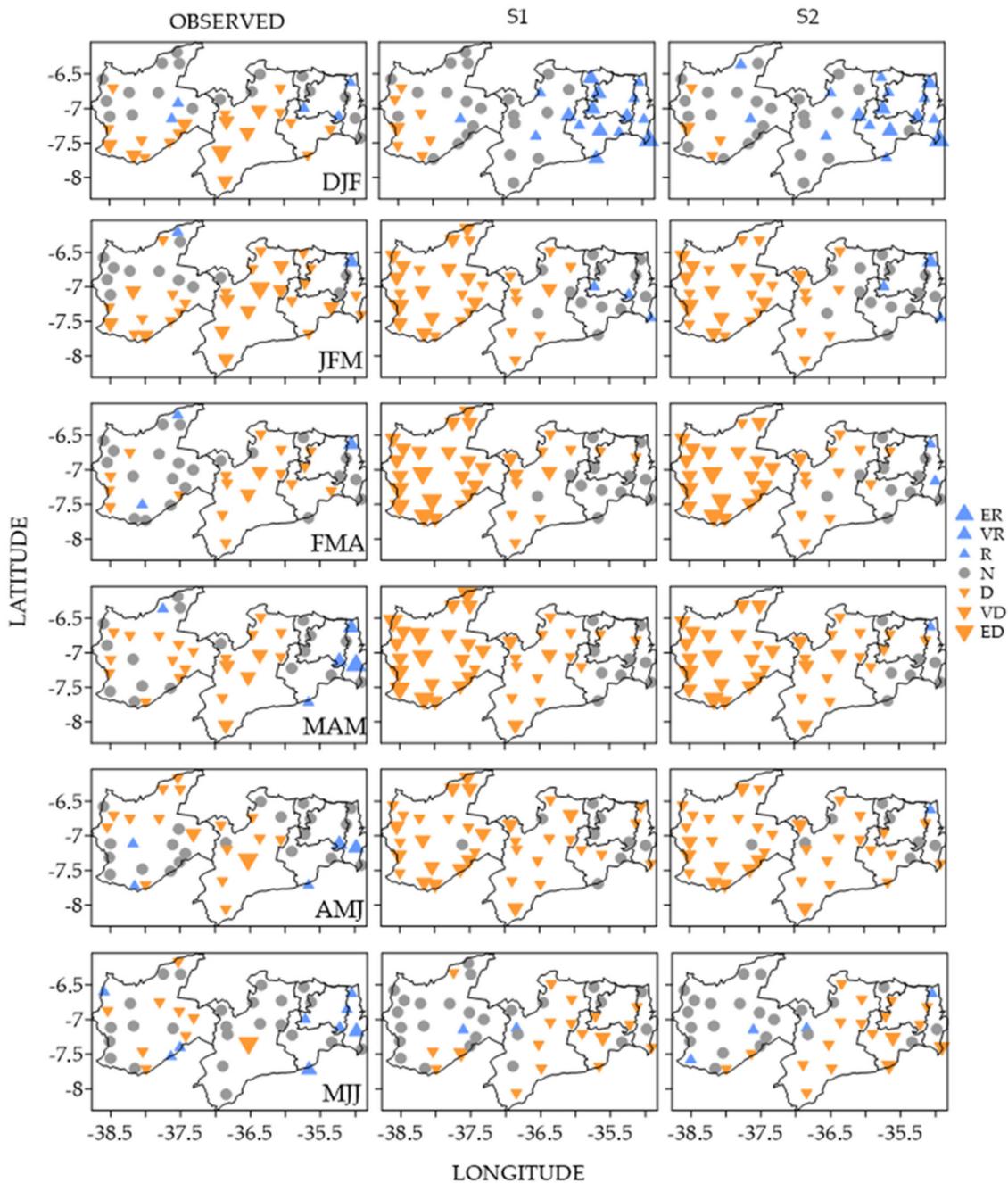
When quantile residues approximately follow the characteristic behavior of normal distribution [102], the rationale validates that the coefficients of the parameters found are capable of diagnosing the influence of the climate indices [103] on the distribution of rainfall and, consequently, its use in the simulation of the seasonal precipitation forecast in the state of Paraíba. Therefore, we used the quantile technique to examine the data and to classify the observed precipitation by percentile intervals (categories) of seasonal rainfall for each of the 45 locations, and to observe the applicability of the GAMLSS model to reproduce seasonal rainfall variability in the post-2015 period. This makes it possible to evaluate in practice (operational work), the quality of the model in estimating the behavior of the rainy season for the next two years (2016 and 2017) in the region, which were climatologically dry. As the ideal period, i.e., with less climate variability and more stable meteorological systems, for forecasting in the NEB is the six-month seasonal rainy (from January to June), given that rainfall during the dry season is insignificant [104], particularly in R4 and R5, for representing the highest percentage of annual rainfall, presenting more weather systems and less climatic variability. The three-month periods used in the approach were divided as DJF (December/January/February), JFM (January/February/March), FMA (February/March/April), MAM (March/April/May), AMJ (April/May/June), MJJ (May/June/July).

Predictions found in both scenarios were satisfactory for the observed rainfall data in the first period of 2016 (Figure 7). Even with the high seasonal and interannual variability of rainfall in the state resulting from mechanisms associated with SST [10], the model was able to highlight the role of climate indices as good predictors. As presented in Figure 7, the predictive capability of the model was able to indicate the general behavior of rainfall for R1, R2 and R3 during the DJF, JFM, FMA, AMJ and MJJ periods. The most reliable predictions for R4 and R5 were found in FMA, MAM, and AMJ.

Figure 8 presents the forecast simulation for 2017, emphasizing the results found between the JFM and AMJ periods. The most reliable rainfall estimates for R1, R2, and R3 were during the periods of FMA, MAM, and AMJ characterized for the proximity of forecasting, as demonstrated by a good association between the sign of observed rainfall and the simulated values. The adjustment of the model to the observed data was also present in the R4 region, in the periods of JFM, FMA, MAM, and AMJ. Subjectively, it is possible to see in Figure 8 coherence between the observed and predicted quantiles in MJJ. In the R5 region, the estimates did not achieve the same success as those found in the R4 region. However, acceptable results for DJF are observed because they indicate approximately the variability of the rainfall volume identified between normal, dry and very dry.



**Figure 7.** Quantiles of values observed and simulated by the GAMLSS model for S1 and S2 scenarios in Paraíba state for 2016 (Legend: ER = Extremely Rainy (▲), VR = Very Rainy (▲), R = Rainy (▲), N = Normal (●), D = Dry (▼), VD = Very Dry (▼) and, ED = Extremely Dry (▼)).



**Figure 8.** Quantiles of values observed and simulated by the GAMLSS model for S1 and S2 scenarios in Paraíba state for 2017 (Legend: ER = Extremely Rainy (▲), VR = Very Rainy (▲), R = Rainy (▲), N = Normal (●), D = Dry (▼), VD = Very Dry (▼) and, ED = Extremely Dry (▼)).

### 3.4. Efficiency Indicators

Possible changes in climate are likely to reduce the NEB rainy season [105], intensifying the meteorological systems that cause extreme rainfall and droughts [106], thus requiring the effectiveness of the models [107]. The reliability of the forecasts using SST anomalies in tropical regions is evident, and the efficacy of the models must be presented by statistical criteria [108]. Due to the adequacy of the scenarios for the first semester of 2016 and 2017, it was necessary to assess the accuracy of the estimates for each homogeneous region of the study. The authors of [109] found that using a statistical model could be the easiest way to predict reservoir capacity and drought events at the regional level.

The performance of the model in estimating the rainfall volume of each homogeneous region of Paraíba through statistical and efficiency indicators can be seen in Tables 5–9. For this purpose, it was chosen to compare the average rainfall of the observed and simulated data for 2016 and 2017.

**Table 5.** Statistical and efficiency indicators for the evaluation of prognostic performance in R1.

Period	Scenario	MAE (mm)	PBIAS (%)	RMSE (mm)	R <sup>2</sup> (%)
DJF	S1	59.56	55.1	60.87	0.65
	S2	71.14	65.9	75.23	0.62
JFM	S1	34.21	25.3	37.54	0.8
	S2	45.11	33.3	45.54	0.75
FMA	S1	12.95	−4.2	15.02	0.96
	S2	1.31	0.5	1.64	0.99
MAM	S1	85.25	−34	86	0.66
	S2	79.65	−31.8	79.65	0.68
AMJ	S1	69.62	−28.2	70.19	0.72
	S2	62.41	−25.3	64.45	0.75
MJJ	S1	49.27	−22	69.05	0.74
	S2	48.91	−18.5	64.17	0.77

MAE = Mean Absolute Error; PBIAS = Percent Bias; RMSE = Root Mean Square Error; R<sup>2</sup> = Coefficient of Determination.

**Table 6.** Statistical and efficiency indicators for the evaluation of prognostic performance in R2.

Period	Scenario	MAE (mm)	PBIAS (%)	RMSE (mm)	R <sup>2</sup> (%)
DJF	S1	35.31	31.5	44.53	0.81
	S2	29.19	26.2	36.89	0.84
JFM	S1	25.27	17.6	30.53	0.32
	S2	19.77	12.7	23.31	0.25
FMA	S1	15.12	−7.1	17.79	0.92
	S2	12.51	−9.5	13.44	0.9
MAM	S1	64.68	−34.6	71.69	0.64
	S2	62.57	−33.5	64.78	0.66
AMJ	S1	47	−25.8	53.4	0.73
	S2	44.15	−24.3	45.72	0.75
MJJ	S1	30.4	−18.4	39.19	0.79
	S2	38.2	−16.7	47.1	0.79

MAE = Mean Absolute Error; PBIAS = Percent Bias; RMSE = Root Mean Square Error; R<sup>2</sup> = Coefficient of Determination.

**Table 7.** Statistical and efficiency indicators for the evaluation of prognostic performance in R3.

Period	Scenario	MAE (mm)	PBIAS (%)	RMSE (mm)	R <sup>2</sup> (%)
DJF	S1	44.86	15.1	45.96	0.83
	S2	37.06	1.8	37.08	0.78
JFM	S1	21.97	17.1	24.71	0.96
	S2	20.34	8.6	21.11	0.97
FMA	S1	10.49	14.8	10.59	0.87
	S2	3.58	5.1	3.62	0.95
MAM	S1	15.91	−16.7	17.83	0.83
	S2	22.7	−23.8	23.55	0.95
AMJ	S1	12.86	−12.8	13.73	0.88
	S2	21.13	−21	21.3	0.79
MJJ	S1	22.67	−7.6	23.79	0.86
	S2	23.02	−16.5	27.82	0.78

MAE = Mean Absolute Error; PBIAS = Percent Bias; RMSE = Root Mean Square Error; R<sup>2</sup> = Coefficient of Determination.

**Table 8.** Statistical and efficiency indicators for the evaluation of prognostic performance in R4.

Period	Scenario	MAE (mm)	PBIAS (%)	RMSE (mm)	R <sup>2</sup> (%)
DJF	S1	26.26	12.3	26.61	0.77
	S2	27.28	11.8	27.6	0.76
JFM	S1	14.86	−0.2	14.86	0.88
	S2	17.6	−3.7	17.69	0.83
FMA	S1	8.55	15.8	10.66	0.86
	S2	5.66	10.6	7.11	0.9
MAM	S1	4.36	−2	4.43	0.98
	S2	3.09	−3.2	3.31	0.97
AMJ	S1	10.03	26.8	12.35	0.86
	S2	9.44	25.9	11.74	0.86
MJJ	S1	13.13	68.5	16.68	0.76
	S2	13.11	68.4	17.11	0.77

MAE = Mean Absolute Error; PBIAS = Percent Bias; RMSE = Root Mean Square Error; R<sup>2</sup> = Coefficient of Determination.

**Table 9.** Statistical and efficiency indicators for the evaluation of prognostic performance in R5.

Period	Scenario	MAE (mm)	PBIAS (%)	RMSE (mm)	R <sup>2</sup> (%)
DJF	S1	15.33	−16.8	20.28	0.82
	S2	20.58	−17.1	24.64	0.81
JFM	S1	52.9	−40.3	54.47	0.59
	S2	55.35	−42.3	59.26	0.56
FMA	S1	47.08	−37.6	52.52	0.61
	S2	49.75	−39.8	52.34	0.59
MAM	S1	44.77	−42.3	46.1	0.57
	S2	43.65	−41.2	44.29	0.59
AMJ	S1	24.41	−9.2	24.94	0.79
	S2	22.77	−3.8	22.87	0.84
MJJ	S1	16.64	48.5	21.79	0.89
	S2	18.84	64.9	24.98	0.8

MAE = Mean Absolute Error; PBIAS = Percent Bias; RMSE = Root Mean Square Error; R<sup>2</sup> = Coefficient of Determination.

The MAE, PBIAS, and RMSE indicators (Table 5) allowed the identification of the limited reliability of the model in estimating rainfall in R1, with only the FMA period being adequate. In other periods, indicators display significant disparity between observed and estimated rainfall.

High values of MAE, PBIAS and RMSE were also observed in the R2 region, as shown in Table 6. However, these values were lower than those shown in Table 5. The FMA seasonal continued to be the most statistically significant for estimating since the associated error between the observed and the simulated values was small, as observed in the MAE, PBIAS and RMSE indicators. As in the R1 region, the values found in the coefficient of determination for the JFM, FMA, AMJ and MJJ periods indicate that the forecast was able to follow the general pattern of rainfall in the region. Once the climate forecast for a given region has been obtained, decision-makers can prepare measures that are resilient to known climatic adversities, corroborating the findings of [98], which underlines the importance of investing in hydro-climatic monitoring systems.

The statistical indicators presented in Table 7 show that the model was capable of capturing rainfall characteristics in the R3 region for the entire first semester. Precipitation estimates for the FMA and MJJ periods stand out as those with the best statistical efficiency, which can be observed in the results of MAE, PBIAS and RSME, with the indication of little associated error and the coefficient of determination to indicate that the model can explain the variability of rainfall.

In addition, the validation of the model demonstrates its applicability qualitatively throughout the semester in the display of the expected climatological behavior, thus enabling the implementation of the climate forecast. This statement is associated with the low value found for PBIAS and the minimally acceptable value of the coefficient of determination

The predictive ability of the model showed efficiency in making estimates in the R4 region, as observed in the MAE, PBIAS, and RMSE indicators (Table 8). The reliability of the prognoses in reproducing the climatological behavior for the region is justified by the determination coefficient, which proves the adequacy in using the GAMLSS model class to generate seasonal precipitation prediction.

DJF and AMJ, as can be seen in Table 9, were seasonal that the model was able to follow at least the variability of rains that occur in R5. In others, it was possible to observe the model's low efficiency through statistical indicators. The predictions managed to follow the general pattern of rainfall behavior observed, as highlighted in the results obtained by the coefficient of determination. Although the indicators highlight the low efficiency in estimating the observed value, it can be seen that, at some locations in Figures 7 and 8, the forecast classification followed the observed rainfall pattern, suggesting that regions with the best predictive characteristics could be filtered when performing these tests by location.

The authors of [110] highlighted the importance of calibration in the modeling of statistical downscaling to obtain quality in the final result. However, they recommend that this modeling cannot be considered as an alternative, but that it should be integrated with other models in an ensemble. The authors of [111] identified that the choice of the appropriate model is not simple and depends on several factors such as location, climatic factors, data quality, among others. They observed a significant improvement in the ability of the models when selecting predictive variables with a higher level of influence in the study area.

#### 4. Discussion

Studies on SST interactions and their impact on NEB rainfall are essential for the assessment of climate change scenarios [106,107] argue that rainfall forecasts have the potential to benefit different sectors of society, but that there is a need for effective implementation of climate services, public policies, and close collaboration between scientists, climate professionals and organized civil society. This information makes it possible to reduce seasonal errors, to filter periodicity with higher statistical significance, and to improve the accuracy of the statistical models. Several authors have discussed the influence of SST in the precipitation variability on NEB [86,90–95].

Seasonal forecasts are crucial for the management of water resources [109], mainly from January to May in Paraíba, a critical period that guarantees the recharge of reservoirs [104] in the R4 and R5 regions. This is due to the frequent problem of water scarcity, which requires a higher predictive capacity of the model; this enables sustainable regional water management, reducing conflicts by improving food and water security [98].

The parameters estimated by the GAMLSS model for precipitation time series allow an understanding of the physical mechanisms associated with precipitation, indicating the climatic indices with the most significant influence on rainfall in the region, as shown in [69]. The presence of AMO, SOI, PDO, TNA and TSA indices was found to be highly representative, with the most relevant AMO, TNA and TSA indices, highlighting the importance of the tropical Atlantic Ocean as the primary regulator of the rainfall regime in the region. These results agree with [17].

The positive TSA phase, associated with the negative TNA phase, contributes to the formation of the north-south thermodynamic gradient [112,113], favoring the intensification of the low TSA, which in turn strengthens the northeast trade winds, displacing the southernmost ITCZ from its average climate position. As observed by [20], this pattern contributes to the occurrence of meteorological systems such as instability lines, easterly wave disturbances, and more frequent convective mesoscale complexes, which cooperate in the distribution of rainfall in all regions, particularly R4 and R5.

The possible explanation for the existence of TNA favoring precipitation in the R4 and R5 regions may be related to the behavior of trade winds [95] because, during the positive TNA phase, the northeast trades are weakened. As a result, rainfall in the R1, R2 and R3 regions is reduced. In this scenario, the increase in consecutive dry days and, thus, the reduction in expected rainfall for some locations can be observed [113]. However, this configuration contributes to the advancement of weather systems by the NEB, to some front and UTCV fragments inside Paraíba during the rainy pre-season, due to the strengthening of southeastern trade winds and cyclonic circulation in eastern South America and the South Atlantic [114].

A description of the behavior of rainfall distribution, which is conditioned by climate indices, is valuable information for water resource management in the state of Paraíba. Resilience measures for social problems associated with weather events, floods, landslides, and drought conditions can, therefore, be planned for the development and implementation of public policies. As the potential of the GAMLSS model is extensive, it can be discussed in future studies, in particular on the subject of environmental science. We also believe that this research can stimulate new applications of statistical modeling, particularly in the preparation of climate forecasts for the northeast region of Brazil.

## 5. Conclusions

It was proposed in this study to create a methodological structure capable of establishing a significant relationship between climatic indices and rainfall in the state of Paraíba, corresponding to 30 years (1986–2015), to develop seasonal rainfall predictions in 2016 and 2017. The results of the GAMLSS model showed the dominant influence of the tropical Atlantic Ocean on rainfall in Paraíba. In addition, the tropical Pacific Ocean also contributes to rainfall behavior in the R4 and R5 regions.

The model emphasized that the SOI, PDO, TNA, TSA, and AMO indices are the main precipitation regulators in Paraíba, classifying them as good predictors. The indices associated with the Niño regions have their characteristics of rainfall contribution, which stand out from the performance of Niño 3 and Niño 3.4.

The model has been shown to have predictive capabilities in all regions of the state during the summer and southern autumns. However, improvements can be applied to the development of the model. The potential contribution of the ZAGA regression model is highlighted in locations where there is no rainfall during the dry season. More significant results are found on a local scale than on a regional level.

The methodology applied can be used as an additional tool for the achievement of climate predictions by decision-makers. It highlights innovation in the application of statistical models for the simulation of rainfall behavior. This will allow the various sectors to promote strategies and planning for water resources in the region, as they proactively manage, with actions to combat, prevent, and respond to climate risks.

**Author Contributions:** Conceptualization, L.G.D., C.A.C.d.S., R.A.d.O., and J.I.B.d.B.; methodology, L.G.D., C.A.C.d.S. and R.A.d.O.; software, L.G.D. and R.A.d.O.; validation, L.G.D., C.A.C.d.S., R.A.d.O. J.I.B.d.B., C.A.G.S. and E.S.P.R.M.; formal analysis, L.G.D., C.A.C.d.S., R.A.d.O. J.I.B.d.B., C.A.G.S., E.S.P.R.M., G.d.O. and N.A.B.; investigation, L.G.D., C.A.C.d.S., R.A.d.O., J.I.B.d.B., C.A.G.S. and E.S.P.R.M.; resources, L.G.D.; data curation, L.G.D.; writing—original draft preparation, L.G.D., C.A.C.d.S., R.A.d.O. and J.I.B.d.B.; writing—review and editing, L.G.D., C.A.C.d.S., R.A.d.O. J.I.B.d.B., C.A.G.S., E.S.P.R.M., G.d.O. and N.A.B.; visualization, L.G.D., C.A.C.d.S. and R.A.d.O.; supervision, L.G.D., C.A.C.d.S., R.A.O., J.I.B.d.B., C.A.G.S., E.S.P.R.M., G.d.O. and N.A.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Brazilian National Council for Scientific and Technological Development (CNPq) (Grants N. 140667/2016-4 and 304493/2019-8) and the Brazilian Federal Agency for the Support and Evaluation of Graduate Education (CAPES)—Finance Code 001 (Pró-Alertas Research Project—Grant No. 88887.091737/2014-01 and 88887.123949/2015-00).

**Acknowledgments:** The authors acknowledge the Brazilian National Council for Scientific and Technological Development (CNPq) for funding the first author's scholarship (Research Project 140667/2016-4), and the Research Productivity Grant (Grant N. 304493/2019-8) for the second author, as well as the Coordination of Superior Level Staff Improvement (CAPES)—Finance Code 001 (Pró-Alertas Research Project—Grant No. 88887.091737/2014-01 and 88887.123949/2015-00).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Walker, G. World Weather. Word Weather. *Mon. Weather Rev.* **1928**, *56*, 167–170. [[CrossRef](#)]
- Bjerknes, J. Atmospheric teleconnections from the equatorial Pacific. *Mon. Weather Rev.* **1969**, *97*, 163–172. [[CrossRef](#)]
- Andreoli, R.V.; Kayano, M.T. A importância relativa do Atlântico Tropical Sul e Pacífico Leste na variabilidade da precipitação do Nordeste do Brasil. *Rev. Bras. Meteorol.* **2007**, *22*, 63–74. [[CrossRef](#)]
- Nogueira, V.S.; Cavalcanti, E.P.; Nogueira, V.F.B.; Rildo, G.M.; Fernandes, A.A. Oscilação interanual da precipitação observada associada aos sistemas convectivos de mesoescala sobre o estado da Paraíba. *Rev. Bras. Geogr. Fís.* **2014**, *7*, 969–978.
- Hastenrath, S.; Heller, L.D. Dynamics of climate hazards in northeast Brazil. *Q. J. R. Meteorol. Soc.* **1977**, *103*, 77–92. [[CrossRef](#)]
- Kousky, V.E.; Kayano, M.T.; Cavalcanti, I.F.A. A review of the Southern Oscillation: Oceanic-atmospheric circulation changes and related rainfall anomalies. *Tellus A* **1984**, *36*, 490–504. [[CrossRef](#)]
- Kayano, M.T.; Rao, V.B.; Moura, A.D. Tropical circulation and the associated rainfall anomalies during two contrasting years. *J. Clim.* **1988**, *8*, 477–488. [[CrossRef](#)]
- Kayano, M.T.; Andreoli, R.V. Relationships between rainfall anomalies over northeastern Brazil and the El Niño–Southern Oscillation. *J. Geophys. Res. Atmos.* **2006**, *111*. [[CrossRef](#)]
- Hastenrath, S. Exploring the climate problems of Brazil’s Nordeste: A review. *Clim. Chang.* **2012**, *112*, 243–251. [[CrossRef](#)]
- Hounsou-Gbo, G.A.; Araujo, M.; Bourlès, B.; Veleda, D.; Servain, J. Tropical Atlantic contributions to strong rainfall variability along the Northeast Brazilian coast. *Adv. Meteorol.* **2015**, 1–13. [[CrossRef](#)]
- Cintra, M.M.; Lentini, C.A.D.; Servain, J.; Araujo, M.; Marone, E. Physical processes that drive the seasonal Evolution of the Southwestern Tropical Atlantic Warm Pool. *Dyn. Atmos. Ocean.* **2015**, *72*, 1–11. [[CrossRef](#)]
- Uvo, C.B.; Repelli, C.A.; Zebiak, S.E.; Kushnir, Y. The relationships between Tropical Pacific and Atlantic SST and Northeast Brazil monthly precipitation. *J. Clim.* **1998**, *11*, 551–562. [[CrossRef](#)]
- Robertson, A.W.; Kirshner, S.; Smyth, P. Downscaling of daily rainfall occurrence over Northeast Brazil using a hidden Markov model. *J. Clim.* **2004**, *17*, 4407–4424. [[CrossRef](#)]
- Andrade, M.M.; Lima, K.C. Projeção climática da frequência de eventos de precipitação intensa no Nordeste do Brasil. *Rev. Bras. Geogr. Fís.* **2013**, *6*, 1158–1173. [[CrossRef](#)]
- Pereira, M.P.S.; Justino, F.; Malhado, A.C.M.; Barbosa, H.; Marengo, J. The influence of oceanic basins on drought and ecosystem dynamics in Northeast Brazil. *Environ. Res. Lett.* **2014**, *9*, 1–8. [[CrossRef](#)]
- Shimizu, M.H.; Ambrizzi, T.; Liebmann, B. Extreme precipitation events and their relationship with ENSO and MJO phases over northern South America. *Int. J. Clim.* **2016**, *37*, 2977–2989. [[CrossRef](#)]
- Marengo, J.A.; Torres, R.R.; Alves, L.M.M. Drought in Northeast Brazil—Past, present, and future. *Appl. Clim.* **2017**, *129*, 1189–1200. [[CrossRef](#)]
- Uvo, C.R.B. A Zona de Convergência Intertropical (ZCIT) e Sua Relação com a Precipitação da Região Norte do Nordeste Brasileiro. Master’s Thesis, National Institute for Space Research (INPE), São Paulo, Brazil, 1989.
- Gan, M.A. Um Estudo Observacional Sobre as Baixas Frias da Alta Troposfera, nas Latitudes Subtropicais do Atlântico Sul e Leste do Brasil. Master’s Thesis, National Institute for Space Research (INPE), São Paulo, Brazil, 1982.
- Cavalcanti, I.F.A.; Ferreira, N.J.; Dias, M.A.F.; Justi da Silva, M.G.A. *Tempo e Clima No Brasil*, 1st ed.; Oficina de Textos: São Paulo, Brazil, 2009.
- Souza, E.B.; Alves, J.M.B.; Repelli, C.A. Um complexo convectivo de mesoescala associado à precipitação intensa sobre Fortaleza—CE. *Rev. Bras. Meteorol.* **1998**, *13*, 1–14.
- Ferreira, A.G.; Mello, N.G.S. Principais sistemas atmosféricos atuantes sobre a região Nordeste do Brasil e a influência dos oceanos Pacífico e Atlântico no clima da região. *Rev. Bras. Clim.* **2005**, *1*, 15–28. [[CrossRef](#)]
- Vianello, R.L.; Alves, A.R. *Meteorologia Básica e Aplicações*, 2nd ed.; Federal University of Viçosa (UFV): Minas Gerais, Brazil, 2012.

24. Moura, A.D.; Shukla, J. On the dynamics of droughts in northeast Brazil: Observations, theory and numerical experiments with a general circulation model. *J. Atmos. Sci.* **1981**, *38*, 2653–2675. [[CrossRef](#)]
25. Marengo, J.A.; Alves, L.M.; Beserra, E.A.; Lacerda, F.F. *Variabilidade e mudanças climáticas no semiárido brasileiro. Recursos Hídricos em Regiões Áridas e Semiáridas*; Instituto Nacional do Semiárido: Paraíba, Brazil, 2011.
26. Sampaio, F.J. Iminência duma grande seca nordestina. *Revista Brasileira Geografia* **1950**, *12*, 3–15.
27. Marinho, M.E.; Rebouças, A.C. *Hidrologia das Secas: Nordeste do Brasil*; Sudene: Pernambuco, Brazil, 1970.
28. CNM—National Confederation of Municipalities (Portuguese, Confederação Nacional de Municípios). Decretações de Anormalidades Causadas por Desastres nos Municípios Brasileiros. *Desenvolvimento Territorial—Proteção e Defesa Civil—Brasília: CNM*. 2018. Available online: <https://www.cnm.org.br/cms/biblioteca/documentos/Decretacoes-de-anormalidades-causadas-por-desastres-nos-Municipios-Brasileiros-10-10-2018-v2.pdf> (accessed on 17 July 2019).
29. Morettin, P.A.; Toloi, C. *Análise de Séries Temporais*, 2nd ed.; Blucher: São Paulo, Brazil, 2006.
30. Bayer, D.M.; Castro, N.M.R. Modelagem e Previsão de Vazões Médias Mensais do Rio Potiribu Utilizando Modelos de Séries Temporais. *RBRH* **2012**, *17*, 229–239. [[CrossRef](#)]
31. Fathian, F.; Mehdizadeh, S.; Sales, A.K.; Safari, M.J.S. Hybrid models to improve the monthly river flow prediction: Integrating artificial intelligence and non-linear time series models. *J. Hydrol.* **2019**, *575*, 1200–1213. [[CrossRef](#)]
32. Mehdizadeh, S.; Fathian, F.; Adamowski, J.D. Hybrid artificial intelligence-time series models for monthly streamflow modeling. *Appl. Soft Comput.* **2019**, *80*, 873–887. [[CrossRef](#)]
33. Rigby, R.A.; Stasinopoulos, D.M. Generalized additive models for location, scale and shape. *J. R. Stat. Soc. C Appl.* **2005**, *54*, 507–554. [[CrossRef](#)]
34. Gao, L.; Huang, J.; Chen, X.; Chen, Y.; Liu, M. Contributions of natural climate changes and human activities to the trend of extreme precipitation. *Atmos. Res.* **2018**, *205*, 60–69. [[CrossRef](#)]
35. Medeiros, E.S.; Lima, R.R.; Olinda, R.A.; Dantas, L.G.; Santos, C.A.C. Space-time kriging of precipitation: Modeling the large-scale variation with model GAMLSS. *Water* **2019**, *11*, 2368. [[CrossRef](#)]
36. Rashid, M.M.; Beecham, S. Simulation of stream with statistically downscaled daily rainfall using a hybrid of wavelet and GAMLSS models. *Hydrol. Sci. J.* **2019**, *64*, 1327–1339. [[CrossRef](#)]
37. Rydén, J. A note on analysis of extreme minimum temperatures with the GAMLSS framework. *Acta Geophys.* **2019**, *67*, 1599–1604. [[CrossRef](#)]
38. Das, J.; Jha, S.; Goyal, M.K. Non-stationary and copula-based approach to assess the drought characteristics encompassing climate indices over the Himalayan states in India. *J. Hydrol.* **2020**, *580*. [[CrossRef](#)]
39. Tan, X.; Gan, T.Y. Non-stationary analysis of annual maximum streamflow of Canada. *J. Clim.* **2015**, *28*, 1788–1805. [[CrossRef](#)]
40. Stasinopoulos, D.M.; Rigby, R.A. Generalized Additive Models for Location Scale and Shape (GAMLSS) in R. *J. Stat. Softw.* **2007**, *23*, 1–46. [[CrossRef](#)]
41. Villarini, G.; Smith, J.A.; Serinaldi, F.; Bales, J.; Bates, P.D.; Krajewski, W.F. Flood frequency analysis for non-stationary annual peak records in an urban drainage basin. *Adv. Water Resour.* **2009**, *32*, 1255–1266. [[CrossRef](#)]
42. Van Ogtrop, F.F.; Vervoort, R.W.; Heller, G.Z.; Stasinopoulos, D.M.; Rigby, R.A. Long-range forecasting of intermittent streamflow. *Hydrol. Earth Syst. Sc.* **2011**, *15*, 3343–3354. [[CrossRef](#)]
43. Voudoris, V.; Gilchrist, R.; Rigby, R.; Sedgwick, J.; Stasinopoulos, D. Modelling skewness and kurtosis with the BCPE density in GAMLSS. *J. Appl. Stat.* **2012**, *36*, 1279–1293. [[CrossRef](#)]
44. López, J.; Francés, F. Non-stationary flood frequency analysis in continental Spanish rivers, using climate and reservoir indices as external covariates. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 3189–3203. [[CrossRef](#)]
45. Gu, X.; Zhang, Q.; Li, J.; Singh, V.P.; Sun, P. Impact of urbanization on nonstationarity of annual and seasonal precipitation extremes in China. *J. Hydrol.* **2019**, *575*, 638–675. [[CrossRef](#)]
46. Apurv, T.; Cai, X. Evaluation of the stationary assumption for meteorological drought risk estimation at the multidecadal scale in contiguous United States. *Water Resour. Res.* **2019**, *55*, 5074–5101. [[CrossRef](#)]
47. IBGE—Brazilian Institute of Geography and Statistics (Portuguese, Instituto Brasileiro de Geografia e Estatística). 2017. Available online: <https://cidades.ibge.gov.br/brasil/pb/panorama> (accessed on 1 January 2018).

48. ANA—National Water Agency of Brazil (Portuguese, Agência Nacional de Águas. 2018. Available online: [http://www.snirh.gov.br/hidroweb/publico/medicoes\\_historicas\\_abas.jsf](http://www.snirh.gov.br/hidroweb/publico/medicoes_historicas_abas.jsf) (accessed on 2 February 2018).
49. AESA—State Water Resources Management Executive Agency (Portuguese, Agência Executiva de Gestão das Águas do Estado da Paraíba. 2018. Available online: <http://www.aesa.pb.gov.br/aesa-website/meteorologia-chuvas/> (accessed on 2 February 2018).
50. CPC/NCEP/NOAA—Climate Prediction Center/National Centers for Environmental Prediction/National Oceanic and Atmospheric Administration. 2018. Available online: <http://www.cpc.ncep.noaa.gov/data/indices/> (accessed on 2 February 2018).
51. ESRL/NOAA—Earth System Research Laboratory/National Oceanic & Atmospheric Administration. 2018. Available online: <https://www.esrl.noaa.gov/psd/data/climateindices/list/> (accessed on 2 February 2018).
52. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. 2019. Available online: <http://cran.r-project.org/index.html> (accessed on 13 December 2019).
53. Dantas, L.G.; Santos, C.A.C.; Olinda, R.A. Tendências anuais e sazonais nos extremos de temperatura do ar e precipitação em Campina Grande—PB. *Rev. Bras. Meteorol.* **2015**, *30*, 423–434. [[CrossRef](#)]
54. Honaker, J.; King, G.; Blackwell, M. Amelia II: A program for missing data. *J. Stat. Softw.* **2011**, *45*, 1–47. [[CrossRef](#)]
55. Dantas, L.G.; Santos, C.A.C.; Olinda, R.A. Reamostragem de séries pluviométricas no estado da Paraíba. *Rev. Bras. Geogr. Fís.* **2016**, *9*, 997–1006. [[CrossRef](#)]
56. Gilabert, M.B.; Cobelas, M.A.; Angeler, D.G. Effects of climatic change on stream water quality in Spain. *Clim. Chang.* **2010**, *103*, 339–352. [[CrossRef](#)]
57. Sheng, H.; Liu, H.; Wang, C.; Guo, H.; Liu, Y.; Yang, Y. Analysis of cyanobacteria bloom in the Waihai part of Dianchi Lake, China. *Ecol. Inf.* **2012**, *10*, 37–48. [[CrossRef](#)]
58. Kruyt, B.; Lehning, M.; Kahl, A. Potential contributions of wind power to a stable and highly renewable Swiss power supply. *Appl. Energy* **2017**, *192*, 1–11. [[CrossRef](#)]
59. Revelle, W. Hierarchical cluster analysis and the internal structure of tests. *Multivar. Behav. Res.* **1979**, *14*, 57–74. [[CrossRef](#)]
60. Milligan, G.W.; Cooper, M.C. A study of the comparability of external criteria for hierarchical cluster analysis. *Multivar. Behav. Res.* **1986**, *21*, 441–458. [[CrossRef](#)]
61. Santos, C.A.G.; Neto, R.M.B.; Silva, R.M.; Costa, S.G.F. Cluster analysis applied to spatiotemporal variability of monthly precipitation over Paraíba state using Tropical Rainfall Measuring Mission (TRMM) data. *Remote Sens.* **2019**, *11*, 637. [[CrossRef](#)]
62. Oliveira, P.T.; Santos e Silva, C.M.S.; Lima, K.C. Climatology and trend analysis of extreme precipitation in subregions of Northeast Brazil. *Appl. Clim.* **2017**, *130*, 77–90. [[CrossRef](#)]
63. Kassambara, A.; Mundt, F. Factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R Package Version 1.0.5. 2017. Available online: <https://CRAN.R-project.org/package=factoextra> (accessed on 13 December 2019).
64. Murtagh, F.; Legendre, P. Ward’s hierarchical agglomerative clustering method: Which algorithms implement ward’s criterion? *J. Classif.* **2014**, *31*, 274–295. [[CrossRef](#)]
65. Von Storch, V.H. Misuses of statistical analysis in climate research. In *Analysis of Climate Variability: Applications of Statistical Techniques*; von Storch, H., Navarra, A., Eds.; Springer: Berlin, Germany, 1995; pp. 11–26. [[CrossRef](#)]
66. Yue, S.; Wang, C.Y. Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test. *Water Resour. Res.* **2012**, *38*, 1–7. [[CrossRef](#)]
67. Blain, G.C. The Mann-Kendall test: The need to consider the interaction between serial correlation and trend. *Acta Sci. Agron.* **2013**, *35*, 393–402. [[CrossRef](#)]
68. Verma, M.K.; Swain, S. Statistical analysis of precipitation over seonath River Basin, Chhattisgarh, India. *Int. J. App. Eng. Res.* **2016**, *11*, 2417–2423. [[CrossRef](#)]

69. Gu, X.; Zhang, Q.; Singh, V.P.; Liu, L.; Shi, P. Spatiotemporal patterns of annual and seasonal precipitation extreme distributions across China and potential impact of tropical cyclones. *Int. J. Clim.* **2017**, *37*, 3949–3962. [[CrossRef](#)]
70. Chan, K.S.; Ripley, B. TSA: Time Series Analysis. R Package Version. 2012. Available online: <https://CRAN.R-project.org/package=TSA> (accessed on 13 December 2019).
71. Villarini, G.; Smith, J.A.; Napolitano, F. Non-stationary modeling of a long record of rainfall and temperature over Rome. *Adv. Water Resour.* **2010**, *33*, 1256–1267. [[CrossRef](#)]
72. Villarini, G.; Serinaldi, F. Development of statistical models for at-site probabilistic seasonal rainfall forecast. *Int. J. Clim.* **2012**, *32*, 2197–2212. [[CrossRef](#)]
73. Zhang, D.D.; Yan, D.H.; Wang, Y.C.; Lu, F.; Liu, S.H. GAMLSS-based non-stationary modeling of extreme precipitation in Beijing-Tianjin-Hebei region of China. *Nat. Hazards* **2015**, *77*, 1037–1053. [[CrossRef](#)]
74. Cribari-Neto, F.; Lucena, S.E.F. Non-nested hypothesis testing inference for GAMLSS modes. *J. Stat. Comput. Sim.* **2016**, *87*, 1189–1205. [[CrossRef](#)]
75. Gao, L.; Huang, J.; Chen, X.; Chen, Y.; Liu, M. Risk of extreme precipitation under non-stationary conditions during the second flood season in the southeastern coastal region of China. *J. Hydrometeorol.* **2017**, *18*, 669–681. [[CrossRef](#)]
76. Rashid, M.M.; Beechman, S. Development of a non-stationary Standardized Precipitation Index and its application to a South Australian climate. *Sci. Total Environ.* **2019**, *657*, 882–892. [[CrossRef](#)]
77. Stasinopoulos, M.D.; Rigby, R.A.; Heller, G.Z.; Voudouris, V.; De Bastiani, F. (Eds.) *Flexible Regression and Smoothing: Using GAMLSS in R*; Chapman and Hall/CRC The R Series: Boca Raton, FL, USA, 2017. [[CrossRef](#)]
78. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [[CrossRef](#)]
79. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6*, 461–464. [[CrossRef](#)]
80. Filliben, J.J. The probability plot correlation coefficient test for normality. *Technometrics* **1975**, *17*, 111–117. [[CrossRef](#)]
81. Van Buuren, S.; Fredriks, M. Worm plot: A simple diagnostic device for modelling growth reference curves. *Stat. Med.* **2001**, *20*, 1259–1277. [[CrossRef](#)] [[PubMed](#)]
82. Tyralis, H.; Papacharalampous, G.; Tantanee, S. How to explain and predict the shape parameter of the generalized extreme value distribution of streamflow extremes using a big dataset. *J. Hydrol.* **2019**, *574*, 628–645. [[CrossRef](#)]
83. Khouakhi, A.; Villarini, G.; Zhang, W.; Slater, L.J. Seasonal predictability of high sea level frequency using ENSO patterns along the U.S. West Coast. *Adv. Water Resour.* **2019**, *131*. [[CrossRef](#)]
84. Pinkayan, S. *Conditional Probabilities of Occurrence of Wet and Dry Years over a Large Continental Area. Hydrology Papers, n. 12*; Colorado State University: Fort Collins, CO, USA, 1966.
85. Xavier, T.M.B.S. *A Técnica Dos Quantis e Suas Aplicações em Meteorologia, Climatologia e Hidrologia, Com Ênfase Para as Regiões Brasileiras*; Thesaurus: Brasília, DF, Brazil, 2002.
86. Kayano, M.T.; Capistrano, V.B.; Andreoli, R.V.; Souza, R.A.F. A further of the tropical Atlantic SST modes and their relations to north-eastern Brazil rainfall during different phases of Atlantic Multidecadal Oscillation. *Int. J. Clim.* **2016**, *36*, 4006–4018. [[CrossRef](#)]
87. Tavares, A.L.; Carmo, A.M.C.; Silva Júnior, R.O.; Souza-Filho, P.W.M.; Silva, M.S.; Ferreira, D.B.S.; Nascimento Júnior, W.R.; Dall’agnol, R. Climate indicators for a watershed in the eastern amazon. *Rev. Bras. Clim.* **2018**, *23*, 389–410. [[CrossRef](#)]
88. Hyndman, R.J.; Fan, Y. Sample quantiles in statistical packages. *Am. Stat.* **1996**, *50*, 361–365. [[CrossRef](#)]
89. Capozzoli, C.R.; Cardoso, A.O.; Ferraz, S.E.R. Padrões de variabilidade de vazão de rios nas principais bacias brasileiras e associação com índices climáticos. *Rev. Bras. Meteorol.* **2017**, *32*, 243–254. [[CrossRef](#)]
90. Alves, J.M.B.; Repelli, C.A. A variabilidade pluviométrica no setor norte do Nordeste e os eventos El Niño-Oscilação Sul (ENOS). *Rev. Bras. Meteorol.* **1992**, *7*, 583–592.
91. Alves, J.M.B.; Souza, R.O.; Campos, J.N.B. Previsão da anomalia de temperatura da superfície do mar (TSM) no atlântico tropical, com a equação da difusão de temperatura. *Rev. Climaná.* **2003**, *1*, 6–19.

92. Moura, G.B.A.; Aragão, J.O.R.; Melo, J.S.P.; Silva, A.P.N.; Giongo, P.R.; Lacerda, F.F. Relação entre a precipitação do leste do Nordeste do Brasil e a temperatura dos oceanos. *Rev. Bras. Eng. Agríc. Ambient.* **2009**, *13*, 462–469. [[CrossRef](#)]
93. Gherardi, D.F.M.; Paes, E.T.; Soares, C.H.; Pezzi, L.P.; Kayano, M.T. Differences between spatial patterns of climate variability and large marine ecosystem in the western South Atlantic. *Pan Am. J. Aquat. Sci.* **2010**, *5*, 310–319.
94. Guerreiro, M.J.S.; Andrade, E.M.; Abreu, I.; Lajinha, T. Long-term variation of precipitation indices in Ceará State Northeast Brazil. *Int. J. Clim.* **2013**, *33*, 2929–2939. [[CrossRef](#)]
95. Hounsou-Gbo, G.A.; Servain, J.; Araujo, M.; Caniaux, G.; Bourlès, B.; Fontenele, D.; Martins, E.S.P.R. SST indexes in the Tropical South Atlantic for forecasting rainy seasons in Northeast Brazil. *Atmosphere* **2019**, *10*, 335. [[CrossRef](#)]
96. Rashid, M.M.; Beecham, S.; Chowdhury, R.K. Simulation of extreme rainfall and projection of future changes using the GLIMCLIM model. *Appl. Clim.* **2016**, *130*, 1–14. [[CrossRef](#)]
97. Rodrigues, J.S. Análise de Diagnóstico em Modelo de Regressão ZAGA e ZAIG. Master's Thesis, University of São Paulo/Federal University of São Carlos, São Paulo, Brazil, 2016.
98. Pousa, R.; Costa, M.H.; Pimenta, F.M.; Fontes, V.C.; Brito, V.F.A.; Castro, M. Climate change and intense irrigation growth in western Bahia, Brazil: The urgent need for hydroclimatic monitoring. *Water* **2019**, *11*, 933. [[CrossRef](#)]
99. Villarini, G.; Vecchi, G.A.; Smith, J.A. Modeling the dependence of tropical storm counts in the north Atlantic basin on climate indices. *Mon. Weather Rev.* **2010**, *138*, 2681–2705. [[CrossRef](#)]
100. Gu, X.; Zhang, Q.; Singh, V.P.; Chen, X.; Liu, L. Nonstationary in the occurrence rate of floods in the Tarim River basin, China, and related impacts of climate indices. *Glob. Planet Chang.* **2016**, *142*, 1–13. [[CrossRef](#)]
101. Tong, E.N.C.; Mues, C.; Thomas, L. A zero-adjusted gamma model for mortgage loan loss given default. *Int. J.* **2013**, *29*, 548–562. [[CrossRef](#)]
102. Chen, S.; Shao, D.; Tan, X.; Gu, W.; Lei, C. An interval multistage classified model for regional inter- and intra-seasonal water management under uncertain and non-stationary condition. *Agric. Water Manag.* **2017**, *191*, 118–128. [[CrossRef](#)]
103. Zhang, Q.; Gu, X.; Singh, V.P.; Xiao, M.; Chen, X. Evaluation of flood frequency under non-stationary resulting from climate indices and reservoir indices in the East River basin, China. *J. Hydrol.* **2015**, *527*, 565–575. [[CrossRef](#)]
104. Delgado, J.M.; Voss, S.; Bürguer, G.; Vormoor, K.; Murawski, A.; Pereira, J.M.R.; Martins, E.; Júnior, F.V.; Francke, T. Seasonal drought prediction for semi-arid northeastern Brazil: Verification of six hydro-meteorological forecast products. *Hydrol. Earth Syst. Sci.* **2018**, *22*, 5041–5056. [[CrossRef](#)]
105. Martins, E.S.P.R.; Coelho, C.A.S.; Haarsma, R.; Otto, F.E.L.; King, A.D.; Van Oldenborgh, G.J.; Kew, S.; Philip, S.; Júnior, F.C.V.; Cullen, H. A multimethod attribution analysis of the prolonged northeast Brazil hydro-meteorological drought (2012–2016). (in “Explaining Extreme Events of 2016 from a Climate Perspective”). *Bull. Am. Meteorol. Soc.* **2018**, *99*, S65–S69. [[CrossRef](#)]
106. Silva, T.L.V.; Veleza, D.; Araujo, M.; Tyaquicã, P. Ocean-atmosphere feedback during extreme rainfall events in eastern Northeast Brazil. *J. Appl. Meteorol. Clim.* **2018**, *57*, 1211–1229. [[CrossRef](#)]
107. Tompkins, A.M.; Lowe, R.; Nissan, H.; Martiny, N.; Roucou, P.; Thomson, M.C.; Nakazawa, T. Predicting climate impacts on health at sub-seasonal to seasonal timescales. In *The Gap between Weather and Climate Forecasting: Sub-Seasonal to Seasonal Prediction*; Vitart, F., Robertson, A., Eds.; Elsevier: Amsterdam, The Netherlands, 2019; pp. 455–477. [[CrossRef](#)]
108. Trenberth, K.E. El Niño Southern Oscillation (ENSO). In *Reference Module in Earth Systems and Environmental Sciences, Encyclopedia of Ocean Sciences*, 3rd ed.; Elsevier: Amsterdam, The Netherlands, 2017; Volume 6, pp. 420–432. [[CrossRef](#)]
109. Pilz, T.; Delgado, J.M.; Voss, S.; Vormoor, K.; Francke, T.; Costa, A.C.; Martins, E.; Bronsert, A. Seasonal drought prediction for semi-arid northeast Brazil: What is the added value of a process-based hydrological model? *Hydrol. Earth Syst. Sci.* **2019**, *23*, 1951–1971. [[CrossRef](#)]
110. Sangelantoni, L.; Ferreti, R.; Redaelli, G. Toward a Regional-Scale Seasonal Climate Prediction System over Central Italy Based on Dynamical Downscaling. *Climate* **2019**, *7*, 120. [[CrossRef](#)]
111. Manzanos, R.; Lucero, A.; Weisheimer, A.; Gutiérrez, J.M. Can bias correction and statistical downscaling methods improve the skill of seasonal precipitation forecasts? *Clim. Dyn.* **2017**, *50*, 1161–1176. [[CrossRef](#)]

112. Alves, J.M.B.; Silva, E.M.; Sombra, S.S.; Barbosa, A.C.B.; Santos, A.C.S.; Lira, M.A.T. Eventos extremos diários de chuva no Nordeste do Brasil e características atmosféricas. *Rev. Bras. Meteorol.* **2017**, *32*, 227–233. [[CrossRef](#)]
113. Nóbrega, J.N. Eventos Extremos de Precipitação nas Mesorregiões do Estado da Paraíba e Suas Relações com a TSM dos Oceanos Atlântico e Pacífico. Master's Thesis, Federal University of Campina Grande, Paraíba, Brazil, 2012.
114. Junior, F.C.V.; Jones, C.; Gandu, A.W. Interannual and intraseasonal variations of the onset and demise of the pre-wet season and the wet season in the Northern Northeast Brazil. *Rev. Bras. Meteorol.* **2018**, *33*, 472–484. [[CrossRef](#)]



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).