

Probabilistic Modeling of Monthly Temperature Historical Series in Mossoró, Northeastern Brazil

Janilson Pinheiro de Assis¹, Roberto Pequeno de Sousa¹, Ben Deivide de Oliveira Batista²,
Paulo César Ferreira Linhares¹, Eudes de Almeida Cardoso¹, José Aluisio de Araújo Paula³
& Ariana Moraes Neves¹

¹ Jitirana Research Group, Department of Agronomic and Forestry Sciences, Federal Rural Semi-Arid University, Mossoró, RN, Brazil

² Department of Physics and Mathematics, Federal University of São João Del Rei, Ouro Branco, MG, Brazil

³ SEBRAE, Mossoró, RN, Brazil

Correspondence: Janilson Pinheiro de Assis, Jitirana Research Group, Department of Agronomic and Forestry Sciences, Federal Rural Semi-Arid University, Mossoró, RN, Brazil. E-mail: janilson@ufersa.edu.br

Received: August 23, 2018

Accepted: September 26, 2018

Online Published: November 15, 2018

doi:10.5539/jas.v10n12p534

URL: <https://doi.org/10.5539/jas.v10n12p534>

Abstract

We fitted the following seven distribution probabilities to the data of monthly average temperature in Mossoró, northeastern Brazil: Normal, Log-Normal, Beta, Gamma, Log-Pearson (Type III), Gumbel, and Weibull. To assess the goodness of fit the empirical distributions to the theoretical distribution, we applied the tests of Kolmogorov-Smirnov, Chi-square, Cramer-von Mises, Anderson-Darling, Kuiper, and Logarithm of Maximum Likelihood, at 10% of probability. The temperature series were obtained from 1970 to 2007. The Normal distribution provided the best fit to the historical series of average monthly temperature. Although the Kolmogorov-Smirnov test showed a very high level of approval, which generated some uncertainty regarding the test criteria, it is the more recommended to studies with approximately symmetric data and small series.

Keywords: temperature, probability distribution, goodness of fit tests

1. Introduction

Studies of the behavior of rainfall, air temperature, relative air humidity, evaporation, direction and speed of the wind, global solar radiation, the occurrence of dew, fog, hail, frost, and snow are essential tools to decision-making related to agricultural and human activities in building and tourism (Mota, 1981). Among these variables, the temperature is one of the leading factors for climatic characterization, because it exerts several effects in the agricultural practice (Araújo et al., 2010a).

The fitting of probability distributions to climatic variables over time aims to understand the meteorological phenomena, clarify their occurrence patterns, and make probabilistic forecasts. Good models provide reasonable predictability of the climatic behavior of a region, becoming critical tools for planning and managing diverse human activities.

A theoretical distribution is an abstract mathematical formula or characteristic shape. Some of these formulae emerge naturally as a consequence of certain kinds of data generating processes, and when applicable, these are mainly plausible candidates for concisely representing variations in a set of data. Even when there is no substantial natural basis behind the choice of a particular theoretical distribution, one may empirically find that such distributions suit very well a set of data (Pedrosa & Gama, 2004).

The simple construction of a frequency histogram for the visualization of a sample of data becomes insufficient to choose which function best represent the data among the several probability density distributions. Some criteria and goodness of fit tests help to verify if a known function suits the distribution of a dataset. There are several distributions of probability for discrete and continuous random variables. Among the ones used for discrete data, we find Bernoulli, Binomial, Negative Binomial, Hypergeometric, Geometric, and Poisson. For continuous variables, we can use the Uniform or Rectangular, Normal, Log-Normal, Gamma, Weibull, Gumbel, Exponential, Beta, Chi-square, t Student, and F Snedecor distributions, among others.

Several studies fitting theoretical probability distributions to climatic variables such as rainfall (Berlato, 1987; Botelho & Morais, 1999; Sampaio et al., 1999; Catalunha et al., 2002; Murta et al., 2005; Dallacort et al., 2011; Rodrigues et al., 2014; Lima et al., 2015; Kist & Virgens Filho, 2015), air temperature (Mota et al., 1999; Assis et al., 2004; Praela-Pantano et al., 2009; Araújo et al., 2010a, 2010b), solar radiation (Buriol et al., 2001), and wind speed (Assis et al., 2017) emphasize the benefits of modeling in planning activities to reduce climatic risks.

Goodness of fit tests, such as Kolmogorov-Smirnov, Chi-square, Cramer-von Mises, Anderson-Darling, Kuiper, Lilliefors, Shapiro-Wilk, and the Logarithm of Maximum Likelihood (Campos, 1983; Assis et al., 1996; Moretin & Bussab, 2004; Cooke, 1993), verify if the empirical values may come from a population with a particular theoretical distribution. In the tests, the null hypothesis states that the distribution is the one specified with their estimated parameters based on the sample data (Assis et al., 1996; Catalunha et al., 2002). The logarithm of the Maximum Likelihood shows a good quality of fit if its value is negative and the lowest possible (Cooke, 1993).

The use of appropriate probabilistic models provides mathematical precision, allowing consistent studies of historical data series. Arruda et al. (1981), assessing a 50 years series data of absolute minimum temperatures of June and July in the region of Campinas, defined and tested the distribution models of extreme values and normal distribution. They concluded that they both models suit the probabilities the dataset. These probabilistic models were also evaluated by Silva et al. (1986) to Lavras, in a historical series of 69 years of daily minimum temperature data for the months from April to September. The distribution of extreme values showed the best fit to the observed data. Camargo et al. (1990), using the model of extreme values for annual absolute minimum temperatures in several locations in the State of São Paulo and Mato Grosso do Sul, identified areas of frost risk.

Through suitable probabilistic or stochastic models based on historical series, the risk levels of absolute minimum temperatures and frosts can be estimated in different periods of the year. Some studies use empirical classification, employing the relative frequency of occurrence of minimum temperatures, to calculate the unconditional probabilities. According to Soares and Dias (1986), such method has the problem of sample size, which may be insufficient to obtain stable probability values. Conrad and Pollak (1950) suggest series of at least 30 years to achieve reliable results.

In a study of the occurrence of low temperatures in Campinas (SP), for the period from 1890 to 1975, Camargo (1977) considered temperatures under a meteorological shelter below 2.5 °C as typical of frost. Ortolani et al. (1981) showed the number of occurrences of minimum temperatures lower than 2 °C, under shelter, from 1962 to 1980, for eight localities of São Paulo. The use of 2 °C as a threshold was based on the average difference between the air temperature in the meteorological shelter and the grass temperature in frost nights, which is 5.6 °C (Fagnani and Pinto, 1981). Considering the air temperature of 2 °C, the leaf has a temperature of -3.6 °C, approaching the values found by Camargo and Salati (1967), Pinto et al. (1977), and Pinto et al. (1978) as the limit for the appearance of damage in coffee trees. Using a relative frequency, with a series of twenty years of data, Soares and Dias (1986) defined for the city of São Paulo the probability of occurrence of daily minimum temperatures, lower than 10 and 15 °C.

The maximum likelihood comprises the primary method to estimate the parameters of probabilistic models. This estimator has the four desirable properties of a good estimator: unbiasedness, consistency, efficiency, and sufficiency, which must satisfy the condition $\alpha > 0$ (by definition) (Thom, 1958; Thom, 1966; Murteira et al., 2001; Catalunha et al., 2002; Bussab & Moretin, 2017; Casella & Berger, 2018).

This work aimed to verify the goodness of fit of average monthly temperature series in the municipality of Mossoró, northeastern Brazil, to the following distributions probabilities: Normal, Log-Normal, Beta, Gamma, Log-Pearson (Type III), Gumbel, and Weibull. For this, we used the criteria of the Kolmogorov-Smirnov, Chi-square, Cramer-von Mises, Anderson-Darling, Kuiper, and Logarithm of Maximum Likelihood, given that the adjusted modeling is most appropriate for infinite populations of the historical series, thus providing relevant information for agriculture in terms of predicting climatic risk, estimating climatic event values and applying agricultural zoning throughout the municipality of Mossoró, RN, Brazil.

2. Material and Methods

Data of average monthly temperature (°C) were used in a historical series of 38 years (1970 to 2007) of data records of the UFERSA (Federal Rural University of Semi-Arid) Meteorological Station, in Mossoró (5°11' S-37°20' W; 18 m of altitude). The region has a mean annual temperature of 27.5 °C and 68.9% of relative humidity (Carmo Filho et al., 1991). According to Köppen's climatic classification, the climate of Mossoró is of the type BShw', that is, hot and dry.

The following probability density function models were fitted: Normal, Log-Normal, Beta, Gamma, Gumbel, Weibull, and Log-Pearson type III (Campos, 1983; Cooke, 1993).

According to Morettin and Bussab (2004), and Assis et al. (1996), the Normal (or Gaussian) distribution serves as a model for many real-life issues. It also appears in many theoretical investigations since many statistical techniques, such as analysis of variance and regression, assume or require data normality.

In the Log-Normal distribution, the logarithms of the random variables are normally distributed. In this work, we used the three-parameter distribution.

The Gamma probability function has two parameters, the shape (α) and the rate (β). Thom (1958), quoted by Miller and Weaver (1968), stated that for β values greater than or equal to 100, the gamma distribution approaches the normal distribution. The rate parameter (β) indicates the degree of dispersion between the data of a series studied.

The Weibull distribution is used in hydrological analyses for extreme events. However, its use in historical series of climatic and biological variables is still little known (Catalunha et al., 2002). The Gumbel probability density distribution is another model widely used in the representation of climatic data, such as for solar radiation and temperature (Assis et al., 2004).

The Log-Pearson Type III distribution is used to represent mean and extreme variables (Sansigolo, 2008).

When fitting a series of data to a probability density distribution, we work with the hypothesis that the distribution suits well that dataset. Some non-parametric analyses provide criteria for testing this hypothesis. We used the following tests assess goodness of fit: Chi-square (χ^2), Kolmogorov-Smirnov, Cramer-von Mises, Kuiper, Anderson-Darling, and the Logarithm of Maximum Likelihood (Worley et al., 1990; Cooke, 1993). These statistics have the power to discriminate fitting in which the other tests fail to point out (Shapiro & Brain, 1981; Cooke, 1993; Campos, 1983; Siegel, 2006).

Some ways of estimating the parameters of the Gamma distribution have been developed, contributing, along with its flexibility of shapes, to its use in several areas. However, the Maximum Likelihood is the primary method for estimating parameters, because its estimators have the four desirable properties of a good estimator: unbiasedness, consistency, efficiency, and sufficiency, which must satisfy the condition $\alpha > 0$ (by definition) (Thom, 1958; Catalunha et al., 2002; Bussab & Morettin, 2017; Casella & Berger, 2018).

3. Results and Discussion

Seven models of probability distribution Normal, Log-Normal, Beta, Gamma, Log-Pearson (Type III), Gumbel and Weibull were fitted using the Kolmogorov-Smirnov, Qui-square, Cramer Von-Mises, Anderson Darling, Küiper and the Maximum Likelihood Logarithm according to Campos (1983), and Cooke et al. (1993), whose results indicated a good fit to these probability distribution functions, most of which were tested with a p 0.10 value. Similar results were obtained by Araújo et al. (2010), and Assis et al. (2004). Considering a significance level of 0.05, that is, with a probability of type I error of 0.05 (the incorrect rejection of a true null hypothesis), few datasets fitted the distributions. While using the Logarithm of Maximum Likelihood, none data series fitted the proposed models (Tables 1, 2, 3 and 4). These result occurred probably due to the little amount of data in the historical series, since the larger the historical series the higher the value of the statistic, i.e., the higher the logarithm of the maximum likelihood function, the better the goodness of fit obtained (Shapiro & Brain, 1981; Cooke, 1993; Sansigolo, 2008). According to Campos (1983), the Kolmogorov-Smirnov test is more suitable for extensive series. Moreover, this test dismisses the grouping of data in class intervals, as in the Chi-squared. So in our study, the Kolmogorov-Smirnov test is the most suitable. Thus, we may infer that among the seven distributions tested, the parameters of any of them could be used to represent the behavior mean fluctuation of the monthly temperature. However, the estimation of parameters of these distributions and the estimation of the probabilities differ in the degree of difficulty. Therefore, it is convenient to verify which of the functions studied has the best fit, which may be the one with less difficulty in obtaining the parameters and still easier in the estimates of probabilities (Cargnelutti Filho et al., 2004).

Table 1. Relative frequency of the number of fit of historical series of monthly average temperature to seven models of the distribution of probability density, Mossoró, northeastern Brazil

| Month | Probability Distributions | | | | | | | | | | | | | |
|-------|---------------------------|--------|------------|--------|------|--------|-------|--------|----------------------|--------|--------|--------|---------|--------|
| | Normal | | Log-Normal | | Beta | | Gamma | | Log-Pearson Type III | | Gumbel | | Weibull | |
| Jan | 5 | 11.11% | 5 | 11.11% | 4 | 10.00% | 4 | 9.09% | 5 | 10.87% | 4 | 8.70% | 5 | 11.36% |
| Feb | 4 | 8.89% | 4 | 8.89% | 4 | 10.00% | 4 | 9.09% | 5 | 10.87% | 4 | 8.70% | 4 | 9.09% |
| Mar | 4 | 8.89% | 4 | 8.89% | 4 | 10.00% | 4 | 9.09% | 4 | 8.70% | 4 | 8.70% | 3 | 6.82% |
| Apr | 3 | 6.67% | 3 | 6.67% | 3 | 7.50% | 3 | 6.82% | 4 | 8.70% | 3 | 6.52% | 3 | 6.82% |
| May | 3 | 6.67% | 3 | 6.67% | 3 | 7.50% | 3 | 6.82% | 4 | 8.70% | 4 | 8.70% | 4 | 9.09% |
| Jun | 3 | 6.67% | 3 | 6.67% | 3 | 7.50% | 3 | 6.82% | 3 | 6.52% | 3 | 6.52% | 3 | 6.82% |
| Jul | 5 | 11.11% | 5 | 11.11% | 3 | 7.50% | 5 | 11.36% | 4 | 8.70% | 5 | 10.87% | 5 | 11.36% |
| Aug | 4 | 8.89% | 5 | 11.11% | 4 | 10.00% | 5 | 11.36% | 5 | 10.87% | 5 | 10.87% | 4 | 9.09% |
| Sep | 4 | 8.89% | 4 | 8.89% | 3 | 7.50% | 4 | 9.09% | 3 | 6.52% | 4 | 8.70% | 3 | 6.82% |
| Oct | 3 | 6.67% | 3 | 6.67% | 3 | 7.50% | 3 | 6.82% | 3 | 6.52% | 3 | 6.52% | 3 | 6.82% |
| Nov | 4 | 8.89% | 3 | 6.67% | 3 | 7.50% | 3 | 6.82% | 3 | 6.52% | 4 | 8.70% | 4 | 9.09% |
| Dec | 3 | 6.67% | 3 | 6.67% | 3 | 7.50% | 3 | 6.82% | 3 | 6.52% | 3 | 6.52% | 3 | 6.82% |
| Total | 45 | 100% | 45 | 100% | 40 | 100% | 44 | 100% | 46 | 100% | 46 | 100% | 44 | 100% |

Table 2. Absolute frequency of the number of fit of historical series of monthly average temperature to seven models of the distribution of probability density, Mossoró, northeastern Brazil

| Probability Distributions | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|---------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Normal | 5 | 4 | 4 | 3 | 3 | 3 | 5 | 4 | 4 | 3 | 4 | 3 |
| Log-Normal | 5 | 4 | 4 | 3 | 3 | 3 | 5 | 5 | 4 | 3 | 3 | 3 |
| Beta | 4 | 4 | 4 | 3 | 3 | 3 | 3 | 4 | 3 | 3 | 3 | 3 |
| Gamma | 4 | 4 | 4 | 3 | 3 | 3 | 5 | 5 | 4 | 3 | 3 | 3 |
| Log-Pearson III | 5 | 5 | 4 | 4 | 4 | 3 | 4 | 5 | 3 | 3 | 3 | 3 |
| Gumbel | 4 | 4 | 4 | 3 | 4 | 3 | 5 | 5 | 4 | 3 | 4 | 3 |
| Weibull | 5 | 4 | 3 | 3 | 4 | 3 | 5 | 4 | 3 | 3 | 4 | 3 |
| Total | 32 | 29 | 27 | 22 | 24 | 21 | 32 | 32 | 25 | 21 | 24 | 21 |

Table 3. P-values of the Chi-square and Kolmogorov-Smirnov tests for the fitting of seven distributions of probability density to the monthly average temperature series. Mossoró, Northeastern Brazil

| Month | Probability distributions and respective P-values | | | | | | | | | | | | | |
|-------|---|----------|------------|----------|-------|----------|-------|----------|-----------------|----------|--------|----------|---------|----------|
| | Normal | | Log-Normal | | Beta | | Gamma | | Log-Pearson III | | Gumbel | | Weibull | |
| | KS | χ^2 | KS | χ^2 | KS | χ^2 | KS | χ^2 | KS | χ^2 | KS | χ^2 | KS | χ^2 |
| Jan | 0.176 | 0.107 | 0.193 | 0.107 | 0.183 | 0.107 | 0.171 | 0.107 | 0.164 | 0.107 | 0.184 | 0.044 | 0.202 | 0.491 |
| Feb | 0.143 | 0.864 | 0.136 | 0.864 | 0.264 | 0.077 | 0.136 | 0.864 | 0.154 | 0.864 | 0.173 | 0.024 | 0.172 | 0.024 |
| Mar | 0.118 | 0.888 | 0.114 | 0.888 | 0.225 | 0.000 | 0.107 | 0.888 | 0.110 | 0.888 | 0.149 | 0.021 | 0.231 | 0.013 |
| Apr | 0.244 | 0.000 | 0.223 | 0.000 | 0.415 | 0.000 | 0.234 | 0.000 | 0.307 | 0.000 | 0.235 | 0.000 | 0.271 | 0.000 |
| May | 0.207 | 0.038 | 0.197 | 0.038 | 0.219 | 0.003 | 0.203 | 0.038 | 0.154 | 0.038 | 0.148 | 0.038 | 0.287 | 0.001 |
| Jun | 0.226 | 0.005 | 0.242 | 0.005 | 0.368 | 0.005 | 0.223 | 0.005 | 0.177 | 0.005 | 0.176 | 0.005 | 0.265 | 0.000 |
| Jul | 0.154 | 0.185 | 0.166 | 0.185 | 0.166 | 0.005 | 0.149 | 0.185 | 0.181 | 0.185 | 0.157 | 0.185 | 0.142 | 0.185 |
| Aug | 0.135 | 0.060 | 0.132 | 0.602 | 0.184 | 0.241 | 0.125 | 0.602 | 0.151 | 0.660 | 0.171 | 0.491 | 0.136 | 0.060 |
| Sep | 0.215 | 0.269 | 0.216 | 0.269 | 0.314 | 0.001 | 0.205 | 0.269 | 0.263 | 0.001 | 0.226 | 0.269 | 0.229 | 0.000 |
| Oct | 0.255 | 0.000 | 0.279 | 0.000 | 0.381 | 0.000 | 0.256 | 0.000 | 0.217 | 0.000 | 0.224 | 0.000 | 0.369 | 0.000 |
| Nov | 0.160 | 0.080 | 0.176 | 0.004 | 0.203 | 0.075 | 0.164 | 0.075 | 0.208 | 0.075 | 0.193 | 0.075 | 0.195 | 0.075 |
| Dec | 0.228 | 0.001 | 0.220 | 0.005 | 0.223 | 0.001 | 0.222 | 0.005 | 0.231 | 0.005 | 0.259 | 0.005 | 0.264 | 0.005 |

Table 4. Classification of the fitting according to the Chi-square (χ^2) and Kolmogorov-Smirnov (KS) tests to seven models of probability density distribution of the monthly average temperature series. Mossoró, northeastern Brazil

| Month | Probability distributions | | | | | | | | | | | | | |
|-------|---------------------------|----------|------------|----------|------|----------|-------|----------|-----------------|----------|--------|----------|---------|----------|
| | Normal | | Log-Normal | | Beta | | Gamma | | Log-Pearson III | | Gumbel | | Weibull | |
| | KS | χ^2 | KS | χ^2 | KS | χ^2 | KS | χ^2 | KS | χ^2 | KS | χ^2 | KS | χ^2 |
| Jan | Aj ¹ | Aj | Aj | Aj | Aj | Aj | Aj | Aj | Aj | Aj | Aj | - | Aj | Aj |
| Feb | Aj | Aj | Aj | Aj | Aj | - | Aj | Aj | Aj | Aj | Aj | - | Aj | - |
| Mar | Aj | Aj | Aj | Aj | Aj | - | Aj | Aj | Aj | Aj | Aj | - | Aj | - |
| Apr | Aj | - | Aj | - | Aj | - | Aj | - | Aj | - | Aj | - | Aj | - |
| May | Aj | - | Aj | - | Aj | - | Aj | - | Aj | - | Aj | - | Aj | - |
| Jun | Aj | - | Aj | - | Aj | - | Aj | - | Aj | - | Aj | - | Aj | - |
| Jul | Aj | Aj | Aj | Aj | Aj | - | Aj | Aj | Aj | Aj | Aj | Aj | Aj | Aj |
| Aug | Aj | - | Aj | Aj | Aj | Aj | Aj | Aj | Aj | Aj | Aj | Aj | Aj | - |
| Sep | Aj | Aj | Aj | Aj | Aj | - | Aj | Aj | Aj | - | Aj | Aj | Aj | - |
| Oct | Aj | - | Aj | - | Aj | - | Aj | - | Aj | - | Aj | - | Aj | - |
| Nov | Aj | - | Aj | - | Aj | - | Aj | - | Aj | - | Aj | - | Aj | - |
| Dec | Aj | - | Aj | - | Aj | - | Aj | - | Aj | - | Aj | - | Aj | - |

Note. ¹AJ = Adjusted distribution; - = Unadjusted distribution (significant at 10% probability). 5% ≤ p value ≤ 10%.

Among the seven distributions tested under a p-value of 0.10, the Normal distribution showed the highest number of fitting. It was evident through the Kolmogorov-Smirnov goodness of fit test, by estimating the parameters of this distribution (mean and standard deviation) for probability calculations within desired range intervals (Sansigolo, 2008). Araújo et al. (2010a, 2010b, 2011), and Almeida et al. (2010) obtained similar results. On the other hand, Araújo et al. (2010b, 2011) obtained temperature adjustments with the Gumbel, Normal, and Log-Normal distributions respectively.

It was verified, then, that the normal distribution was made in the average scale of 25, 50, 75 and 95% of trend of occurrence (Table 5), Percentage of equal or lesser degree of equal or less than those calculated, according to Table 5, there is a 95% probability that the value of the monthly average for the month of January is not higher than 34.8 °C. Such value can be interpreted as if it were every twenty years, eighteen presenting a value less than or equal to 34.8 °C.

Table 5. Different levels of cumulative probability of occurrence of the average monthly air temperature, quantile value of the distribution and return period according to the Normal or Gaussian distribution for Mossoró, RN, Brazil

| Monthly scale for all months of the year | P(X ≤ a) = P(x); Levels of probability (%) | | | |
|--|--|------|------|------|
| | 25 | 50 | 75 | 95 |
| Jan | 26.5 | 29.3 | 31.6 | 34.8 |
| Fev | 24.2 | 25.8 | 27.4 | 29.6 |
| Mar | 24.5 | 26.8 | 29.1 | 32.5 |
| Abr | 24.9 | 26.7 | 28.5 | 31.5 |
| Mai | 25.8 | 27.5 | 29.2 | 31.5 |
| Jun | 27.8 | 25.0 | 27.8 | 32.2 |
| Jul | 24.7 | 27.0 | 29.3 | 32.7 |
| Ago | 25.8 | 28.6 | 31.4 | 35.4 |
| Set | 25.0 | 26.9 | 28.8 | 31.5 |
| Out | 27.2 | 29.4 | 31.6 | 34.8 |
| Nov | 25.2 | 27.0 | 29.5 | 32.6 |
| Dez | 26.8 | 28.8 | 30.8 | 33.6 |
| Return period (in years) | 1.3 | 2.0 | 4.0 | 20.0 |

Note. Return period or recurrence time (T): Average time measured in years when a given atmospheric event must be matched or exceeded at least once.

Since the reason for a precise diet is known, the choice of a favorable area can be made, that is, the average control of annual and diurnal temperature differences are almost all parts of the world. A soil surface, with or without vegetation, is the main receiver of solar radiation and atmospheric energy, and is also a radiation emitter. The network balance, which is variable throughout the day and year, promotes daily and actions in soil and air temperatures. Among the climatic elements, the propagation temperature of the species is great and the physiological processes occurring in animals and mainly plants (Valeriano & Picini, 2000). Growth and plant species are strongly influenced by this element, since agricultural activity is closely related to temperatures at the air temperature (Valeriano & Picini, 2000). Temperature is essential for climate and basic information for climatic zoning of all crops, as well as for characterization of local climates. The diversity index is directly linked to the temperature of a given region, for example, when rainfall occurs differently but is negatively influenced by agricultural production in the region. By means of the climatic water balance, with parameters of correctness, temperature and evapotranspiration, an available effect can be estimated in the soil and, therefore, the potential of the region for the agricultural crops (Hillel, 2003).

In medium-sized cities, such as Mossoró-RN, the densification caused by the verticalization, depending on the locality, may imply mesoclimatic changes, mainly, in relation to the island heat phenomenon.

As one of the natural components that undergoes more changes thanks to the process of urbanization is the climate, and Mossoró has peculiar climatic characteristics, as high annual temperatures to be located in an area of tropical climate equatorial also classified of semiarid does if necessary an investigation relating the local urban growth and the reflections of the thermal field (Mendonça & Dani-Oliveira, 2007).

According to Saraiva (2010) in his study on the climate of Mossoró, it has been proven that the city generates its own climate, resulting from the interference of all factors (little vegetation, large flow of motor vehicles, constructions etc.) that are processed on the urban boundary layer and that act to change the climate on a local scale. In the semi-arid region, this configuration becomes even more pronounced, since the high incidence of solar radiation causes in its cities temperatures quite high.

4. Conclusions

The average monthly temperature data of Mossoró, RN, Brazil, fitted the Normal, Log-Normal, Beta, Gamma, Log-Pearson (Type III), Gumbel, and Weibull probability density distributions. The Normal distribution showed the best goodness of fit. The Kolmogorov-Smirnov, Chi-squared, Cramer-von Mises, Anderson-Darling, and Kuiper fitting criteria were similar, and in this case, all can be used as indicators for model ranking. The Logarithm of Maximum Likelihood was a poor indicator for fitting. The Normal probability distribution performed well in estimating the occurrence of the average monthly air temperature. The construction of probability tables is a valuable tool in the study of the thermal behavior of this region. The city of Mossoró has a peculiar climate, predominating a great part of the year a dry season, presenting high temperatures, and another part of the year (autumn) that presents/displays cooler temperatures in the rainy period. The behavior of mean air temperature did not vary significantly over the 38-year series. No change was observed in the temperature values, which continue with the mean of 27.6 °C.

References

- Araujo, E. M., Oliveira, J. B. M. G. S., Viana, P. C., & Alves, A. S. (2010b). Análise da aderência de distribuições de probabilidade aos dados de temperatura máxima e mínima do ar em Iguatu-Ce. *Revista Caatinga*, 23(3), 104-109.
- Araújo, E. M., Silva, I. N., Oliveira J. B., Cavalcante Junior E. G., & Almeida, B. M. (2010a). Aplicação de seis distribuições de probabilidade a séries de temperatura máxima em Iguatu-Ce. *Revista Ciências Agrônômica*, 41(1), 36-45.
- Araújo, E. M., Silva, I. N., Oliveira, J. B., Cavalcante Junior, E. G., & Almeida, B. M. (2011). Probabilidades quinzenais de ocorrência de temperaturas máximas do ar na cidade de Iguatu-Ce. *Revista Brasileira de Agricultura Irrigada*, 5(1), 48-53. <https://doi.org/10.7127/rbai.v5n100039>
- Arruda, H. V., Pinto, H. S., & Penteadó, R. S. (1981). Modelos probabilísticos para a interpretação da ocorrência de temperaturas mínimas na região de Campinas-SP. *Sociedade Brasileira de Agrometeorologia*, 2, 143-145.
- Assis, F. N., Arruda, H. V., & Perreira, A. R. (1996). *Aplicações de estatística à climatologia: Teoria e prática* (p. 161). Pelotas: Ed. Universitária UFPel.

- Assis, J. P., Neto, D. D., Manfron, P. A., Martin, T. N., Sparovek, G., & Tim M. T. C. (2004). Ajuste de séries históricas de temperatura e radiação solar global diária às funções densidade de probabilidade normal e log-normal, em Piracicaba, SP. *Revista Brasileira de Agrometeorologia*, 12(1), 113-121.
- Assis, J. P., Sousa, R. P., Maracajá, P. B., Linhares, P. C. F., Medeiros, A. C., Moreira, A. R., Martins, D. J. B., & Almeida, T. H. S. (2017). Probability distributions of wind speed historical data series in Mossoro, Northeastern Brazil. *International Journal of Development Research*, 7(12), 17503-17508.
- Berlato, M. A. (1987). Modelo de relação entre o rendimento de grãos de soja e o déficit hídrico para o Estado do Rio Grande do Sul (93 f. Tese (Doutorado), Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP).
- Botelho, V. A., & Morais, A. R. (1999). Estimativas dos parâmetros da distribuição gama de dados pluviométricos do Município de Lavras, Estado de Minas Gerais. *Ciência e Agrotecnologia*, 23(1), 697-706.
- Buriol, G. A., Heldwein, A., Estefanel, B., Matzenauer, V., & Marcon, R. I. A. (2001). Probabilidade de ocorrência de valores de radiação solar prejudiciais ao cultivo do pepineiro na região do baixo vale do Taquari, RS. *Pesquisa Agropecuária Gaúcha*, 7(2), 97-104.
- Bussab, W. O., & Morettin, P. A. (2017). *Basic statistics* (9th ed.). Saraiva, São Paulo. Retrieved from <https://www.ime.usp.br/~pam/scriptsR.html>
- Camargo, A. P. (1977). *Frequência das geadas excepcionais como as de julho de 1975* (pp. 1-3). Seminário Internacional sobre Climatologia no Hemisfério Sul, Campinas.
- Camargo, A. R., & Salati, E. (1967). Determinación de la temperatura letal para hojas de café en noches de heladas. *Café*, 8(3), 12-15.
- Camargo, M. B. P., Pedro Junior, M. J., Alfonsi, R. R., Ortolani, A. A., Brunini, O., & Chiavegato, O. M. P. (1990). Guia agrometeorológico do agricultor irrigante: n°4. *Probabilidade de ocorrência de geadas nos Estados de São Paulo e Mato Grosso do Sul* (Boletim técnico, 136, p. 10). Campinas, Instituto Agronômico. Retrieved from http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0006-87051993000200008
- Camargo, M. B. P., Pedro Júnior, M. J., Alfonsi, R. R., Ortolani, A. A., & Brunini, O. (1993). Probabilidade de ocorrência de temperaturas mínimas absolutas mensais e anual no estado de São Paulo. *Bragantia*, 52(2), 161-168. <https://doi.org/10.1590/S0006-87051993000200008>
- Campos, H. (1983). *Non-parametric statistics* (4th ed., p. 49). Piracicaba: ESALQ/USP.
- Cargnelutti Filho, A., Matzenauer, R., & Trindad, J. K. (2004). Ajustes de funções de distribuição de probabilidade à radiação solar global no Estado do Rio Grande do Sul. *Pesquisa Agropecuária Brasileira*, 39(12), 1157-1166. <https://doi.org/10.1590/S0100-204X2004001200001>
- Carmo-Filho, F., Sobrinho, J. E., & Maia Neto J. M. (1991). *Dados climatológicos de Mossoró: Um município semi-árido nordestino* (p. 121). Mossoró, RN: UFERSA.
- Casella, G., Berger, R. L. (2018). *Statistical Inference* (2nd ed., p. 612). Cengage Learning, São Paulo.
- Catalunha, M. J., Sedyama, G. C., Leal, B. G., Soares, C. P. B., & Ribeiro, A., (2002). Aplicação de cinco funções densidade de probabilidade a séries de precipitação pluvial no Estado de Minas Gerais. *Revista Brasileira de Agrometeorologia*, 10(1), 153-162.
- Conrad, V., & Pollak, L. W. (1950). *Methods in climatology* (p. 459). Cambridge, Harvard University Press. Retrieved from <http://www.hup.harvard.edu/catalog.php?isbn=9780674187856>
- Cooke, R. A. (1993). VTFIT: A routine for fitting homogenous Probability density functions. *User documentation* (p. 21). Blacksburg: Department of agricultural Engineering. Virginia, Polytecnic Institute. Retrieved from citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.496...
- Cooke, R. A., Mostaghimi, S., & Woeste, F. E. (1993). VTFIT: A microcomputer-based routine for fitting probability distribution functions to data. *American Society of Agricultural Engineers*, 9(4), 401-408.
- Dallacort, R., Martins, J. A., Inoue, M. H., Freitas, P. S. L., & Coletti, A. J. (2011). Distribuição das chuvas no município de Tangará da Serra, médio norte do Estado de Mato Grosso, Brasil. *Acta Scientiarum Agronomy*, 33(2), 193-200. <https://doi.org/10.4025/actasciagron.v33i2.5838>
- De Camargo, M. B. P., Pedro Júnior, M. J., Alfonsi, R. R., Ortolani, A. A., & Brunini, O. (1993). Probabilidade de ocorrência de temperaturas mínimas absolutas mensais e anual no estado de São Paulo. *Bragantia*, 52(2), 161-168. <https://doi.org/10.1590/S0006-87051993000200008>

- Fagnani, M. A., & Pinto H. S. (1981). Simulação de temperaturas de folhas de cafeeiros em noites sujeitas a geadas de irradiação. *Sociedade Brasileira de Agrometeorologia*, 139-142.
- Hillel, D. (2003). *Introduction of environmental soil physics* (p. 494). New York: United Kingdom. <https://doi.org/10.1016/B978-012348655-4/50015-0>
- Kist, A., & Virgens Filho, J. S. (2015). Análise probabilística da distribuição de dados diários de chuva no estado do Paraná. *Revista Ambiental Água*, 10(1), 172-181.
- Lima, R. B., Oliveira, C. P., Silva, D. A., Sousa, C. S. C., & Souza, R. N. (2015). *Modelagem probabilística da precipitação pluviométrica para o estado de Pernambuco*. Congresso Técnico Científico da Engenharia e da Agronomia CONTECC. Centro de Eventos do Ceará-Fortaleza-CE 15 a 18 de Setembro de 2015.
- Mendonça, F., & Danni-Oliveira, I. M. (2007). *Climatologia: Noções básicas e climas do Brasil* (p. 158). São Paulo: Oficina de Texto.
- Miller, M. E., & Weaver, C. R. (1968). Monthly and annual precipitation probabilities for climatic divisions in Ohio. *Research Bulletin*, 1005, 11.
- Morettin, P. A., & Bussab, W. O. (2004). *Estatística básica* (p. 526). São Paulo: Saraiva. Retrieved from <https://www.saraiva.com.br/estatistica-basica-8-ed-2013-4982219.html>
- Mota, F. S. (1981). *Agricultural Meteorology* (5th ed., p. 376). São Paulo, Nobel. Retrieved from <https://www.bdpa.cnptia.embrapa.br/.../busca?...%22MOTA...S...%22MOTA...S...>
- Mota, F. S., Rosskoff, J. L. C., & Silva, J. B. (1999). Probabilidade de ocorrência de dias com temperaturas iguais ou superiores a 35 °C no florescimento de arroz no Rio Grande do Sul. *Revista Brasileira de Agrometeorologia*, 7(2), 147-149.
- Murta, R. M., Teodoro, S. M., Bonom, O. P., & Chaves, M. A. (2005). Precipitação mensal em níveis de probabilidade pela distribuição gama para duas localidades do sudoeste da Bahia. *Ciências Agrotecnologia*, 29(5), 988-994. <https://doi.org/10.1590/S1413-70542005000500011>
- Murteira, B., Ribeiro, C. A., Silva, J. A., & Pimenta, C. (2001). *Introdução à estatística* (p. 688). McGraw-Hill, Portugal. Retrieved from <https://www.fnac.pt/Introducao-a-Estatistica-MURTEIRA...E-RIBEIRO.../a318157>
- Ortolani, A. A., Camargo, M. B. P., & Angelloci, L. R. (1981). Frequência de geadas no Estado de São Paulo. *Sociedade Brasileira de Agrometeorologia*, 146-148. Retrieved from http://www.scielo.br/scielo.php?script=sci_nlinks&ref=000035&pid=S0006...Ing...
- Pedrosa, A. C., & Gama, S. M. A. (2004). *Introdução computacional à probabilidade e estatística* (p. 608). Porto Editora Ltda, Porto, Portugal. Retrieved from <http://www.portoeditora.pt>
- Pinto, H. S., Camargo, M. B. P., Nobrega, M., & Ortolani, A. A. (1978). A Um sistema computacional para levantamento de prejuízos causados por geadas à cafeicultura do Estado de São Paulo. *IBC-GERCA*, 145-147.
- Pinto, H. S., Tarifa, J. R., Alfonsi, R. R., & Pedro Júnior, M. J. (1977). Estimation of frost damage in coffee trees in the State of São Paulo, Brazil. *American Meteorological Society*, 37-38.
- Prela-Pantano, A., Rolin, G. S., & Camargo, M. B. P. (2009). Probabilidade de ocorrência de temperaturas mínimas menores que 5 °C na região do médio Paranapanema. *Bragantia*, 68(1), 279-284. <https://doi.org/10.1590/S0006-87052009000100030>
- Rodrigues, J. A., Silva, A. P. C. M., & Santos Filho, J. (2014). Uso de distribuições de probabilidade na modelagem de intensidade de secas ocorridas em laranjeiras do sul. *Revista Brasileira Biometria*, 32(4), 570-583.
- Sampaio, S. C., Corrêa, M. M., Souza, M. R., Guimarães, J. C., & Silva, A. M. (1999). Precipitação provável para o município de Lavras/MG, utilizando a distribuição Log-Normal. *Ciência e Agrotecnologia*, 23(2), 382-389.
- Sansigolo, C. A. (2008). Distribuição de extremos de precipitação diária, temperatura máxima e mínima e velocidade do vento em Piracicaba. *Revista Brasileira de Meteorologia*, 23(3), 341-346. <https://doi.org/10.1590/S0102-77862008000300009>

- Saraiva, A. L. B. C. (2014). *O clima urbano de Mossoró (RN): O subsistema termodinâmico* (234 f. Dissertação (Mestrado em Geografia), Universidade do Espírito Santo, Vitória, ES). Retrieved from http://portais4.ufes.br/posgrad/teses/tese_7595_analuiza.pdf
- Shapiro, S. S., & Brain, C. W. (1981). A review of distributional testing procedure and development of a censored sample distributional test. *Statistical distribution in scientific work* (V. 5). Dordrecht: D. Reidel Publishing Co. https://doi.org/10.1007/978-94-009-8552-0_1
- Siegel, S. (2006). *Non-parametric statistics for the behavioral sciences* (2nd ed., p. 448).
- Silva, I., Castro Neto, P., & Silveira, J. V. (1986). Época e probabilidade de ocorrência de temperaturas mínimas abaixo de dado valor, para a região de Lavras, MG. *Ciência e Prática*, 10(2), 210-219.
- Soares, J. R., & Dias, M. F. S. (1986). Probabilidade de ocorrência de alguns eventos meteorológicos extremos na cidade de São Paulo. *Revista Brasileira de Meteorologia*, 2(1), 67-75.
- Thom, H. C. S. (1958). A note on the gama distribution. *Monthly Weather Review*, 86(2), 117-122. [https://doi.org/10.1175/1520-0493\(1958\)086<0117:ANOTGD>2.0.CO;2](https://doi.org/10.1175/1520-0493(1958)086<0117:ANOTGD>2.0.CO;2)
- Thom, H. C. S. (1966). *Some methods of climatological analysis* (Technical Notes, 81, p. 50). Roma: FAO. Retrieved from <https://www.worldcat.org/.../some-methods-of-climatological-anal...>
- Valeriano, M. M., & Picini, A. G. (2000). Uso de Sistema de Informações Geográfica para a geração de mapas de médias mensais de temperatura do Estado de São Paulo. *Revista Brasileira de Agrometeorologia, Santa Maria*, 8(2), 255-262.
- Worley, J. W., Bollinger, J. A., Woeste, F. E., & Kline, K. S. (1990). Graphic distribution analysis (GDA). *American Society of Agricultural Engineers*, 6(3), 367-371.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).