

ADEQUAÇÃO DE MODELOS PROBABILÍSTICOS À EVAPOTRANSPIRAÇÃO DE REFERÊNCIA NO SUBMÉDIO DO VALE DO RIO SÃO FRANCISCO

EDGO JACKSON PINTO SANTIAGO¹; FRANK GOMES-SILVA¹; ANTONIO SAMUEL ALVES DA SILVA¹; JOSÉ RAMON BARROS CANTALICE¹; MOACYR CUNHA FILHO¹ E JOSÉ DOMINGOS ALBUQUERQUE AGUIAR¹

¹ Departamento de Estatística e Informática-DEINFO, Programa de Pós-Graduação em Biometria e Estatística Aplicada-PPGBEA, Universidade Federal Rural de Pernambuco-UFRPE, Rua Dom Manoel de Medeiros s/n, Dois Irmãos, CEP: 52.171.900, Recife, Pernambuco, Brasil. edgoj@hotmail.com, franksinatrags@gmail.com, antonio.sasilva@ufrpe.br, cantalice21@hotmail.com, moacyr2006@gmail.com, aguiar.domingos@gmail.com.

1 RESUMO

A evapotranspiração consiste no processo de perda de água do solo, da planta, e é fundamental para produção vegetal, constituindo uma das principais variáveis agrometeorológicas. Apesar disso, são escassos trabalhos que relacionam adequabilidade de distribuições de probabilidade a dados de evapotranspiração. O objetivo desse trabalho foi testar a aderência de diferentes distribuições de probabilidade a dados de evapotranspiração de referência, selecionando as mais adequadas para este fim. Esse estudo foi realizado com dados de evapotranspiração de referência obtidos pelas estações meteorológicas da Universidade Federal do Vale do São Francisco (UNIVASF) em Petrolina, PE e Juazeiro, BA. Foram ajustadas as distribuições Gama, Weibull, Log-Normal, Beta, Exponencial, Log-Logística e Log-Logística Exponenciada. Os maiores p-valores foram obtidos para as distribuições Log-Logística e Log-Logística Exponenciada, possivelmente devido à leve assimetria positiva destas aos dados de evapotranspiração. Pelo teste da razão de verossimilhanças, a distribuição Log-Logística Exponenciada adequou-se mais aos meses de janeiro, agosto e dezembro em Juazeiro e Petrolina, somando-se a esta última o mês de novembro. As distribuições Log-Logística e Log-Logística Exponenciada foram as mais adequadas para modelar a evapotranspiração. A partir dessas distribuições, foram estimados valores de evapotranspiração para diferentes níveis de probabilidade, sendo janeiro o mês com maior demanda hídrica provável.

Palavras-chave: transpiração, evaporação, demanda hídrica, distribuição log-logística, irrigação.

SANTIAGO, E. J. P.; GOMES-SILVA, F.; SILVA, A. S. A.; CANTALICE, J. R. B.; CUNHA FILHO, M.; AGUIAR, J. D. A.

ADJUSTMENT OF PROBABILISTIC MODELS TO THE REFERENCE EVAPOTRANSPIRATION IN THE SUB-MEDIUM OF SÃO FRANCISCO RIVER VALLEY

2 ABSTRACT

Evapotranspiration is the process of water loss from soil and plant surfaces, and it is essential for plant production, constituting one of the main agrometeorological variables. Nevertheless,

there are few studies that relate the adequacy of probability distributions to evapotranspiration data. The objective of this work was to test the adherence of different probability distributions to reference evapotranspiration data by selecting the most suitable ones for this purpose. This study was carried out with daily evapotranspiration reference data obtained by the meteorological stations of the Federal University of Vale of São Francisco (UNIVASF) in Petrolina, PE and Juazeiro, BA. The Gamma, Weibull, Log-Normal, Beta, Exponential, Log-Logistics and Exponentiated Log-Logistics distribution were adjusted. The highest p-values were obtained for the Log-Logistics and Exponentiated Log-Logistics distributions. The highest p-values were obtained for the Log-Logistics and Exponentiated Log-Logistics distributions, possibly due to the slight positive asymmetry of those to the evapotranspiration data. By testing the likelihood ratio, the Exponentiated Log-Logistics distribution was more suitable for the months of January, August and December in Juazeiro and Petrolina, adding to the latter the month of November. The Log-Logistics and Exponentiated Log-Logistics distributions were the most suitable to model evapotranspiration. From these distributions, evapotranspiration values were estimated for different levels of probability, with January being the month with the highest probable water demand.

Keywords: transpiration, evaporation, water demand, log-logistics distribution, irrigation.

3 INTRODUCTION

Evapotranspiration is a fundamental phenomenon for crop production and constitutes one of the main variables in agrometeorological research. It is relevant to the hydrological cycle, irrigation management, and water resource management. Evapotranspiration refers to the process of water loss from the soil and plants to the atmosphere and is therefore a highly relevant parameter for the design and management of irrigation systems (OLIVEIRA et al. 2017).

The importance of knowing the amount of water required by crops for the correct planning, sizing, and management of any irrigation system is highlighted by Fernandes, Fraga Júnior, and Takay (2011). Knowledge of the actual water requirements of crops is extremely important, as it impacts the management of water applied through irrigation, aiming at its replenishment (SANTIAGO et al., 2016). Despite this, the variability of meteorological elements during the period of maximum water demand of irrigated crops results in considerable dispersion of estimated

evapotranspiration values, which requires an analysis of the frequency distribution of estimated values for the purpose of sizing irrigation projects. This is because some criteria for sizing irrigation systems consider the probability of evapotranspiration and precipitation occurrence.

Castro and Leopoldo (1995) and Doorenbos and Pruitt (1997) reported that, in most irrigated regions, probability levels are between 75% and 80%. Passos, Raposo, and Mendes (2017) recommend a 75% probability level because it offers greater reliability; however, they state that a 90% probability level can be used for sizing crops that are highly sensitive to water deficits. Importantly, the choice of probability level adopted in sizing an irrigation system, whenever possible, should consider the crop's water requirements and economic value in terms of losses resulting from reduced production due to water deficits or increased costs associated with implementing greater depths, as well as the available water capacity of the soil under which it will be irrigated.

Probability distributions applied to evapotranspiration are still scarce in various

regions of the country. In this context, the beta, normal, log-normal, and gamma distributions have emerged as the most widely used distributions. However, there are still no studies on the suitability of probabilistic models for evapotranspiration data in the Sub-Middle São Francisco River Valley region. Therefore, the objective of this study was to test the suitability of different probability distributions for reference evapotranspiration data, select the most appropriate distributions for this purpose, and estimate probable reference evapotranspiration values for the Petrolina, Pernambuco, Juazeiro, and Bahia regions at different probability levels.

4 MATERIALS AND METHODS

This study was carried out with daily data (2015--2019) of reference evapotranspiration - E_{To} (mm day⁻¹) obtained from the automatic meteorological stations of the Federal University of São Francisco Valley (UNIVASF), Juazeiro campus, BA (latitude: 09°26'56"S,

longitude: 40°31'27"W, altitude: 356 m) and Petrolina, PE (latitude: 09°19'28"S, longitude: 40°33'34"W, altitude: 393 m) located in the submiddle region of the São Francisco River Valley.

The region's climate characteristics reveal average minimum temperatures ranging from 18.4 to 22.2°C and average maximum temperatures ranging from 29.6 to 33.9°C. The lowest temperatures are recorded in July, and the highest temperatures are recorded in November. The region's climate, according to the Köppen classification, is BSwH, making it a semiarid climate. The average annual precipitation is approximately 530 mm, with the rainy season occurring between November and April.

To estimate the reference evapotranspiration values (E_{To}), the Gamma (G), Weibull (W), Log-Normal (LN), Beta (B), Exponential (E), Log-Logistic (LL) and Exponentiated Log-Logistic (LLE) probability distributions were adjusted, whose densities are given as follows:

$$f_G(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}; x > 0 \quad (1)$$

$$f_W(x) = \frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^\alpha}; x > 0 \quad (2)$$

$$f_{LN}(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\log x - \mu}{\sigma}\right)^2}; x > 0 \quad (3)$$

$$f_B(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}; 0 < x < 1 \quad (4)$$

$$f_E(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}}; x \geq 0 \quad (5)$$

$$f_{LL}(x) = \frac{\beta}{\alpha} \left(\frac{x}{\alpha}\right)^{\beta-2} \left(1 + \left(\frac{x}{\alpha}\right)^\beta\right)^{-2}; x > 0 \quad (6)$$

$$f_{LLE}(x) = \frac{\delta\beta}{\alpha} \left(\frac{x}{\alpha}\right)^{\beta-2} \left(1 + \left(\frac{x}{\alpha}\right)^\beta\right)^{-(\delta-1)} \left(1 + \left(\frac{x}{\alpha}\right)^\beta\right)^{-2}; x > 0 \quad (7)$$

In Equations (1)–(7), x represents the reference evapotranspiration (mm) on a daily scale, and $\alpha > 0$, $\beta > 0$, $\delta > 0$, $\sigma > 0$ and $-\infty < \mu < \infty$ are parameters of the respective probabilistic models.

The aforementioned distributions were chosen because they have already been applied to other hydrological phenomena. Parameter estimates for all distributions were made via the maximum likelihood method.

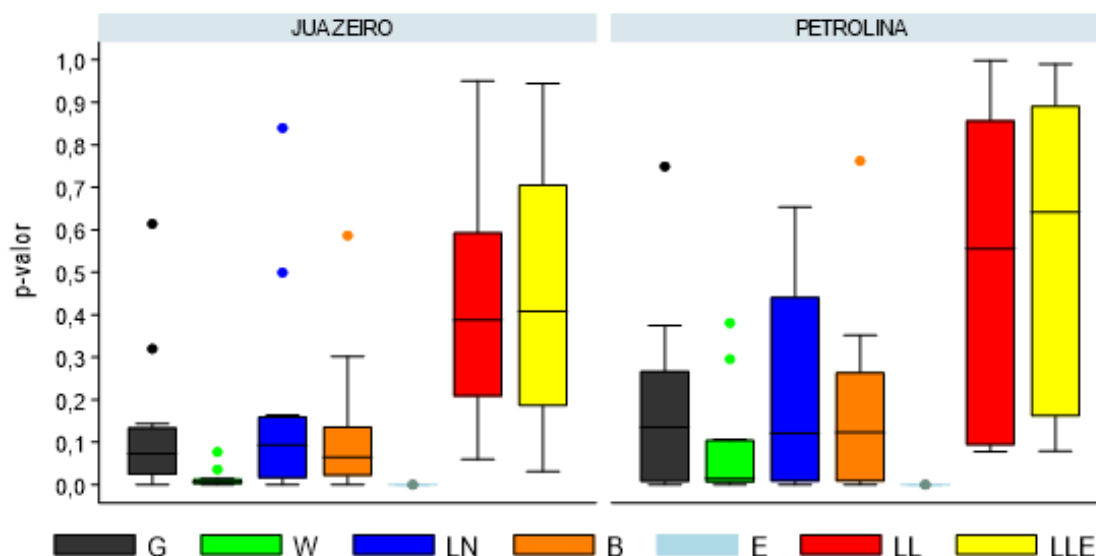
The Kolmogorov–Smirnov goodness-of-fit test was applied at a maximum probability of 5% type I error. The distributions were compared via the p values obtained via the goodness-of-fit test, selecting the most suitable candidate distributions to model the data, those with the highest p values. Furthermore, the Akaike information criterion (AIC) and the Bayesian criterion (BIC) were also applied, as were the likelihood ratio test. Probable reference evapotranspiration estimates were

obtained for probability levels of 10, 20, 30, 40, 50, 60, 70, 80, 90, and 95%, using the distributions that best fit the data according to the highest p value obtained in the goodness-of-fit test.

5 RESULTS AND DISCUSSION

The boxplots of the p values obtained by the goodness-of-fit test, when testing the hypothesis that the proposed theoretical distributions adequately represent the empirical distribution of the data in both cities, are presented in Figure 1. Six of the seven distributions produced a p value capable of ensuring the absence of sufficient evidence to reject the hypothesis tested, i.e., a p value > 0.05 . Among these, LL and LLE stand out for presenting higher mean and median values in both Petrolina and Juazeiro. On the other hand, the E distribution did not adhere to the reference evapotranspiration data.

Figure 1. Boxplot of p values obtained via the Kolmogorov–Smirnov test at a maximum probability of Type I error of 5%. Each box shows the interquartile range; the lines parallel to the x-axis represent the maximum, median, and minimum, respectively; and the balls indicate extreme values.



The good results of the LL and LLE distributions may be due to the slight positive skewness of these distributions in relation to the ETo data (Figure 2). This is because the shape of the LL density function can, like that of the LN, be bell shaped and symmetric, but it can be asymmetric with

thicker grouts. The same can occur with the LLE but with greater flexibility due to the additional parameter (δ). These facts provided adequate representation of the data and yielded p values greater than 0.05, which consequently led to greater acceptance of these distributions.

Figure 2. Reference evapotranspiration frequencies superimposed by the probability density functions studied in the municipalities of Petrolina, PE, and Juazeiro, BA, from 2015-2019.

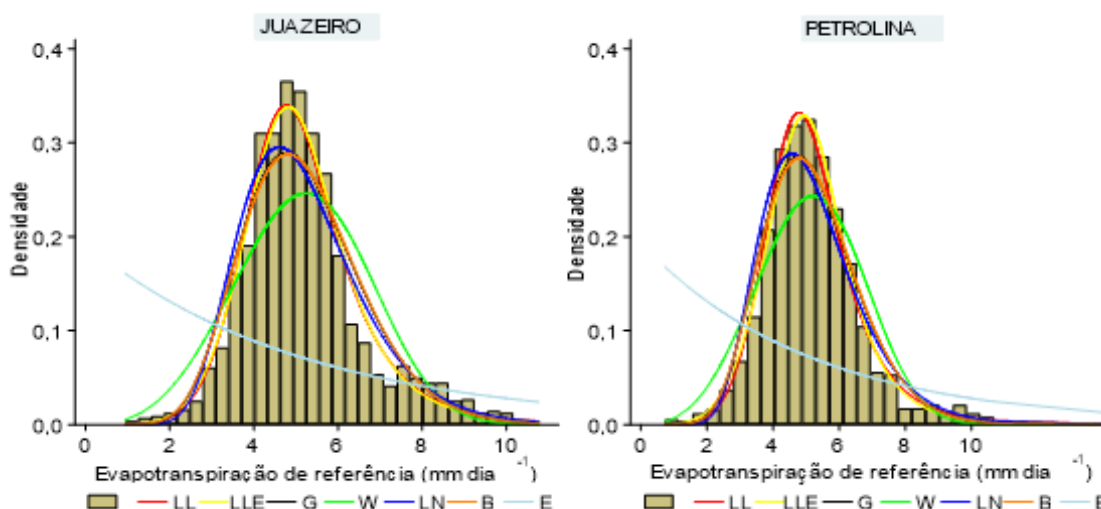


Table 1 presents the p values resulting from the goodness-of-fit test. With the exception of the E distribution, the other distributions fit the data. According to the highest p value, the LL distribution and its exponentiated version (LLE) alternated as the ones that most frequently produced the best goodness-of-fit results. This behavior was observed for the two municipalities studied, likely due to their proximity, as they are neighbors, separated only by the São

Francisco River, and therefore experience synoptic similarities in atmospheric conditions.

The LL distribution generated more robust evidence (higher p values) for nonrejection of the null hypothesis at seven months in Juazeiro and five months in Petrolina. The same number was obtained for the LLE distributions in Petrolina and Juazeiro.

Table 1. p - Monthly values of the Kolmogorov–Smirnov goodness-of-fit test at a maximum probability of type I error of 5%.

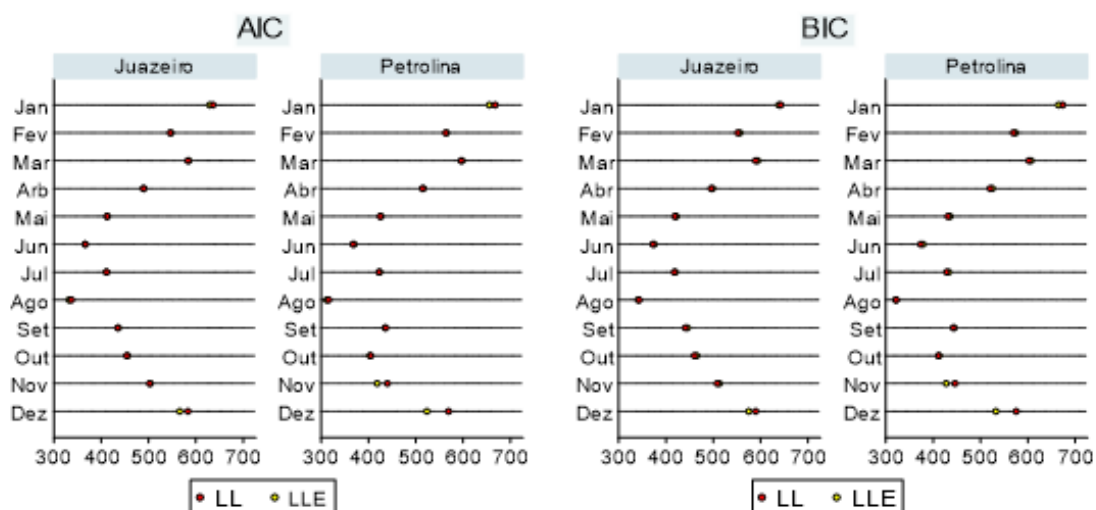
City	Month	Probability distributions						
		G	W	LN	B	AND	LL	LLE
Juazeiro	Jan	0.127	0.013	0.022	0.143	<0.001	0.253	0.235
	Feb	0.067	0.004	0.157	0.058	<0.001	0.415	0.415
	Sea	0.002	<0.001	0.009	0.001	<0.001	0.105	0.097
	Apr	0.001	<0.001	0.006	0.001	<0.001	0.161	0.136
	May	0.614	0.036	0.840	0.587	<0.001	0.950	0.841
	June	0.320	0.012	0.499	0.302	<0.001	0.911	0.944
	Jul	0.079	0.001	0.164	0.073	<0.001	0.686	0.664
	Aug	0.124	0.077	0.064	0.131	<0.001	0.378	0.748
	Set	0.044	0.001	0.093	0.039	<0.001	0.313	0.345
	Out	0.051	0.001	0.104	0.046	<0.001	0.396	0.402
	Nov	0.143	0.006	0.094	0.132	<0.001	0.502	0.482
	Ten	<0.001	0.014	<0.001	<0.001	<0.001	0.060	0.031
Petrolina	Jan	0.050	0.007	0.008	0.045	<0.001	0.078	0.148
	Feb	0.070	0.003	0.203	0.058	<0.001	0.649	0.635
	Sea	0.001	<0.001	0.007	0.001	<0.001	0.087	0.079
	Apr	0.002	<0.001	0.010	0.002	<0.001	0.093	0.090
	May	0.375	0.045	0.654	0.352	<0.001	0.858	0.930
	June	0.308	0.016	0.520	0.291	<0.001	0.927	0.928
	Jul	0.202	0.006	0.364	0.188	<0.001	0.858	0.857
	Aug	0.227	0.107	0.133	0.238	<0.001	0.603	0.856
	Set	0.200	0.013	0.108	0.191	<0.001	0.509	0.520
	Out	0.749	0.296	0.574	0.762	<0.001	0.999	0.990
	Nov	0.011	0.381	0.003	0.013	<0.001	0.498	0.647
	Ten	<0.001	0.104	<0.001	<0.001	<0.001	0.093	0.172

Gamma (G), Weibull (W), log-normal (LN), beta (B), exponential (E), log-logistic (LL), log-logistic exponentiated (LLE) distributions.

Figure 3 presents the AIC and BIC goodness-of-fit criteria for the LL and LLE distributions, as they are nested models and present the highest p values. The AIC values were very close for both distributions, with

the LL model obtaining lower values in the months of February, March, April, May, September, October, and November in Juazeiro and in the first half of the year in Petrolina.

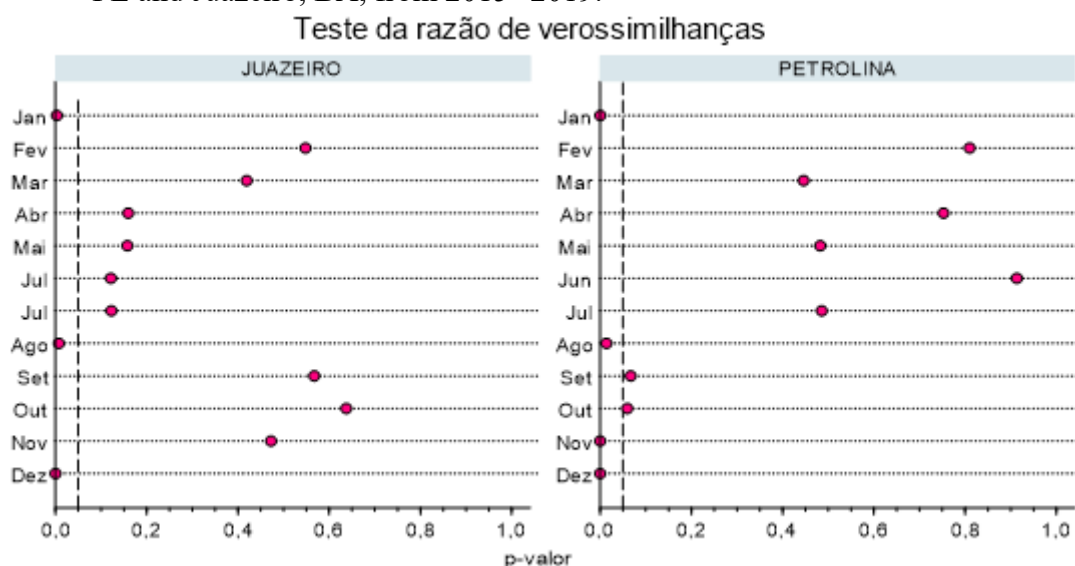
Figure 3. AIC and BIC values for the LL and LLE probability distributions in Petrolina, PE and Juazeiro, BA, from 2015--2019.



For the BIC values, the LLE distribution produced lower values in January, August and December in Juazeiro

and Petrolina, adding to the latter month of November. The same result was obtained via the likelihood ratio test (Figure 4).

Figure 4. Likelihood ratio test to compare the LL and LLE probability distributions in Petrolina, PE and Juazeiro, BA, from 2015--2019.



The Bayesian criterion (BIC) is particularly useful for model selection because it penalizes those with a greater

degree of freedom, that is, a greater number of parameters. In practice, according to this criterion, both the LL and LLE distributions

were adequate, with neither one standing out as hegemonic, as both, compared within each month and city, presented lower values in some months and not in others. Nevertheless, in months when one distribution did not fit, the other excelled, and vice versa.

The estimates of the LL and LLE distribution parameters, as well as the ETo estimates at different probability levels, revealed that the highest average ETo rates (Table 2) occurred from October to January (spring/summer) in both cities. This is due to the high supply of solar radiation, low relative humidity, moderate winds, and the

absence or scarcity of rainfall in the region during this period of the year.

Among the meteorological elements that affect the evapotranspiration rate, both average air temperature and solar radiation are the variables with the greatest influence on reference evapotranspiration. According to Lemos Filho et al. (2010), relative humidity and wind are also factors. According to Santiago et al. (2016), the Juazeiro and Petrolina regions are characterized by high solar radiation and, consequently, greater availability of sensible heat for the evapotranspiration process, especially during dry periods.

Table 2. Estimates of the parameters of the LL and LLE distributions as well as the reference evapotranspiration (mm day⁻¹) at different probability levels.

Evapotranspiration (mm day ⁻¹) at different probability levels:															
City	Month	Estimates			Average	Probability level (%) - p(X ≤ xi)									
		α	β	δ		10	20	30	40	50	60	70	80	90	95
Reference evapotranspiration (mm day ⁻¹)															
Juazeiro	Jan	5.70	5.38	-	5.8	3.8	4.4	4.9	5.3	5.7	6.2	6.7	7.4	8.6	9.9
	Feb	5.03	5.60	0.87	5.1	3.2	3.7	4.1	4.5	4.8	5.2	5.7	6.3	7.2	8.3
	Sea	4.95	5.79	-	5.3	3.4	3.9	4.3	4.6	4.9	5.3	5.7	6.3	7.2	8.2
	Apr	4.76	7.22	-	5	3.5	3.9	4.2	4.5	4.8	5.0	5.4	5.8	6.5	7.2
	May	4.40	8.74	-	4.5	3.4	3.8	4.0	4.2	4.4	4.6	4.8	5.2	5.7	6.2
	June	3.72	7.81	1.74	4.2	3.3	3.5	3.7	3.9	4.1	4.3	4.5	4.8	5.3	5.8
	Jul	4.36	8.76	-	4.5	3.4	3.7	4.0	4.2	4.4	4.6	4.8	5.1	5.6	6.1
	Aug	5.12	16.51	0.51	4.7	3.9	4.2	4.5	4.6	4.8	4.9	5.1	5.3	5.6	5.9
	Set	5.24	9.16	1.15	5.5	4.3	4.6	4.9	5.1	5.4	5.6	5.9	6.2	6.8	7.3
	Out	5.63	9.82	1.11	5.8	4.6	5.0	5.3	5.5	5.7	5.9	6.2	6.6	7.1	7.7
	Nov	5.81	8.25	-	5.9	4.5	4.9	5.2	5.5	5.8	6.1	6.4	6.9	7.6	8.3
	Ten	5.58	6.50	-	5.6	4.0	4.5	4.9	5.2	5.6	5.9	6.4	6.9	7.8	8.8
Petrolina	Jan	6.86	7.01	0.47	5.9	3.4	4.2	4.8	5.3	5.8	6.2	6.7	7.4	8.4	9.3
	Feb	4.86	5.01	-	5.2	3.1	3.7	4.1	4.5	4.9	5.3	5.8	6.4	7.5	8.7
	Sea	4.98	5.60	-	5.3	3.4	3.9	4.3	4.6	5.0	5.4	5.8	6.4	7.4	8.4
	Apr	4.69	6.50	-	4.9	3.3	3.8	4.1	4.4	4.7	5.0	5.3	5.8	6.6	7.4
	May	4.24	7.85	1.23	4.5	3.4	3.7	4.0	4.2	4.4	4.6	4.9	5.2	5.8	6.3
	June	4.33	8.35	-	4.2	3.3	3.7	3.9	4.1	4.3	4.5	4.8	5.1	5.6	6.2
	Jul	4.17	7.79	1.24	4.5	3.4	3.7	3.9	4.1	4.3	4.5	4.8	5.1	5.7	6.3
	Aug	5.15	17.46	0.53	4.8	4.0	4.3	4.6	4.7	4.9	5.0	5.2	5.3	5.6	5.9
	Set	5.67	11.24	0.67	5.4	4.2	4.6	4.9	5.2	5.4	5.6	5.8	6.2	6.6	7.1
	Out	5.70	11.74	-	5.8	4.8	5.1	5.4	5.6	5.8	6.0	6.2	6.5	7.0	7.4
	Nov	6.62	18.65	0.33	5.9	4.6	5.1	5.5	5.7	6.0	6.2	6.4	6.6	7.0	7.3
	Ten	6.68	15.24	0.26	5.5	3.7	4.4	4.9	5.3	5.6	5.9	6.2	6.5	7.0	7.4

Legend: α , β and δ are parameters of the distributions: Gamma (G), Weibull (W), Log-Normal (LN), Beta (B), Exponential (E), Log-Logistic (LL) and Exponentiated Log-Logistic (LLE).

Notably, the average ETo values had a probability of occurrence between 50% and 60%, meaning that the average reference evapotranspiration is generally expected to occur with a frequency ranging from one every two days to six out of ten days. In this context, using average ETo values to estimate the water consumption of a crop of interest in the study region could lead to an accumulation of water deficit and drastically reduce its production. Furthermore, the use

of average values can compromise agricultural climate risk zoning by masking water deficiency risks. When used as a reference for the most critical period and maximum demand for hydraulic sizing of irrigation systems, it can compromise it, as it will not meet the actual water demand of the agricultural crop.

The use of the average for dimensioning can lead to the undersizing of irrigation systems, resulting in losses

(COAN; BACK; BONETTI, 2014). According to Silva et al. (2013), the average value normally occurs between 40% and 50% probability, and for Passos, Raposo and Mendes (2017), the average value should not be adopted as a parameter in the planning of irrigation systems.

For crops of great economic value, such as grapes, the hydraulic dimensioning of irrigation systems should consider January as the period of maximum water demand, for which ETo values of up to 8.6 and 8.4 mm day⁻¹ are expected, with a 90% probability for the municipalities of Juazeiro and Petrolina, respectively.

6 CONCLUSION

Exponentiated and log-logistic distributions were shown to be the most suitable for modeling reference evapotranspiration and making estimates of it at different probability levels in the municipalities of Juazeiro, BA, and Petrolina, PE.

The month of January presented the highest probable water demand in the Juazeiro, BA, Petrolina, and PE regions.

7 ACKNOWLEDGMENTS

This work was carried out with the support of the Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES) - Financing Code 001.

8 REFERENCES

- CASTRO, R.; LEOPOLDO, PR Adjustment of the incomplete gamma distribution in the estimation of probable rainfall for the 15- and 10-day periods in the city of São Manuel (SP). **Energia na Agricultura**, Botucatu, v. 10, n. 1, p. 20-28, 1995.
- COAN, BP; BACK, AJ; BONETTI, AV Probable monthly and annual precipitation in the state of Santa Catarina. **Brazilian Journal of Climatology**, Curitiba, v. 15, n. 10, p. 122-142, 2014.
- DOORENBOS, J.; PRUITT, WO **Crop water requirements**. Rome: FAO, 1997. 212 p.
- FERNANDES, ALT; FRAGA JÚNIOR, EF; TAKAY, BY Evaluation of the Penman -Piche method for estimating reference evapotranspiration in Uberaba, MG. **Brazilian Journal of Agricultural and Environmental Engineering**, Campina Grande, v. 15, n. 3, p. 270-276, 2011.
- LEMOES FILHO, LC; CARVALHO, LG; EVANGELISTA, AWP; ALVES JUNIOR, J. Spatial analysis of the influence of meteorological elements on reference evapotranspiration in Minas Gerais. **Brazilian Journal of Agricultural and Environmental Engineering**, Campina Grande, v. 14, n. 12, p. 1294-1303, 2010.
- OLIVEIRA, GM; SANTIAGO, EJP; LEITÃO, MMVBR; ROCHA, RC; GONÇALVES, IS; SILVA, RR Estimation of reference evapotranspiration for protected environment. **Irriga & Inovagri**, Botucatu, v. 1, n. 1, p. 21-30, 2017.

PASSOS, MLV; RAPOSO, AB; MENDES, TJ Estimation of the distribution of probable rainfall at different levels of probability of occurrence. **Brazilian Journal of Irrigated Agriculture** , Fortaleza, v. 11, n. 1, p. 1106-1115, 2017.

SANTIAGO, EJP; OLIVEIRA, GM; SILVA, RR; OLIVEIRA, VEA; GONÇALVES, IS Adjustment of the Hargreaves and Samani equation from lysimetric data for the municipality of Juazeiro, BA. **Irriga & Inovagri** , Botucatu, v. 1, n. 1, p. 108-114, 2016.

SILVA, ES; ZAVISLAK, FD; DALLACORT, R.; CARVALHO, MAC; ARAUJO, DV Distribution of probability of rainfall in the municipality of Sapezal, MT. **Encyclopedia Biosfera** , Goiânia, v. 9, n. 17, p. 1112-1122, 2013.