DOI: 10.5965/223811711722018260

Revista de Ciências Agroveterinárias 17(2): 2018
Universidade do Estado de Santa Catarina

Probability distributions of frequency analysis of rainfall at the southwest region of Paraná State, Brazil

Distribuições de probabilidade para análise de frequência de chuva no sudoeste do Paraná

Frederico Márcio Corrêa Vieira*, Juliane Mayara Casarim Machado, Edgar de Souza Vismara & Jean Carlo Possenti

Universidade Tecnológica Federal do Paraná, Dois Vizinhos, PR, Brasil. *Author for correspondence: fredvzoo@gmail.com.

Submission: 12/01/2017 | Acceptance: 09/05/2018

ABSTRACT

The study of monthly rainfall probabilities is of great importance due to the increasing occurrences of extreme events in different regions of Brazil. However, the rainfall distribution at the southwest region of Paraná State, Brazil, is still unknown. Thus, the aim of this work is to assess the probabilistic distribution of rainfall frequency at Dois Vizinhos, in the southwest of Paraná State, Brazil. A probabilistic analysis was performed using a historic 40-year rainfall dataset (1973-2012). The gamma, Weibull, normal log, and normal probability distributions were compared. The distribution adherence was performed through Akaike Information Criterion, and the R statistical software was used for estimation. The results showed that the gamma and Weibull distributions were most suitable for probabilistic fitting. Based on this, the average annual rainfall for Dois Vizinhos (PR) was found to be 2,010.6 mm. Moreover, we found that throughout the year, October has the highest rainfall occurrence probability, with an 86% rainfall probability of above 150 mm and 64% rainfall probability above 200 mm.
KEYWORDS: probability distribution, rainfall, biometeorology.

RESUMO

O estudo de probabilidades mensais de precipitação pluvial torna-se importante em virtude do aumento de ocorrências de eventos extremos em diversas regiões do país. Entretanto, poucas informações são conhecidas a respeito da distribuição de chuva na região do sudoeste do Paraná. Com base no exposto objetivou-se por meio desta pesquisa analisar a distribuição probabilística mensal da precipitação pluvial na região de Dois Vizinhos, sudoeste do estado do Paraná. Analisou-se a série histórica de precipitação pluvial de 40 anos (1973-2012). Foram comparadas as seguintes distribuições de probabilidade: gama, Weibull, log normal e normal. A aderência das distribuições aos dados pluviométricos da região foi verificada através do critério de Akaike (AIC). As análises foram realizadas utilizando o software estatístico R. Por meio dos resultados observou-se que as distribuições de gama e Weibull foram as mais adequadas em termos de ajustes probabilísticos. Com base nas distribuições, a média anual de chuva na região de Dois Vizinhos é de 2.010,6 mm e o mês de outubro registra a maior probabilidade de ocorrência de precipitação, sendo 86% de probabilidade de chuvas acima de 150 mm e 64% acima de 200 mm.
PALAVRAS-CHAVE: distribuição de probabilidade, chuva, biometeorologia.

INTRODUCTION

Rainfall is listed as one of the principal meteorological and climatic elements, and it is of great importance because of its impact on human activities. When it comes to causal agents of water scarcity in different regions of the country and the world, as well as factors influencing the occurrence of extreme climate events, rain and its intensity are highly discussed in the fields of climatology and meteorology (WMO 2013).

Excessive rainfall may lead to landslides, flashfloods, or flooding in areas at risk, whereas drought and silting up of rivers lead to water scarcity (SOUZA et al. 2012). With regards to agriculture, rainfall directly interferes with water balance, temperature, and humidity of air and soil, thus affecting the growth and development of the crops (LEITE et al. 2011). Therefore, understanding rainfall occurrence dynamics affects changes in the utilization of land for urban purposes, hydrological studies in cities, and strategic agricultural planning in the principal regions of the country.

According to the Köppen classification, there is a predominance of two climate types in the state of Paraná: namely Cfa, which is concentrated in the north, west, southwest, and Vale do Ribeira; and Cfb, which is concentrated in the south. The proportions of occurrence of these climate types are 61.7% and 37%, respectively. Both types are characterized as humid subtropical, with Cfa being defined as having hot summers with driest month precipitation of above 40 mm (ALVARES et al. 2013). Furthermore, these authors reported that tropical climate types Am and Aw, or tropical monsoon and tropical savanna with dry winter, respectively, are also observed in the Paraná region, albeit with a tiny proportion of occurrences, which are 0.4% and 0.8%, respectively. Regarding the seasonal rainfall distribution, SAMPAIO et al. (2006) stated that in Paraná State, the highest record of ten-day rainfall usually occurs from October to February, whereas the smallest amount of rainfall usually occurs in the months of July and August. As Paraná State is a rainy region, we further highlight that the practice of irrigation should only be partially carried out to supply the water necessity of the crops.

In this context, for a better understanding of the historical dataset on rainfall variation with high precision in this scientific scope, it is necessary to use a mixed probabilistic approach. This consists of fitting different probability distributions and, consequently, the performance of goodness-of-fit tests to verify which of the models obtained the best fit to the data. In studies on monthly rainfall probability, the gamma and Weibull distributions are most frequently used due to the positive asymmetry characteristic of the curves, which guarantees values that are different to zero (SAMPAIO et al. 2007, SILVA et al. 2007, SHARMA & SINGH 2010). Other authors also found satisfactory results with incomplete and mixed gamma distributions (MOREIRA et al. 2010, DALLACORT et al. 2011). However, only one study, with a descriptive approach, has been carried out in southwest region of Paraná State (POSSENTI et al. 2007). Thus, probabilistic studies are necessary to deepen the analysis on the rainfall variability in this region, which is of national importance for the agriculture and livestock industries. In this study, different probability distributions were evaluated, with the aim of analyzing the monthly rainfall variation in the region of Dois Vizinhos, Paraná State, Brazil.

MATERIAL AND METHODS

We analyzed 40 years of historical data, which were from 1973 to 2012. The precipitation dataset we used in our study was obtained through the Coasul Agroindustrial Cooperative (25° 45’ 00’’ latitude South and 53° 03’ 25’’ longitude West; altitude 509 m). According to the Köppen classification, the Cfa climate type, which is characterized as humid subtropical, predominates in the region of Dois Vizinhos City (ALVARES et al. 2013). A pluviometer was installed on a stand 1.5 m above the ground. Data collection was carried out on a daily basis, at 09:00 h, as standardized by the World Meteorological Organization.

Initially, monthly and annual rainfall frequencies were obtained for the region; and from these frequencies, a graphic analysis of the historical dataset was performed. Mann Kendall Trend Test was applied to the series to analyze for monotonic trends, whether positive or negative, in the rainfall values. The null hypothesis was tested for the absence of monotonic trend, whereas the alternate hypothesis was tested for the presence of such tend.

After this stage, gamma, Weibull, log normal, and normal probability distributions were compared. These distributions were chosen as per previous studies (SILVA et al. 2007, RODRIGUES et al. 2013, SILVA et al. 2013).

The gamma function is a generalization of the factorial function, with high applicability in meteorological studies. It presents curves with positive asymmetry, considering that the normal distribution consists of its extreme:

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} X^{\alpha-1} \cdot e^{\left(-\frac{X}{\beta}\right)} \qquad [1]$$

where *β* refers to the scale parameter. *α* is the shape parameter of the distribution, and *Γ(α)* is the gamma function of the alpha parameter.

The Weibull function is one of the most used functions in the fields of climatology and meteorology. Its application describes families of univariate distributions of positive and negative asymmetry, as in:

$$f(x) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} e\left[-\left(\frac{x}{\beta}\right)^{\alpha}\right] \qquad [2]$$

where σ is the shape parameter, and β is the scale parameter.

The log normal distribution is the function whose logarithmic transformation of variable *x* results in a normal distribution. Moreover, its formula is similar to that of normal distribution, as in:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e\left[\frac{-(\ln(x)-\mu)^2}{2\sigma^2}\right] \quad [3]$$

in which x > 0, and μ is the mean of x and σ is the standard deviation of x.

The normal function has the following probability density function:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e\left[\frac{-(x-\mu)^2}{2\sigma^2}\right] \quad [4]$$

where $-\infty<x<+\infty$. The model parameters were already defined.

For each of these distributions, the parameters were fitted by maximum likelihood method and the monthly rainfall values were taken into consideration. Subsequently, the distributions were compared to verify the best fit through the Akaike Information Criterion (AIC). This method is one of the main methodologies used in the models of probability distribution studies (RODRIGUES et al. 2013). According to AKAIKE (1974), the criterion was developed based on decision theory, as in:

$$AIC = -2\log L(\hat{\theta}) + 2k \quad [5]$$

in which $L(\hat{\theta})$ represents the maximum magnitude of the support function, and *k* denotes the number of parameters to be estimated by the model. The classification criterion was determined from the lowest obtained value of AIC attributed to each model. As the AIC only classifies the models, the Kolgomorov-Smirnov test was also applied to each fitted model to check the adherence of the data to the candidate model (MARSAGLIA et al. 2003). The test is carried out of the null that the distribution function which generated the data is the candidate model.

After this assessment, the two best models were used in the determination of the monthly precipitation probabilities for the region with greater than 100, 150, and 200 mm precipitation. These values were chosen with respect to the rainfall characteristics of the climate types found in the state of Paraná, (SAMPAIO et al. 2007, KIST & VIRGENS FILHO 2015). All the statistical analyses in our study were performed using the R statistical software (R DEVELOPMENT CORE TEAM 2014).

RESULTS AND DISCUSSION

In accordance with the descriptive results, the average annual rainfall for the Dois Vizinhos (PR) region is 2,010.6 mm. The rainiest months are October and January, with 243 mm and 193 mm of average annual rainfall, respectively. However, the months with the least amount of rain are July and August, with 129 mm and 113 mm of average annual rainfall, respectively (Figure 1).

In general, rainfall is well distributed throughout the year, thus characterizing the Dois Vizinhos region as humid subtropical. This distribution is in line with the results of SAMPAIO et al. (2006), in which the decrease in rainfall in defined periods were only reported in west and north of the state. JERSZURKI et al. (2015) stated that November to March are the rainiest months in the Telêmaco Borba Region, Paraná State. However, the driest months in the said region were July and August, with precipitation values of 92 mm and 61 mm, respectively. The dry and rainy months described by the previously mentioned authors were similar to those of the present study. However, in our study, the values found for the driest months in Dois Vizinhos City (128.7) were higher than those of previous results (113.5 mm). Such incidence is explained by the different climate types of the two cities, as Telêmaco Borba region is a transition between Cfa and Cfb climates (ALVARES et al. 2013).

From the annual distribution between 1973 and 2012, it is possible to observe a great variability during the period (Figure 2). This variability is random because according to the results of Mann-Kendall test, there is no positive or negative trend in the monthly and annual rainfall series. For the average monthly rainfall (Figure 1), the tests resulted in a p value of 0.8981; and for the total annual rainfall series (Figure 2), the p-value was 0.7006.

The year of 1998 was the rainiest year in the historical dataset, with 2,885.7 mm of average annual rainfall, and was followed by the year 1983, which had an average annual rainfall of 2,687.7 mm. The driest years were 1978 (1,253.4 mm) and 1985 (1,396.9 mm). According to LEITE et al. (2011), such oscillation in the frequency of rainfall can be attributed to the El Niño and La Niña phenomena. They also reported that one of the consequences of El Niño, with respect to the abundant precipitation in the south of South America, is the provocation of extreme events and catastrophes in these regions, such as what occurred in 1983 and 1998. The results of the present study are in accordance with these findings, given that such water excesses were not reached in other periods. On the other hand, the driest months do not correspond to the effects of the La Niña phenomenon, which promotes severe droughts in South America (CPTEC 2016). According to the series of occurrences, the only registered La Niña event was in 1985, which was then classified as weak, as well as the El Niño in 1978. According to LEITE et al. (2011), the low rate of rainfall may be related to the presence of stationary masses of hot dry air in the region, which provoked a reduction in rainfall during the period. In contrast, FISCH & VALÉRIO (2005) and OLIVEIRA et al. (2015) found no correlation between El Niño and La Niña phenomena and rainfall distribution. However, in the regions of the referred studies, center-west and southeast, respectively, the relationship between El Niño/La Niña and rainfall variability is not evident, unlike in the South, as observed in the present study.

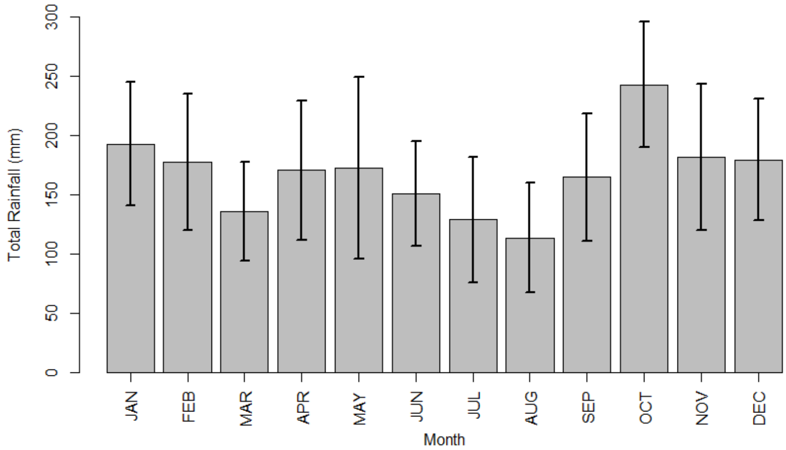

Figure 1. Monthly means and standard deviations of rainfall at the Dois Vizinhos Region from 1973 to 2012.

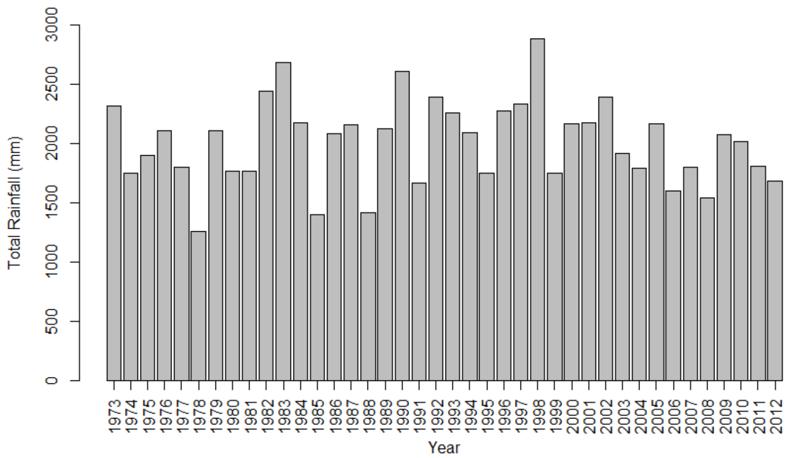

Figure 2. Annual totals and standard deviations of rainfall at the Dois Vizinhos Region from 1973 to 2012.

In the analysis of the functions, all the candidate distributions exhibited satisfactory adherence to the data (Table 1). Based on the AIC, the two distributions with best fit were the gamma and Weibull distributions. These two functions showed comparable results; however, the Weibull function demonstrated better results than the gamma function in the months of January, April, June, September, and December. On the other hand, the gamma function was superior to the Weibull function in the months of March, July, October, and November. The two functions had similar results in February, May, and August. The log-normal function showed satisfactory results in four months and were comparable with those of gamma function. Among the functions, the normal function demonstrated the most suboptimal results, having two months with good fit that were both equal to the Weibull function (Table 2).

Table 1. Parameter estimates of the Gamma, Weibull, Normal, and Log-Normal probability distribution functions for each month from the years 1973 to 2012.

Months	Distributions											
	Gamma			Weibull			Normal			Log-Normal		
	α	β	AIC	α	β	AIC	μ	σ	AIC	μ	σ	AIC
Jan.	3.8256	0.0198	477.3	2.2969	217.8565	474.7*	192.7550	89.0116	476.6	5.1251	0.5780	483.7
Feb.	2.7891	0.0157	480.4*	1.8803	200.8519	479.5*	177.5600	98.9179	485.1	4.9895	0.6617	483.6
Mar.	3.8133	0.0282	449.2*	2.0217	153.3924	451.9	135.4600	71.5257	459.1	4.7719	0.5331	448.9*
Apr.	2.4147	0.0142	481.2	1.7279	190.8519	479.5*	170.5750	100.3569	486.2	4.9180	0.8456	497.5
May	1.5206	0.0088	492.2*	1.3039	187.8860	492.1*	172.5050	131.6188	507.9	4.7869	0.9486	496.2
Jun.	2.5105	0.0166	470.2	1.9953	168.6362	464.5*	150.7950	75.0727	463.0*	4.8043	0.8030	484.3
Jul.	2.4919	0.0194	457.8*	1.5558	144.0183	460.8	128.6575	91.0532	478.4	4.6433	0.6766	457.7*
Aug.	1.3181	0.0116	460.8*	1.2850	121.6800	459.3*	113.4650	78.9439	467.0	4.3064	1.1625	474.1
Sep.	2.4864	0.0151	477.6	1.8317	185.3380	475.2*	164.6950	91.9137	479.2	4.8897	0.7392	484.5
Oct.	7.2056	0.0297	474.1*	2.8759	272.5451	476.2	242.9925	90.4177	477.9	5.4220	0.3831	474.5*
Nov.	3.4546	0.0190	475.9*	1.8443	205.0826	479.6	181.8300	105.4420	490.2	5.0514	0.5634	475.7*
Dec.	3.1355	0.0175	477.7	2.1108	202.0501	474.1*	179.2625	87.8891	475.6*	5.0210	0.6699	487.1

*best-fits according to the Akaike Information Criterion (AIC).

Table 2. Results of the Kolgomorov-Smirnov tests of the Gamma, Weibull, Normal, and Log-Normal probability distribution functions fitted for each of the months from 1973 to 2012, where D is the statistics of the test and p value is the type II error.

	Distributions							
Months	Gamma		Weibull		Normal		Log-Normal	
	D	p value	D	p value	D	D	p value	D
Jan.	0.1102	0.6610	0.0932	0.8359	0.1014	0.7556	0.1395	0.3681
Feb.	0.0885	0.8769	0.0836	0.9142	0.1106	0.6711	0.1004	0.7660
Mar.	0.0802	0.9356	0.1053	0.7141	0.1518	0.2726	0.0677	0.9854
Apr.	0.1273	0.4964	0.0912	0.8634	0.1106	0.6711	0.1922	0.0909
May	0.0794	0.9585	0.0958	0.8465	0.1696	0.1888	0.1077	0.7287
Jun.	0.1571	0.2638	0.1298	0.4940	0.1257	0.5366	0.2059	0.0618
Jul.	0.0906	0.8892	0.1121	0.6817	0.1739	0.1675	0.0750	0.9754
Aug.	0.1249	0.5045	0.1029	0.7397	0.1182	0.5745	0.1636	0.1993
Sep.	0.0927	0.8405	0.0764	0.9557	0.1078	0.6867	0.1383	0.3783
Oct.	0.0951	0.8185	0.0829	0.9186	0.0920	0.8470	0.0930	0.8384
Nov.	0.0824	0.9434	0.1010	0.7970	0.1371	0.4240	0.0908	0.8876
Dec.	0.1295	0.4593	0.0907	0.8583	0.0855	0.9006	0.1653	0.1902

The results found in the present study are similar to those obtained by SILVA et al. (2007), wherein they stated that the two probability functions, gamma and Weibull functions, are suitable for rainfall modeling. The gamma distribution has a high potential of best fit to the regional rainfall data due to its low magnitude as compared to other distributions, thus giving it a good acceptability in goodness-of-fit tests (CATALUNHA et al. 2002). RAMOS et al. (2015) stated that in periods with a higher occurrence of rainfall, the values of *α* (shape parameter) tend to be higher due to the asymmetry of dry periods, which are inversely proportional. The results of the present study are in line with those of the previously mentioned authors, given that the highest value of α found in the dataset was in the month of October (7.21), which conferred superiority on the gamma function in relation to Weibull function. In the other rainy months (November and March), the values were higher in relation to the dry months.

The same can be observed for the Weibull function, which, together with the gamma function, obtained the highest number of months with goodness-of-fit as confirmed by the AIC. According to VIEIRA et al. (2010), this distribution has significant fit for rainy months especially when given the absence of null values commonly found in dry seasons. As the Dois Vizinhos region is characterized by rainfall every month, this probabilistic distribution has a desirable effect in rainfall variability studies.

Despite the difference between the two functions in the analyzed months and concerning the fit of the models and their parameters, only a little difference was observed between gamma and Weibull functions in the estimation of monthly rainfall probabilities (Figure 3).

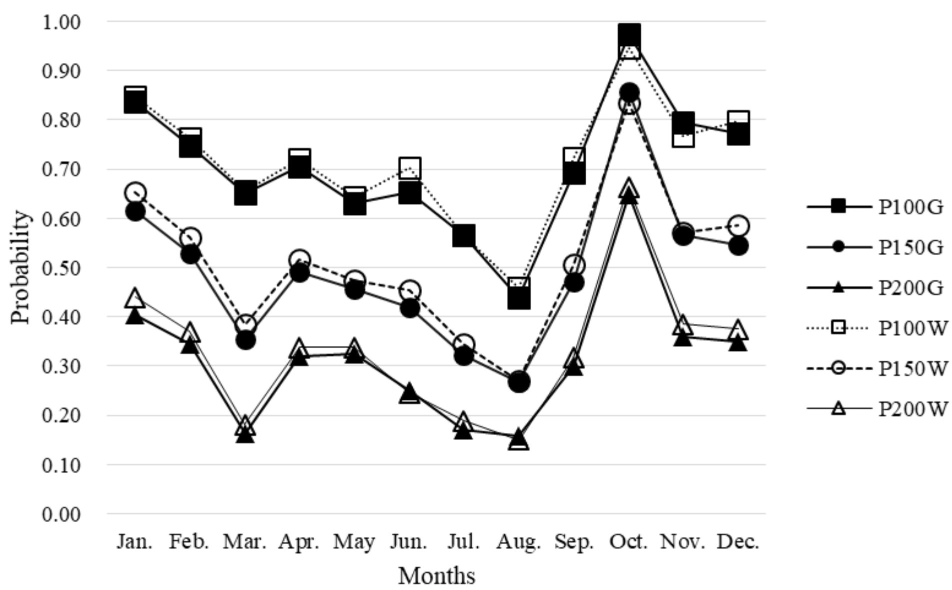

Figure 3. Probabilities of monthly rainfall occurrences greater than 100 mm (P100), 150 mm (P150), and 200 mm (P200) as estimated by the gamma (G) and Weibull functions (W) in the City of Dois Vizinhos, Paraná State, Brazil.

Thus, a high probability of rainfall can be noted for the month of October, with an 86% and 64% probability of rainfall above 150 mm and 200 mm, respectively. In contrast, the driest was found to be August, with 44% probability of rainfall above 100 mm. The results corroborate those of SAMPAIO et al. (2007), wherein the lowest rainfall rates for the same period were reported in most of the regions of the state, except the coastal municipalities. In the months with intermediate rainfall, the probability of rainfall was above 50% for heights between 100 and 150 mm.

CONCLUSION

The gamma and Weibull distributions are the most suitable functions to be used for rainfall estimation in the Dois Vizinhos region. Through the monthly rainfall probabilities, it was possible to observe that the region has a regular rainfall distribution throughout the year, with October being the rainiest month and August as the driest month on the climatological dataset.

ACKNOWLEDGEMENTS

We would like to thank the COASUL Agroindustrial Cooperative for providing us the meteorological data we used in this study.

REFERENCES

AKAIKE H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 19: 716-723.

ALVARES CA et al. 2013. Köppen's climate classification map for Brazil. Meteorologische Zeitschrift 22: 711-728.

CATALUNHA MJ et al. 2002. Aplicação de cinco funções densidade de probabilidade a séries de precipitação pluvial no Estado de Minas Gerais. Revista Brasileira de Agrometeorologia 10: 153-162.

CPTEC. 2016. El Niño e La Niña. Centro de Previsão do Tempo e Estudos Climáticos, São José dos Campos, Brazil. Available at: http://enos.cptec.inpe.br/. Access on: Feb. 2016.

DALLACORT R et al. 2011. Distribuição de chuvas no município de Tangará da Serra, médio norte do Estado de Mato Grosso, Brasil. Acta Scientiarum. Agronomy 33: 193-200.

FISCH G & VALÉRIO MC. 2005. Variabilidade intra e interanual da precipitação em Taubaté-SP associado aos eventos El Niño e La Niña. Revista Biociências 11: 19-29.

JERSZURKI D et al. 2015. Distribuição de probabilidade e movimento temporal da precipitação na região de Telêmaco Borba, Paraná. Revista Brasileira de Ciências Agrárias 10: 110-116.

KIST A & VIRGENS FILHO JS. 2015. Análise probabilística da distribuição de dados diários de chuva no estado do Paraná. Revista Ambiente & Água 10: 172-181.

LEITE ML et al. 2011. Análise da frequência e da intensidade das chuvas em Ponta Grossa, Estado do Paraná, no período entre 1954 e 2001. Acta Scientiarum. Technology 33: 57-64.

MARSAGLIA G et al. 2003. Evaluating Kolmogorov's Distribution. Journal of Statistical Software 8: 1-4.

MOREIRA PSP et al. 2010. Distribuição e probabilidade de ocorrência de chuvas no município de Nova Maringá-MT. Revista de Ciências Agro-Ambientais 8: 9-20.

OLIVEIRA NL et al. 2015. Influência do El Niño e La Niña no número de dias de precipitação pluviométrica no Estado de Mato Grosso. Ciência e Natura 37: 284-297.

POSSENTI JC et al. 2007. Distribuição da precipitação pluvial em Dois Vizinhos, Paraná, Brasil. In: Seminário de sistemas de produção agropecuária, Gerência de Pesquisa e Pós-Graduação. Anais... Dois Vizinhos: UTFPR. p.140-142.

R DEVELOPMENT CORE TEAM. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Available at: http://www.R-project.org. Access on: Sep. 1, 2015.

RAMOS HC et al. 2015. Precipitação pluvial de Diamantino-MT em anos de ocorrência de El Niño, La Niña e neutros. Revista Brasileira de Meteorologia 30: 71-80.

RODRIGUES JA et al. 2013. Funções densidade de probabilidade para a estimativa de precipitação mensal. Semina: Ciências Exatas e Tecnológicas 34: 3-8.

SAMPAIO SC et al. 2006. Estimativa e distribuição da precipitação mensal provável no Estado do Paraná. Acta Scientiarum. Human and Social Sciences 28: 267-272.

SAMPAIO SC et al. 2007. Estimativa e distribuição de precipitações decendiais para o Estado do Paraná. Irriga 12: 38-53.

SHARMA MA & SINGH JB. 2010. Use of Probability Distribution in Rainfall Analysis. New York Science Journal 3: 40-49.

SILVA JC et al. 2007. Análise de distribuição de chuva para Santa Maria, RS. Revista Brasileira de Engenharia Agrícola e Ambiental 11: 67-72.

SILVA IN et al. 2013. Distribuição de frequência da chuva para a região Centro-Sul do Ceará, Brasil. Revista Ciência Agronômica 44: 481-487.

SOUZA WM et al. 2012. Classificação da precipitação diária e impactos decorrentes dos desastres associados às chuvas na cidade do Recife-PE. Revista Brasileira de Geografia Física. 2: 250-268.

VIEIRA JPG et al. 2010. Estudo da precipitação mensal durante a estação chuvosa em Diamantina, Minas Gerais. Revista Brasileira de Engenharia Agrícola e Ambiental 14: 762-767.

WMO. 2013. The global climate 2001 – 2010: a decade of climate extremes. World Meteorological Organization. Geneva, Switzerland. Available at: http://www.unep.org/ pdf/wmo_report.pdf. Access on: Feb. 2016.