

XII ENCONTRO NACIONAL DE ÁGUAS URBANAS

DISTRIBUIÇÕES PROBABILÍSTICAS PARA PRECIPITAÇÕES DO MUNICÍPIO DE ARCOVERDE-PE

*Jaime Joaquim da Silva Pereira Cabral¹; Moacyr Magalhães Cavalcanti Neto² & Willames de
Albuquerque Soares³ & Wrias Elmom Firmino Lino⁴*

RESUMO – Existem diversas preocupações em relação aos recursos hídricos, assim, para o devido controle de tal recurso faz-se necessário o conhecimento das probabilidades de precipitações e estes eventos podem ser estudados através de distribuições probabilísticas. Sendo assim, o presente trabalho tem como objetivo avaliar qual o modelo de distribuição de probabilidade que melhor se adere ao conjunto de médias pluviométricas anuais na cidade de Arcoverde-PE, através das distribuições tipo Beta, Weibull, Normal, Gama e Log-normal. Para verificar o ajuste das médias anuais às distintas funções densidade de probabilidade foram realizados os testes de aderência Kolmogorov-Smirnov (K-S), Anderson-Darling (AD) e Qui-quadrado (χ^2), onde adotou-se o nível de significância de 5% para o cálculo dos valores críticos de cada teste. A distribuição Beta foi a que melhor aderiu aos dados de chuva para a série de medias anuais de precipitação, ao passo que as distribuições Gama e Log-normal ocuparam a última colocação para a mesma série de dados.

ABSTRACT– There are several concerns regarding water resources, so for proper control of this resource it is necessary to know the probability of precipitation and these events can be studied through probabilistic distributions. Therefore, the present work has as objective to evaluate which model of probability distribution best fit to the set of annual rainfall averages in the city of Arcoverde-PE, through the Beta, Weibull, Normal, Range and Log-normal distributions. The Kolmogorov-Smirnov (KS), Anderson-Darling (AD) and Chi-square (χ^2) fit tests were used to verify the fit of the annual averages to the different probability density functions, where the significance level of 5 % for the calculation of the critical values of each one. The Beta distribution was the one that best fit to the rainfall data for the series of annual rainfall averages, whereas the Gama and Log-normal distributions occupied the last place for the same dataset.

Palavras-Chave – Distribuições de probabilidade. Teste de aderência. Precipitação pluviométrica.

1) Universidade de Pernambuco. Rua Benfica, 455, Recife-PE, (81) 21268223, jcabral.upe@gmail.com.
2) Universidade de Pernambuco. Rua Benfica, 455, Recife-PE, (82) 99304-0904, mmcen@poli.br.
3) Universidade de Pernambuco. Rua Benfica, 455, Recife-PE, (81) 3184-7540, was@poli.br.
4) Universidade de Pernambuco. Rua Benfica, 455, Recife-PE, (81) 97120-2983, wefl@poli.br

1. INTRODUÇÃO

A água está presente em diversas atividades humanas, tratando-se do recurso natural mais importante para a vida em nosso planeta. O aumento populacional está atrelado a este recurso natural, sendo de grande valia para atividades econômicas, como por exemplo, agrícolas e industriais, que acrescentadas ao constante desenvolvimento urbano, atribuem a este recurso grande importância. Contudo, o elevado consumo dos recursos hídricos tem causado preocupações. (TUNDISI e MATSUMURA-TUNDISI, 2011).

Outro aspecto relevante sobre importância da água pode ser considerado no tocante ao desenvolvimento das cidades. No Brasil o processo de ocupação ocorre, na sua maioria, de forma desordenada, sem um planejamento adequado, o constante desenvolvimento urbano contribui com o aumento de áreas impermeáveis, diminuindo a infiltração das águas pluviais no solo e aumentando o escoamento da mesma, ocasionando pontos de inundações em diferentes partes da cidade.

Segundo Silva et al. (2015) em um grande número de países diversas investidas têm sido tomadas no intuito de aliviar os impactos causados pela urbanização, principalmente sobre o aspecto do aumento dos picos de vazão e a poluição das águas pluviais. Outro importante fator a ser considerado é a distribuição das chuvas, que não se dá de forma uniforme.

Em vista da situação, faz-se necessário o melhoramento da eficiência das redes pluviométricas, que devem gerar dados confiáveis, que assegurem qualidade e quantidade satisfatórias para futuros estudos. A análise minuciosa da variabilidade temporal e espacial das precipitações é o elemento que guiará uma boa configuração das redes pluviométricas (SALGUEIRO e MONTENEGRO, 2004). Segundo Assis et al (2004) os fenômenos meteorológicos podem ser estudados através de distribuições de variáveis, para previsibilidade aceitável do comportamento climático de uma área e determinação de padrões de ocorrência, se tratando de uma ferramenta bastante útil para planejamento e gestão de variadas atividades humanas.

Diante disto, este trabalho tem como objetivo avaliar qual o modelo de distribuição de probabilidade melhor se adere ao conjunto de médias pluviométricas anuais da cidade de Arcoverde-PE.

2. MATERIAL E MÉTODOS

O município de Arcoverde está e está localizado na mesorregião Sertão e na Microrregião Sertão do Moxotó do Estado de Pernambuco, limitando-se a norte com Estado da Paraíba, a sul com Buíque e Pedra, a leste com Pesqueira, e a oeste com Sertânia. Distante de 256 km da capital Recife, cujo acesso é feito pela BR-232, com uma área territorial de aproximadamente 324 km² e densidade demográfica de 196,05 hab/km² (IBGE, 2010). Segundo o Ministério da Integração Nacional, o

munícipio faz parte da região do Semiárido Brasileiro que é caracterizada pelo clima seco, com poucas chuvas e elevada evapotranspiração.

Foram utilizados para análise das precipitações, médias pluviométricas anuais de 1973 a 2017, da estação pluviométrica Arcoverde (Tabela 1), obtidos através do banco de dados meteorológicos para ensino e pesquisa do Instituto Nacional de Meteorologia – INMET (2018).

Tabela 1 - Estação Pluviométrica Arcoverde

Estação	Nº da Estação	Longitude	Latitude	Altitude (m)	Período (anos)
Arcoverde	82890	-37,08°	-8,41°	680,70	1973-2017

Para o estudo, serão testadas as distribuições Beta, Weibull, Log-normal, Normal e Gama.

a) Distribuição Beta: De acordo com Souza et al. (2010) é definida no intervalo (0,1). Seu gráfico pode apresentar diferentes formas dependendo dos dois parâmetros de forma alfa e beta. Com a seguinte função densidade probabilidade:

$$F(x) = \frac{x^{\alpha-1} \cdot (1-x)^{\beta-1}}{B(\alpha, \beta)} \text{ onde, } B(\alpha, \beta) = \int_0^1 x^{\alpha-1} \cdot (1-x)^{\beta-1} dx \quad (1)$$

Onde, $0 \leq X \leq 1$ e $\alpha, \beta > 0$.

b) Distribuição Weibull: Conforme Catalunha et al. (2002) sua função de densidade de probabilidade é habitualmente exibida de acordo com a equação:

$$F(x) = \frac{\gamma}{\beta} \left(\frac{x-\alpha}{\beta}\right)^{\gamma-1} \cdot e^{-\left(\frac{x-\alpha}{\beta}\right)^{\gamma}}, \text{ para } x \geq \alpha \quad (2)$$

Sendo,

$F(x) = 0$ para outros intervalos;

$\alpha \geq 0, \beta > 0, \gamma > 0$ são os parâmetros da distribuição.

c) Distribuição Log-normal: Uma amostra obedece a distribuição Log-normal quando o logaritmo dos seus valores corresponde a distribuição normal. Essa distribuição tem sido bastante empregada em hidrologia com o propósito de modelar cheias anuais (CLARKE, 1993). A função densidade da distribuição Log-normal a três parâmetros pode ser representada pela expressão:

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

Onde,

$F(x)$ - função densidade de probabilidade da variável;

e - base do logaritmo neperiano;

x - valor da variável aleatória;

μ - média dos logaritmos da variável x ;

σ - desvio-padrão dos logaritmos da variável x ;

a - limite inferior da amostra.

d) Distribuição normal: Também denominada de distribuição Gaussiana, é a distribuição de probabilidade contínua mais usada, tem uma elevada importância em análise matemática, pois algumas técnicas estatísticas assumem ou exigem normalidade de dados. Porém, a heterogeneidade de amostras climatológicas pode não seguir esta distribuição (ESSENWANGER, 1986). É comumente apresentada pela equação:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \text{ para } -\infty < x < +\infty \quad (4)$$

Onde,

μ - média;

σ - desvio padrão.

e) Distribuição Gama: A variável aleatória contínua x , assumindo valores positivos, tem uma distribuição gama com parâmetros $\alpha > 0$ e $\beta > 0$, se sua função densidade de probabilidade dada por:

$$F(x) = \frac{x^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)} e^{-\frac{x}{\beta}}, \text{ para } 0 < x < \infty \quad (5)$$

Sendo,

$F(x)$ - probabilidade de ocorrer um valor menor ou igual a x ;

x - variável aleatória contínua;

$\Gamma(\alpha)$ - função Gama do parâmetro alfa;

α - parâmetro de forma da variável aleatória x ;

β - parâmetro de escala da variável aleatória x ;

e - base do logaritmo neperiano.

Onde, $\Gamma(\alpha)$ é a função gama, importante em muitas áreas da Matemática, dada por:

$$\Gamma(\alpha) = \int_0^{\infty} e^{-x} x^{\alpha-1} dx, \alpha > 0 \quad (6)$$

Para verificar o ajuste das médias anuais às distintas funções densidade de probabilidade foram utilizados os testes de aderência Kolmogorov-Smirnov (K-S), Anderson-Darling (AD) e Qui-quadrado (χ^2).

Segundo Lira e Macedo (2017) o teste de Kolmogorov-Smirnov é amplamente utilizado na estatística e se caracteriza por ser aplicável nas distribuições contínuas, além de avaliar a hipótese de uma amostra observada aderir ou não a uma distribuição teórica específica. Ainda segundo Lira e Macedo (2017) o teste do Qui-quadrado é uma das distribuições mais utilizadas em estatística

inferencial e serve para avaliar quantitativamente a relação entre o resultado de um experimento e a distribuição esperada para o fenômeno.

O teste Anderson-Darling é útil para examinar se a amostra de dados decorre de uma determinada distribuição (STEPHENS, 1974).

Os valores críticos utilizados para rejeitar ou aceitar uma hipótese de que a probabilidade de ocorrência dos valores dos dados de sua amostra pode ser estimada com base numa distribuição de probabilidade parametrizada, dependerá do nível de significância adotado, do tamanho da amostra e dos parâmetros relacionados a cada tipo de distribuição.

Com a obtenção das curvas de probabilidade acumulada real *versus* teórica pode-se visualmente verificar a aproximação ou não das probabilidades reais e teóricas e a partir dos resultados dos testes de aderência, é possível determinar qual a distribuição que melhor representa as médias de precipitações anuais da cidade de Arcoverde.

Para cada teste de aderência, será definida qual a melhor a distribuição a partir do *ranking* da distância do valor encontrado em relação ao valor crítico do teste, onde a distribuição que obtiver o valor mais distante do crítico será considerado a que melhor se adere, e conseqüentemente, aquela com valor mais próximo ao crítico a pior.

Para definição, entre todos os testes de aderência, de qual melhor distribuição se adere aos dados da amostra, será aplicado o método de pontuação por colocação, onde serão estipuladas pontuações 5, 4, 3, 2 e 1 para as colocações 1º, 2º, 3º, 4º e 5º lugar respectivamente de cada teste. Construindo assim, um *ranking* das distribuições a partir do somatório final de pontuação de cada uma, onde a primeira colocada será aquela que obtiver mais pontos e a última a que obtiver a menor pontuação.

3. RESULTADOS E DISCUSSÃO

A análise das médias anuais de pluviosidade pode ser observada na Tabela 2, onde, nota-se um baixo valor de média de precipitação anual, observa-se também um baixo valor de desvio padrão, indicando uma pequena variação de chuvas ao longo dos anos, e possuem valores máximos e mínimos de 103,43 mm e 14,08 mm, respectivamente.

Tabela 2 - Informações dos dados coletados.

Estação	Tamanho da amostra	Média (mm)	Máximo (mm)	Mínimo (mm)	Desvio Padrão	Assimetria
Arcoverde	45	54,65	103,43	14,08	21,07	0,21

Foi possível também, a partir da análise dos dados, determinar o histograma de frequência da amostra através da definição das classes, Figura 1.

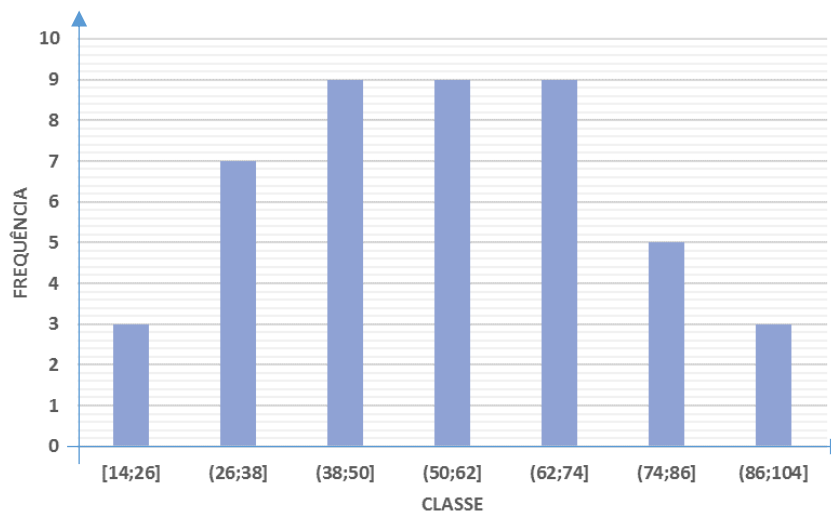


Figura 1 - Histograma de frequência da amostra.

Para a aplicação das distribuições Beta, Weibull, Log-normal, Gama e Normal foram realizados cálculos das probabilidades acumuladas, para cada dado da nossa amostra, utilizando os parâmetros calculados, Tabela 3, e construído suas respectivas curvas de distribuição de probabilidade teórica sobrepondo com a curva de distribuição de probabilidade real, ilustrado na Figura 2.

Tabela 3 - Parâmetros para a distribuição Beta, Weibull, Log-normal, Gama e Normal.

Distribuição	Alfa (α)	Beta (β)	a	b	Desvio Padrão (σ)	Média (μ)	Desvio Padrão dos Logaritmos da Variável	Média dos Logaritmos da Variável
Beta	2,14	2,82	9,98	113,61	-	-	-	-
Weibull	2,70	60,49	-	-	-	-	-	-
Log-normal	-	-	-	-	-	-	0,43	3,92
Gama	6,72	8,13	-	-	-	-	-	-
Normal	-	-	-	-	21,07	54,65	-	-

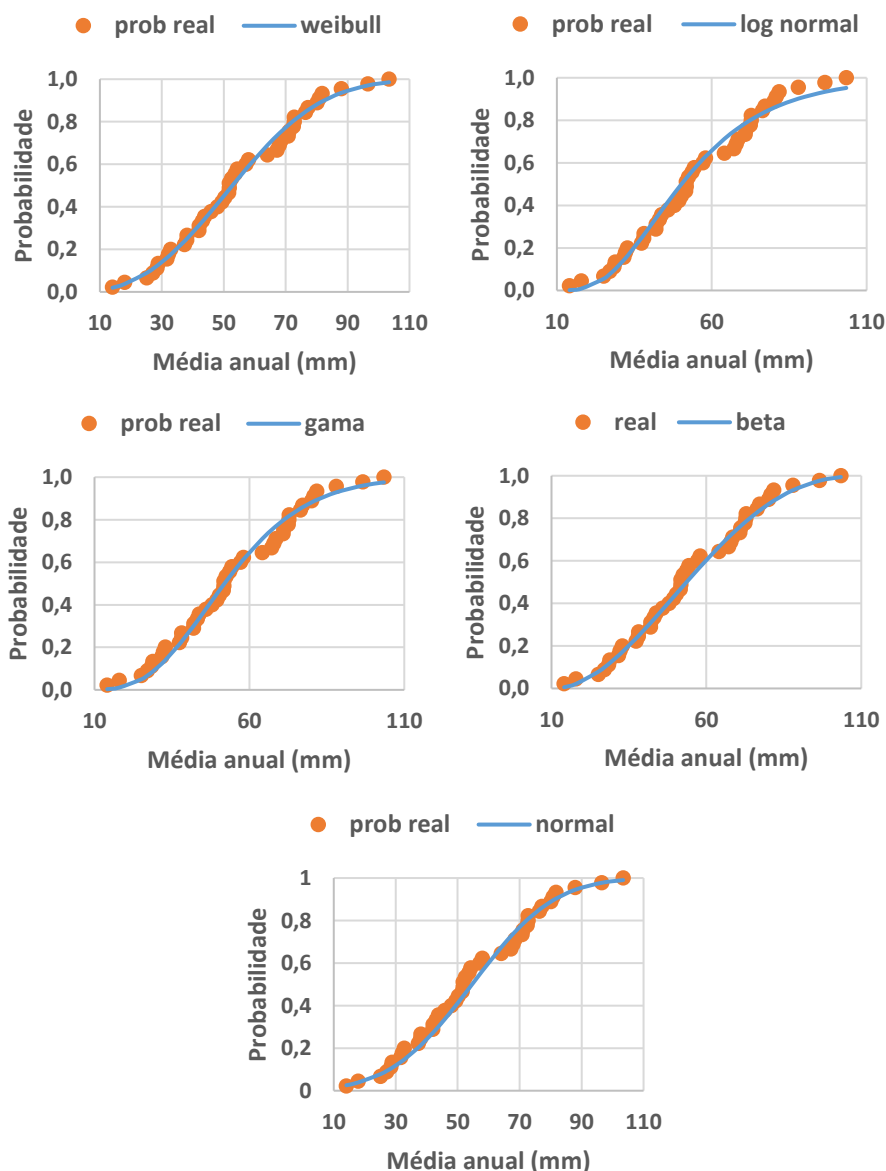


Figura 2 - Sobreposição de curvas de probabilidade acumulada real *versus* teórica

Visualmente é possível verificar a proximidade das curvas teóricas com os pontos da probabilidade real, mas não é possível definir qual delas mais se adere aos dados da amostra. Para confirmar esta possível aproximação e qual delas melhor representa as médias anuais de pluviosidade foram realizados testes de aderência para cada tipo de distribuição de probabilidade teórica. Assim, foi adotado o nível de significância de 5% (0,05) para o cálculo dos valores críticos de cada teste. Ao realizar o teste K-S obteve-se os resultados apresentados na Tabela 4.

Tabela 4 - Resultados para o teste de aderência K-S para nível de significância $\alpha=0,05$.

Distribuição	Dn,máx	Dcrítico	Resultado	Distância em relação ao crítico (Dcrítico – Dn,máx)	Ranking
Beta	0,0656	0,2027	não-rejeitado	0,1371	1°
Weibull	0,0682	0,2027	não-rejeitado	0,1346	2°
Log-normal	0,0822	0,2027	não-rejeitado	0,1205	3°
Gama	0,0856	0,2027	não-rejeitado	0,1172	5°
Normal	0,0837	0,2027	não-rejeitado	0,1191	4°

Assim, para o teste de aderência K-S, todas as distribuições não obtiveram rejeição, uma vez que seu valores de Dn,máx foram menores que o Dcrítico, e a distribuição que melhor representa os dados da amostra é a distribuição de probabilidade Beta, onde o seu valor de Dn,máx é o mais distante do crítico entre todas as distribuições. Ou seja, seus valores de probabilidade foram os que mais se aproximaram do real.

Dando continuidade, foi realizado o teste de aderência Qui-quadrado (X^2) para as distribuições e encontrados os seguintes resultados, apresentados na Tabela 5.

Tabela 5 - Resultados para o teste de aderência Qui-quadrado para nível de significância $\alpha=0,05$.

Distribuição	X^2	X^2 crítico	Resultado	Distância em relação ao crítico (X^2 crítico – X^2)	Ranking
Beta	0,5496	9,4877	não-rejeitado	8,9382	2°
Weibull	0,4752	9,4877	não-rejeitado	9,0126	1°
Log-normal	3,1182	9,4877	não-rejeitado	6,3695	5°
Gama	1,4758	9,4877	não-rejeitado	8,0119	4°
Normal	0,5761	9,4877	não-rejeitado	8,9117	3°

Os dados foram analisados da mesma forma que para o teste K-S e verificou-se que todas as distribuições também não obtiveram rejeição, para este teste de aderência a distribuição que obteve maior valor de distância em relação ao crítico foi a de Weibull, sendo assim, a distribuição de probabilidade que maior teve aderência as médias anuais de chuva para o teste de aderência Qui-quadrado. Por fim, foi feito o teste de Anderson-Darling (AD), onde os resultados estão apresentados na Tabela 6.

Tabela 6 - Resultados para o teste de aderência Anderson-Darling (AD) para nível de significância $\alpha=0,05$.

Distribuição	A^2 corr	A^2 crítico	Resultado	Distância em relação ao crítico (A^2 crítico – A^2 corr)	Ranking
Beta	0,1955	0,7570	não-rejeitado	0,5615	1°
Weibull	0,2446	0,7570	não-rejeitado	0,5124	2°

Distribuição	A ² corr	A ² crítico	Resultado	Distância em relação ao crítico (A ² crítico – A ² corr)	Ranking
Log-normal	0,5678	0,7570	não-rejeitado	0,1892	5°
Gama	0,3948	0,7570	não-rejeitado	0,3622	4°
Normal	0,3086	0,7570	não-rejeitado	0,4484	3°

Para este teste não existiram distribuições com rejeição, assim como nos testes anteriores. Percebe-se que mais uma vez a distribuição Beta foi a que melhor representou os dados amostrais por possuir seu valor de A²corr com maior distância em relação ao A²crítico.

Para definir o ranking final das distribuições de probabilidade, quanto a sua aderência em relação aos dados da amostra, aplicou-se o método de pontuação por colocação e os resultados estão apresentados na Tabela 7.

Tabela 7 - Resultado final para o método de pontuação por colocação.

Distribuição	1°	2°	3°	4°	5°	Pontuação Total	Ranking
Beta	10	4	0	0	0	14	1°
Weibull	5	8	0	0	0	13	2°
Normal	0	0	6	2	0	8	3°
Gama	0	0	0	4	1	5	4°
Log-normal	0	0	3	0	2	5	4°

A distribuição probabilística que obteve melhor resultado e mais se aderiu aos dados pluviométrico foi a Beta, seguida da Weibull e da Normal. No estudo de Silva et al. (2015), para seus dados de precipitação a distribuição que melhor se aderiu foi a de Gumbel. Marques et al. (2014) obtiveram melhor aderência para Gumbel, seguido da Generalizada de Valor Extremo (GEV) e Gama.

4. CONCLUSÃO

Com base na análise de dados das médias de precipitações anuais do município de Arcoverde e a aplicação dos testes de aderência para as variadas distribuições de probabilidade, conclui-se que os dados podem ser representados por todas. Tendo a distribuição probabilística Beta é a que melhor representa a série de dados.

REFERÊNCIAS

ASSIS, J. P. et al. (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, v. 12, n. 01, pp. 113-121.

- CATALUNHA, M. J., SEDIYAMA, G. C., LEAL, B. G., SOARES, 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, pp. 153-162.
- CLARKE, R. T. (1993). “*Hidrologia Estatística*”, in *Hidrologia Ciência e Aplicação*. Ed. UFRGS, Porto Alegre - RS, pp. 659-698.
- ESSENWANGER, O.M. (1986). “*World Survey of Climatology. General Climatology 1B*”. Elements of Statistical Analysis, pp. 424.
- Instituto Brasileiro de Geografia e Estatística – IBGE. Disponível em: <<https://cidades.ibge.gov.br/brasil/pe/arcoverde/panorama>>. Acesso em: 25 abr. 2018.
- Instituto Nacional de Meteorologia - INMET. Disponível em: <<http://www.inmet.gov.br/portal/index.php?r=home2/index>>. Acesso em: 25 abr. 2018.
- MARQUES, R. F. de P. V. et al (2014). “*Performance of the probability distribution models applied to heavy rainfall daily events*”. Ciência e Agrotecnologia, v. 38, n.4, pp. 335-342.
- SALGUEIRO, J. H. P. B.; MONTENEGRO, S. M. G. L. (2004). “*Geoestatística aplicada à variabilidade espacial e padrões de precipitação na bacia do rio Ipojuca em Pernambuco*” in Anais do VII Simpósio de Recursos Hídricos do Nordeste, São Luis, Nov. 2004.
- SILVA, T. F.; PAIVA, A. L. R.; SANTOS, S. M. (2015). “*Análise Estatística e Tendência das Precipitações no Município de Caruaru – PE*” in Anais do XXI Simpósio Brasileiro de Recursos Hídricos, Brasília, Nov. 2015, pp. 1-8.
- SILVA, I. N. et al (2015). “*Precipitação provável para a região centro-sul do Ceará, Brasil*”. Revista Brasileira de Geografia Física, v. 8, n. 3.
- SOUZA, I. F. et al (2010). “*Comparação entre distribuições de probabilidades da precipitação mensal no estado de Pernambuco*”. Scientia Plena, v. 6, pp. 1-10.
- STEPHENS, M. A. (1974). “*FED para a Bondade de Estatística Fit e algumas comparações*”. Jornal da Associação Americana Estatística, v. 69, p. 730-737.
- TUNDISI, J.G.; MATSUMURA-TUNDISI, T. (2011). *Recursos hídricos no século XXI*. Oficina de Textos São Paulo - SP, 328p.