

ADEQUABILIDADE DE SETE DISTRIBUIÇÕES DE PROBABILIDADE APLICADAS A UMA SÉRIE DE PRECIPITAÇÕES MÁXIMAS, EM BARREIRAS (BA)

EULER CIPRIANI VICTORINO¹, THALINE BIGHI SILVEIRA DA SILVA¹, GUILHERME HENRIQUE DA SILVA¹, ANA LUIZA SANTOS FERREIRA¹, LUIZ GONSAGA DE CARVALHO¹

¹ Departamento de Engenharia, UFLA, Lavras, MG. Fone: (35) 3829-1481

Apresentado no
XLIII Congresso Brasileiro de Engenharia Agrícola - CONBEA 2014
27 a 31 de julho de 2014- Campo Grande- MS, Brasil

RESUMO: O presente trabalho teve como objetivo comparar o desempenho das distribuições de probabilidade de Gumbel, Fréchet, Log 2p, Log 3p, Gama 2p, Gama 3p e Weibull aplicada aos dados de precipitação máxima diária anual na cidade de Barreiras, Estado da Bahia, e indicar qual a mais adequada. Os dados utilizados são da estação de número 1244011, consultada através do Hidroweb no site da Agência Nacional de Águas (ANA). A série histórica compreende os anos entre 1937 e 2013, resultando em 74 anos de dados (excluídas as falhas). Os parâmetros das distribuições foram estimados pelo método dos momentos. A avaliação da qualidade dos ajustes foi feita a partir de três testes de aderência: Kolmogorov-Smirnov, Anderson-Darling e Qui-Quadrado, sendo este último adotado como indicador de precisão do ajuste. As distribuições Log 3P, Gama 2P, Gama 3P e Weibull foram as únicas que consideradas adequadas perante todos os testes de aderência. Dentre essas a distribuição Weibull foi a mais precisa, por apresentar menor valor de Qui Quadrado calculado.

PALAVRAS-CHAVE: distribuições de probabilidade, precipitações máximas, testes de aderência

SUITABILITY OF SEVEN PROBABILITY DISTRIBUTIONS APPLIED TO MAXIMUM PRECIPITATION SERIES, AT BARREIRAS (BA)

ABSTRACT: This study aimed to compare the performance of the probability distributions of Gumbel, Fréchet, Log (2 and 3 parameters), Gamma (2 and 3 parameters) and Weibull applied to the annual maximum daily precipitation data for the city of Barreiras, Bahia state, and indicate the most appropriate of them. The data used are from the station number 1244011, accessed through HIDROWEB in the Agência Nacional de Águas (ANA) website. The historical series comprises the years between 1937 and 2013, resulting in 74 years of data (excluding failures). The method of moments estimated the distributions parameters. The quality assessment of the adjustments was taken from three tests of adherence: Kolmogorov - Smirnov, Anderson - Darling and Chi-Square, the latter used as an indicator of need adjustment being. Distributions Log 3P, Gamma 2P, Gamma 3P and Weibull were the ones considered adequate by all tests of adherence. Among these, the Weibull distribution was more precise, as it presented lower value of Chi-Square calculated.

KEYWORDS: probability distributions, maximum precipitation, adherence tests

INTRODUÇÃO: A água é um mineral presente em toda a natureza, que se renova constantemente através dos processos físicos do ciclo hidrológico. É de grande importância o conhecimento do comportamento e da distribuição das precipitações, a fim de fornecer informações que possam reduzir os impactos causados por oscilações no regime pluviométrico em uma certa região, a atividades como o emprego da irrigação ou o dimensionamento de obras hidráulicas (ANDRADE et al., 1998).

Sendo a ocorrência de precipitação pluvial máxima um processo aleatório, que não permite uma previsão determinística com grande antecedência, o ajuste de um modelo probabilístico que melhor descreva o processo se faz necessário (HARTMANN; MOALA; MENDONÇA, 2011). Segundo MURTA et al. (2005) e BACK (2001), o procedimento que costuma ser adotado para a estimativa de

chuva de projeto consiste no ajuste de uma distribuição teórica aos dados observados e após a comprovação da aderência dos dados à distribuição, extrapolam-se os valores de precipitações extremas, associadas a uma certa probabilidade.

Segundo estudos realizados por SANSIGOLO (2008), em que o autor comparou os ajustes de diversas distribuições de probabilidade às precipitações máximas diárias coletadas em Piracicaba (SP), a distribuição Gumbel se mostrou mais adequada para descrever e extrapolar extremos de precipitação diária para longos períodos de retorno naquela região. HARTMANN et al. (2011) também obtiveram bons resultados no ajuste da distribuição Gumbel aos dados de precipitações máximas mensais em Presidente Prudente (SP), representando com bastante fidelidade o regime de chuvas dessa região. Já RIBEIRO et al. (2007) ajustaram diferentes distribuições de probabilidade aos dados de precipitações decendiais e mensais na região de Barbacena (MG) e concluíram que a distribuição Gama se ajustou melhor à série histórica.

Sendo assim, esse trabalho foi realizado com o objetivo de comparar diferentes distribuições de probabilidade, sendo elas Gumbel, Fréchet, Log 2p, Log 3p, Gama 2p, Gama 3p e Weibull, e indicar a que se ajusta melhor aos dados de precipitação máxima diária mensal na região de Barreiras, no Estado da Bahia.

MATERIAL E MÉTODOS: Os dados utilizados nesse trabalho correspondem às series históricas de precipitação (mm) na região de Barreiras (BA), obtidos através do Hidroweb, no website da Agência Nacional de Águas (ANA). A estação consultada foi a de identificação 1244011, situada a uma latitude de -11:9:16 e longitude de -45:0:33. A série histórica é composta de dados compreendidos entre os anos de 1937 e 2013, de onde foram extraídas as máximas diárias de cada ano e classificadas em ordem crescente. Nesse período houve alguns anos de faltas de dados, sendo que no total foram aproveitados 74 anos de dados. As distribuições aplicadas foram Gumbel para máximos, Fréchet, Log a 2p, Log 3p, Gama 2p, Gama 3p e Weibull, sendo que seus parâmetros foram estimados com base no Método dos Momentos (Tabela 1).

Na avaliação da adequabilidade das distribuições foram aplicados os testes de aderência de Kolmogorov-Smirinov (KS), Anderson-Darling (AD) e Qui-Quadrado (χ^2) como 5% de nível de significância. Para a aplicação do teste KS, calcula-se as diferenças entre as frequências observadas e as frequências estimadas pela distribuição de probabilidade desejada, em seguida compara-se a o maior valor obtido com um valor tabelado em função do tamanho da amostra (n) e do nível de significância (α). A hipótese H_0 testada é de que a distribuição é adequada, e é aceita caso o valor calculado seja menor que o tabelado.

O teste de qui-quadrado consiste em agrupar os dados em classes de frequência e obter a soma do quadrado das diferenças entre as frequências observadas e teóricas (equação 1). Esse valor é comparado com χ^2 tabelado em função dos graus de liberdade e do nível de significância. Há duas maneiras de se determinar os graus de liberdade da distribuição, número de classes menos um ou número de classes menos o número de parâmetros menos um. Foi utilizado um valor intermediário de graus de liberdade. Segundo NAGHETTINI; PINTO (2007), o teste qui-quadrado é um ótimo indicador de precisão por refletir o quadrado médio do erro.

$$\chi^2_{calc} = \sum_{i=1}^n \frac{(f_{obs\ i} - f_{teórico\ i})^2}{f_{teórico\ i}} \quad (1)$$

O teste de Anderson-Darling busca ponderar melhor as caudas das distribuições, onde os testes KS e χ^2 perdem força (NAGHETTINI & PINTO, 2007). A estatística do teste segue a equação 2 abaixo:

$$AD^2 = -N - \sum_{i=1}^N \frac{(2i - 1) \{ \ln(F(x_i)) + \ln(1 - F(x_{N-1+1})) \}}{N} \quad (2)$$

O valor calculado de AD^2 é então comparado a um valor tabelado em função do nível de significância. Caso AD^2 seja maior que o valor tabelado, a distribuição não é adequada para a série de dados observados.

RESULTADOS E DISCUSSÃO: Dentre as distribuições testadas, apenas a Weibull, Log 3P, Gama 2P e Gama 3P foram consideradas adequadas à série de precipitações máximas diárias anuais (Tabela 1). Esses resultados demonstram a importância em se escolher um teste de aderência adequado para analisar os ajustes, evidenciado pelos resultados do teste KS, que rejeitou apenas a distribuição de Fréchet, em contraste com o teste AD que rejeitou 3 das distribuições ajustadas (Gumbel, Fréchet e Log 2P).

TABELA 1. Resultado dos testes de aderência aplicados a sete distribuições de probabilidade ajustadas a série de precipitações máximas diárias anuais em Barreiras (BA)

Distribuições	Kolmogorov-Smirnov (KS)	Qui-Quadrado (χ^2)	Anderson-Darling (AD)
Gumbel	0,102 A	12,54 R	1,159 R
Fréchet	0,340 R	67,15 R	23,46 R
Weibull	0,060 A	3,12 A	0,357 A
Log2P	0,116 A	7,53 A	0,849 R
Log3P	0,062 A	2,77 A	0,237 A
Gama 2p	0,081 A	5,17 A	0,420 A
Gama 3p	0,074 A	6,25 A	0,489 A

R: rejeitado a nível de significância de 5%; A: adequado a nível de significância de 5%.

Os gráficos apresentados nas figuras 1 e 2 representam visualmente a qualidade das distribuições testadas. Na figura 1 foram plotados os ajustes das distribuições consideradas adequadas em todos os testes, e na figura 2 as distribuições reprovadas. Sendo o teste Qui-Quadrado um bom indicador de precisão das distribuições (NAGHETTINI; PINTO, 2007), dentre as distribuições que se adequaram, a distribuição Log 3P foi considerada a mais precisa por apresentar um menor valor de Qui-Quadrado calculado.

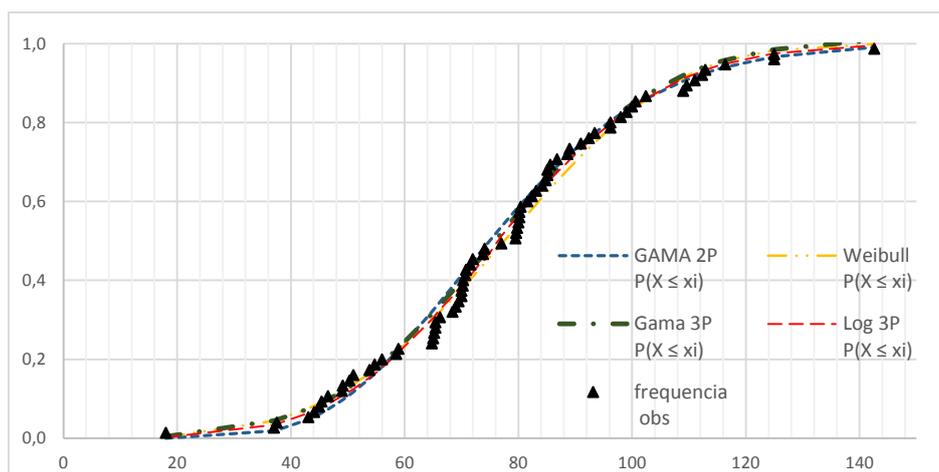


FIGURA 1. Gráfico das distribuições de probabilidade que tiveram ajuste aprovado em todos os testes de aderência aplicados.

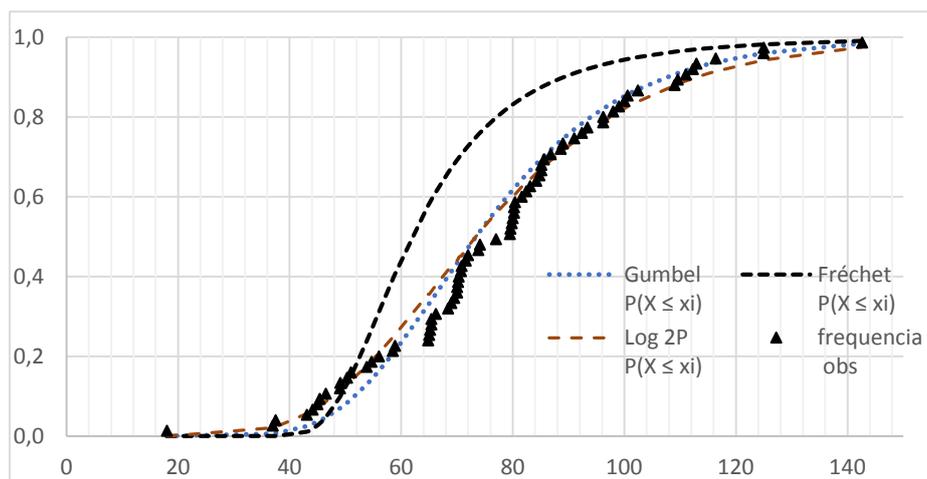


FIGURA 2. Gráfico das distribuições de probabilidade que tiveram ajuste reprovado em pelo menos um dos testes de aderência aplicados

CONCLUSÕES: As distribuições Weibull, Log 3P, Gama 2P e Gama 3P apresentaram ajustes estatisticamente adequados aos dados de precipitações máximas diárias na cidade de Barreiras (BA) sendo que Log 3P apresentou um ajuste mais preciso em relação às outras.

REFERÊNCIAS

- BACK, Á. J. Seleção de distribuição de probabilidade para chuvas diárias extremas do Estado de Santa Catarina. **Revista Brasileira de Meteorologia**, v. 16, p. 211-222, 2001.
- HARTMANN, M.; MOALA, F. A.; MENDONÇA, M. A. Study of the maximum annual precipitation in Presidente Prudente. **Revista Brasileira de Meteorologia**, v. 26, n. 4, p. 561-568, 2011.
- MURTA, R. M. et al. Precipitação pluvial mensal em níveis de probabilidade pela distribuição gama para duas localidades do sudoeste da Bahia. **Ciência e Agrotecnologia**, v. 29, n. 5, p. 988-994, 2005.
- NAGHETTINI, M.; PINTO, E. **Hidrologia Estatística**. 2007.
- RIBEIRO, B. T. et al. Comparação de distribuições de probabilidade e estimativa da precipitação provável para região de Barbacena, MG. **Ciência e Agrotecnologia**, v. 31, p. 1297-1302, 2007.
- SANSIGOLO, C. A. Distribuições de extremos de precipitação diária, temperatura máxima e mínima e velocidade do vento em Piracicaba, SP (1917-2006). **Revista Brasileira de Meteorologia**, v. 23, n. 3, p. 341-346, 2008.