Journal of Hydrology: Regional Studies 4 (2015) 123–133

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

journal homepage: www.elsevier.com/locate/ejrh

Multiparameter probability distributions for heavy rainfall modeling in extreme southern Brazil

Samuel Beskow[a,*], Tamara L. Caldeira[b], Carlos Rogério de Mello[c], Lessandro C. Faria[a], Hugo Alexandre S. Guedes[d]

[a] *Center of Technological Development/Water Resources Engineering, Federal University of Pelotas, 01 Gomes Carneiro Street, CEP 96010-610 Pelotas, RS, Brazil*

[b] *Center of Technological Development/Post-Graduate Program in Water Resources, Federal University of Pelotas, 01 Gomes Carneiro Street, CEP 96010-610, Pelotas, RS, Brazil*

[c] *Soil and Water Engineering Group, Federal University of Lavras, C.P. 3037, CEP 37200-000 Lavras, MG, Brazil*

[d] *Engineering Center/Civil Engineering, Federal University of Pelotas, 987 Benjamin Constant Street, CEP 96010-020 Pelotas, RS, Brazil*

A R T I C L E   I N F O

Article history:
Received 13 October 2014
Received in revised form 15 May 2015
Accepted 8 June 2015
Available online 23 June 2015

Keywords:
Extreme rainfall events
Probabilistic modeling
Kappa
GEV

A B S T R A C T

Study region: The study was conducted in the Rio Grande do Sul state – Brazil.
Study focus: Studies about heavy rainfall events are crucial for proper flood management in river basins and for the design of hydraulic infrastructure. In Brazil, the lack of runoff monitoring is evident, therefore, designers commonly use rainfall intensity–duration–frequency (IDF) relationships to derive streamflow-related information. In order to aid the adjustment of IDF relationships, the probabilistic modeling of extreme rainfall is often employed. The objective of this study was to evaluate whether the GEV and Kappa multiparameter probability distributions have more satisfying performance than traditional two-parameter distributions such as Gumbel and Log-Normal in the modeling of extreme rainfall events in southern Brazil. Such distributions were adjusted by the L-moments method and the goodness-of-fit was verified by the Kolmogorov–Smirnov, Chi-Square, Filliben and Anderson–Darling tests.
New hydrological insights for the region: The Anderson–Darling and Filliben tests were the most restrictive in this study. Based on the Anderson–Darling test, it was found that the Kappa distribution presented the best performance, followed by the GEV. This finding provides evidence that these multiparameter distributions result, for the region of study, in greater accuracy for the generation of intensity–duration–frequency curves and the prediction of peak streamflows and design hydrographs. As a result, this finding can support the design of hydraulic structures and flood management in river basins.

* Corresponding author. Tel.: +55 5339211240.
E-mail addresses: samuel.beskow@ufpel.edu.br (S. Beskow), tamaraleitzkecaldeira@gmail.com (T.L. Caldeira), crmello@deg.ufla.br (C.R. de Mello), lessandro.faria@ufpel.edu.br (L.C. Faria), hugo.guedes@ufpel.edu.br (H.A.S. Guedes).

http://dx.doi.org/10.1016/j.ejrh.2015.06.007

1. Introduction

Studies of extreme rainfall events have great relevance for water resources management, as they provide insight into the understanding of the hydrological behavior of a given watershed under the flooding point of view. In this context, engineers commonly need to establish flood control techniques and to estimate the peak streamflows or hydrographs for the design of hydraulic structures. Because it is more frequent the existence of rainfall monitoring, a good option for ungauged watersheds is to estimate a design precipitation or a design hyetograph and then, using rainfall-runoff modeling, to derive the aforementioned streamflow characteristics.

The rainfall monitoring enables registration of its temporal variability and the probabilistic modeling of extreme values, thus making it possible to estimate intensity–duration–frequency (IDF) curves. Such curves can be employed in ungauged watersheds to estimate peak streamflows through empirical models and to derive design hyetographs and hydrographs according to well-known hydrologic methods.

Daily or sub-daily rainfall data sets are used for the modeling of heavy rainfall events, the latter being the most suitable, because it allows the determination of rainfall intensities associated with different durations. However, due to the lack of sub-daily data, studies conducted in Brazil have commonly applied daily rainfall data sets through annual maximum daily rainfall series (Silva and Clarke, 2004; Souza et al., 2012; Aragão et al., 2013; Franco et al., 2014; Caldeira et al., 2015).

There are numerous probabilistic models applied to continuous random variables, such as annual maximum daily rainfall. In Brazil, the fit of more simplified theoretical probability models has been commonly observed, such as 2 and 3 parameter Log-Normal distribution, and Asymptotic Extreme Value Type I, also known as Gumbel (Silva et al., 2002; Sansigolo, 2008; Santos et al., 2009; Back et al., 2011; Souza et al., 2012; Mello and Viola, 2013; Caldeira et al., 2015). However, several studies associated with heavy rainfall events have sought also evaluate other probability distributions, such as the 2-parameter Gamma (Franco et al., 2014), Generalized Extreme Value (Durrans and Kirby, 2004; Nadarajah and Choi, 2007; Blain and Camargo, 2012; Blain and Meschiatti, 2014), Kappa (Parida, 1999; Park and Jung, 2002; Norbiato et al., 2007; Ahmad et al., 2013; Blain and Meschiatti, 2014), Wakeby (Park et al., 2001; Blain and Meschiatti, 2014), Generalized Logistic (Norbiato et al., 2007; Hailegeorgis et al., 2013; Rahman et al., 2013) and Generalized Pareto (Hailegeorgis et al., 2013; Rahman et al., 2013).

Nevertheless, a continuous random variable can be represented by more than a probabilistic model. The choice of the model that best fits the data series is performed by nonparametric tests seeking to evaluate relationship between the observed and theoretical frequencies. In hydrology, it can be highlighted the goodness-of-fit tests of Kolmogorov–Smirnov, Chi-Square, Filliben and Anderson–Darling (Sansigolo, 2008; Ben-Zvi, 2009; Back et al., 2011; Franco et al., 2014).

The methodology of statistical inference (estimation of parameters) exerts influence upon the adjustment quality of a probabilistic model. There are various methods of statistical inference such as the method of moments, maximum likelihood and L-moments. Naghettini and Pinto (2007) report that the method of moments is the simplest, but can produce low quality estimators, especially for probability distributions with three or more parameters, when compared to the maximum likelihood method. This, in turn, may be considered more efficient than the previous, yielding estimators of lower variance, however, it involves equations which are generally nonlinear and implicit. According to these authors, the L-moments method results in parameter estimators comparable in quality to those produced by the maximum likelihood method, being frequently more accurate for small samples which are commonly used in hydrologic studies.

Given the above, the objectives of this study were: (i) to evaluate the performance of multiparameter probability distributions versus those commonly used in hydrologic studies applied to the modeling of extreme rainfall events in the extreme south of Brazil; and (ii) to define the most suitable probabilistic model, based on the analysis of goodness-of-fit tests, for the random variable evaluated in this study.

2. Material and methods

2.1. Study area

The state of Rio Grande do Sul, whose area is about 282,000 km^2, is located in southern Brazil, approximately between 27° and 34° South latitude and 49° and 58° West longitude, and is bounded by the state of Santa Catarina (northeast), in Brazil, Argentina (northwest), Uruguay (southwest) and the Atlantic Ocean (southeast) (Fig. 1). The population is 10,693,929 inhabitants, distributed in 497 municipalities in such a way that 85.1% of them live in urban areas and 14.9% live in rural areas (IBGE, 2010).

Sparovek et al. (2007) conducted a study on climate in Rio Grande do Sul state using Köppen classification, reporting that such state has humid climate in all seasons of the year. Also, these researchers defined the climate as rainy temperate of the *Cfa* and *Cfb* Type, whose average temperature of the hottest month is, respectively, above and below 22 °C. The rainfall is well distributed throughout the months, with no concentration of rainfall in a given season (Mello et al., 2013).

2.2. Hydrological data

The data sets used in this study were obtained from the hydrometeorological database of the National Water Agency (ANA), through the HidroWeb – Hydrological Information System platform, and correspond to time series of daily rainfall in the state of Rio Grande do Sul.

Daily rainfall records from 342 rain gauges were used in this study (Fig. 1). Only historical series which afforded at least 10 years of records were employed. Other studies on extreme rainfall events in Brazil (Back, 2001; Santos et al., 2009; Souza et al., 2012; Aragão et al., 2013; Caldeira et al., 2015) made use of historical series following a minimum limit of 10–15 years of observations.

2.3. Probability density function (PDF)

The records observed on a daily basis allowed to compute annual maximum daily rainfall series, which were adjusted to the Gumbel, two-parameter Log-Normal (2P-LN), Generalized Extreme Value (GEV) and Kappa probability distributions, also applied by various researchers (Parida, 1999; Park

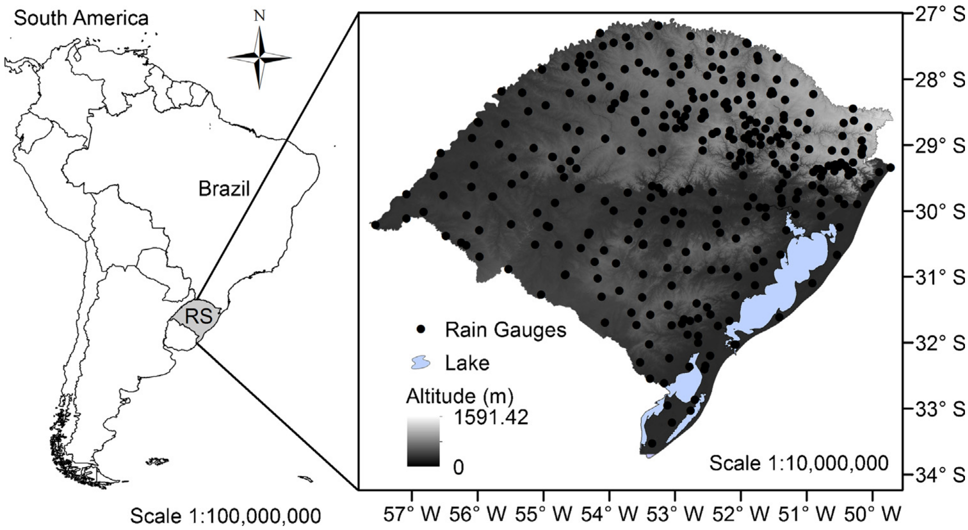

Fig. 1. Spatial distribution of rain gauges analyzed in this study.

and Jung, 2002; Durrans and Kirby, 2004; Ben-Zvi, 2009; Hailegeorgis et al., 2013) for probabilistic modeling of extreme rainfall events.

The PDF of the Gumbel distribution is expressed by:

$$f(x) = \alpha \cdot \exp^{-\alpha\cdot(x-\mu)-\exp^{-\alpha\cdot(x-\mu)}} \tag{1}$$

where x is a continuous random variable, that in this application corresponds to the annual maximum daily rainfall, α is the scale parameter and μ is the location parameter of the probability distribution in question.

The PDF of the 2P-LN distribution, whose random variable logarithms present normal distribution, is given by:

$$f(x) = \frac{1}{x\sigma_{\ln(x)}\sqrt{2\pi}} \exp^{-(1/2)[(\ln(x)-\mu_{\ln(x)})/\sigma_{\ln(x)}]^2} \tag{2}$$

where x is the annual maximum daily rainfall and e $\mu_{\ln(x)}$ and $\sigma_{\ln(x)}$ are the probability distribution parameters.

The GEV distribution, which includes three asymptotic forms of extreme values, has its PDF expressed by:

$$f(x) = \frac{1}{\alpha} \cdot \left((1-k) \cdot \left(\frac{x-\beta}{\alpha}\right)\right)^{1/(k-1)} \cdot \exp\left(-\left((1-k) \cdot \left(\frac{x-\beta}{\alpha}\right)\right)^{1/k}\right) \tag{3}$$

where x is the annual maximum daily rainfall, k is the shape parameter ranging from $-\infty$ and $+\infty$, α is the scale parameter, which varies between 0 and $+\infty$, and β is the position parameter having values between $-\infty$ and $+\infty$.

For $k > 0$, the GEV distribution is bounded above according to Eq. (4), and for $k < 0$, the lower limit of GEV is given by Eq. (5).

$$\omega = \beta + \frac{\alpha}{k} \tag{4}$$

$$\varepsilon = \beta + \frac{\alpha}{k} \tag{5}$$

With $-\infty < x < +\infty$ and $k = 0$, the GEV refers to the Type I asymptotic distribution of extreme values or the Gumbel distribution, $\varepsilon \leq x \leq +\infty$ and $k < 0$, the Type II asymptotic distribution of extreme values, or Fréchet and $-\infty < x < \omega$ and $k > 0$, the Type III asymptotic distribution of extreme value or Weibull.

The PDF of Kappa distribution (Hosking, 1994) is given by:

$$f(x) = \frac{1}{\alpha} \cdot \left[1 - \frac{k(x-\varepsilon)}{\alpha}\right]^{1/(k-1)} \cdot F(x)^{1-h} \tag{6}$$

where x is the annual maximum daily precipitation, α and ε are scale and position parameters, respectively, k and h are the shape parameters and $F(x)$ is the probability cumulative function (PCF) represented by:

$$F(x) = \left\{1 - h\left[1 - \frac{k(x-\varepsilon)}{\alpha}\right]^{1/k}\right\}^{1/h} \tag{7}$$

For $k > 0$, the random variable is bounded above at $\varepsilon + \alpha/\mathrm{k}$, if $k \leq 0$, x has no upper limit, if $h > 0$, x has a lower limit given in $\varepsilon + \alpha(1 - \mathrm{h}^{-\mathrm{k}})/\mathrm{k}$, if $h \leq 0$ and $k < 0$, x is bounded below at $\varepsilon + \alpha/\mathrm{k}$, and for $h \leq 0$ and $k \geq 0$, x has no lower bound.

Different values for the k and h shape parameters of the Kappa distribution may, in special cases, include the PDF and PCF distributions, such as Generalized Pareto ($h = 1$ and $k \neq 0$), GEV ($h = 0$ and $k \neq 0$), Generalized Logistic ($h = -1$ and $k \neq 0$), Exponential ($h = 1$ and $k = 0$), Gumbel ($h = 0$ and $k = 0$), Logistic ($h = -1$ and $k = 0$), Normal ($h = 1$ and $k = 1$) and Inverse Exponential ($h = 0$ and $k = 1$).

2.4. L-moments technique

The estimation of the probability distribution parameters was performed using the L-moments method. Valverde et al. (2004) reported that the L-moments method is similar to the conventional method of moments, however, the former derives from probability weighted moments. This method applies linear combinations of the asymmetry, the kurtosis and the coefficient of variation, thus producing a more reliable statistical system. Parida (1999) corroborates this argument when reports that this method produces more reliable estimates, particularly for small samples, besides being more robust, as is not influenced by the presence of outliers.

According to Hosking (1994), the first L-moment (λ_1) is the location parameter of the distribution and is equivalent to the arithmetical mean of the random variable; the second L-moment (λ_2) is the scale parameter of the probability distribution, related to the variance; and the coefficients τ_3 and τ_4 are a function of the probability distribution shape parameters.

To estimate the Gumbel distribution parameters by the L-moments method, Eqs. (8) and (9) were employed; the 2P-LN distribution parameters were estimated using Eqs. (10) and (11) (Mello and Silva, 2013).

$$\alpha = \frac{\ln(2)}{\lambda_2} \tag{8}$$

$$\mu = \lambda_1 - \frac{0.5772}{\alpha} \tag{9}$$

$$\sigma = 2\frac{Z}{\sqrt{2}} \tag{10}$$

$$\mu = \ln(\lambda_1) - \frac{\sigma^2}{2} \tag{11}$$

where Z is the standard normal variable corresponding to the probability $((\lambda_2/\lambda_1)+1)/2$.

The methodology for the calculation of sample L-moments, as well as to estimate the parameters of GEV and Kappa distributions, followed the requirements of the algorithm developed by Hosking (2005), in FORTRAN language, adapted to the Delphi platform for the environment of the System of Hydrological Data Acquisition and Analysis software (SYHDA) (Beskow et al., 2013a).

A particularity of Kappa distribution parameter estimates by the L-moments method, identified by Hosking (1994) and implemented in the Hosking algorithm (2005), is in the parameters k and h of the probabilistic model, which is restricted to the following conditions: $k > -1$, $h < 0$ and $h \cdot k > -1$, $h > -1$, and $k + (h/1.38) > -1$.

2.5. Goodness-of-fit of the probability density function

The adequacy of the probabilistic models to the annual maximum daily rainfall series was verified by the Kolmogorov–Smirnov (KS), Chi-Square (χ^2), Anderson–Darling (AD) and Filliben tests, all under the H_0 hypothesis that the data samples follow probability distribution tested at a significance level of 5%. Such tests have been widely used to verify adjustments of probability distributions in the field of hydrology (Sansigolo, 2008; Ben-Zvi, 2009; Back et al., 2011; Franco et al., 2014; Caldeira et al., 2015), especially the last two because they give more importance to the behavior of the PDF in the distribution tails.

The KS is based on the greatest difference, in absolute value, between the theoretical and empirical cumulative probabilities. The maximum absolute error between the theoretical and observed frequencies is compared to a tabulated critical value which depends on the sample size and the significance level. If the maximum absolute error is inferior to the tabulated critical value, it means that H_0 hypothesis is accepted (Back, 2001).

The χ^2 statistic is derived from the sum of squared differences between the observed and theoretical frequencies and its value is compared to a tabulated value, which varies as a function of the degrees of freedom and the level of significance. The calculated χ^2 can be regarded as an indicator of the PDF fit accuracy, since it corresponds to the mean square error. If the χ^2 calculated is less than the tabulated χ^2, the H_0 hypothesis is accepted (Franco et al., 2014).

Despite being quite similar to the KS test, the AD test (D'Agostino and Stephens, 1986) gives greater importance to the tails of the distributions, presenting potential in the adjustment verification of asymptotic series. In this test, the calculated value is multiplied by a correction factor, which is a function of the probability distribution. Finally, this calculated value is compared to a statistically null value, which is a function of the probability distribution and the level of significance.

The Filliben test takes into account the linear correlation coefficient between the observed records and theoretical quantiles. This coefficient is afterwards compared to a critical value, which must be lower for the H_0 hypothesis not to be rejected (Naghettini and Pinto, 2007). Given the nonexistence of critical values of the Filliben test for the Kappa distribution, tabulated critical values of the Gumbel distribution were employed. The same methodology was used to evaluate adequacy by the AD test, since this only makes it available tabulated critical values for the Normal, Log-Normal, Weibull and Gumbel distributions.

Besides the analysis of adjustments of the probabilistic models, it was also investigated in this study the influence of the theoretical probability model on estimated maximum annual daily rainfall events associated with recurrence intervals of 20, 50 and 100 years. Considering a time series set that exhibited more satisfactory adjustment for a given probability distribution, we estimated the maximum (RAE_{max}), minimun (RAE_{min}) and average (RAE_{ave}) Relative Absolute Errors that would occur if these same series were fit to other models.

2.6. Software used

The importation and manipulation of data for the preparation of the annual maximum daily precipitation series, as well as the estimation of L-moments, the adjustment of the probabilistic models and adequacy tests were conducted with the aid of a computer application called System of Hydrological Data Acquisition and Analysis (SYHDA) (Beskow et al., 2013a).

The SYHDA is a computer application that has been developed in the Hydrology and Hydrological Modeling Laboratory, of the Center of Technological Development/Water Resources Engineering at Federal University of Pelotas. It was designed to support the acquisition and analysis of rainfall and streamflow time series with the purposes of: (i) integrating to the Lavras Simulation of Hydrology (LASH) hydrologic model (Beskow et al., 2011, 2013b; Viola et al., 2013), thus facilitating the compilation of databases for the simulation and exploration of output results; and (ii) independently assisting in decision making regarding water resources (Beskow et al., 2013a).

3. Results and Discussion

The results of adequacy tests applied to the adjustment of probability distributions are shown in Table 1. They allow for a quantitative assessment of the number of distributions that were appropriate according to each test considering a significance level of 5%.

Table 1
Percentage of historical series from Rio Grande do Sul state, representing annual maximum daily rainfall, adjusted by the 2P-LN, Gumbel, GEV and Kappa probability distributions according to different adequacy tests.

Probability distribution	Kolmogorov–Smirnov (KS)		Chi-square (χ^2)		Anderson–Darling (AD)		Filliben	
	S	UnS	S	UnS	S	UnS	S	UnS
2P-LN	99.71	0.29	97.66	2.34	88.60	11.40	83.92	16.08
Gumbel	99.71	0.29	97.95	2.05	86.55	13.45	91.81	8.19
GEV	100.00	0.00	95.32	4.68	97.08	2.92	97.95	2.05
Kappa	90.94	9.06	85.09	14.91	76.90	23.10	98.54	1.46

S = suitable; UnS = unsuitable.

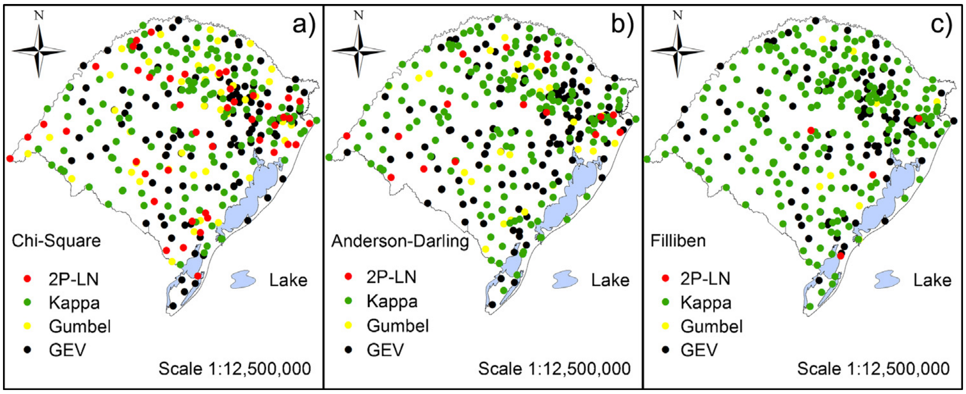

Fig. 2. Most adequate probability distribution, at a significance level of 5%, according to the adequacy tests of Chi-Square (a), Anderson–Darling (b) and Filliben (c) for each studied rain gauge in Rio Grande do Sul state.

Considering the results obtained with the KS adequacy test (Table 1), it was found that 9.06% of the series did not have an appropriate fit by the Kappa distribution, corresponding to 31 rain gauges. However, the KS test is qualitative and only allows to conclude about the suitability of the probability distributions tested, not offering a sufficient basis to compare the fit between different distributions (Mello and Silva, 2013). It is worth noting that the theoretical distribution should be completely known in this test, i.e., its parameters should not be estimated from the sample elements. When the theoretical distribution is unknown, Monte Carlo simulations demonstrate that undue rejection of the tested hypothesis (Type II error) (Naghettini and Pinto, 2007) may occur.

Although it is a not very rigorous nonparametric test (Table 1), given the number of acceptances of the null hypothesis, the KS test has been widely used to verify adequacy of the maximum rainfall series to the probability distributions (Back, 2001; Back et al., 2011; Aragão et al., 2013; Caldeira et al., 2015).

Likewise, it is evident in Table 1 that, according to the χ^2 adequacy test, a high percentage of series were not suited to Kappa distribution. This fact can be attributed to the degrees of freedom of the statistic measure, since it is defined by the number of parameters of the probabilistic model. The number of parameters causes, for the same data series, the value of tabulated χ^2 to be more restrictive when analyzing a distribution with a higher number of parameters, such as Kappa distribution. However, the comparative scenario among the four distributions (Fig. 2a) allowed to observe that the lowest values of the χ^2 statistic were obtained from the Kappa probabilistic model. According to this analysis, Kappa distribution was adjusted satisfactorily to 133 historical series, followed by the GEV, Gumbel and 2P-LN distributions, which presented the lowest values of the χ^2 statistic for, respectively 114, 47 and 44 rain gauges.

The number of historical series not suitable for the Kappa probabilistic model was considerable for the AD adequacy test (Table 1). Nevertheless, when the comparative scenario among the four distributions was plotted considering the calculated AD values (Fig. 2b), it was found that the Kappa had a more satisfactory fit in 48.24% of the series, followed by the GEV distribution with 36.84% of the series. The Gumbel and 2P-LN probability distributions resulted in better adjustment for only 27 and 17 rain gauges, respectively.

Adjustments of the Kappa and GEV probabilistic models were also verified through the Filliben test (Fig. 2c). Using such test, the results pointed to a more satisfactory fit for the first model in 224 rain gauges and the second in 103 rain gauges, corresponding to 95.61% of the series. Considering the Gumbel and 2P-LN distributions, both were satisfactorily fitted to 8 and 4 rain gauges, respectively.

According to Naghettini and Pinto (2007), the ability to discern between truthfulness or not of a hypothesis through adequacy tests is greatly reduced in the lower and upper tails, especially regarding the KS and χ^2 statistics, and this mainly occurs because of the low number of records in samples.

A limitation of the χ^2 test, depending on the sample size of continuous random variable, is in the number and amplitude of the histogram classes. When analyzing the results in this study, it was found that the time series containing a limited number of records concentrated high frequencies observed in a given class. This made the sum of the errors between the observed and theoretical frequencies to tend toward zero, which could result in undue acceptance of the null hypothesis, i.e., Type I errors.

Analyzing the Gumbel, Gamma and GEV distributions to extreme rainfall series in the Verde river basin, in the state of Minas Gerais – Brazil, Franco et al. (2014) found that the KS and χ^2 tests were not accurate when evaluating asymptotic series, recommending that they be avoided to reduce Type I errors. Sansigolo (2008) adjusted various probability distributions to maximum and minimum temperature, wind speed and maximum daily rainfall data, in the city of Piracicaba (São Paulo). The adequacy of these models was verified by KS and χ^2 tests and, to evaluate the extreme tails of the distributions, the percentile–percentile and quantile–quantile plots were used. According to Sansigolo (2008), χ^2 and KS were suitable only to the central part of the distribution. Ben-Zvi (2009), in a study of extreme rainfall events through partial series, concluded that the use of robust adequacy tests, particularly AD, was important.

The above corroborates the discussion of this present work because: (i) KS test did not allow to establish the best distribution fit to the series analyzed; and (ii) χ^2 test, despite allowing a comparative study, did not result in the desired accuracy when applied to verify the suitability of asymptotic distributions. Franco et al. (2014) reported that the AD test stood out for assigning greater weight to the tails of the distributions, and the Filliben test was the most rigorous among those employed in their study. Considering the findings obtained in this study and the statement above affirmed by Franco et al. (2014), AD and Filliben tests were established as criteria to support the indication of a probability distribution. A brief analysis on the number of acceptances of the null hypothesis (Table 1) made it possible to infer that the performance of the AD test was superior to the Filliben test in this study, because the former was more restrictive than the latter.

The AD test indicated a more satisfactory fit of the Kappa distribution, followed by the GEV, to the annual maximum daily rainfall series. In contrast, there has been commonly found in the Brazilian scientific literature applications mainly of the Gumbel distribution for asymptotic series with parameters estimated by the method of moments and adequacy verified by the KS test.

Souza et al. (2012) used the Gumbel probabilistic model in the adjustment of different extreme rainfall equations for 74 municipalities in the state of Pará. In the state of Mato Grosso do Sul, Santos et al. (2009) fitted intensity–duration–frequency equations through the Gumbel probabilistic model from the annual maximum daily rainfall series at 109 rain gauges. The mapping of heavy rainfall in Minas Gerais state was carried out by Mello and Viola (2013) by employing the Gumbel distribution.

According to Back (2001), many authors assume the hypothesis that the analyzed historical series follow the Gumbel distribution without testing it or verifying if another distribution produces more reliable results. The KS test results obtained by Silva et al. (2002) to model the heavy rainfall events in the state of Bahia through sub-daily records, indicated that among the Gumbel, 2 and 3 parameter Log-Normal, Pearson and Log-Pearson distributions, the former gave the best fit. Silva et al. (2003) adjusted intensity–duration–frequency equations from sub-daily rainfall records in the state of Tocantins, and found results that corroborate Silva et al. (2002), since they employed KS to test the same probability distributions and found that the Gumbel presented the most satisfactory fit.

However, Aragão et al. (2013) applied the Gumbel and Weibull distributions to annual maximum rainfall series from 48 stations in the state of Sergipe and could conclude that Weibull provided the best results. Ben-Zvi (2009) tested the distributions of Gumbel, Log-Normal and Generalized Extreme Value to determine intensity–duration–frequency curves for rain gauges of Israel, noting that the latter model presented the most appropriate adjustment.

Rahman et al. (2013) applied the L-moments method successfully to estimate the parameters of the GEV, Generalized Logistic and Generalized Pareto distributions for modeling of the annual maximum daily rainfall from 68 stations in Bangladesh. In a study by Back (2001), it was found that the series that had low skewness and kurtosis were better adjusted to the 3-parameter Log-Normal distribution, while the series with high skewness and kurtosis resulted in more satisfactory adjustment to Log-Pearson distribution, followed by the 2-parameter Log-Normal. In the study of Franco et al. (2014)

Table 2
Basic descriptive statistics of the probability distribution parameters, considering the suitable series according to the AD test at a significance level of 5%.

Distribution	Number of stations	Parameter	Minimum	Maximum	Mean	Standard deviation	Coefficient of variation (%)
Kappa	263	ε	−311.12	114.90	70.71	58.10	82.16
		α	6.57	943.13	40.84	108.72	266.18
		k	−0.33	3.03	0.23	0.41	178.99
		h	−1.00	2.44	0.20	0.78	394.57
GEV	332	α	7.19	38.51	20.97	5.35	25.50
		β	47.47	109.64	81.38	10.62	13.05
		k	−0.47	0.65	0.04	0.18	482.11
2P-LN	303	$\mu_{\ln(x)}$	4.00	4.84	4.49	0.13	2.90
		$\sigma_{\ln(x)}$	0.13	0.63	0.27	0.06	21.09
Gumbel	296	μ	47.31	111.41	81.37	10.54	12.96
		α	0.01	0.13	0.05	0.02	29.27

in Minas Gerais state, they concluded from adequacy tests that the GEV, with parameters estimated by the L-moments method, was that with the best fit. Blain and Meschiatti (2014) compared the performance of GEV, Kappa and Wakeby distributions to model maximum annual rainfall of 1, 2 and 3 days through data from São Paulo state. They concluded that the Kappa, with parameters estimated by the L-Moments method, and GEV with parameters estimated by the maximum likelihood and L-moments methods, presented the best adjustments.

Table 2 displays the results obtained by a statistical analysis on the parameters estimated for each historical series of daily rainfall records, following the adjustment suggested by the AD test. It can be inferred from the coefficient of variation that the Kappa distribution presented, in general, greater variability among the parameters. This fact can be tied to the suitable fit of this distribution and to the better representation of the frequency distribution of the sample data.

The analysis of the shape parameters (k and h) estimated for the Kappa distribution allowed to identify that only one series characterized the PDF as a Generalized Pareto distribution and 28.36% of the series characterized the Generalized Logistic distribution, where $h = -1$ and $k \neq 0$. In similar studies, involving monsoon rainfall modeling in India (Parida, 1999) and Pakistan (Ahmad et al., 2013), it was observed that 50% of the rainfall series analyzed in the first case were framed in this particularity, while in the second case, there was no occurrence.

Due to the results of adequacy tests applied in this study, as well as the variability of the parameters of probabilistic models (Table 2), it was conducted an analysis of the influence of the probability distribution on the estimated maximum annual daily rainfall associated to different recurrence intervals. Table 3 summarizes the maximum (RAE_{max}), minimum (RAE_{min}) and average Relative Absolute Errors (RAE_{ave}) that occurred when estimating a quantile by a probabilistic model, compared to the same quantile estimated by the model that obtained the most satisfactory adjustment according to the AD test.

The results obtained from the AD test (Table 3) and the Kappa probabilistic model indicated more satisfactory adjustment to 165 out of the 342 series analyzed. It was noted that such series, upon being adjusted to the Gumbel and 2P-LN distributions, tended, on average, to under or overestimate the annual maximum daily rainfall by approximately 11%, when it was associated with a recurrence interval of 100 years. In other words, the quantile of 1% probability of exceedance in this case would be estimated by Gumbel and 2P-LN with an approximate RAE_{med} of 11% compared to the same quantile for the Kappa probabilistic model. Also noteworthy are the noticeable RAE_{max} values produced by 2P-LN when the Kappa was the most suitable distribution according to the AD test at a significance level of 5%.

The results presented here provided evidences of how important the adjustment of various theoretical probability models is, from the most simplified to the multiparameter, in order to get the one that best represents the frequency distribution of the sample data. Furthermore, the use of robust

Table 3
Absolute errors in the estimation of the annual maximum daily rainfall associated with a recurrence interval (RI), through a probabilistic model, in relation to the best adjusted distribution, according to results reported by AD adequacy test.

Best PDF	RI (years)	RAE_{max} (%)			RAE_{ave} (%)			RAE_{min} (%)		
		20	50	100	20	50	100	20	50	100
Kappa	Kappa	–	–	–	–	–	–	–	–	–
	GEV	4	10.579	16.380	0.5975	2.731	4.987	0.0015	0.040	0.050
	2P-LN	127.156	190.349	241.921	4.391	8.215	11.589	0.008	0.033	0.013
	Gumbel	18.712	33.673	44.900	3.883	7.751	11.287	0.064	0.049	0.009
GEV	Kappa	64.674	81.184	93.758	9.111	11.936	15.109	0.033	0.006	0.048
	GEV	–	–	–	–	–	–	–	–	–
	2P-LN	22.984	38.990	51.810	4.152	8.830	12.538	0.101	0.008	0.038
	Gumbel	250.879	312.763	356.799	5.520	9.962	13.809	0.019	0.023	0.157
2P-LN	Kappa	10.108	19.423	26.179	3.104	7.120	10.944	0.173	0.036	0.853
	GEV	4.665	9.597	13.348	1.243	2.055	2.840	0.028	0.026	0.220
	2P-LN	–	–	–	–	–	–	–	–	–
	Gumbel	8.093	9.310	11.223	3.224	3.993	4.886	0.224	0.219	0.014
Gumbel	Kappa	60.593	94.475	129.865	11.986	17.999	24.229	0.142	0.494	1.126
	GEV	5.934	20.107	36.386	2.317	4.255	6.2329	0.032	0.090	0.119
	2P-LN	6.891	8.961	10.627	2.194	3.525	4.533	0.048	0.171	0.051
	Gumbel	–	–	–	–	–	–	–	–	–

adequacy tests for asymptotic series elucidated that the adjustment of probabilistic models should be analyzed with caution. Table 3 corroborates this discussion, since there was a large discrepancy between the quantiles estimated by a distribution that did not fit properly, when these were compared to those that were estimated by a distribution with a more satisfactory adjustment.

The use of inadequate probabilistic models may culminate in the estimate of unreliable quantiles and inappropriate intensity–duration–frequency equations. These distortions can occur due to a sum of errors such as the unsatisfactory choice of the probability distribution and errors intrinsic to the methodology. Such factors may reveal that spurious rainfall estimates may impact infrastructure and best management practice.

4. Conclusions

Based on results presented here, it can be concluded that:

(i) The Kappa, GEV, Gumbel and 2P-LN probability distributions, with parameters estimated by the L-moments method, were suitable for the modeling of extreme rainfall events in the state of Rio Grande do Sul – Brazil, however, the Kappa was found to be the most indicated, followed by the GEV distribution.
(ii) The AD adequacy test was the most suitable for assessing the goodness-of-fit of the probability models to the historical series analyzed, as this proved to be more restrictive than Filliben, KS, and χ^2 tests.
(iii) The multi-parameter distributions can lead to greater accuracy for the generation of intensity–duration–frequency curves as well as the estimation of peak streamflows and design hyetographs and hydrographs, thus supporting the design of hydraulic structures and flood management in ungauged watersheds.

References

Ahmad, I., Shah, S.F., Mahmood, I., Ahmad, Z., 2013. Modeling of monsoon rainfall in Pakistan based on Kappa distribution. Sci. Int. 25 (2), 333–336.
Aragão, R., Santana, G.R., Costa, C.E.F.F., Cruz, M.A.S., Figueiredo, E.E., Srinivasan, V.S., 2013. Chuvas intensas para o estado de Sergipe com base em dados desagregados de chuva diária. Rev. Bras. Eng. Agr. Amb. 17 (3), 243–252.

Back, A.J., 2001. Seleção de distribuição de probabilidade para chuvas diárias extremas do estado de Santa Catarina. Rev. Bras. Meteorol. 16 (2), 211–222.
Back, A.J., Hen, A., Oliveira, J.L.R., 2011. Heavy rainfall equations for Santa Catarina, Brazil. Rev. Bras. Ciênc. Solo 35, 2127–2134.
Ben-Zvi, A., 2009. Rainfall intensity–duration–frequency relationships derived from large partial duration series. J. Hydrol. 367, 104–114.
Beskow, S., Mello, C.R., Norton, L.D., Silva, A.M., 2011. Performance of a distributed semi-conceptual hydrological model under tropical watershed conditions. Catena 86 (3), 160–171.
Beskow, S., Correa, L.L., Mahl, M., Simões, M.C., Caldeira, T.L., Nunes, G.S., Hund, E.L., Faria, L.C., Mello, C.R., 2013a. Desenvolvimento de um sistema computacional de aquisição e análise de dados hidrológicos. In: XX Simpósio Brasileiro de Recursos Hídricos, 17–22 Novembro 2013 Bento Gonçalves. CD-rom.
Beskow, S., Norton, L.D., Mello, C.R., 2013b. Hydrological prediction in a tropical watershed dominated by oxisols using a distributed hydrological model. Water Resour. Manag. 27, 341–363.
Blain, G.C., Camargo, M.B.P., 2012. Probabilistic structure of an annual extreme rainfall series of a coastal area of the state of São Paulo, Brazil. Eng. Agríc. 32 (3), 552–559.
Blain, G.C., Meschiatti, M.C., 2014. Using multi-parameters distributions to assess the probability of occurrence of extreme rainfall data. Rev. Bras. Eng. Agríc. Amb. 18 (3), 307–313.
Caldeira, T.L., Beskow, S., Mello, C.R., Faria, L.C., Souza, M.R., Guedes, H.A.S., 2015. Modelagem probabilística de eventos de precipitação extrema no estado do Rio Grande do Sul. Rev. Bras. Eng. Agríc. Amb. 19 (3), 197–203.
D'Agostino, R.B., Stephens, M.A., 1986. Goodness-of-fit Techniques. Marcel Dekker, New York.
Durrans, S.R., Kirby, J.T., 2004. Regionalization of extreme precipitation estimates for the Alabama rainfall atlas. J. Hydrol. 295, 101–107.
Franco, C.S., Marques, R.F.P.V., Oliveria, A.S., Oliveira, L.F.C., 2014. Distribuição de probabilidades para precipitação máxima diária na Bacia Hidrográfica do Rio Verde, Minas Gerais. Rev. Bras. Eng. Agríc. Amb. 18 (7), 735–741.
Hailegeorgis, T.T., Thorolfsson, S.T., Alfredsen, K., 2013. Regional frequency analysis of extreme precipitation whit consideration of uncertainties to update IDF curves for the city of Trondheim. J. Hydrol. 498, 305–318.
Hosking, J.R.M., 1994. The four-parameter Kappa distribution. IBM J. Res. Dev. 38 (3), 251–258.
Hosking, J.R.M., 2005. FORTRAN Routines for Use With the Method of L-moments. Version 3.04, Rep. No. RC 20525 (90933). IBM Research Division, T.J. Watson Research Center, Yorktown Heights, NY.
IBGE – Intituto Brasileiro de Geografia e Estatística, 2010. Censo demográfico 2010 [online], available from: http://censo2010.ibge.gov.br/estadosat (accessed 02.09.14).
Mello, C.R., Silva, A.M., 2013. Hidrologia: princípios e aplicações em sistemas agrícolas. UFLA, Lavras.
Mello, C.R., Viola, M.R., 2013. Mapeamento de chuvas intensas no estado de Minas Gerais. Rev. Bras. Ciênc. Solo. 37, 37–44.
Mello, C.R., Viola, M.R., Beskow, S., Norton, L.D., 2013. Multivariate models for annual rainfall erosivity in Brazil. Geoderma 202–203, 88–102.
Nadarajah, S., Choi, D., 2007. Maximum daily rainfall in South Korea. J. Earth Syst. Sci. 116 (4), 311–320.
Naghettini, M., Pinto, E.J.A., 2007. Hidrologia estatística. CPRM, Belo Horizonte.
Norbiato, D., Borga, M., Sangati, M., Zanon, F., 2007. Regional frequency analysis of extreme precipitation in the eastern Italian Alps and the August 29, 2003 flash flood. J. Hydrol. 345, 149–166.
Parida, B.P., 1999. Modelling of Indian summer monsoon rainfall using a four-parameter Kappa distribution. Int. J. Climatol. 19, 1389–1398.
Park, J.S., Jung, H.S., 2002. Modelling Korean extreme rainfall using a Kappa distribution and maximum likelihood estimate. Theor. Appl. Climatol. 72, 55–64.
Park, J.S., Jung, H.S., Kim, R.S., Oh, J.H., 2001. Modelling summer extreme rainfall over the Korean Peninsula using Wakeby distribution. Int. J. Climatol. 21, 1371–1384.
Rahman, M.M., Sarkar, S., Najafi, M.R., Rai, R.K., 2013. Regional extreme rainfall mapping for Bangladesh using L-moment technique. J. Hydrol. Eng. 18, 603–615.
Sansigolo, C.A., 2008. Distribuições de extremos de precipitação diária, temperatura máxima e mínima e velocidade do vento em Piracicaba, SP (1917–2006). Rev. Bras. Meteorol. 23 (3), 341–346.
Santos, G.G., Figueiredo, C.C., Oliveira, L.F.C., Griebeler, N.P., 2009. Intesidade-duração-frequência de chuvas para o Estado de Mato Grosso do Sul. Rev. Bras. Eng. Agríc. Amb. 13 (Suppl.), 899–905.
Silva, B.C., Clarke, R.T., 2004. Análise estatística de chuvas intensas na Bacia do Rio São Francisco. Rev. Bras. Meteorol. 19 (3), 265–272.
Silva, D.D., Gomes Filho, R.R., Pruski, F.F., Pereira, S.B., Novaes, L.F., 2002. Chuvas intensas no Estado da Bahia. Rev. Bras. Eng. Agríc. Amb. 6 (2), 362–367.
Silva, D.D., da Pereira, S.B., Pruski, F.F., Gomes Filho, R.R., Lana, A.M.Q., Baena, L.G.N., 2003. Equações de Intensidade-Duração-Frequência da precipitação pluvial para o estado de Tocantis. Eng. Agric. 11 (1–4), 7–14.
Souza, R.O.R., Scaramussa, P.H.M., Amaral, M.A.C.M., Pereira Neto, J.A., Pantoja, A.V., Sadeck, L.W.R., 2012. Equação de chuvas intensas para o Estado do Pará. Rev. Bras. Eng. Agríc. Amb. 16 (9), 999–1005.
Sparovek, G., Van Lier, Q.J., Dourado Neto, D., 2007. Computer assisted Köppen climate classification: a case study for Brazil. Int. J. Climatol. 37, 257–266.
Valverde, A.E.L., Leite, H.G., Silva, D.D., Pruski, F.F., 2004. Momentos-L: teoria e aplicação em hidrologia. Rev Árvore 28 (6), 927–933.
Viola, M.R., Mello, C.R., Beskow, S., Norton, L.D., 2013. Applicability of the LASH Model for hydrological simulation of the Grande River Basin, Brazil. J. Hydrol. Eng. 18, 1639–1652.