

Establishment of IDF-curves for precipitation in the tropical area of Central Africa – comparison of techniques and results

B. Mohymont¹, G. R. Demarée¹, and D. N. Faka²

¹Department of Meteorological Research and Development, Royal Meteorological Institute of Belgium, Ringlaan 3, B-1180 Brussels, Belgium

²National Agency of Météorology and Télédétection by Satellite (METTELSAT), Kinshasa, Congo

Received: 25 September 2003 – Revised: 26 February 2004 – Accepted: 23 March 2004 – Published: 28 May 2004

Abstract. The establishment of Intensity-Duration-Frequency (IDF) curves for precipitation remains a powerful tool in the risk analysis of natural hazards. Indeed the IDF-curves allow for the estimation of the return period of an observed rainfall event or conversely of the rainfall amount corresponding to a given return period for different aggregation times.

There is a high need for IDF-curves in the tropical region of Central Africa but unfortunately the adequate long-term data sets are frequently not available. The present paper assesses IDF-curves for precipitation for three stations in Central Africa. More physically based models for the IDF-curves are proposed. The methodology used here has been advanced by Koutsoyiannis et al. (1998) and an inter-station and inter-technique comparison is being carried out.

The IDF-curves for tropical Central Africa are an interesting tool to be used in sewer system design to combat the frequently occurring inundations in semi-urbanized and urbanized areas of the Kinshasa megapolis.

1 Introduction

The purpose of this study is mainly to produce IDF-curves for precipitation for three different climatological stations in Congo in Central Africa. These stations are respectively: Kinshasa–Binza, Kinshasa–Ndjili and Yangambi. Indeed long-term and high temporal resolution time-series of extreme values of precipitation intensities are rather rare in that part of the world and there is a high need for that type of information.

In the second section the rainfall climates at the stations are described with emphasis on the occurrence of extreme precipitation events while the non-recording and recording raingauges and their locations used in this paper are described in the Sect. 3. In Sects. 5–7, a number of statisti-

cal hypotheses in the context of the annual extreme values, such as the independence of the data and the selection of the probability distribution function, are dealt with in some detail. For the assessment of the IDF-curves of precipitation, the framework of Koutsoyiannis et al. (1998) is used. In that innovative paper new ideas were advanced and, in particular, three essentially different techniques were proposed to assess the IDF-curves of precipitation. The implementation of these techniques in the present paper has shown that the three techniques provide similar results what concerns the errors (see Sect. 8). This result was obtained as well as for short-term (typically 15 to 20 years) as for medium-term (typically 30 to 50 years) data sets of yearly extremes in climatological problems. Section 9 contains the conclusions of the study.

In recent studies, various authors are tempting to relate the IDF-relationship to the synoptic meteorological conditions in the area of the stations (see e.g. Llasat, 2001). In the Sect. 9 of the present paper the results obtained for the three rainfall stations from tropical Central Africa are compared. The Ukkel (Belgium) rainfall intensity data are used here solely as a basis for comparison (Demarée, 1985; Buishand and Demarée, 1990; Willems, 2000). The comparison between the IDF-curves from three tropical rainfall stations from Central Africa with the maritime temperate station of Ukkel enabled to have a better insight in some of the physical aspects of the modelling.

2 The rainfall climate and its extreme at Kinshasa and Yangambi (Congo)

The precipitation climate for the Kinshasa region of Congo is that of a tropical warm and humid climate. There are two distinct seasons: a rainy season from September to the end of May with two extreme months of precipitation, November and April, and with less precipitation in January. The dry season ranges from June till September. In the first half of September, it starts to rain again while the rainfall rates gradually increase in the second half of the month. The yearly

Table 1. Number of raindays at Kinshasa-Binza, reference period 1961–1970 (Crabbé, 1980).

January	February	March	April	May	June	July	August	September	October	November	December	Year
11	11	14	14	10	1	1	1	4	10	18	13	106

Table 2. Average number of raindays at the station of Yangambi–km5 for the reference period 1931–1980 (Ruwet et al., 1985).

January	February	March	April	May	June	July	August	September	October	November	December	Year
9	9	14	15	15	12	13	14	16	18	17	13	165

number of rainy days is about one hundred; this consists of about one day upon two in March, April, November, one day upon three in January, February, May and October, and none in June, July and August. Table 1 gives the average number of raindays at Kinshasa-Binza for the reference period 1961–1970 (see Crabbé, 1980).

The Kinshasa climate is noted (Aw4)_S in the Köppen climate typology. The class A defines the tropical humid climates. The letters (w) and (S) define the presence of a dry period in the southern winter and the number 4 stands for the length of the dry season (Bultot, 1950).

Large amounts have been registered by recording rain-gauges in the Kinshasa region: 34 mm in 10 min, 55 mm in 20 min, 65 mm in 30 min and 80 mm in 1 h were observed in the reference periods considered in this paper. Fortunately, the heavy rainfall events do not last longer than half-an-hour to an hour but may be followed by a light rain that may last several hours. The heavy rainfall events usually produce locally inundations in the paved parts of the town and where roads converge and may have disastrous consequences in the peripheral hills (Crabbé, 1980).

The station of Yangambi is located in the climatic zone Af in the Köppen climate classification. This zone is defined as having a monthly rainfall amount larger than 60 mm for its driest month. This is typically the domain of the equatorial ombrophile forest (Bultot, 1950). Yangambi is located in the Central Basin of the Congo River, in the Upper Congo Region, along the Congo River about 75 km westward of the city of Kisangani (see Fig. 1). The station does not experience a real dry season; however, one notes two rainfall minima, one in the month of January, the other one in the month of July. Table 2 gives the average number of raindays at the station Yamgambi-km5 for the reference period 1931–1980 (see Ruwet et al., 1985).

Finally, the station of Uccle, a Brussels suburb, is located in the central part of Belgium and has a maritime temperate climate. Rainfall showers due to convective activity occur in the summer months while in the winter long-lasting precipitation of frontal origin is predominant.

3 Description of the raingauges

The different climatological stations of Kinshasa and some peripheral stations have been briefly described by Crabbé (1980).

The station of Kinshasa-Binza is located at 15°15' E–4°22' S and at an elevation of 440 m. The station is at the headquarters of the “Direction générale de l'Agence Nationale de Météorologie et de Télédétection par Satellite (METTELSAT)”, on the hill of Ngaliema, and commonly named Binza-météo, in the southwest of the Congo capital. The station functions since the 1 June 1955. The relief of the station is slightly undulating. The large number of trees at a distance of at least 200 m all around the site has a wind breaking effect and the neighbouring buildings are sufficiently away. The neighbourhood is practically exempt of important obstacles. The presently available time-series covers the time span 1961 to 2001 for the monthly rainfall totals and 1977–2000 for the rainfall intensity data.

The station of Ndjili is located 15°22' E–4°22' S, at an elevation of 310 m, at the international airport of Kinshasa at the east of the same town. The station was initiated in a period preliminary to the founding of the airport and has been definitely put in function on the 1 March 1959 and functions since then without interruption. The presently available time-series is 1961–2001 for the monthly rainfall depths and 1977 to 1993 for the rainfall intensities.

The station of Kinshasa-Ouest (previously named Léopoldville-Ouest) was situated at an elevation of 358 m on the hill of Ngaliema in the southwest of the town of Kinshasa. The station has functioned from 1929 to 1960 under the administration of the “Ministère de l'Agriculture”. The station Kinshasa-Ouest was installed on a shelf of the hill at a distance of 5 km from Kinshasa-Binza. The raingauge was of the Mini-Agri type and with an opening of 4 dm² (Vandenplas, 1943).

The time series of the daily precipitation extremes at Kinshasa-Binza (1929–2000) is constituted of the merging of the corresponding data from the two stations on the Ngaliema hill. From the point of view of the statistical analysis of

the annual extremes, the homogeneity of the complete series may be inferred.

The precipitation intensity data are deduced from the rainfall charts originating from siphoning type recording rainfall gauges with the rim of the receiver at a level of 1.50 m above the ground, with an opening of 2 dm² and siphoning at an amount of 10 mm. The time of the siphoning is of the order of 10 s, which may induce considerable measurement errors during heavy rainfall episodes. The analysis of the rainfall charts is carried out on a 5 min time step at the exact time of the start of the rainfall event. The maximum rainfall depths were searched for the durations of 5, 10, 20, 30, 60 min. The yearly maxima were deduced from the monthly maxima. Besides the rainfall intensity data, a non-recording raingauge of the IRM type with an opening of 1 dm² provided the daily precipitation depths. The daily rainfall depths were taken for the time period from 6:00 UT to 6:00 UT the next day. These values were then associated with the rainfall intensities in 1440 min.

The station of Yangambi-Km 5 is located at 00°49' N–24°29' E at an elevation of 470 m. The Yangambi station is the central station of the “Institut National pour l'Étude et la Recherche agronomiques (INERA)” which was founded in 1933 under the name of “Institut National pour l'Étude agronomique du Congo belge (INEAC)”. The climatological observations as well as other activities were interrupted at Yangambi in the first days of December 1964 due to the Simba rebellion. The observations could be resumed, after a revision of the stations, on 1 October 1965. In order to cope with such a long period of missing data, an artificial year 1964/1965 was produced by taking 11 months in 1964 and adding the month of December 1965 to produce a complete set of 12 months (Crabbé, 1970).

The recording raingauge in the INEAC network was of the siphoning system of the type “O.N.M. – Office National de Météorologie” and produced by Richard at Paris. M. Frère has ameliorated the instrument. The receiving area is the same as for the ordinary raingauge. The rim of the recipient is at 1.50 m above the ground level. The rainfall chart is changed at 08:00 MLT (Mean Local Time). However, in the case of heavy precipitation, the chart may be changed between 08:00 and 09:15 MLT at the latest. Consequently, the daily time span corresponding to the recording rain gauge extends between 08:15 MLT, the time of the provoked siphoning on the day of its placement, until approximately 08:15 MLT the next day. Therefore, the differences between daily precipitation amounts of the non-recording raingauge and from the recording raingauge have been noted. The time step used in the manual processing of the precipitation charts is 15 min. Therefore, the precipitation amounts produced by the recording raingauge need to be seen as clock-time data (Crabbé, 1971).

The ordinary raingauge used for daily readings is of the “Miniagri” type that was quite common in Congo, Rwanda and Burundi. This model is used since 1911. The opening of the recipient measures 400 cm² and its rim is at 65 cm above ground level. The precipitation is collected in a large reser-

voir in such a way that heavy rains can be easily collected. The observation hours are at 06:00 and 18:00 MLT. The observation read at 06:00 MLT being part of the ecological day of the preceding day is transcribed on the previous day while the observation read at 18:00 MLT is transcribed on the day itself (Crabbé, 1971).

Finally, only to allow for comparison of the IDF-curves, the extreme precipitation data at the station of Uccle near Brussels, Belgium, are considered. At the site of the climatological park of the Royal Meteorological Institute of Belgium, a siphoning recording raingauge of the Hellmann-Fuess type functions without noticeable interruptions since 1 March 1898. All raincharts have digitised at a time step of 10-min clock-time producing a long-term high temporal resolution time-series of more than 51/4 million data (Demarée, 2003). However, the data set used in this study is a particular case of that set corresponding to the 1934–1983 period. The main difference with the longer series is that all data of the shorter series have an arbitrary starting point. During rainfall events the highest intensive periods of 10, 20, 30, 60, 120, 360, 720 and 1440 min were looked for at a monthly time base (Réménieras, 1972). Furthermore, the extreme instantaneous intensities were graphically determined and were labelled for practical reasons as 1-min intensities. The latter data provided interesting indications for the flattening of the IDF-curves for durations going towards zero.

4 The IDF-curves

A set of Intensity-Duration-Frequency (IDF) curves constitutes a relation between the intensity (more precisely, the mean intensity) of precipitation (measured in mm/h), the duration or the aggregation time of the rainfall (in min) and the return period of the event. The return period of an event (here the rainfall amount or depth) indicates how rare/how frequent is this event and is defined by the inverse of the annual exceedance probability (the precise mathematical definition is given by formula (9) in the Sect. 8).

Denote by i the rainfall intensity (mm/h), d the duration of the rainfall (min) and T the return period (years). The IDF-relation is then expressed mathematically as follow

$$i = f(T, d). \quad (1)$$

The rainfall intensity is a function of the variables T and d . We consider in this paper only expressions of the type:

$$i = \frac{a(T)}{b(d)}. \quad (2)$$

In Eq. (2), the dependency in d and the dependency in T are modelled by two separate equations $a(T)$ and $b(d)$. The relation (2) is moreover parameterised:

$$i = \frac{a(T, \mathbf{p1})}{b(d, \mathbf{p2})} \quad (3)$$

where $\mathbf{p1}$ and $\mathbf{p2}$ are two vectors of parameters. The estimation of the IDF-curves results in the estimation of those

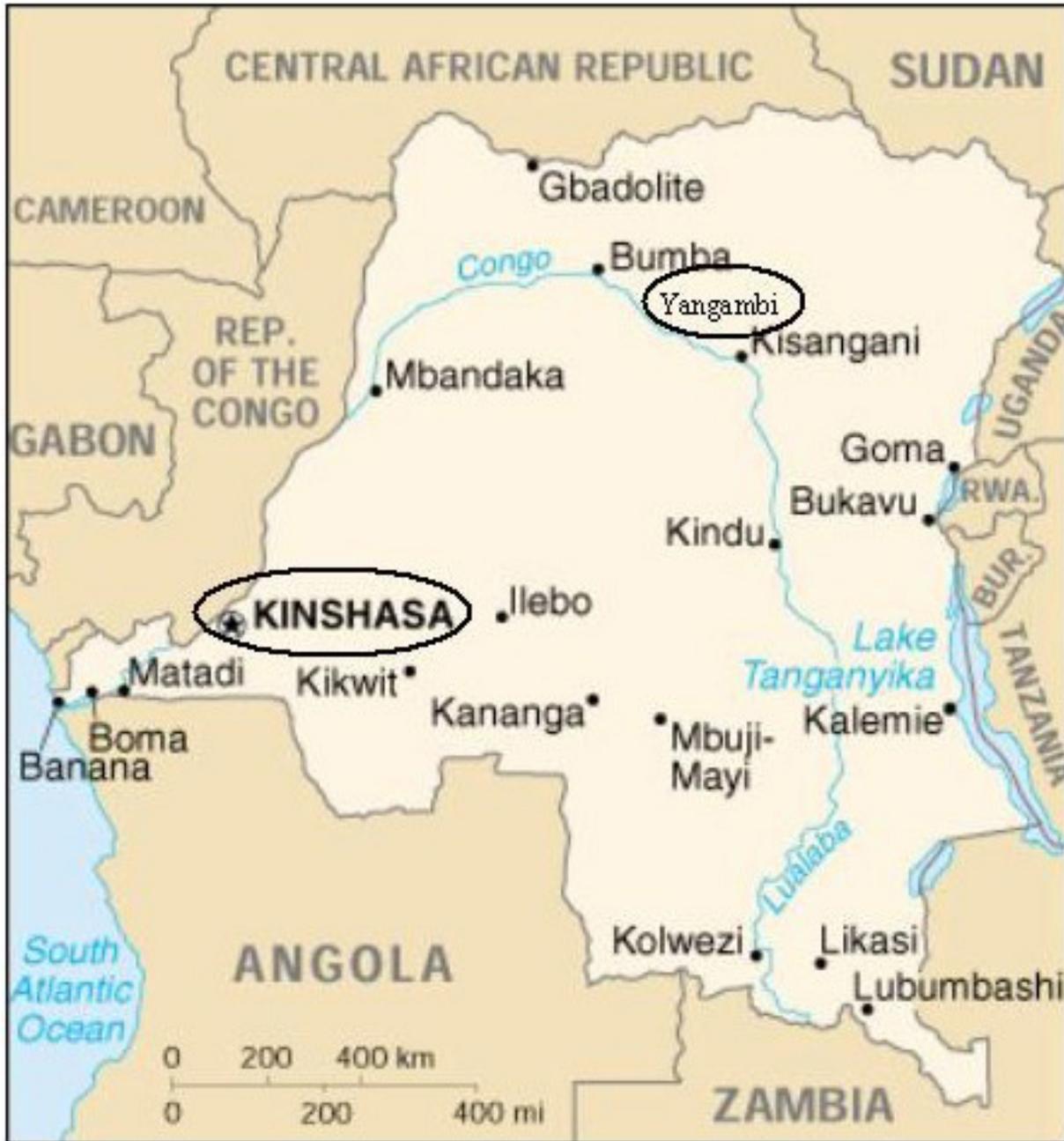


Fig. 1. Map of the Democratic Republic of Congo. The locations of the three studied stations are inside the ovals (Two stations at Kinshasa and one station at Yangambi).

parameters. The IDF-relation forms a group of parallel decreasing curves. The intensity decreases with the duration d and increases with the return period T . Usually, the denominator of the relation (3) is chosen equal to

$$b(d) = (d + \theta)^\eta \quad (4a)$$

where θ and η are two parameters to be estimated (Koutsoyiannis et al., 1998). This function $b(d)$ is purely empirical. On the other hand, variants of the relation such as

$$b(d) = d^\eta + \theta \quad (4b)$$

have been used in Demarée (1985) and in others. In this paper, we will use the relation

$$b(d) = \theta \left(1 + \frac{d}{\theta}\right)^\eta \quad (4c)$$

in order to have parameters with physical dimensions (with θ expressed in min and η having no dimension).

One can prove mathematically that it is useless to introduce empirically the numerator of relation (3). Its form can in fact be calculated mathematically (Koutsoyiannis et al., 1998) from the cumulative distribution function of the

maximum annual values of rainfall amounts. It can be shown from extreme value theory that this distribution is very likely to be the GEV-distribution or its simplified form, the Gumbel distribution (these two distributions arise naturally as limit distributions when dealing with maximum values in blocks of i.i.d. random variables).

If this distribution is the Gumbel distribution then it can be shown that the numerator of relation (3) has the following form :

$$a(T) = \lambda \left\{ \frac{\psi}{\lambda} - \ln \left(-\ln \left(1 - \frac{1}{T} \right) \right) \right\} \tag{5a}$$

There are, in this case, two parameters (λ and ψ) to estimate. If the distribution is a Jenkinson or General Extreme Value (GEV)-distribution, then

$$a(T) = \lambda \left\{ \frac{\psi}{\lambda} + \frac{\left[-\ln \left(1 - \frac{1}{T} \right) \right]^{-\kappa} - 1}{\kappa} \right\} \tag{5b}$$

There are thus, in this latter case, three parameters (λ , ψ and κ) to estimate. We discuss in Sect. 5 the selection of a Gumbel vs. a GEV distribution. An introduction to extreme value theory and its applications can be found in Beirlant et al. (1996).

5 Independence of the annual maximum values

Let us fix a duration d . We have then a sample of n data points or n precipitation depths (in our case, we have $n=33$ annual maximum precipitation depths for different durations at the station Yangambi-Km 5; $n=24$ at the station Binza; $n=17$ at the station Ndjili; and $n=50$ at the station Uccle, Belgium) denoted x_1, x_2, \dots, x_n . These data are supposed to be independent. We can test this assumption by computing the autocorrelation function for each duration and each station. As an example, Fig. 2 represents the graph of the autocorrelation function corresponding to the duration of one hour at the station Yangambi. This graph has been computed for different lags comprised between 0 and approximately $n/2$. The autocorrelation function has been computed using the Spearman statistic. We have used this statistic because it is a non-parametric one that is not based on the assumption of normality of the variables. The lines represent the 95% confidence limits to test the hypothesis of the nullity of the autocorrelation function in the case of a simple test (continuous line) or in the case of a multiple test (dotted line) testing simultaneously the hypothesis for all values of the autocorrelation function comprised between 0 and $n/2$. It is observed that the values of the autocorrelation function are small in absolute value (close to zero) and inside the 95% of confidence limits shown by the continuous and the dotted lines. So, the hypothesis of zero autocorrelation values can be accepted at the 5% significance level. This means that the series of the annual maximum rainfall values at the station Yangambi for a duration of one hour is likely to be constituted by independent values. The same analysis has been made for the other

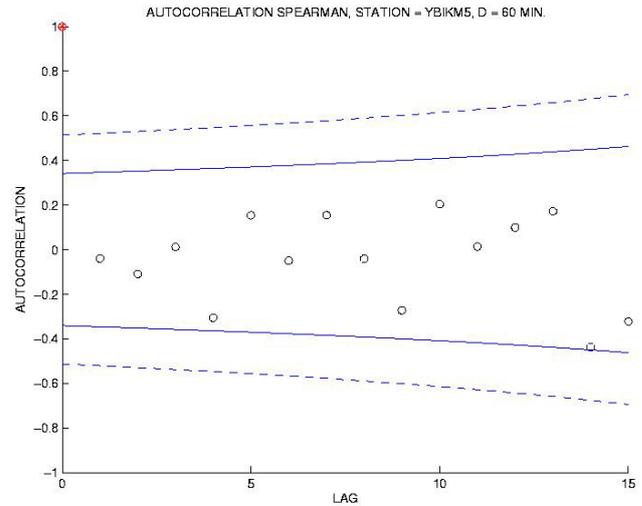


Fig. 2. Graph of the autocorrelation function of the annual maximum precipitation depths for the duration of 1 h at the station Yangambi, Congo.

durations and for the other stations with the same conclusions.

6 The underlying probability distribution: Gumbel vs. GEV

Different probability distributions can fit our data. Two choices seem reasonable here: the Gumbel probability distribution and the GEV probability distribution. The Gumbel distribution function has the following form:

$$F(x) = \exp \left(-\exp \left(-\frac{(x - \psi)}{\lambda} \right) \right) \tag{6a}$$

where ψ and λ are respectively the location and scale parameters. The mathematical form of the GEV distribution function is the following:

$$F(x) = \exp \left\{ - \left[1 + \kappa \left(\frac{x - \psi}{\lambda} \right) \right]^{-1/\kappa} \right\} \text{ for } \kappa \neq 0 \tag{6b}$$

defined on an adequate domain. Notice that for $\kappa=0$ the GEV distribution turns into the Gumbel distribution. The parameter κ is called the shape parameters. To estimate these parameters the L-Moments estimators are used (Hosking et al., 1985; Stedinger et al., 1993). As an illustration, Table 3a (respectively Table 3b) contains the numerical results of the L-Moments type estimates $\hat{\psi}$ and $\hat{\lambda}$ for the Gumbel distribution function (respectively the L-Moments type estimates $\hat{\psi}$, $\hat{\lambda}$, and $\hat{\kappa}$ for the GEV distribution function) of the annual maximum precipitation depths (expressed in 0.1 mm) corresponding to different durations at the station of Binza, Congo.

The problem is to choose between the Gumbel and the GEV distribution in order to model our data. The solution

Table 3a. Numerical results of the parameter estimates for a Gumbel distribution function corresponding to the different durations at the station of Binza, Congo.

Duration (min)	$\hat{\psi}$ (0.1 mm)	$\hat{\lambda}$ (0.1 mm)
5	122.33	25.35
10	208.03	36.12
20	314.74	49.24
30	399.84	73.68
60	569.32	88.02
1440	833.15	189.28

Table 3b. Numerical results of the parameter estimates for a GEV distribution function corresponding to the different durations at the station of Binza, Congo.

Duration (min)	$\hat{\psi}$ (0.1 mm)	$\hat{\lambda}$ (0.1 mm)	$\hat{\kappa}$
5	122.48	25.66	0.01
10	207.32	34.62	-0.04
20	318.8	56.21	0.17
30	400.45	74.89	0.02
60	575.48	98.86	0.14
1440	831.37	185.6	-0.02

to this problem can be found by means of well-chosen statistical tests. The Kolmogorov-Smirnov and the χ^2 (or Chi-square) tests are two non-parametric tests that are well known and whose objective is to test if a given sample comes from a distribution fixed in advance. We can test if the annual maximum values of rainfall amounts fallen in Yangambi, Kinshasa-Binza, Kinshasa-Ndjili and Uccle follow either a Gumbel or a GEV distribution. To test these assumptions the Kolmogorov-Smirnov and the χ^2 tests were carried out for each duration, each station and each hypothesis (Gumbel or GEV). The results of these tests are given in Tables 4a and 4b for the station Yangambi. It is noted that the test statistic never exceeds the limiting 95% value and this for each of the two tests. Consequently, for this station, the assumptions made cannot be rejected (at the 95% level) and the assumption according to which the data follow either a Gumbel distribution or a GEV distribution for each aggregation time is accepted. The same conclusions can be drawn for the other stations for which the results of the statistical tests are similar. So we cannot decide, by means of the statistical tests used, which distribution to choose. Both seem equally good. Obviously, the reason for this is the short lengths of our series of annual values.

Because the simplest descriptions are the best, we propose here to work with the Gumbel distribution which has only two parameters (the GEV distribution have three parameters) and which has a simpler mathematical expression. It is noted, however, that recent studies (Wilks, 1993; Kout-

Table 4a. Numerical results for the Kolmogorov-Smirnov test and the χ^2 test for the annual maximum values of the rainfall depths for the different durations at Yangambi, Congo. The distribution tested is the Gumbel distribution.

Kolmogorov-Smirnov test		χ^2 Test	
Number of points	33	Number of bins	6
95% limit	0.23	95% limit	7.81
Duration (min)	Statistic	Duration (min)	Statistic
1	0.08	15	3.55
30	0.14	30	5.36
45	0.15	45	5.73
60	0.12	60	1.73
120	0.10	120	3.55
1440	0.10	1440	2.45

soyiannis and Baloutsos, 2000; Coles et al., 2003; Koutsoyiannis, 2003) have expressed scepticism for the appropriateness of the Gumbel distribution for rainfall extremes, showing that it underestimates seriously the largest extreme rainfall amounts. After the observations made in these studies, it is not advisable to extrapolate the IDF curves based on the Gumbel distribution to return periods much greater than the record lengths.

To validate the choice of the Gumbel distribution we make a third test: the test on the shape parameter κ (Stedinger et al., 1993). The null hypothesis of this test is that the shape parameter κ of the fitted GEV distribution is equal to zero. The results of this test are given in Table 4c for the station Binza. It is noted that the test statistic $Z=\hat{\kappa}\sqrt{n/0.5633}$ always stays within the limiting 95% bounds (-1.96 and 1.96) and this for each duration. Consequently, for this station, the null hypothesis can be accepted. The same conclusions can be drawn for the other stations for which the results of the statistical tests are similar.

7 Extreme value distributions of the daily precipitation depths

For the daily precipitation depths, longer time series are available for Uccle (typically from the 1880s till present) and for Binza (from 1929 to 2000). For Yangambi, daily rainfall amounts are available since 1930 but only the series from 1950 onwards has been used. The annual maximum rainfall amounts are taken from the daily precipitation records generated by the non-recording raingauge. We show in this section the QQ-plots (quantile-quantile plots) corresponding to the daily annual extremes for the stations Uccle, Yangambi and Binza.

A QQ-plot (see for example Beirlant et al., 1996) is a plot showing the observations (in our case the daily annual maximum precipitation depths) versus a well-chosen

Table 4b. Numerical results for the Kolmogorov-Smirnov test and the χ^2 test for the annual maximum values of the rainfall depths for the different durations at Yangambi, Congo. The distribution tested is the GEV distribution.

Kolmogorov-Smirnov test		χ^2 Test	
Number of points	33	Number of bins	6
95% limit	0.23	95% limit	5.99
Duration (min)	Statistic	Duration (min)	Statistic
15	0.08	15	3.91
30	0.09	30	2.82
45	0.12	45	2.82
60	0.11	60	2.09
120	0.09	120	2.82
1440	0.10	1440	3.91

Table 4c. Numerical results for the test on the shape parameter κ for the annual maximum values of the rainfall depth for the different durations at Binza, Congo.

Test on the shape parameter κ	
Number of points	24
95% limits	-1.96 and 1.96
Duration (min)	Statistic
5	0.09
10	-0.29
20	1.09
30	0.12
60	0.94
1440	-0.14

reduced variate (function of the probability). The function is chosen in such a way that the observations approach a straight line when the distribution fitted corresponds to the true distribution type. QQ-plots can be used to see visually if our sample of daily maximum values is likely to be drawn from the fitted distribution (i.e. if the fit is good enough). In this way, the conclusions drawn from several tests in the previous section can be verified. QQ-plots can also be used to visualise the relation existing between the values (depths) of the observations and their frequency of occurrence or return period.

In our case, the fitted distribution is the Gumbel distribution or the GEV distribution (but we have seen why the Gumbel distribution was preferred). The cumulative distribution function of these distributions can be rewritten as:

$$F(x) = \exp(-\exp(-y(x))) \tag{7}$$

where

$$y(x) = \frac{1}{\lambda}(x - \psi) \Leftrightarrow x(y) = \psi + \lambda y$$

for the Gumbel distribution and

$$y(x) = \frac{1}{\kappa} \ln \left[1 + \kappa \left(\frac{x - \psi}{\lambda} \right) \right] \Leftrightarrow$$

$$x(y) = \psi - \lambda \left(\frac{1 - \exp(\kappa y)}{\kappa} \right)$$

for the GEV distribution. The variable y is chosen as reduced variate because it is linear dependent on x for the Gumbel distribution. The reduced probability y_i corresponding to the i -th biggest value of a sample of size n is obtained by using the following formulas :

$$y_i = -\ln(-\ln(1 - p_i)) \tag{8a}$$

$$p_i = \frac{i - 0.44}{n + 0.12} \tag{8b}$$

This last expression is called the Gringorten formula and gives the optimal plotting position for the case of the Gumbel distribution.

Figures 3a–c show the six QQ-plots corresponding to the daily time series for the stations Uccle, Yangambi and Binza and for each of the models tested (Gumbel and GEV). The confidence intervals of the quantity $\hat{x}(y)$ were obtained by using standard formulas for the Gumbel case or by using a non-parametric resampling technique (bootstrap) for the GEV case. A good introduction on the bootstrap theory can be found in Efron and Tibshirani (1993). We see that the results given by the model GEV are very close to the results given by the model of Gumbel. So, two parameters seem enough to describe our data. This confirms what we have seen in Sect. 5. We also see that the confidence intervals are smaller for the model of Gumbel than for the model GEV. We observe finally that the levels and the slopes of the continuous curves (representing the model) are smaller for the station of Uccle than for the other two stations.

8 Methods for the establishment of the IDF-curves for precipitation

The IDF-curves for precipitation will be established according to three different methodologies (Koutsoyiannis et al., 1998). The techniques and the numerical results will be discussed below.

8.1 The classical technique

The classical technique for establishing the IDF-curves of precipitation has got three steps (see Réminieras, 1972; Ven Te Chow, 1964; Demarée, 1985). The first one consists of fitting a probability distribution function to the data and this for each duration. In the Sect. 5, it was shown that the Gumbel distribution function was a good candidate. In a second step

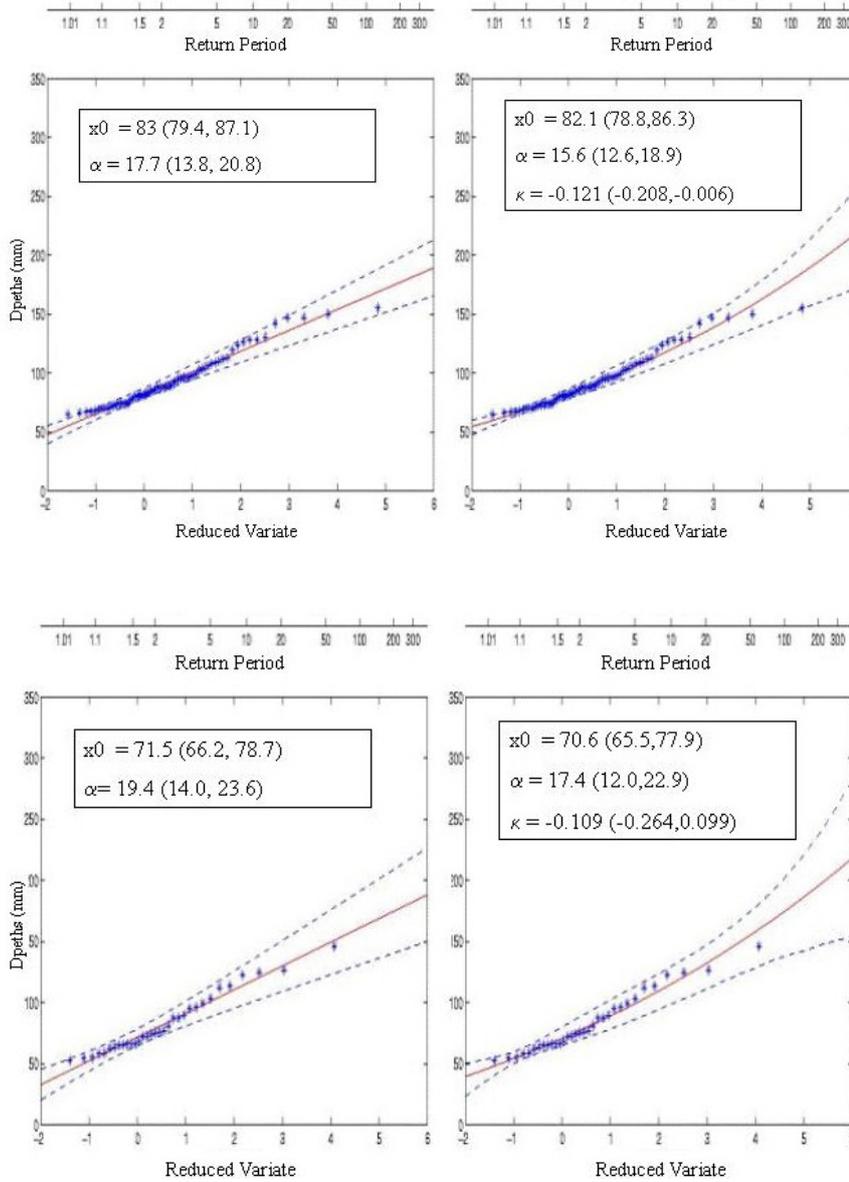


Fig. 3a. QQ-plots corresponding to the daily ($d=24$ h) time series for the stations of Kinshasa-Binza. The fitted distribution is either the Gumbel distribution (left) or the GEV distribution (right). For each subfigure, the observations are shown by mean of stars ('*') and the model is shown by mean of a continuous line. Broken lines represents 95% confidence intervals on the model. The values of the estimated model parameters are shown as well as the 95% confidence intervals on these values. These intervals were calculated using a non-parametric bootstrap technique.

Fig. 3b. QQ-plots corresponding to the daily ($d=24$ h) time series for the stations of Yangambi. The fitted distribution is either the Gumbel distribution (left) or the GEV distribution (right). For each subfigure, the observations are shown by mean of stars ('*') and the model is shown by mean of a continuous line. Broken lines represents 95% confidence intervals on the model. The values of the estimated model parameters are shown as well as the 95% confidence intervals on these values. These intervals were calculated using a non-parametric bootstrap technique.

the quantiles for each duration and for a given set of return periods are calculated by using the probability function derived in the first step. In the third step, finally the IDF-curves are obtained by performing a global non-linear regression on the quantiles given a criterion function. The three steps procedure allows for the estimation of the four parameters of the Eq. (3).

The procedure may be resumed as follows:

- Gumbel fitting for each given duration. This step has been done in Sect. 6 and the results, the estimated parameters, are shown in Tables 3a and b (for a particular station)
- Estimation of the quantiles for each given duration. Suppose that the data follow a distribution F . The quantile x_T^* having a return period T is defined by the value

x^* that verifies the expression

$$F(x_T^*) = 1 - \frac{1}{T} \quad (9)$$

In the case of a Gumbel distribution, by inverting Eq. (9) the following expression for x_T^* is found:

$$x_T^* = \lambda \left\{ \frac{\psi}{\lambda} - \ln \left(-\ln \left(1 - \frac{1}{T} \right) \right) \right\} \quad (10)$$

and for the corresponding mean rainfall intensity $i_{T,d}^*$:

$$i_{T,d}^* = \frac{x_T^*}{d}. \quad (11)$$

Empirical quantiles can then be estimated by replacing the parameters ψ and λ in Eq. (10) by their estimates $\hat{\psi}$

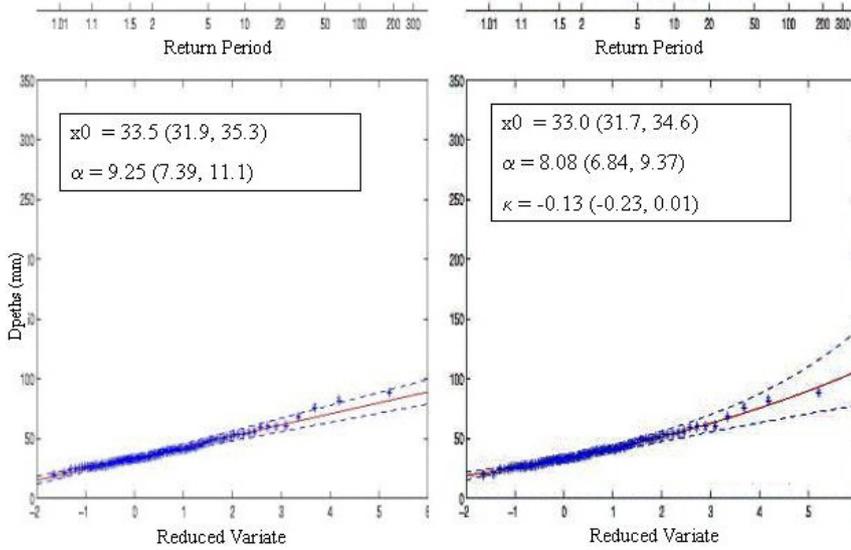


Fig. 3c. QQ-plots corresponding to the daily ($d=24$ h) time series for the stations of Uccle. The fitted distribution is either the Gumbel distribution (left) or the GEV distribution (right). For each subfigure, the observations are shown by mean of stars (**) and the model is shown by mean of a continuous line. Broken lines represents 95% confidence intervals on the model. The values of the estimated model parameters are shown as well as the 95% confidence intervals on these values. These intervals were calculated using a non-parametric bootstrap technique.

and $\hat{\lambda}$. The sets of T -values considered in this method are equal to $\{2, 5, 10, 20, 50, 100, 200\}$ years, $\{2, 5, 10, 20, 50, 75, 100\}$ years and $\{2, 5, 10, 20, 35, 50, 75\}$ years for, respectively, the stations of Uccle, of Yangambi and each of the two stations situated in Kinshasa. The sets of d -values considered in this method (and in the other two) are equal to $\{1, 10, 20, 30, 60, 120, 360, 720, 1440\}$ min, $\{15, 30, 45, 60, 120, 1440\}$ min and $\{5, 10, 20, 30, 60, 1440\}$ min for, respectively, the stations of Uccle, of Yangambi and each of the two stations situated in Kinshasa.

- Global non-linear least squares estimation on the empirical quantiles. The regression is performed on the empirical quantiles $i_{T,d}^*$. The estimated parameter values are given in Table 3. The regression function used is given by the expression (12). To perform the regression one minimizes the objective function given by the expression (13). Relative errors are preferred instead of absolute errors in order to give an “more equal” weight to each error associated with each experimental quantile.

$$i_{T,d}(\psi, \lambda, \theta, \eta) = \frac{\lambda \left\{ \frac{\psi}{\lambda} - \ln \left(-\ln \left(1 - \frac{1}{T} \right) \right) \right\}}{\theta \left(1 + \frac{d}{\theta} \right)^\eta} \quad (12)$$

$$F = \sum_{T,d} \left(1 - \frac{i_{T,d}(\psi, \lambda, \theta, \eta)}{i_{T,d}^*} \right)^2 \quad (13)$$

In the next two subsections two less classical methods (Koutsoyiannis et al., 1998) for assessing the IDF-curves are explained.

8.2 Method in two stages

Koutsoyiannis et al. (1998) propose a robust technique for assessing the IDF-curves of precipitation in which firstly the

parameters of the function $b(d)$ are estimated followed in a second stage by the estimation of the parameters of the function $a(T)$. For the notations, we assume that the data set consisting of k groups each one containing the intensity values of a particular duration $d_j, j=1, \dots, k$ (the intensity values are simply the depths values divided by the duration). Let us denote by n_j the length of the group j , and by i_{jl} the intensity values of this group. In our case and for each station, all groups have the same length.

The intensities i_{jl} are samples of the random variables $I_j := I(d_j)$ where $l=1, \dots, n_j$ denotes the rank of the value i_{jl} in the group j of intensity values arranged in descending order.

The underlying hypothesis of this robust technique is that the distribution functions of the random variables $Y_j = I_j b(d_j)$ of all k groups corresponding to the different durations d_j are identical. The Kruskal-Wallis technique tests the null hypothesis H_0 that the group medians are the same (see for example Siegel et al., 1988). Assuming that the parameters θ and η of the function $b(d)$ intervening in Eq. (4c) are known, it is then possible to compute all values $y_{jl} = i_{jl} b(d_j)$. The overall number of data points is given by $m = \sum_{j=1}^k n_j$ where the summation is across the k samples. The Kruskal-Wallis test statistic is given by

$$k_{KW} = \frac{12}{m(m+1)} \sum_{j=1}^k n_j \left(\bar{r}_j - \frac{m+1}{2} \right)^2 \quad (14)$$

where the ranks r_{jl} are assigned to the values y_{jl} (with respect to the global sample). The average rank of the n_j values of the j -th group is denoted by \bar{r}_j . However, it appears from the relation $y_{jl} = i_{jl} b(d_j)$ that the ranks r_{jl} will depend on the numerical values of the parameters θ and η of the function $b(d)$. From which it results that the estimation problem is reduced to minimizing the test statistic k_{KW} as a function of the parameters θ and η . A derivative free numerical search

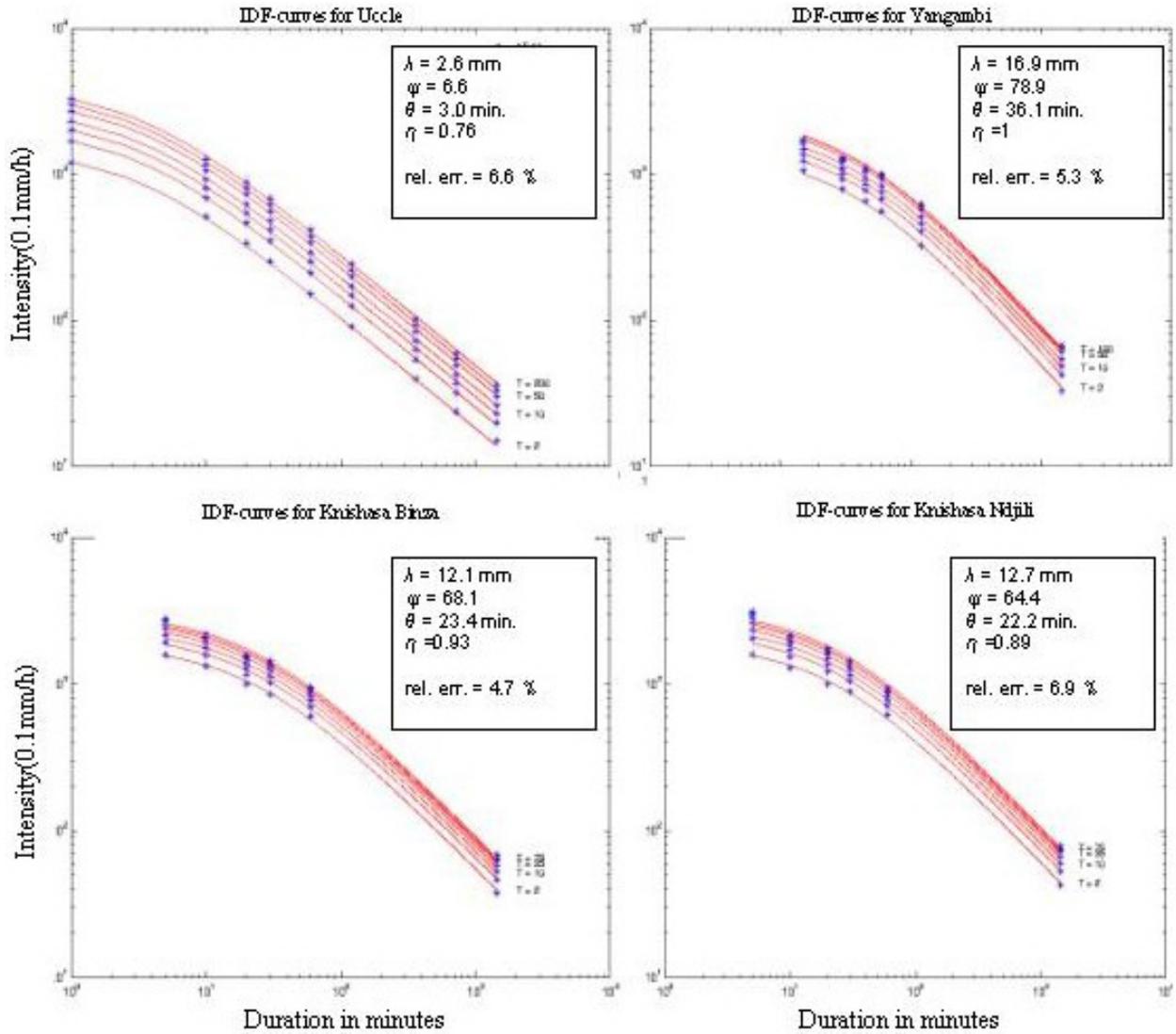


Fig. 4. The four sets of IDF-curves, shown in logarithmic scale, and obtained via the one-step method Top-left: Uccle, top-right: Yangambi, bottom-left: Binza, bottom-right: Ndjili. The stars (*) represent the empirical quantiles whereas the continuous lines represent the sets of IDF-curves. The values of the parameters are displayed on the sub-figures: param. 1 stands for, param. 2 stands for, param. 3 stands for and param. 4 stands for whereas the fitting error is displayed on the line below these four parameters.

technique is most appropriate solving the problem. The minimizing technique can be speeded up when only results in the region of higher intensities are needed by using only a part of the data values of each group instead of the complete series. In this application only the highest half of the intensity values for each duration were used.

Using the optimized values $\hat{\theta}$ and $\hat{\eta}$, all values of $y_{jl}=i_{jl}b(d_j)$ form a unique sample. Selecting a probability distribution, in our case the Gumbel distribution function, and an appropriate estimation technique (in our case the L-moments technique), the parameters ψ and λ of the function $a(T)$ are then estimated from this sample. Finally, the estimates $\hat{i}_{T,d}(\hat{\psi}, \hat{\lambda}, \hat{\theta}, \hat{\eta})$ for the extreme intensities are given

by the expression:

$$\hat{i}_{T,d}(\hat{\psi}, \hat{\lambda}, \hat{\theta}, \hat{\eta}) = \frac{a(T, \hat{\psi}, \hat{\lambda})}{b(d, \hat{\theta}, \hat{\eta})} = \frac{\hat{\lambda} \left\{ \frac{\hat{\psi}}{\hat{\lambda}} - \ln \left(-\ln \left(1 - \frac{1}{T} \right) \right) \right\}}{\hat{\theta} \left(1 + \frac{d}{\hat{\theta}} \right)^\eta} \quad (15)$$

8.3 Method in one stage

In this technique (Koutsoyiannis et al., 1998), all parameters of both functions $a(T)$ and $b(d)$ are estimated in one single step minimizing the total square error of the fitted IDF-relationship to the data. More precisely, to each data value i_{jl} an empirical return period using the Gringorten plotting

Table 5. Estimated parameter values obtained with the three methods of estimation for the four stations considered. The errors given by the last column are computed according to formula (17) for all the three methods.

Station	Method	$\hat{\lambda}$ (mm)	$\hat{\psi}$ (mm)	$\hat{\psi}/\hat{\lambda}$	$\hat{\theta}$ (min)	$\hat{\eta}$	Error
Uccle	3 steps	2.8	7.4	2.67	3.4	0.78	0.068
Uccle	2 steps	2.6	7.2	2.79	3.2	0.77	0.068
Uccle	1 step	2.6	6.6	2.55	3	0.76	0.066
Yangambi	3 steps	14.5	78.3	5.38	34.7	0.98	0.059
Yangambi	2 steps	16.9	79.9	4.74	35.5	1	0.055
Yangambi	1 step	16.9	78.9	4.68	36.1	1	0.053
Binza	3 steps	10.1	57.8	5.73	18.6	0.89	0.052
Binza	2 steps	12.1	66.2	5.47	22.4	0.93	0.048
Binza	1 step	12.1	68.1	5.63	23.4	0.93	0.047
Ndjili	3 steps	9.2	49.3	5.37	14.9	0.84	0.072
Ndjili	2 steps	12.7	63.5	5.01	21.6	0.9	0.069
Ndjili	1 step	12.7	64.4	5.08	22.2	0.89	0.069

position

$$T_{jl} = \frac{n_j + 0.12}{l - 0.44} \tag{16}$$

is assigned. So, for each data value we have a triplet of numbers (i_{jl}, T_{jl}, d_j) . On the other hand, for a specific form of $a(T)$, the modelled intensity is given by

$$\hat{i}_{jl} = \frac{a(T_{jl})}{b(d_j)} \tag{17}$$

and the corresponding relative error:

$$e_{jl} = 1 - \frac{\hat{i}_{jl}}{i_{jl}}. \tag{18}$$

In our case, the specific form of $a(T)$ is again given by the Eq. (5a) and the form of $b(d)$ is given by the Eq. (4c). The overall mean squared relative error is given by

$$e^2 = \frac{1}{k} \sum_{j=1}^k \frac{1}{n_j} \sum_{l=1}^{n_j} e_{jl}^2. \tag{19}$$

The estimation problem is again reduced to an optimization problem that can be defined as:

$$\text{minimize } e = e(\psi, \lambda, \theta, \eta). \tag{20}$$

Again, a numerical search technique for optimization that makes no use of derivatives is appropriate for this problem. As in the previous method, the minimizing technique can be speeded up by using only a part of the data values of each group. In this application only the highest half of the intensity values for each duration were used.

8.4 Results

The estimated parameter values obtained with the three methods of estimation and for the four stations considered are given in Table 5.

In this table, we can first notice the units of the parameters. The parameter $\hat{\theta}$ is expressed in minutes because we chose to express the duration d in minutes. The parameter $\hat{\lambda}$, as well as the parameter $\hat{\psi}$ are expressed in mm so that the left member of Eq. (3) is expressed in mm/min. The last parameter, $\hat{\eta}$, has no unit, by definition. In order to compare the three methods we have calculated, for each of them, the fitting error given by formula (19) which is also the objective function of the one-step method. It is noted that, as expected, the one step method minimizes the fitting error. Figure 4 shows the four sets of IDF-curves obtained via the one-step method for each of the four stations considered.

9 Comparison of the results for the three methods and for the four stations

We see from Table 5 that, for each station, all methods give approximately the same error, the difference between the worst and the best method being not so big. In fact, although the three methods are based on totally different philosophies, they provide very similar results. For all the stations and for the classical method, the fitting error between the IDF-curves and the empirical quantiles is of the order of 5 to 7%, which is very low. For each of the four stations, the estimated parameter values for the three techniques are numerically very similar. In fact, the difference in the graphs corresponding to the three techniques is hardly discernable. Note that the methods in two or in three stages do not try to minimize the fitting error (17); they have in fact different objective functions.

Each of the four parameters of the estimated IDF-curves has got a signification. If we look at these curves in log-log scale then the parameter λ plays a role in the level of the curves: if all the other parameters are fixed, multiplying this parameter by two is equivalent to multiplying the level by two. The combined parameter ψ^* defined as ψ/λ corresponds to the spreading of the different curves. The smaller ψ^* is, the larger is the spreading. The parameter θ points to

Table 6. Extreme depths calculated for the four stations and for two different couples of duration and return periods. The estimated parameters result from the method in 3 steps.

Station	Method	depths (mm)	depths (mm)
		d = 1 h T = 20 years	d = 24 h T = 20 years
Uccle	3 steps	28.1	59.0
Yangambi	3 steps	78.5	127.8
Binza	3 steps	78.5	140.0
Ndjili	3 steps	79.4	157.8

the position where the IDF-curves begin to become straight lines. This point is roughly situated at $d=3$. Finally, η is proportional to the slope of the curves.

When looking at Table 5, it is noted that the slope of the IDF-curves for precipitation for the station of Uccle ($\eta=0.76$) is smaller than the slopes of the two curves of the two stations situated in Kinshasa (these two slopes are similar, $\eta=0.89$ and $\eta=0.93$) which in turn are smaller than the slope of the Yangambi IDF-curves ($\eta=1$). It is also noted that the spreading of the IDF-curves is similar for all the African stations but is much larger in the case of the Belgian station. This fact is clearly visible in Fig. 4. For example, the ratio between the extreme quantity corresponding to a return period of 50 years and the extreme quantity corresponding to a return period of 2 years is equal to 2.83 for Uccle but is only equal to 2.09 for Yangambi (and equals respectively to 2.05 and 2.09 for Binza and Ndjili). We also remark that the parameter θ is very low for Uccle (about 3 min) compared to the other stations (from 15 to about 23 min for Binza and Ndjili, but reaching 36 min for Yangambi). Finally, the parameter λ is also much lower for the station Uccle than for the other stations (about 4 to 5 times lower).

The high values of the parameter η obtained in this study for the Central African stations, even approaching the value 1 for the station Yangambi, retain our special attention. Menabde et al. (1999) using a simple scaling/fractal model has obtained values ranging from 0.65 to 0.75 for South African stations. However the simple scaling model also leads to the IDF-curve expression

$$i = \frac{a(T)}{d^\eta} \quad (21)$$

where the parameter θ out of the formula (4a) is equal to zero. The latter model is not appropriate for the data of the Central African stations as the data present a rather long bending aspect represented by high values of the parameter θ . On the contrary, the station of Uccle has a short bending part of the IDF-curves giving rise to a rather long straight line aspect, yielding a value of the parameter θ of about 3 min. For the latter part, the Menabde theory can be easily applied. In the case of the Central African stations, due to the bending,

the value of η can be seen as the slope of the tangent to the IDF-curves at the data points of 24 h.

We can also look at the extreme depths calculated by mean of the four estimated parameters for a well-chosen return period and a well-chosen duration. Table 6 shows the extreme depths obtained for the four stations with the four parameters calculated by mean of the three-steps method and where we have chosen a return period of 20 years and a duration of either one hour or one day. We see that the extreme depths obtained are much higher for the African stations than for the Belgian station (about 2 to 2.5 times larger). Moreover the calculated depths for a duration of one hour are very similar for the three Central African stations (about 79 mm) but we observe noticeable differences between the estimated depths for these stations and for a duration of one day. These differences can be explained by the different values of the slopes (described by the parameter η) of the IDF-curves for the different Central African stations.

10 Conclusions

It was shown by applying different statistical tests that the annual maximum values of precipitation at Yangambi, Binza, Ndjili (all three in Congo) and Uccle (Belgium), follow a probability distribution function that can be chosen to be the Gumbel distribution function for the durations considered. Based upon this preliminary result, IDF-relationships covering different ranges of durations from 1, 5 or 15 min till 1440 min (1 day) and using only 4 model parameters were assessed by three different methods. In all three methods, the separation of the functions $a(T)$ and $b(d)$ is maintained and $a(T)$ is taken as being the inverse function of the Gumbel distribution function. Although, the underlying philosophies are different, particularly in their optimisation processes, the three methods yielded similar results.

In this paper, IDF-curves for Yangambi, Binza and Ndjili (Congo), have been assessed. The relationship characterizes the extreme rainfall at different places of Congo. Rainfall intensity information is rather rare in the tropical area of Central Africa since few high-frequency long-term extreme rainfall statistics have been processed in the region. IDF-curves for Uccle (Belgium) have also been assessed. The values of the obtained parameters were used to compare the IDF-curves of the four stations.

Physically, high values of the parameter η for the Central African stations can be related to the short duration of the length of the heavy rainfall. In Central Africa, short but heavy rains occur once a day, so that the rainfall depth is fairly constant for durations ranging from, say, 4 to 24 h. This property implies a value of η in that range close to 1.

It was tempted to provide graphical interpretations of the model parameters of the implemented model for the IDF-curves. In this procedure the differences between IDF-curves and their parameter values in the tropical area of Central Africa and Uccle, Belgium, were shown evidencing the difference in the extreme precipitation regimes.

Acknowledgements. The authors are particularly indebted to M. Crabbé who provided the manuscript with the extreme rainfall data at Yangambi. A warm tribute should be given to B. Totiwe T'Essabe at the "Institut National pour l'Étude et la Recherche agronomiques" for producing the rainfall statistics at Yangambi. Without his continuing efforts the establishment of the IDF-curve of precipitation at Yangambi would not have been possible. D. N. Faka thanks the "Commission Universitaire au Développement" of the C.I.U.F. for having provided a scholarship under the Direction of F. Ronday at the Université de Liège, Belgium. The authors wish to thank D. Koutsoyiannis and P. Willems for their constructive remarks during the review process of this paper.

Edited by: F. Castelli

Reviewed by: D. Koutsoyiannis and P. Williams

References

- Beirlant, J., Teugels, J. L., and Vynckier, P.: Practical analysis of extreme values, Leuven University Press, 1–137, 1996.
- Buishand, T. A. and Demarée, G. R.: Estimation of the annual maximum distribution from samples of maxima in separate seasons, *J. of Stochastic Hydrology and Hydraulics*, 4, 89–103, 1990.
- Bultot, F.: Cartes des régions climatiques du Congo Belge établies d'après les critères de Köppen, Publ. de l'Institut National pour l'Étude Agronomique du Congo Belge (INÉAC), Bruxelles, Bureau Climatologique, Communication 2, 1–15, 1950.
- Bultot, F.: Etude statistique des pluies intenses en un point et sur une aire au Congo Belge et au Ruanda-Urundi, Publ. de l'Institut National pour l'Étude Agronomique du Congo Belge, Bruxelles, Bureau Climatologique, Communication 11, 1–90, 1956.
- Chow, V. T.: Handbook of Applied Hydrology, McGraw-Hill Book Company, 1–1450, 1964.
- Coles, S., Pericchi, L. R., and Sisson, S.: A fully probabilistic approach to extreme rainfall modelling, *J. of Hydrology*, 273(1–4), 35–50, 2003.
- Crabbé, M.: L'Écoclimat de Yangambi (0°49' N, 24°29' E, 470 m) de 1960 à 1965, République Démocratique du Congo, Publications de l'Institut National pour l'Étude agronomique du Congo (INÉAC), Bruxelles, 1–80, 1970.
- Crabbé, M.: Recueil d'instructions relatives au réseau d'écoclimatologie, Quatrième édition revue et complétée, Inst. Nat. pour l'Étude agronomique du Congo (INÉAC), Kinshasa, 1971.
- Crabbé, M.: Climat de Kinshasa d'après les observations centrées sur la période 1931–1970, AGCD, Bruxelles, 1–120, 1980.
- Demarée, G.: Intensity-Duration-Frequency Relationship of Point Precipitation at Uccle, Reference Period 1934–1983, Institut Royal Météorologique de Belgique, Publications Série A, 116, 1–52, 1985.
- Demarée, G. R.: Le pluviographe centenaire du plateau d'Uccle: son histoire, ses données et ses applications, *La Houille Blanche*, 4, 1–8, 2003.
- Efron, B. and Tibshirani, R. J.: An Introduction to the Bootstrap, New York, Chapman and Hall, 1–436, 1993.
- Koutsoyiannis, D.: On the appropriateness of the Gumbel distribution for modelling extreme rainfall, Proceedings of the ESF LESC Exploratory Workshop, Hydrological Risk: recent advances in peak river flow modelling, prediction and real-time forecasting, Assessment of the impacts of land-use and climate changes, October 2003, European Science Foundation, National Research Council of Italy, University of Bologna, Bologna, (available online at <http://www.itia.ntua.gr/e/docinfo/590/>), 2003.
- Koutsoyiannis, D. and Baloutsos, G.: Analysis of a long record of annual maximum rainfall in Athens, Greece, and design rainfall inferences, *Natural Hazards*, 22(1), 31–51, 2000.
- Koutsoyiannis, D., Kozonis, D., and Manetas, A.: A mathematical framework for studying rainfall intensity-duration-frequency relationships, *J. of Hydrology*, 206, 118–135, 1998.
- Llasat, M.-C.: An objective classification of rainfall events on the basis of their convective features: application to rainfall intensity in the northeast of Spain, *Intern. J. of Climatology*, 21, 1085–1400, 2001.
- Menabde, M., Seed, A., and Pegram, G.: A simple scaling model for extreme rainfall, *Water Resour. Res.*, 35(1), 335–339, 1999.
- Réménieras, G.: L'Hydrologie de l'Ingénieur, Collection du Centre de Recherches et d'Essais de Chatou (3e édition), Eyrolles, 1–456, 1972.
- Ruwet, A., Sengele, N., Agana, P. and Totiwe, T.: Paramètres moyens et extrêmes principaux du climat des stations du réseau INERA, Tome 1. République du Zaïre, Inst. Nat. pour l'Étude et la Recherche agronomiques (INERA), Sect. de Yangambi, 1–199, 1985.
- Siegel, S. and Castellan Jr., N. J.: Nonparametric Statistics for the Behavioral Sciences, McGraw-Hill Book Company. XXIII, 1–329, 1988.
- Stedinger, J. R., Vogel, R. M., and Foufoula-Georgiou, E.: Frequency Analysis of Extreme Events, in: Handbook of Hydrology, edited by Maidment, D. R., McGraw-Hill Book Company, 18, 18.1–18.66, 1993.
- Vandenplas, A.: La pluie au Congo Belge, Institut Royal Météorologique de Belgique, Mémoires, XVI, 1–130, 1943.
- Wilks, D. S.: Comparison of three-parameter probability distributions for representing annual extreme and partial duration precipitation series, *Water Resour. Res.*, 29(10), 3543–3549, 1993.
- Willems, P.: Compound intensity/duration/frequency-relationships of extreme precipitation for two seasons and two storm types, *J. of Hydrology*, 233, 189–205, 2000.