COMPARISON OF THREE METHODS OF ESTIMATING RAINFALL FREQUENCY PARAMETERS ACCORDING TO THE DURATION OF ACCUMULATION

M. Slimani
Institut de Mecanique de Grenoble.
BP 68X
38402 Saint Martin d'Heres Cedex.
France

T. Lebel
ORSTOM-Hydrologie.
ZOLAD, MiniParc
34100 Montpellier.
France

ABSTRACT. The 'Gradex' method estimates high return period discharges by assuming that the two marginal increases in rainfall and flood volumes are equal for return periods greater than a hundred years. This requires an asymptotically exponential decay of the rainfall distribution, the gradex being the gradient of this exponential. Among several statistical relationships satisfying this requirement, the Gumbel distribution is widely acknowledged to properly fit the extreme value distribution of observed daily rainfalls. However, difficulties may arise for small time steps (from less than one hour up to twelve hours).
Among the many reasons for these difficulties, seasonal variation of the distribution parameters and the mixing of rainstorms issuing from different weather conditions are the most frequent. Rainfall distributions should therefore be analyzed using data from a homogeneous season. However, the mixing of weather patterns remains whatever the season considered, and it is important to find a way to estimate the gradex parameter even when the experimental distribution of extreme values does not allow a good Gumbel model fit. One solution is to fit a two component negative exponential distribution to the complete data set. Such a theoretical distribution is asymptotically parallel to a Gumbel distribution, thus providing an estimate of the gradex.

In this communication, three gradex estimators are compared from bias and efficiency standpoints : the moment and maximum likelihood methods used to fit a Gumbel distribution to an extreme value data set, and the moment method used to fit a two component exponential distribution to the complete data set of observed rainfalls. This comparison can be used to determine the most suitable estimator for subsequent regional regression and mapping of rainfall frequency parameters(see : linear relations between rainfall and morphometric parameters by Slimani and Obled, this symposium).

V. P. Singh (ed.), Hydrologic Frequency Modeling, 277–291.

separate books, FLOOD HYDROLOGY, REGIONAL FLOOD FREQUENCY ANALYSIS, and APPLICATION OF FREQUENCY AND RISK IN WATER RESOURCES, which are being published simultaneously. Arrangement of these books under fc different titles was a natural consequence of the diversity of techr cal material discussed in the papers. These books can be treated almost independently, although some overlap does exist between them.

This book contains eight sections encompassing major aspects of hydrologic frequency analysis. Each section starts usually with an invited state-of-the-art paper, followed by contributed papers. Beginning with a discussion of hydrologic frequency analysis and its relevance, the papers go on to discuss univariate flood frequency models, mixed distributions, rainfall frequency analysis, applicatio of entropy in flood frequency analysis, methods of parameter estimation, selection of flood frequency models, and multivariate stochastic models.

The book will of of interest to researchers as well as those engaged in practice of Civil Engineering, Agricultural Engineering, Hydrology, Water Resources, Earth Resources, Forestry and Environmental Sciences. The graduate students as well as those wishing to conduct research in flood hydrology will find this book to be of particular value.

I wish to take this opportunity to express my sincere appreciation to all the members of the Organizing and Advisory Committees, and the Louisiana State University administration for their generous and timely help in the organization of the Symposium. A lack of space does not allow me to list all of them by name here. Numerous other people contributed to the Symposium in one way or another. The authors, including the invited keynote speakers, contributed to the Symposium technically and this book is a result of their efforts. The session chairmen administered the sessions in a positive and professional manner. The referees took time out from their busy schedules and reviewed the papers. Graduate students assisted in smooth conduct of the sessions. I owe my sincere gratitude to all of these individuals.

If the success of a Symposium is measured in terms of the quality of participants and presentations then most people would agree that this was a very successful Symposium. A very large number of internationally well-known people, who have long been recognized for their contributions and have long been at the forefront of hydrologic research came to participate in the Symposium. More than 35 countries, covering the five continents and most of the countries of the world active in hydrologic research, were represented. It is hoped that long and productive personal associations will develop as a result of this Symposium.

March 1987
Baton Rouge, Louisiana

Vijay P. Singh
Symposium Director

1. INTRODUCTION

This communication is one outcome of a continuous effort aimed at providing engineers in relevant disciplines with means of estimating high return period floods in the **CEVENNES** (fig.1), France's most problematic region in this respect. The upland watersheds in this region are relatively small (a few hundred square kilometers), but rugged topography and complex meteorological phenomena combine to produce flash-floods threatening human life and property, as already observed in October 1958 (33 casualties) and September 1980 (a dozen casualties and damage amounting to more than one billion francs). Despite the avaibility of only one or two long series of stream flows in this region, the **GRADEX** method is believed to give reliable estimates of the risk. The method relies on the analysis of rainfall data assumed to display an asymptotically exponential decay behavior. The GRADEX is the gradient of this exponential decay. The ultimate goal of the study is to provide those in charge of hydraulic development projects with a map of the gradex (g) and rainfalls for given return periods at various durations of accumulation (see communication by Slimani and Obled, this symposium). For this, reliable estimates of g are needed, and the present paper is devoted to comparing some such possible estimates.

Accepting the exponential decay assumption, an estimation of g may be reached by studying either the parent distribution series (PDS) or the series of maxima computed over a given period (month, season, year). Under the hypothesis that the compound Poisson process (CPP) (eg see Todorovic and Zelenhasic, 1970) is a valid model for successive rainfall depths accumulated over a given duration, the asymptotic exponential decay of the PDS involves the asymptotic exponential decay of the extreme value series (EVS). While it is more convenient to deal with EVS, since the samples are smaller and complete rainfall data processing is not necessary, the heavy right hand tail of some EVS sometimes casts doubt on the validity of the classical Gumbel distribution (EV1) as a suitable model of rainfall extreme value distributions. This is particularly the case for short duration rainfall all around the Mediterranean basin since the weather patterns accounting for monthly or seasonal maxima vary greatly

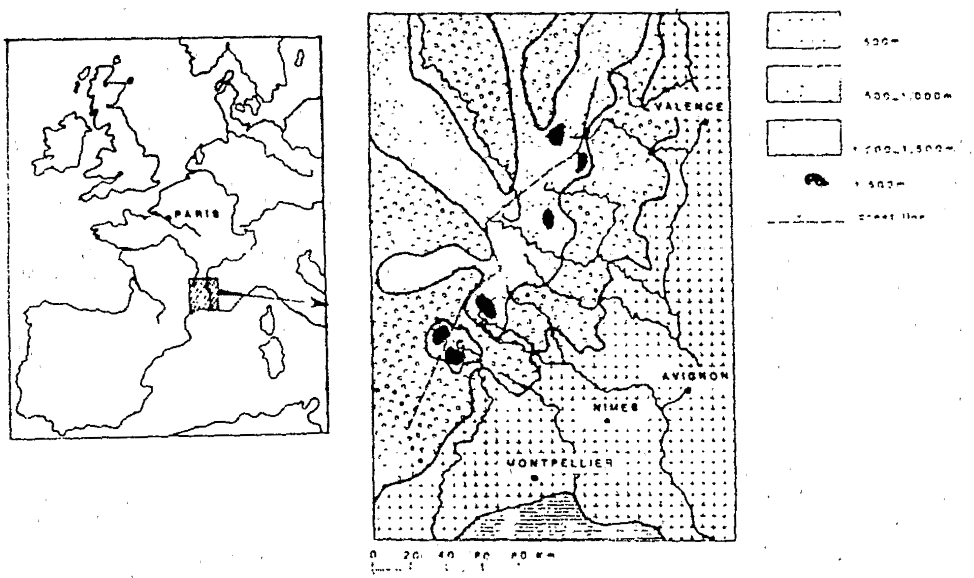

Figure 1. The Cevennes Region of France.

throughout the year. Accordingly, the year must be carefully divided into homogeneous seasons before any attempt to analyze such series. In this part of France, the large majority of severe weather patterns occurs in fall, which essentially means September, October, November. Of course, the distribution of interest is that of annual rainfall, but when almost every yearly maximum occurs during the same season, clearly the distribution of yearly maxima rapidly becomes parallel to that of the concerned season maxima(Guillot and Duband, 1967).

However, large skewness in EVS may be found, even when working on homogeneous season data, and in such cases, the EV1 distribution does not properly fit observed data. The estimation of the skewness coefficient and the detection of outliers has been given great attention by hydrologists for the past twenty years, mainly with respect to flood series analysis (e.g. Bernier, 1967; Todorovic and Rousselle 1971; Matalas et al.,1975; Gupta et al., 1976; Kottegoda, 1984). By comparison, relatively few papers have been concerned with the question of how to deal with outliers in rainfall EVS, using PDS information. To our knowledge, D.Duband (1967) was the first to consider a two component exponential distribution as a model of the underlying distribution of daily rainfall depths, in order to make better use of the information contained within the rainfall data. Since the coefficient of one component is greater than that of the other, the asymptotic behavior of the model is similar to that of a single exponential; as a consequence, the derived distribution of extremes is an EV1 distribution, following Gumbel's asymptotic theory (e.g. Gumbel, 1958). It must therefore be emphasized that, if the underlying distribution is truly a two component exponential distribution, the study of PDS or EVS should lead to the same estimates of g, not taking into account model and sampling errors. Such errors are of dramatic importance, since g conditions the extrapolation of flood series to high return period values. Hence, a comparative study of three estimates was carried out, two of them applying to EVS and the other to PDS.

2. THE GRADEX METHOD OF ESTIMATING HIGH RETURN PERIOD FLOODS

Pointing out that very often available flood series are inadequate for reliable extrapolation of the experimental distribution towards extreme flood flows, Guillot and Duband (1967) proposed a method using rainfall data analysis to extrapolate flood series distributions. In flood conditions, precipitation and flood volumes computed over a common integration time tg :

$$R_g = \int_{t_1}^{t_1+t_g} R(t)\,dt \quad and \quad Q_g = \int_{t_2}^{t_2+t_g} Q(t)\,dt$$

can be plotted on a graph as in figure 2. If the integration time is close to the base time of the overland flow hydrograph (tb), the portion of the rainfall Dg = Rg-Qg not reaching the main stream is related to the overall basin retention during the total time (t1, t2 + tg). This retention Dg is a random quantity tending towards an upper limit as long as the rainfall intensity is large compared to the percolation rate to the ground water tables.

Most hydrologists now agree that for moderately permeable watersheds of a few hundred to a few thousand square kilometers, it is physically sound and reasonably safe to assume that, as the total basin saturation is approached, any additional precipitation Pg tends to give rise to a nearly equal increase in the discharge Qg. Thus the asymptotes of the quantile curves of the joint distribution (Rg, Qg) (figure 2) are straight lines with slopes of one (Qg = Rg), whatever the quantile considered. The total amount of precipitation needed to

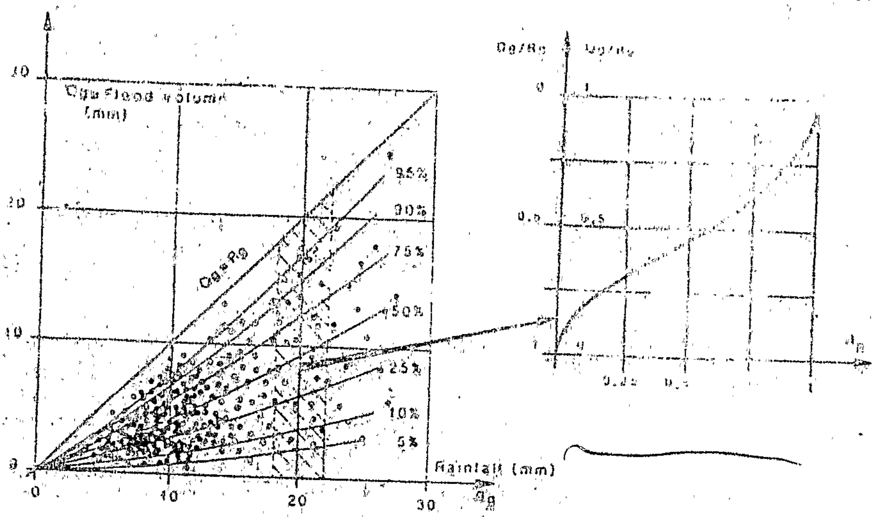

Figure 2. Conditional distribution of flood volumes with [illegible] rainfall [illegible]
Dg = Qg- Pg is the retention.

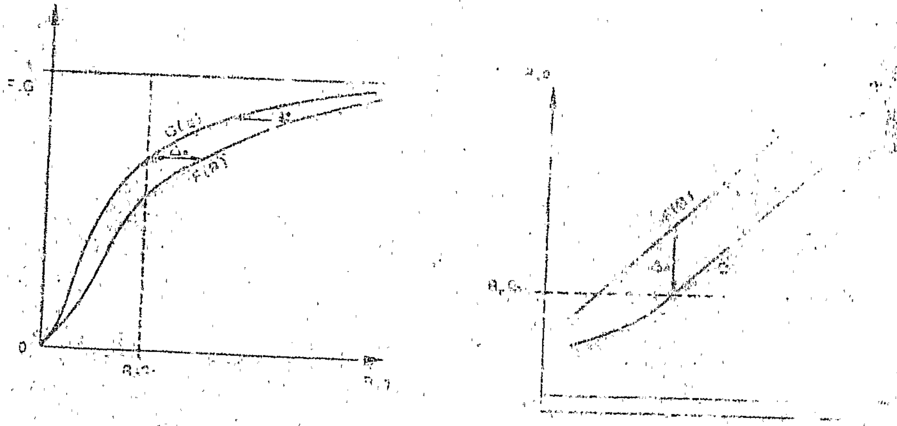

Figure 3. Parallelism of rainfall and discharge [illegible] when plotted on a [illegible] probability paper. The equidistance of the distributions [illegible] and F(R), beyond the [illegible] (Ro, qo) in arithmetic coordinates(left) involves the parallelism of the asymptotes [illegible] coordinates(right).

reach saturation depends on the initial dryness of the soil, as does the distance between each asymptote and the curve Q = R.

When large samples of rainfall-flood data are available, the conditional probability is obtained from the experimental cumulative distributions Hp(Qg), and the theoretical flood volume cumulative distribution is given by:

$$G(Q_g) = \int_{q=0}^{q=Qg} \int_{R=0}^{R=\infty} h_p(q)\ dF(R)\ dq$$

Furthermore, a simple and operationally important consequence results from the assumptions of the method. If the rainfall distribution is asymptotically an exponential, the extrapolated asymptote of G(Qg) is translated from the asymptote of F(Rg) in a direction parallel to the variate (Rg,Qg) axis. In particular, on conventional Gumbel probability paper, these two asymptotes are parallel straight lines (fig. 3).To sum up, while it is very unlikely that a few decades of flood measurements allow a reliable estimation of the direction of extrapolation, the gradex method assumes that the numerous long series of rainfall data available almost anywhere in the world can provide a safe extrapolation means. Indeed, extrapolation of flood series using a model fitted to flood data only can lead to tremendous overestimation as it occurs when fitting a Frechet or a Galton model to the flood data sample. The use of log space (Z = Log(Q)) distributions has also been shown to be dangerous and counterproductive by Landwher et al.(1978); the fact that a log-type distribution is often a better description of the observed reality does not imply that it is suitable for extrapolation purposes. Accordingly, Waylen et al.(1982) proposed a two component exponential distribution to fit annual flood series (AFS) accounting for the mixture of snowmelt and rain generated floods throughout the year. Rossi et al.(1984) proposed a similar distribution for annual flood series of central and southern Italy, presenting a large skewness. Although this distribution closely reproduces the observed skewness, Rossi et al. emphasized the great uncertainty related to the parameter estimates of the outlying component (i.e. the one determining the extrapolation) when a single 40 year AFS is used, which means that in many cases the available discharge measurements can only to be used to give a good estimate of the starting point (Q0) of the extrapolation(figure 3).

For all its advantages when applied to relevant watersheds, the gradex method presents some practical difficulties which are not central to the basic concept, but which need some clarification and further study, as pointed out by Beran(1981). Some are related to the time of integration and the peak to volume adjustment factor Others are rainfall related problems, among which the transformation of point rainfall statistics into areal rainfall is crucial. An initial procedure was presented by Lebel (1984), yielding good results for small upland catchments, but not yet extended to larger watersheds. It is also important to ascertain the best way to estimate the point gradex from rainfall series. Indeed an important characteristic of g is its spatial structure (derived from the spatial stucture of rainfall) allowing the mapping of this parameter. Such maps are needed by hydraulics engineers as a basic design tool and their preparation requires the interpolation of homogeneous point values of g, which in turn means a regionally dependable estimation procedure. Also, gradex maps are needed for several rainfall durations, since the time of integration tg is related to watershed areas. This requirement gives rise to another set of difficulties since recording raingage series are fewer and often shorter than their raindepth counterparts, and outliers are frequently more numerous in hourly rainfall than in daily rainfall series.

3. ESTIMATING THE GRADEX FROM EXTREME VALUE SERIES

A physically based explanation was given by Giraud (1959) to support the idea that, at least for durations equal to or smaller than 24 hours,the asymptotic behavior of rainfall distributions is exponential. While this has been widely accepted for decades, discussion continues as to which procedure provides the best estimate of the slope g of this exponential. The simplest method is to assume that the consecutive rainfalls PDS has a single exponential distribution, from which it is a straightforward task to derive the associated EVS distribution. Obviously, the single exponential cannot hold for every PDS, and it may be more general to write the equation for values above a threshold S0; the number of events $\{R>r\}$ for N independent observations of R is then writen as :

$$N_R(r) = \exp(-\alpha * (r-r_o)) \qquad \text{for all } r>S0 \qquad (1)$$

At the threshold:

$$N_R(S0) = \exp(-\alpha * (S0-r_o)) \qquad (2)$$

and below the threshold, $N_R(r)$ is unknown. Further assuming that :

$$P_1 = P(r<S0) \qquad (3a)$$

$$P_2 = 1-P_1 = P(r>S0), \qquad (3b)$$

NP_2 values above S0 are observed on the average for N independent observations of R. Thus, $NP_2 = N_R(S0)$, and from (2), we obtain:

$$r_o = S0 + \text{Log}(NP_2)/\alpha \qquad (4)$$

Now let X be the maximum of the N independent observations. Then:

$$F_x(r) = \{F_R(r)\}^N = \{1-N_R(r)/N\}^N \qquad r>S0 \qquad (5a)$$

$$F_x(S0) = P_1^N \qquad r = S0 \qquad (5b)$$

Assuming N to be large with respect to N_R values, we obtain :

$$F_x(r) \# \exp(-N_R(r)) \qquad \text{for } S0<r, \qquad (6)$$

which substituting (1) and (4) for $N_R(r)$, gives:

$$F_x(r) = \exp\{-\exp(-\alpha * (r-r_o)\} \qquad (7)$$

with r related to S0 and N by expression (4). Expression (7) is the EV1 distribution, and $1/\alpha = g$ is the gradex. It holds for $r>S0$, while for $r<S0$ the distribution is unknown, but very likely different from that given by (6 and 7).

If the basic model (1) is plausible, it has certain practical consequences when fitting an EV1 model to EVS:

1- There is a ($100*P_1^N$) per cent chance that a given value of the EVS does not belong to the EV1 distribution ($r<S0$).

2- The upper part of the distribution is EV1, and a separation appears at S0. The gradex g is equal to $1/\alpha$.

3- As N increases, the probability of having an EVS value below the threshold S0 decreases, and the better is approximation (6).

To illustrate these statements, consider a PDS of daily rainfalls and an EVS of monthly maxima (N1 = 30), with a recurrence rate of 93% for the event {r<5 mm}(S0 = 5mm). Thus $P = (.93)^{K} = .142$, and hence, on the average, the fourteen lowest values out of a hundred do not belong to the EV1 distribution. To get rid of this left hand tail, an EVS of season maxima (N2 = 90) may be chosen. The probability of a low value outlier then falls to .001, and the EV1 model fits the sample data very well (fig.4). Of course, the recording period must be long enough to allow the reliable fitting of any model. Thirty values is widely acknowledged to be a minimum. Therefore, short rainfall series displaying a break in the point swarm of monthly maxima call for another kind of modeling if a dependable estimate of g is to be computed. This is also the case for distributions with several high value outliers. This generally denotes the mixing of two underlying populations (R1 and R2), the recurrence rate of R2 being small compared to that of R1, but R2 yielding much heavier rainfalls. If the two distributions of the maxima X1 and X2 of each population are denoted respectively F_{X1} and F_{X2} and if they are independent, the resulting cumulation function of the rainfll X is given by:

$$F(X) = pF_{X1}(X) + (1-p)F_{X2}(X),$$

where p is the probability of occurrence of X1. In the CEVENNES region, frontal and convective rains are often mixed in the same rainfall event in fall and it proves difficult to separate data originating in the first distribution from that originating in the second distribution. On the one hand, table 1 clearly shows that long and intense rainfalls are not always associated (compare event # 31 with event # 28). On the other hand, events #2 and 23, among others, were long and had strong hourly maxima around 60 mm. Hence, while some events belong to only one of the above distributions, others result from the addition of two physical processes.To get around this difficulty, it was decided to test the suitability of a direct model of the complete PDS, when the series was available.

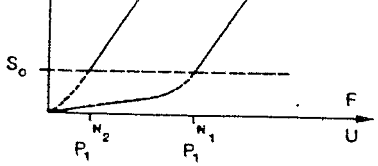

Figure 4 Elimination of low value outliers. When the number of events over which the maxima are calculated is increased, the break in the lower part of the EVS distribution disappears

4. MODELING OF PARENT DISTRIBUTION SERIES

Let the PDS be a sequence of N consecutive rainfalls, Ri, $i = 1,N$, for a given duration of accumulation, including zero values. Duband(1967) assessed that the sum of two exponentials represented a good model for fitting daily rainfall distributions of the central part of France :

$$F_R(r) = 1 - \gamma \exp(-r/g) - \beta \exp(-r/c), \quad r >= 0, \tag{8}$$

with $g > c$. (9)

Table.1 Main rainfall events of the ten past ten years. Rainfall depth in mm.

Event number	Beginning Date	Duration (approximate)	Mean Depth° for Event	Event Maximum	Hourly Maximum
1	19 sept 1971	30 h	23.1	68.4	28.5
2	4 sept. 1972	66 h	85.0	227.5	57.0
3	10 oct. 1972	80 h	21.4	241.7	90.0
4	25 oct. 1972	36 h	100.8	246.3	37.0
5	20 sept. 1973	34 h	34.6	272.0	80.0
6	1 oct. 1973	40 h	135.2	183.0	35.0
7	2 nov. 1973	90 h	94.2	352.0	36.5
8	16 sept. 1974	90 h	148.6	371.8	65.0
9	14 nov. 1974	103 h	64.8	219.0	19.0
10	17 nov. 1974	40 h	35.3	82.5	40.0
11	14 sept. 1975	56 h	63.4	176.0	70.0
12	29 sept. 1975	30 h	61.9	161.5	57.0
13	28 aug. 1976	90 h	148.0	392.0	60.0
14	11 sept. 1976	40 h	110.7	400.5	85.0
15	27 sept. 1976	52 h	57.7	145.5	40.5
16	30 sept. 1976	53 h	46.8	120.7	33.0
17	10 oct. 1976	92 h	78.8	178.8	34.5
18	23 oct. 1976	160 h	175.1	404.9	23.0
19	8 nov. 1976	65 h	180.2	371.4	42.5
20	28 jul. 1977	70 h	81.8	131.0	19.0
21	25 aug. 1977	48 h	87.9	266.5	49.0
22	5 oct. 1977	75 h	121.5	355.7	35.2
23	19 oct. 1977	120 h	273.3	596.5	61.0
24	25 oct. 1977	72 h	43.7	117.5	30.0
25	4 oct. 1979	36 h	50.7	118.0	71.0
26	6 oct. 1979	27 h	51.3	121.0	26.0
27	10 oct. 1979	160 h	198.6	566.0	38.5
28	22 oct. 1979	160 h	228.0	375.0	16.5
29	15 aug. 1980	22 h	35.7	86.7	51.0
30	26 aug. 1980	30 h	49.5	90.7	45.0
31	20 sept. 1980	32 h	176.8	591.0	100.0
32	15 oct. 1980	46 h	141.1	204.2	60.0

° Mean value of 59 raingages.

Expressions (8) and (9) imply that, for large values of r,

$$F(r) \# 1 - \gamma \exp(-r/g),$$

thus, the EVS are asymptotically EV1 distributed. This function (8) was fitted using a method similar to the method of moments, involving the estimation of four parameters γ, β, g,c, assuming at first the invariance of the coefficient of variation σ/μ, and fixing the relation:

$$\gamma = \mu^2 / \sigma^2 ,$$

where μ and σ are the mean and the standard deviation of the sample. This gives:

$$g = (\mu\gamma + (\gamma\beta(K\theta - \mu^2))^{1/2})/\gamma \quad (10a)$$

$$c = (\mu - \gamma a)/\beta , \quad (10b)$$

with $\beta = 1 - F(0) - \gamma$; $\theta = 1 - F(0)$; $K = (\mu^2 + \sigma^2)/2$ (10c)

5. COMPARISON OF THREE GRADEX ESTIMATES

Three estimates of the fall gradex were first compared:
* g_1: slope of the conventional moment fitting of the EV1 model to EVS.
* g_2: slope of the maximum likelihood fitting of the EV1 model.
* g_3: parameter of the model (8) computed using formula (10a) above.

Among the 54 recording raingages available in 1982, 25 were selected to carry out this comparison.

5.1. Gumbel Distribution Fitted to Period Maxima by the Method of Moments.

The method was tested on rainfall maxima over periods of 10 days, 15 days, one month, one year (i.e. one season) and two years (two seasons), with special attention paid to possible changes at any breaks in the point swarm. The aim was to determine the minimum sampling period required for correct estimation of the Gumbel parameters by the moments. As a matter of fact, when the empirical distribution does not correspond perfectly to a Gumbel distribution and when a curve or break occurs in the point swarm, this method gives a gradex which is biased by the values at the bottom of the distribution in the second order moment. Several stations offering long duration records were used, such as Montpellier Bel-Air (1922-1980), allowing the determination of maxima over periods of up to three years.The resulting lines of fit for maxima calculated over 5-day, 15-day, 1-month, 1-year, 2-year periods are shown for hourly and daily rainfalls in figure 5, revealing a break starting from the first 5-day period. This break becomes less and less distinct as the period length increases, becoming virtually negligible for annual maxima. For some other stations, monthly maxima series are almost as good as annual maxima series. In figure 5, each EVS displays a straight line asymptpotic behavior, the lines being roughly parallel to one another. Expression (4) allows computation of the theoretical distance between the asymptotes; if one asymptote is obtained with a sample of N1 independent observations, and the other with a sample of N2 independent observations, then:

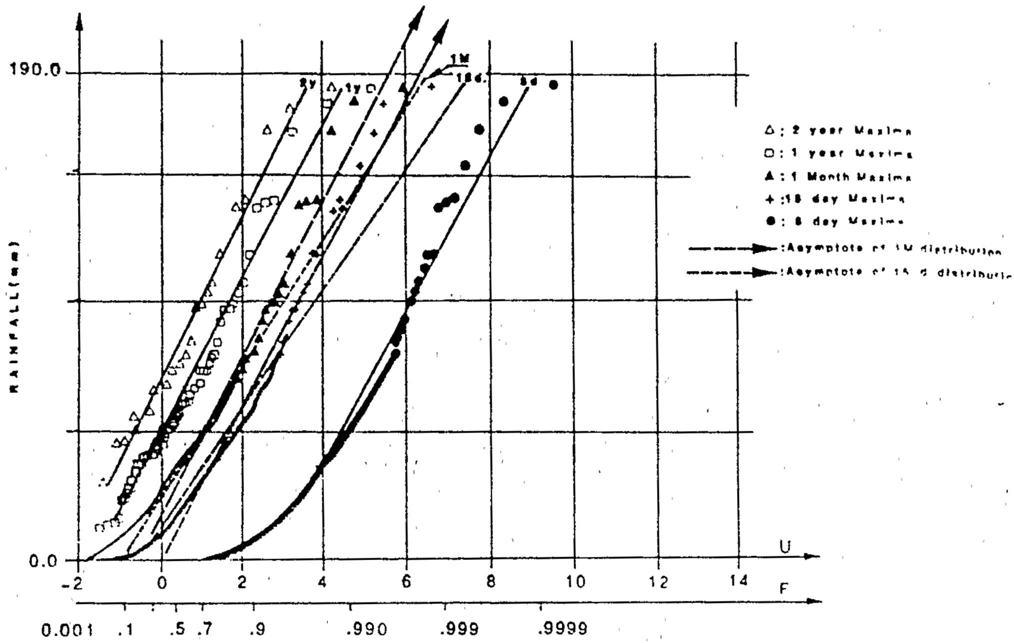

Figure 5. EVS distributions of daily rainfalls at Montpellier Bel-air. The maxima are calculated over 5-day, 15-day, 1-month, 1-year, and 2-year periods.

$$r_o(N1) - r_o(N2) = (\text{Log } N1 - \text{Log } N2)/\alpha = g * \text{Log}(N1/N2),$$

which is the distance along the rainfall axis. The ratio N1/N2 is equal to 3 when the period P1 is the month and the period P2 the season. The experimental distances measured in figure 6 are very close to these theoretical values. Even though the general asymptotic behavior is the same for each EVS, the EV1 model appears to be suitable only for periods greater than a season in the case of Montpellier Bel-Air and computing g_1 from the monthly maximum series would lead to an underestimation of the gradex.

5.2. Gumbel Distribution Fitted by the Method of Maximum Likelihood.

The estimate is given by the numerical resolution of the following system:

$$\bar{x} = (g_2 + \sum_{i=1}^{n} x_i \, exp(-y_i)) / \sum_{i=1}^{n} exp(-y_i)$$

$$r_o = -g_2 \, Log\left((\sum_{i=1}^{n} exp(-y_i))/n \right) \qquad (11)$$

where x is the mean of the sample x_i and $y_i = x_i/2$.

The fit is only good if the sample corresponds perfectly to a Gumbel distribution. Indeed, as soon as there is a slight break, the line fitted by this method shows relatively large deviations with respect to high values (fig.6) and the distribution parameters are underestimated. The same finding prompted Fiorentino and Gabriele (1984) to propose a correction. If g_2 is the gradex estimated by the maximum likelihood method, then the corrected parameters may be expressed as:

$$g'_2 = g_2/(1-.8/n), \text{ and}$$

$$r'_0 = g'_2 * \mathrm{Log}(n/\sum_{i=1} \exp[-y_i]) - .7 * g'_2/n$$

This correction is based on the results of tests on samples simulated according to the Gumbel distribution with $g=1$ and $r_0 = 0$. To evaluate its performance, we tested it on experimental samples. The correction unfortunately never exceeds 2%, which is insufficient to correct the bias of the M.L. method on our rainfall samples. This was to be expected, since the correction is based on the hypothesis that the population is exactly EV1 distributed, while we are searching for reliable estimates of g for approximately EV1 distributed EVS .

Another approach was taken by Revfeim (1983). In the procedure described in section 3 it was shown that the EV1 model is the limiting form of the cumulative distribution function $[F(R)]^N$, when N tends to infinity, if the PDS model is an exponential. Alternatively, the rain may be defined as a succession of rainfall events, the number of which in a fixed

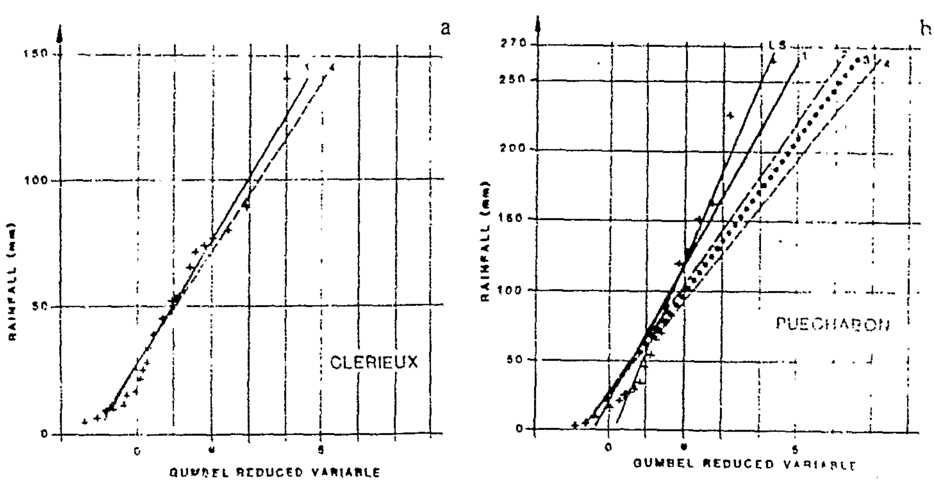

Figure 6: Fitting the EV1 distribution to monthly maximum distributions.
a: twelve-hour rainfall; Clerieux raingage.
b: two-hour rainfall; Puechabon raingage.
LS: least square fitting to the upper part of the observed distribution.
1: moment fitting; 4,2,3: maximum likelihood fitting using expreesion 11 and corrections proposed by Revfeim (1983) and Fiorentino(1984) respectively.period of time has a Poisson

distribution. If the rainfall depth of each event has an exponential distribution, the extreme value model is also an EV1 distribution, the coefficients of which can be related to the parameters of the underlying process, i.e the mean of the process (μ) and the recurrence rate(ρ). Refveim (1983) proposed maximum likelihood estimates ($\hat{\mu}$, $\hat{\rho}$) of these parameters, from which the estimated parameters of the EV1 distribution are readily obtained as:

$$g''_2 = .78\, S_x = \hat{\mu},$$

where S_x is the standard deviation of the EVS, and:

$$r''_0(\rho) = \mu \operatorname{Log} \hat{\rho} = g''_2 \operatorname{Log} \hat{\rho}$$

Note that the gradex g''_2 is the mean of the underlying process, but this process is not the sequence of every consecutive rainfall considered in the first approach described above. Hence, the method proposes new estimates of the EV1 coefficients, but no additional mean to test the suitability of the model itself for a given EVS. Figure 6 shows how a few low value outliers would lead to a biased estimate of g, especially for methods sensitive to these low values, as maximum likelihood methods generally are.

After devoting a great deal of time and effort to M.L. methods, it appears that they can be ruled out for fitting an EV1 distribution to rainfall data in the Mediterranean region, at least for time steps ranging from 1 to 24 hours.

5.3. Two Component Exponential Function Fitted to all Successive Rainfalls.

Figure 5 shows that increasing the period over which the maxima are determined eliminates the break; however the sample size is reduced (divided by 3 in the case of one season instead of one month). The risk is therefore to have insufficient sample size when using seasonal maxima, even though the sample follows a virtual Gumbel distribution. On the other hand, if the period is decreased, the break becomes increasingly distinct, leading to a limiting case consisting of two lines. This is in particular the case when the period taken is equal to the rainfall duration considered (i.e all consecutive rainfalls are taken). On Gumbel paper, these two lines require a special distribution function corresponding to the sum of two exponentials. This function accounts for both tails of the distribution with the low values represented by the second exponential and the values above a certain threshold dominated by the first exponential. The quality of the fitting of this function to maxima at low durations using expressions 10 a,b,c was poor. In order to improve short time step estimates, an experimental study of the functions $\gamma = f(t)$ and $\gamma = f(F(0))$ was carried out, allowing a function CK(t,F(0)) to be established, by which $\gamma = \mu^2 / \sigma^2$ can be multiplied :

$$CK(t,F(0)) = 1/(1-\operatorname{Log}(t/24)) \qquad (12)$$

For t= 24 hours, CK(t,F(0)) = 1 and its application is not possible for durations greater than 24 hours since it has been established empirically.

With this correction the fit of the two component exponential model was improved for all 25 stations, especially at short time steps (figure 7). For Montpellier Bel-air (52 years of recording available over the period 1922-1980), it is noteworthy that the gradex obtained by the moment fitting of the EV1 model to seasonal maxima and by the corrected fitting of the two component exponential model are very close. It was also observed that on most EVS series displaying a break in the low values,a least square fitting to the upper part of the point swarm (figure 6) provided better gradex estimations than the moment fitting.

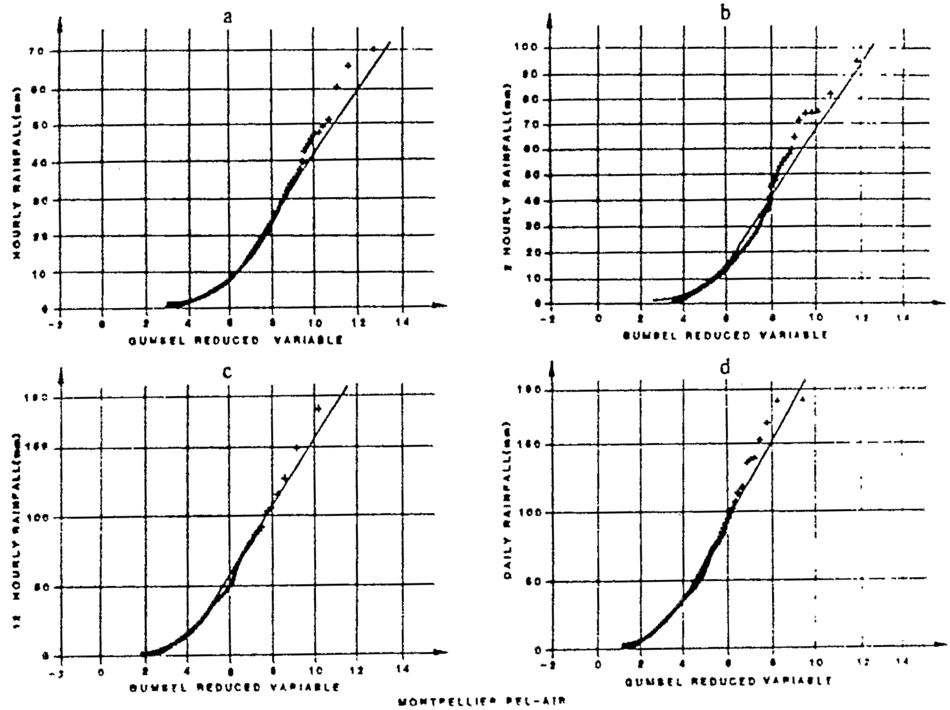

Figure 7 Fitting the two component exponential model to the rainfalls of Montpellier Bel-air. The parameters of the model are computed using the correction function CK(t,F(0)).

Table 2 Average ratios of the three estimates, computed on 25 stations. g_1 is the conventional moment estimate, while g_2 g_3 and g'_3 are calculated using expressions 11 10a and 12 respectively.

Duration (hours)		1	2	4	6	12	24
Ratios	g_1/g_2	1.15	1.15	1.20	1.20	1.30	1.40
	g_1/g_3	1.30	1.25	1.20	1.05	0.85	0.80
	g_1/g'_3	1.12	1.08	1.04	1.03	0.84	0.80

6. COMPARISON OF THE THREE METHODS

Table 2 shows that:

a) Maximum likelihood estimates (computed with expression (11)) give for all time steps consistently lower estimations of g than the other two estimates.

b) For time steps greater than 6 hours, the EV1 model gives lower estimates than the two component exponential model. Experiments performed on long duration series indicate that the estimates provided by the later model are more reliable.

c) For short time steps, none of the method is really satisfactory. The estimate g_1 is good when no break occurs in the point swarm, but underestimates the gradex in cases such as the one in figure 6.

7. CONCLUSION

The fitting of a two component exponential model to all consecutive rainfalls is the method that uses the most hydrometeorological information and appears to be the best as long as the necessary data is available. The method proved to be especially suitable when corresponding extreme value series cannot be well fitted by a Gumbel distribution. In cases where only extreme value series can be used, it was observed that maximum likelihood estimates could not compete with less sophisticated but more efficient moment fitting or least square fitting estimates.

When a break or a curve occurs in the point swarm, the least squares fit accounts for this break, and the upper part of the empirical distribution corresponds to a Gumbel distribution. If this part contains enough values, this method yields gradex estimates close to those of the two component exponential model. However, the robustness of the least squares estimate is poor when there are several high value outliers. The main advantage of working on parent distribution series rather than on extreme value series is therefore the minimization of the influence of these high value outliers, observed in rainfall in the Mediterranean area.

ACKNOWLEDGEMENTS

The study was funded by the Gard "Direction Departementale de l'Equipement" (DDE) . The Division Technique Generale of Electricite De France and the Service Regionale d'Amenagement des Eaux of the Languedoc Roussillon Region provided indispensable support.

REFERENCES

Beran,M. 1981. The "gradex" method, letter to the editor. IAHS bulletin Vol. 26, 2 pp. 223-224.

Bernier,J., 1967. Sur la theorie du renouvellement et son application en hydrologie. Bulletin de la Direction des Etudes et Recherches, Electri-cite de France, Serie B Vol.10.

Duband,D., 1967. Distribution of the sum of daily precipitations. in: International Hydrology Symposium, vol. 1 pp. 516-526. Colorado State University, Fort Collins, Colorado.

Fiorentino, M., Gabriele, S. 1984. A Correction for the Bias of Maximum-Likelihood Estimators of Gumbel Parameters. Journal of Hydrology. Vol. 73, pp.39-49.

Giraud, M. 1959. Initiation aux Processus Aleatoires. Dunod Editor, Paris.

Guillot,P.,Duband,D., 1967. The Gradex to Compute Extreme Flood Probability from the Observation of Rainfall. in: International Hydrology Symposium, vol. 1 pp. 506-515. Colorado State University, Fort Collins, Colorado.

Guillot,P., Duband,D., 1968. La methode du Gradex pour le calcul de la probabilite des crues a partir des pluies. Xmes journees de l'hydraulique, Societe Hydrotech. de France, Paris 1968.

Gumbel,E.J., 1958. Statistics of Extremes, Columbia University Press, New York.

Gupta,V.K., Duckstein,L., andPeebles,R.W. 1976. On the Joint Distribution of the Largest Flood and its Time Occurrence. Water Resources Research, Vol. 12(2), pp.295-304.

Kottegoda, N.T., 1984. Investigation of Outliers in Annual Maximum Flow Series. Journal of Hydrology, Vol. 72, pp. 105-137.

Landwher, J.M., Matalas, N.C., and Wallis, J.R. 1978. Some Comparison of Flood Statistics in Real and Log Space. Water Resources Research, Vol. 14(5), pp.902-920.

Lebel,T., 1984. Moyenne Spatiale de la Pluie sur un Bassin Versant: Estimation Optimale, Generation stochastique et Gradex des Valeurs Extremes. Thesis, Institut National Polytechnique de Grenoble, France.

Matalas, N.C., Slack, J.R., and Wallis, J.R. 1975. Regional Skew in Search of a Parent. Water Resources Research, Vol. 11(6), pp.815-826.

Revfeim, K.J.A., 1983. On the Analysis of Extreme Rainfalls. Journal of Hydrology, Vol. 62 pp. 107-117.

Rossi,F., Fiorentino, M., and Versace, P. 1984. Two Component Extreme Value Distribution for Flood Frequency Analysis. Water Resources Research, Vol. 20(7), pp. 847-856.

Todorovic,P. and Zelenhasic,E. 1970. A Stochastic Model for Flood Analysis. Water Resources Research, Vol. 6(6), pp. 1641-1648.

Todorovic,P. and Rousselle, J. 1971. Some Problems of Flood AnalysisWater Resources Research, Vol. 7(5), pp. 1144-1150.

Waylen,P.,and Woo,M.K. 1982. Prediction of annual floods generatedby mixed processes. Water Resources Research Vol.18(4), pp.1283-1286

HYDROLOGIC FREQUENCY MODELING

Proceedings of the International Symposium
on Flood Frequency and Risk Analyses,
14–17 May 1986, Louisiana State University, Baton Rouge, U.S.A.

Edited by

VIJAY P. SINGH

Department of Civil Engineering,
Louisiana State University, Baton Rouge, U.S.A.

D. REIDEL PUBLISHING COMPANY

A MEMBER OF THE KLUWER ACADEMIC PUBLISHERS GROUP

DORDRECHT / BOSTON / LANCASTER / TOKYO