Technical Assistance
HYDROLOGY PROJECT
World Bank & Government of The Netherlands funded

Training module # SWDP - 12

How to analyse rainfall data

New Delhi, February 2002

CSMRS Building, 4th Floor, Olof Palme Marg, Hauz Khas,
New Delhi – 11 00 16 India
Tel: 68 61 681 / 84 Fax: (+ 91 11) 68 61 685
E-Mail: hydrologyproject@vsnl.com

DHV Consultants BV & DELFT HYDRAULICS

with
HALCROW, TAHAL, CES, ORG & JPS

Table of contents

1. Module context

While designing a training course, the relationship between this module and the others, would be maintained by keeping them close together in the syllabus and place them in a logical sequence. The actual selection of the topics and the depth of training would, of course, depend on the training needs of the participants, i.e. their knowledge level and skills performance upon the start of the course.

2. Module profile

Title	:	How to analyse rainfall data
Target group	:	Assistant hydrologists, Hydrologists, Data Processing Centre Managers
Duration	:	Three sessions of 60 min
Objectives	:	After the training the participants will be able to: 1.Carry out prescribed analysis of rainfall data
Key concepts	:	• Homogeneity of rainfall data sets • Basic statistics • Annual maximum & exceedence series • Fitting of frequency distribution • Frequency and duration curves • Intensity-duration-frequency curves • Depth-area-duration curves
Training methods	:	Lecture, software
Training tools required	:	Board, computer
Handouts	:	As provided in this module
Further reading and references	:	

3. Session plan

No	Activities	Time	Tools
1	***General*** • Analysis of rainfall data	3 min	 OHS 1
2	***Checking data homogeneity*** • General	3 min	 OHS 2
3.	***Computation of basic statistics*** • Basic statistics • Example 3 (a) – Statistics for monthly rainfall series • Example 3 (b) - Plot of frequency distribution	4 min	 OHS 3 OHS 4 OHS 5
4.	***Fitting of frequency distributions*** • General • Available distribution • Analysis results • Example 4 (a) – Basic statistics • Example 4 (b) – Goodness of fit test • Example 4 (c) – Annual maximum of daily data series • Example 4 (d) – Annual maximum of monthly data series • Example 4 (e) – Annual data series	10 min	 OHS 6 OHS 7 OHS 8 OHS 9 OHS 10 OHS 11 OHS 12 OHS 13
5.	***Frequency-duration curves*** • Definition – Frequency curves • Example 5.1 (a) – Frequency curves • Definition – Duration curves • Example 5.1 (b) – Duration curves • Definition – Average duration curves • Example 5.1 (c) – Average duration curve	10 min	 OHS 14 OHS 15 OHS 16 OHS 17 OHS 18 OHS 19
6.	***Intensity-duration-frequency IDF-curves*** • Definition and computational procedure • Illustration of first step in computational procedure • Illustration of second step in computational procedure • Illustration of third step in computational procedure • Mathematical forms of IDF-curves • T = 50 years Iso-pluvials for 1-hour maximum rainfall • T = 50 years Iso-pluvials for 24-hour rainfall • Definition of partial duration series • Annual maximum and annual exceedance series • Conversion of return period between annual maximum and annual exceedance series • Ratio of return periods for annual exceedance and annual maximum series • Example 7.1 (a) – EV1-fit to hourly rainfall extremes • Example 7.1 (b) – EV1-fit to 1-48 hr rainfall extremes • Example 7.1 (c) – IDF analysis tabular output • Example 7.1 (d) – IDF analysis tabular output continued • Example 7.1 (e) – IDF curves linear scale (ann. maximum series) • Example 7.1 (f) – IDF curves log-log scale (ann. maximum series) • Example 7.1 (g) – IDF analysis for ann. exceedances, tabular output	45 min	 OHS 20 OHS 21 OHS 22 OHS 23 OHS 24 OHS 25 OHS 26 OHS 27 OHS 28 OHS 29 OHS 30 OHS 31 OHS 32 OHS 33 OHS 34 OHS 35 OHS 36 OHS 37

	• Example 7.1 (h) – IDF curves linear scale (ann. exceedances) • Example 7.1 (i) – IDF curves log-log scale (ann. exceedances) • Example 7.1 (j) – Generalisation of IDF curves in mathematical formula • Example 7.1 (k) – Test on quality of fit of formula to IDF-curves		OHS 38 OHS 39 OHS 40 OHS 41
7.	***Depth-area-duration analysis*** • Definition and computational procedure • Example 8.1 (a) – Division of basin in zones • Example 8.1 (b) – cumulation of rainfall per zone • Example 8.1 (c) - cumulation of rainfall in cumulated zones • Example 8.1 (d) – Drawing of DAD-curves • Areal reduction factor, definition and variability • Example of ARF's for different durations • Time distribution of storms, pitfalls • Example of time distribution of storms	15 min	 OHS 42 OHS 43 OHS 44 OHS 45 OHS 46 OHS 47 OHS 48 OHS 49 OHS 50
	Exercise • Compute basic statistics for a monthly rainfall data series • Use normal distribution fitting for maximum monthly rainfall data series • Derive frequency duration curves for a ten daily rainfall data series • Derive IF-curves for various durations • Derive IDF-curves for annual maximum series • Derive IDF-curves for annual exceedances • Repeat IDF analysis for other stations in Bhima basin • Fit regression equation to IDF-curves	 10 min 10 min 10 min 10 min 5 min 5 min 20 min 20 min	

4. Overhead/flipchart master

5. Handout

Add copy of the main text in chapter 7, for all participants

6. Additional handout

These handouts are distributed during delivery and contain test questions, answers to questions, special worksheets, optional information, and other matters you would not like to be seen in the regular handouts.

It is a good practice to pre-punch these additional handouts, so the participants can easily insert them in the main handout folder.

7. Main text

Contents

How to analyse rainfall data

1 General

- **The purpose of hydrological data processing software is not primarily hydrological analysis. However, various kinds of analysis are required for data validation and further analysis may be required for data presentation and reporting.**

- **The types of processing considered in this module are:**

  - checking data homogeneity
  - computation of basic statistics
  - annual exceedance rainfall series
  - fitting of frequency distributions
  - frequency and duration curves

- **Most of the hydrological analysis for purpose of validation will be carried out at the Divisional and State Data Processing Centres and for the final presentation and reporting at the State Data Processing Centres.**

2 Checking data homogeneity

For statistical analysis rainfall data from a single series should ideally possess property of homogeneity - i.e. properties or characteristics of different portion of the data series do not vary significantly.

Rainfall data for multiple series at neighbouring stations should ideally possess spatial homogeneity.

Tests of homogeneity is required for validation purposes and there is a shared need for such tests with other climatic variables. Tests are therefore described in other Modules as follows:

Module 9 **Secondary validation of rainfall data**
- **A. Spatial homogeneity testing**
- **B. Consistency tests using double mass curves**

Module 10 **Correcting and completing rainfall data**
- **A. Adjusting rainfall data for long-term systematic shifts - double mass curves**

Module 17 **Secondary validation of climatic data**
- **A. Single series tests of homogeneity, including trend analysis, mass curves, residual mass curves, Student's t and Wilcoxon W-test on the difference of means and Wilcoxon-Mann-Whitney U-test to investigate if the sample are from same population.**
- **B. Multiple station validation including comparison plots, residual series, regression analysis and double mass curves.**

3 Computation of basic statistics

Basic statistics are widely required for validation and reporting. The following are commonly used:

- **arithmetic mean**
- **median - the median value of a ranked series X_i**
- **mode - the value of X which occurs with greatest frequency or the midlle value of the class with greatest frequency**
- **standard deviation - the root mean squared deviation S_x :**

$$S_x = \sqrt{\frac{\sum_{i=1}^{N} (X_i - \bar{X})^2}{N-1}}$$

- **skewness and kurtosis**

In addition empirical frequency distributions can be presented as a graphical representation of the number of data per class and as a cumulative frequency distribution. From these selected values of exceedence probability or non-exceedence probability can be extracted, e.g. the daily rainfall which has been excceded 1%, 5% or 10% of the time.

<u>Example 3.1</u>

Basic statistics for monthly rainfall data of MEGHARAJ station (KHEDA catchment) is derived for the period 1961 to 1997. Analysis is carried out taking the actual values and all the months in the year. The results of analysis is given in Table 3.1 below. The frequency distribution and the cumulative frequency is worked out for 20 classes between 0 and 700 mm rainfall and is given in tabular results and as graph in Figure 3.1. Various decile values are also listed in the result of the analysis.

Since actual monthly rainfall values are considered it is obvious to expect a large magnitude of skewness which is 2.34 and also the sample is far from being normal and that is reflected in kurtosis (=7.92). The value of mean is larger then the median value and the frequency distribution shows a positive skew. From the table of decile values it can may be seen that 70 % of the months receive less than 21 mm of rainfall. From the cumulative frequency table it may be seen that 65 percent of the months receive zero rainfall (which is obvious to expect in this catchment) and that there are very few instances when the monthly rainfall total is above 500 mm.

Table 3.1: Computational results of the basic statistics for monthly rainfall at MEGHARAJ

```
First year  =  1962     Last year   =  1997

Actual values are used

Basic Statistics:
=================
     Mean              = .581438E+02
     Median            = .000000E+00
     Mode              = .175000E+02
     Standard deviation= .118932E+03
     Skewness          = .234385E+01
     Kurtosis          = .792612E+01
     Range             = .000000E+00 to .613500E+03
     Number of elements=         420
```

Decile	Value
1	.000000E+00
2	.000000E+00
3	.000000E+00
4	.000000E+00
5	.000000E+00
6	.000000E+00
7	.205100E+02
8	.932474E+02
9	.239789E+03

Cumulative frequency distribution and histogram

Upper class limit	Probability	Number of elements
.000000E+00	.658183	277.
.350000E+02	.729543	30.
.700000E+02	.769981	17.
.105000E+03	.815176	19.
.140000E+03	.836584	9.
.175000E+03	.867507	13.
.210000E+03	.881779	6.
.245000E+03	.903187	9.
.280000E+03	.917460	6.
.315000E+03	.934110	7.
.350000E+03	.950761	7.
.385000E+03	.955519	2.
.420000E+03	.967412	5.
.455000E+03	.981684	6.
.490000E+03	.986441	2.
.525000E+03	.988820	1.
.560000E+03	.995956	3.
.595000E+03	.995956	0.
.630000E+03	.998335	1.
.665000E+03	.998335	0.
.700000E+03	.998335	0.
		0.

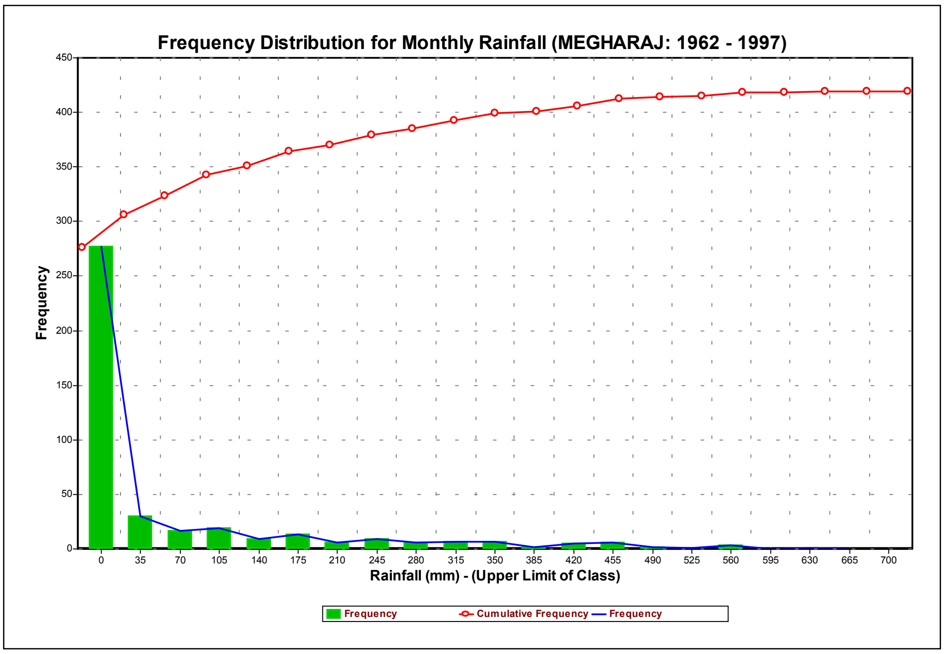

Figure 3.1: Frequency and cumulative frequency plot of monthly rainfall at MEGHARAJ station

4 Annual exceedance rainfall series

The following are widely used for reporting or for subsequent use in frequency analysis of extremes:

- maximum of a series. The maximum rainfall value of an annual series or of a month or season may be selected using HYMOS. All values (peaks) over a specified threshold may also be selected. Most commonly for rainfall daily maxima per year are used but hourly maxima or N-hourly maxima may also be selected.
- minimum of a series. As the minimum daily value with respect to rainfall is frequently zero this is useful for aggregated data only.

5 Fitting of frequency distributions

A common use of rainfall data is in the assessment of probabilities or return periods of given rainfall at a given location. Such data can then be used in assessing flood discharges of given return period through modelling or some empirical system and can thus be applied in schemes of flood alleviation or forecasting and for the design of bridges and culverts.

Frequency analysis usually involves the fitting of a theoretical frequency distribution using a selected fitting method, although empirical graphical methods can also be applied. The fitting of a particular distribution implies that the rainfall sample of annual maxima were drawn from a population of that distribution. For the purposes of application in design it is assumed that future probabilities of exceedence will be the same as past probabilities. However there is nothing inherent in the series to indicate whether one distribution is more likely to be appropriate than another and a wide variety of distributions and fitting procedures has been recommended for application in different countries and by different agencies. **Different distributions can give widely different estimates, especially when extrapolated or when an outlier (an exceptional value, well in excess of the second**

largest value) occurs in the data set. Although the methods are themselves objective, a degree of subjectivity is introduced in the selection of which distribution to apply.

These words of caution are intended to discourage the routine application and reporting of results of the following methods without giving due consideration to the regional climate. Graphical as well as numerical output should always be inspected. Higher the degree of aggregation of data, more normal the data will become.

The following frequency distributions are available in HYMOS:

- Normal and log-normal distributions
- Pearson Type III or Gamma distribution
- Log-Pearson Type III
- Extreme Value type I (Gumbel), II, or III
- Goodrich/Weibull distribution
- Exponential distribution
- Pareto distribution

The following fitting methods are available for fitting the distribution:

- modified maximum likelihood
- method of moments

For each distribution one can obtain the following:

- estimation of parameters of the distribution
- a table of rainfalls of specified exceedance probabilities or return periods with confidence limits
- results of goodness of fit tests
- a graphical plot of the data fitted to the distribution

Example 5.1

Normal frequency distribution for rainfall data of MEGHARAJ station (KHEDA catchment) is fitted for three cases: (a) annual maximum values of daily data series, (b) annual maximum values of monthly data series and (c) actual annual rainfall values.

Figure 5.1 shows the graphical fitting of normal distribution for annual maximum values of daily series. The scatter points are the reduced variate of observed values and a best fit line showing the relationship between annual maximum daily with the frequency of occurrence on the basis of normal distribution. The upper and lower confidence limit (95 %) are also shown, the band width of which for different return periods indicate the level of confidence in estimation.

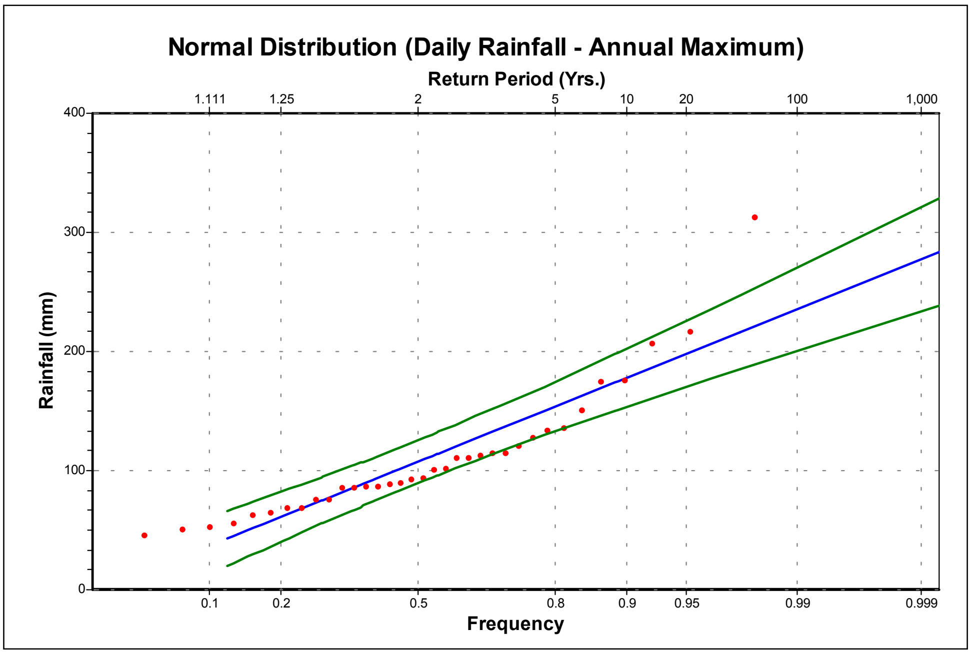

Figure 5.1: Normal distribution for annual maximums of daily rainfall at MEGHARAJ station

Figure 5.2 and Figure 5.3 shows the normal distribution fit for annual maximum of monthly data series and the actual annual values respectively. The level of normality can be seen to have increased in the case of monthly and yearly data series as compared to the case of daily data.

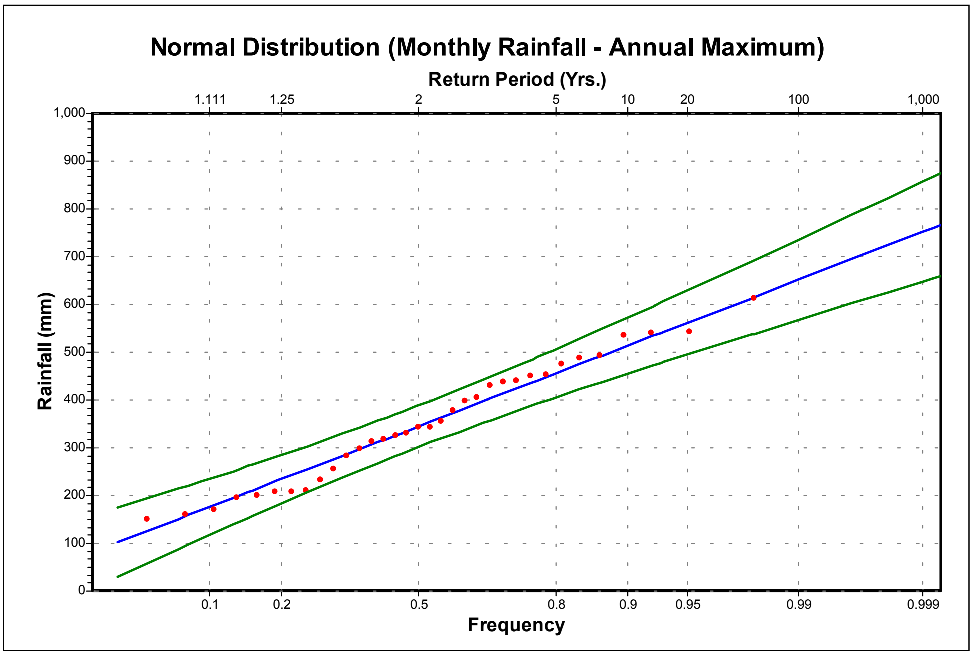

Figure 5.2: Normal distribution for annual maximums of monthly rainfall at MEGHARAJ station

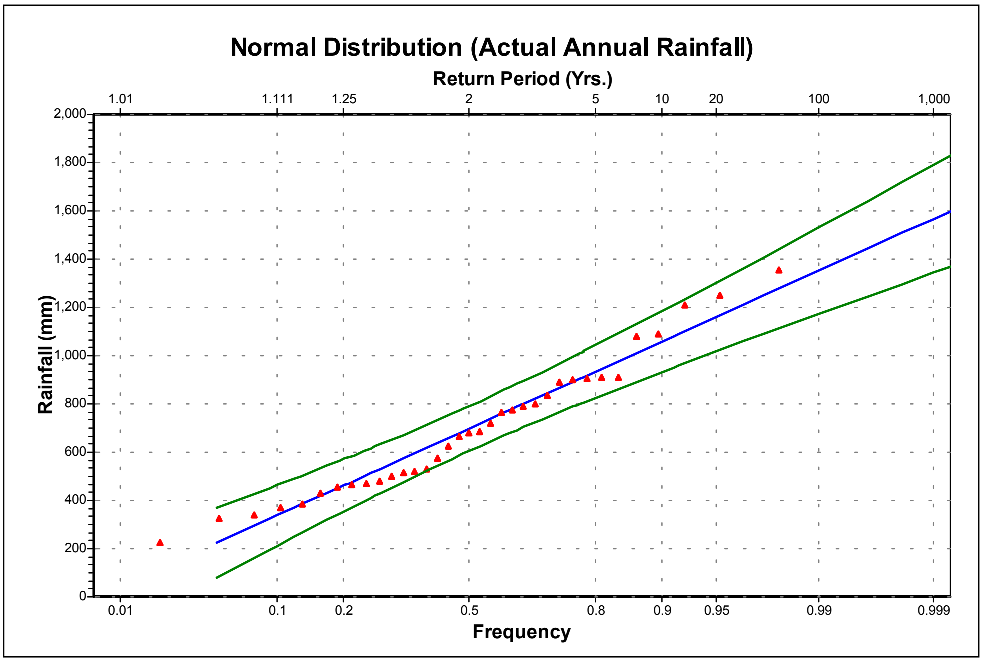

Figure 5.3: Plot of normal distribution for actual annual rainfall at MEGHARAJ station

The analysis results for the case of normal distribution fitting for actual annual rainfall is given in Table 5.1. There are 35 effective data values for the period 1962 to 1997 considered for the analysis. The mean annual rainfall is about 700 mm with an standard deviation of 280 mm and skewness and kurtosis of 0.54 and 2.724 respectively. The observed and theoretical frequency for each of the observed annual rainfall value is listed in the increasing order of the magnitude. The results of a few tests on good of fit is also given in table. In the last, the rainfall values for various return periods from 2 to 500 years is given alongwith the upper and lower confidence limits.

Table 5.1: Analysis results for the normal distribution fitting for annual rainfall at MEGHARAJ

```
First year    =  1962
Last year     =  1997
```

<u>Basic statistics:</u>

```
Number of data      =    35
Mean                =    697.725
Standard deviation  =    281.356
Skewness            =       .540
Kurtosis            =      2.724
```

Nr./ year	observation	obs.freq.	theor.freq.p	return-per.	st.dev.xp	st.dev.p
13	225.700	.0198	.0467	1.05	73.7978	.0256
11	324.300	.0480	.0922	1.10	65.2262	.0383
18	338.000	.0763	.1005	1.11	64.1160	.0401
25	369.500	.1045	.1217	1.14	61.6525	.0443
5	383.300	.1328	.1319	1.15	60.6155	.0460

Nr./ year	observation	obs.freq.	theor.freq.p	return-per.	st.dev.xp	st.dev.p
3	430.500	.1610	.1711	1.21	57.2863	.0517
17	456.000	.1893	.1951	1.24	55.6434	.0546
24	464.500	.2175	.2036	1.26	55.1224	.0554
23	472.500	.2458	.2117	1.27	54.6448	.0563
20	481.300	.2740	.2209	1.28	54.1341	.0571
8	500.000	.3023	.2411	1.32	53.1019	.0588
4	512.900	.3305	.2556	1.34	52.4337	.0599
2	521.380	.3588	.2654	1.36	52.0147	.0606
29	531.500	.3870	.2773	1.38	51.5363	.0614
7	573.800	.4153	.3298	1.49	49.8067	.0641
10	623.500	.4435	.3960	1.66	48.3757	.0663
30	665.500	.4718	.4544	1.83	47.7129	.0672
31	681.000	.5000	.4763	1.91	47.5996	.0674
27	686.000	.5282	.4834	1.94	47.5784	.0674
1	719.100	.5565	.5303	2.13	47.6261	.0673
34	763.500	.5847	.5924	2.45	48.2011	.0665
33	773.000	.6130	.6055	2.53	48.3988	.0662
22	788.000	.6412	.6258	2.67	48.7632	.0657
19	799.000	.6695	.6406	2.78	49.0705	.0652
15	833.800	.6977	.6857	3.18	50.2575	.0634
21	892.000	.7260	.7551	4.08	52.9196	.0591
6	900.200	.7542	.7641	4.24	53.3571	.0584
26	904.000	.7825	.7683	4.32	53.5647	.0581
28	911.500	.8107	.7763	4.47	53.9833	.0574
16	912.000	.8390	.7768	4.48	54.0117	.0573
12	1081.300	.8672	.9136	11.58	66.0629	.0370
32	1089.500	.8955	.9181	12.21	66.7472	.0359
14	1210.300	.9237	.9658	29.20	77.5678	.0209
9	1248.000	.9520	.9748	39.61	81.1752	.0170
35	1354.000	.9802	.9902	101.67	91.7435	.0086

Results of Binomial goodness of fit test

```
variate dn = max(|Fobs-Fest|)/sd=     1.4494 at Fest= .2773
prob. of exceedance P(DN>dn)    =      .1472
number of observations          =    35
```

Results of Kolmogorov-Smirnov test

```
variate dn = max(|Fobs-Fest|)   =      .1227
prob. of exceedance P(DN>dn)    =      .6681
```

Results of Chi-Square test

```
variate = chi-square           =     5.2000
prob. of exceedance of variate =      .2674
number of classes              =     7
number of observations         =    35
degrees of freedom             =     4
```

Values for distinct return periods

Return per.	prob(xi<x) p	value x	st. dev. x	confidence lower	intervals upper
2	.50000	697.725	47.558	604.493	790.957
5	.80000	934.474	55.340	825.987	1042.961
10	.90000	1058.347	64.184	932.521	1184.173
25	.96000	1190.401	75.692	1042.014	1338.788
50	.98000	1275.684	83.867	1111.271	1440.096
100	.99000	1352.380	91.565	1172.876	1531.885
250	.99600	1444.011	101.085	1245.846	1642.177
500	.99800	1507.611	107.852	1296.180	1719.043

6 Frequency and duration curves

A convenient way to show the variation of hydrological quantities through the year, by means of frequency curves, where each frequency curve indicates the magnitude of the quantity for a specific probability of non-exceedance. The duration curves are ranked representation of these frequency curves. The average duration curve gives the average number of occasions a given value was not exceeded in the years considered. The computation of frequency and duration curves is as given below:

6.1 Frequency Curves

Considering "n" elements of rainfall values in each year (or month or day) and that the analysis is carried out for "m" years (or months or days) a matrix of data X_{ij} {for i=1,m and j=1,n} is obtained. For each j = j_0 the data $X_{I,jo}$, {for i=1,m} is arranged in ascending order of magnitude. The probability that the i^{th} element of this ranked sequence of elements is not exceeded is:

$$F_i = \frac{i}{m+1}$$

The frequency curve connects all values of the quantity for j=1,n with the common property of equal probability of non-exceedance. Generally, a group of curves is considered which represents specific points of the cumulative frequency distribution for each j. Considering that curves are derived for various frequencies F_k {k=1,n_f}, then values for rainfall $R_{k,j}$ is obtained by linear interpolation between the probability values immediately greater (F_I) and lesser (F_{i-1}) to n_k for each j as:

$$R_{k,j0} = R_{1-1} + (R_i - R_{i-1}) \frac{F_k - F_{i-1}}{F_i - F_{i-1}}$$

6.2 Duration Curves

When the data $R_{k,j}$, k=1,n_f and j=1,n is ranked for each k, the ranked matrix represents the duration curves for given probabilities of non-exceedance.

When all the data is considered without discriminating for different elements j {j=1,n} and are ranked in the ascending order of magnitude, then the resulting sequence shows the average duration curve. This indicates how often a given level of quantity considered will not be exceeded in a year (or month or day).

Example 6.1

A long-term monthly rainfall data series of MEGHARAJ station (KHEDA catchment) is considered for deriving frequency curves and duration curves. Analysis is done on the yearly basis and the various frequency levels set are 10, 25, 50, 75 and 90 %.

Figure 6.1 shows the frequency curves for various values (10, 25, 50, 75 and 90%) for each month in the year. Monthly rainfall distribution in the year 1982 is also shown superimposed on this plot for comparison. Minimum and maximum values for each month of the year in the plot gives the range of variation of rainfall in each month. Results of this frequency curve analysis is tabulated in Table 6.1.

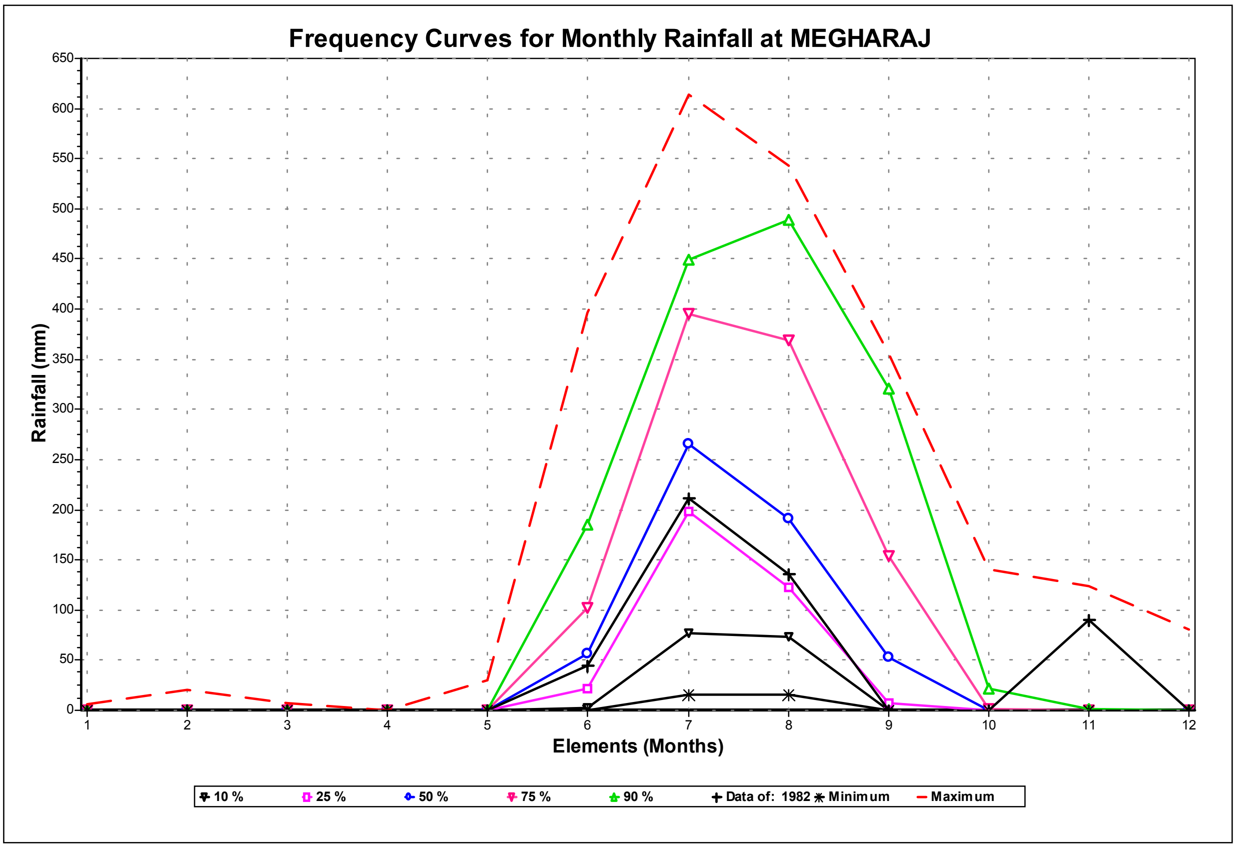

Figure 6.1: Monthly frequency curves for rainfall at MEGHARAJ station

Figure 6.2 shows the plot of duration curves for the same frequencies. The plot gives values of monthly rainfall which will not be exceeded for certain number of months in a year with the specific level of probability. The results of analysis for these duration curve is given in Table 6.2.

The average duration curve, showing value of rainfall which will not be exceeded "on an average" in a year for a certain number of months is given as Figure 6.3.

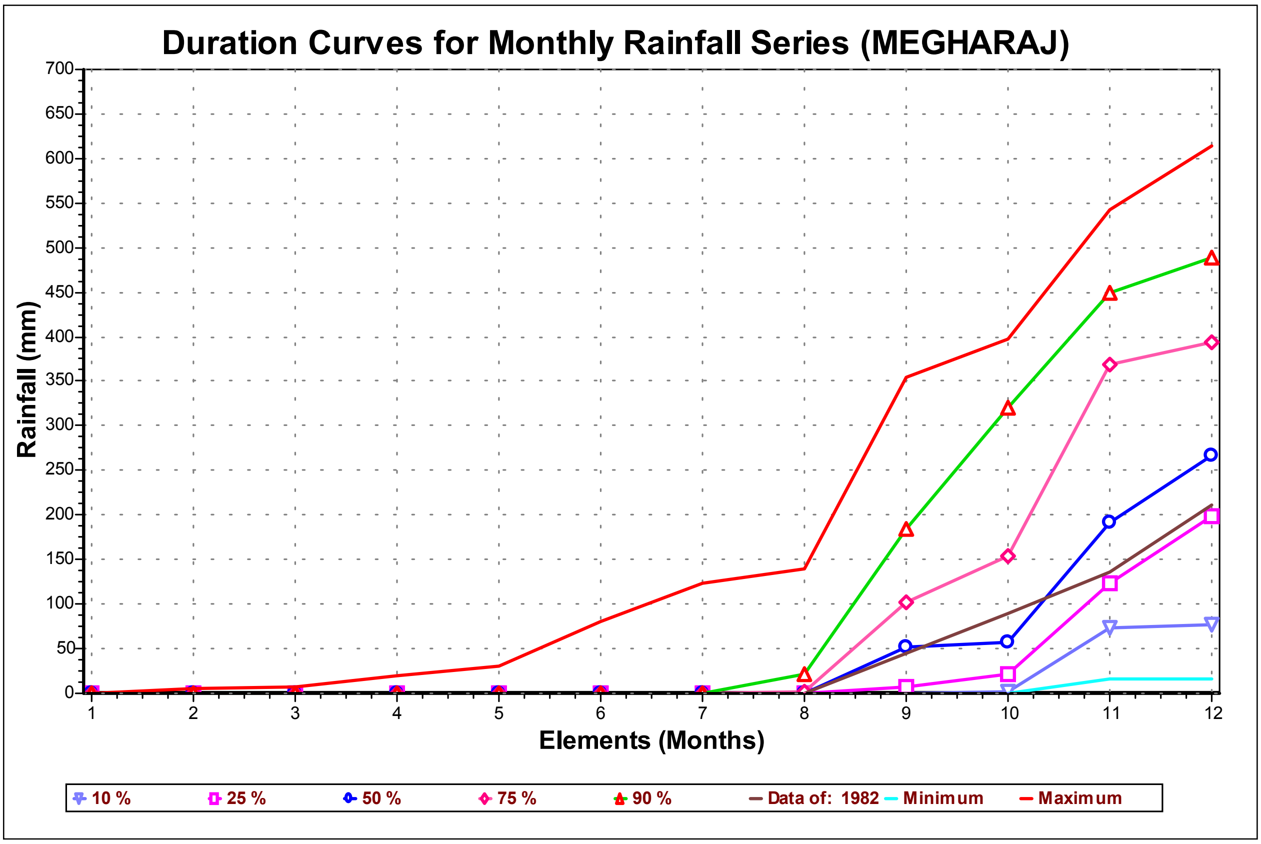

Figure 6.2: Monthly duration curves for rainfall at MEGHARAJ station

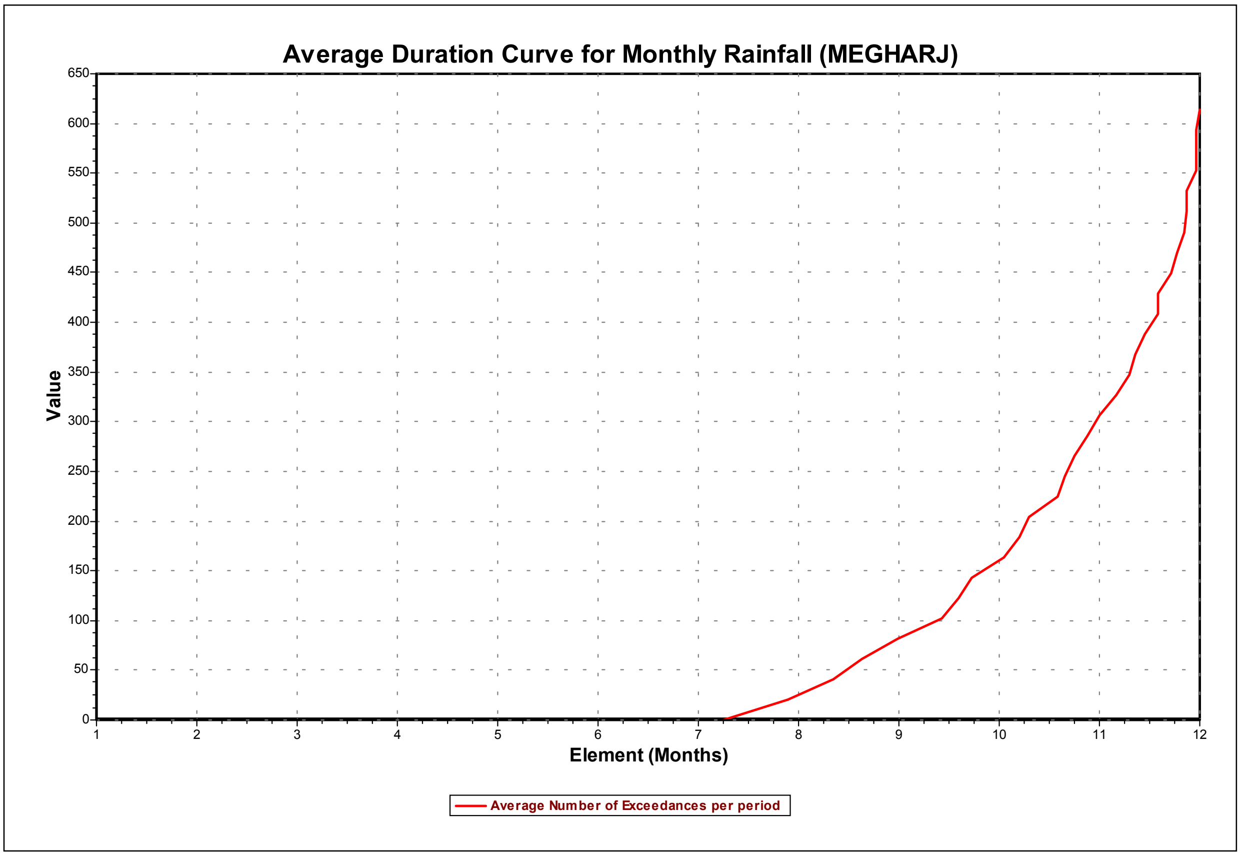

Figure 6.3: Average monthly duration curves for rainfall at MEGHARAJ station

Table 6.1: Results of analysis for frequency curves for monthly data for MEGHARAJ station (rainfall values in mm)

Element	No. of Data	Frequency					Year 1982	Min.	Max.
		0.1	0.25	0.5	0.75	0.9			
1	29	0	0	0	0	0	0	0	6
2	29	0	0	0	0	0	0	0	20
3	29	0	0	0	0	0	0	0	7
4	29	0	0	0	0	0	0	0	0
5	29	0	0	0	0	0	0	0	30
6	33	2	22	57	101.54	184.8	45	0	397
7	36	77.1	198	266	394.75	448.93	211	15.5	613.5
8	36	73.64	122.77	190.75	368.87	488.92	135.3	15.3	543
9	36	0	6.63	52.5	154	320.9	0	0	355.09
10	30	0	0	0	1.37	21.05	0	0	140
11	29	0	0	0	0	0.8	90	0	124
12	29	0	0	0	0	0	0	0	80

Table 6.2: Results of analysis for duration curves for monthly data for MEGHARAJ station (rainfall values in mm)

No. of Elements	Frequency					Year 1982	Min.	Max.
	0.1	0.25	0.5	0.75	0.9			
1	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	6
3	0	0	0	0	0	0	0	7
4	0	0	0	0	0	0	0	20
5	0	0	0	0	0	0	0	30
6	0	0	0	0	0	0	0	80
7	0	0	0	0	0.8	0	0	124
8	0	0	0	1.37	21.05	0	0	140
9	0	6.63	52.5	101.54	184.8	45	0	355.09
10	2	22	57	154	320.9	90	0	397
11	73.64	122.77	190.75	368.87	448.93	135.3	15.3	543
12	77.1	198	266	394.75	488.92	211	15.5	613.5

Table 6.3: Results of analysis for average duration curves for monthly data for MEGHARAJ station (rainfall values in mm)

1	Rainfall Value	0	20.45	40.9	61.35	81.8	102.25	
	No. of Exceedances	7.25	7.89	8.34	8.63	8.98	9.43	
2	Rainfall Value	122.7	143.15	163.6	184.05	204.5	224.95	
	No. of Exceedances	9.59	9.72	10.04	10.2	10.3	10.59	
3	Rainfall Value	245.4	265.85	286.3	306.75	327.2	347.65	
	No. of Exceedances	10.65	10.75	10.88	11.01	11.17	11.29	
4	Rainfall Value	368.1	388.55	409	429.45	449.9	470.35	
	No. of Exceedances	11.36	11.45	11.58	11.58	11.71	11.78	
5	Rainfall Value	490.8	511.25	531.7	552.15	572.6	593.05	613.5
	No. of Exceedances	11.84	11.87	11.87	11.97	11.97	11.97	12

7 Intensity-Frequency-Duration Analysis

If rainfall data from a recording raingauge is available for long periods such as 25 years or more, the frequency of occurrence of a given intensity can also be determined. Then we obtain the intensity-frequency-duration relationships. Such relationships may be established for different parts of the year, e.g. a month, a season or the full year. The procedure to obtain such relationships for the year is described in this section. The method for parts of the year is similar.

The entire rainfall record in a year is analysed to find the maximum intensities for various durations. Thus each storm gives one value of maximum intensity for a given duration. The largest of all such values is taken to be the maximum intensity in that year for that duration. Likewise the annual maximum intensity is obtained for different duration. Similar analysis yields the annual maximum intensities for various durations in different years. It will then be observed that the annual maximum intensity for any given duration is not the same every year but it varies from year to year. In other words it behaves as a random variable. So, if 25 years of record is available then there will be 25 values of the maximum intensity of any given duration, which constitute a sample of the random variable. These 25 values of any one duration can be subjected to a frequency analysis. Often the observed frequency distribution is well fitted by a Gumbel distribution. A fit to a theoretical distribution function like the Gumbel distribution is required if maximum intensities at return periods larger than can be obtained from the observed distribution are at stake. Similar frequency analysis is carried out for other durations. Then from the results of this analysis graphs of maximum rainfall intensity against the return period for various durations such as those shown in figure 7.1 can be developed.

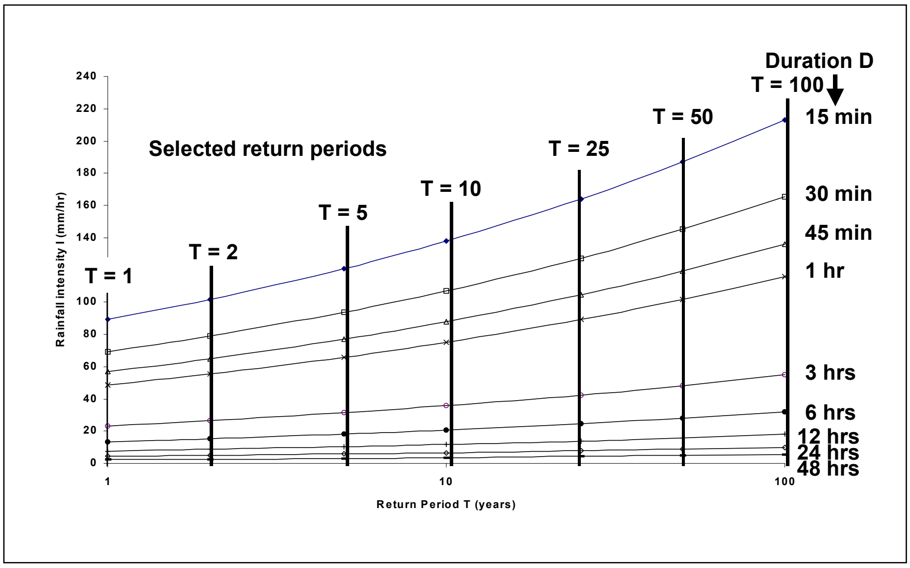

Figure 7.1: Intensity-frequency-duration curves

By reading for each duration at distinct return periods the intensities intensity-duration curves can be made. For this the rainfall intensities for various durations at concurrent return periods are connected as shown in Figure 7.2

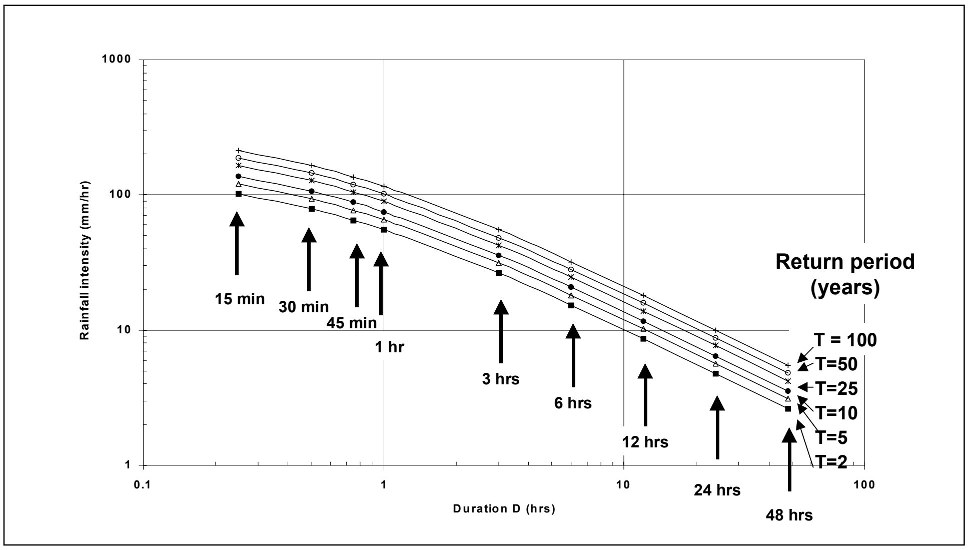

Figure 7.2: Intensity-frequency-duration curves for various return periods

From the curves of figure 7.2 the maximum intensity of rainfall for any duration and for any return period can be read out.

Alternatively, for any given return period an equation of the form.

$$I = \frac{c}{(D+a)^b} \tag{7.1}$$

can be fitted between the maximum intensity and duration

where I = intensity of rainfall (mm/hr)
D = duration (hrs)
c, a, b are coefficients to be determined through regression analysis.

One can write for return periods T_1, T_2, etc.:

$$I = \frac{c_1}{(D+a_1)^{b_1}}\ ; I = \frac{c_2}{(D+a_2)^{b_2}}; \text{etc} \tag{7.2}$$

where c_1, a_1 and b_1 refer to return period $T_1$and c_2, a_2 and b_2 are applicable for return period T_2, etc. Generally, it will be observed that the coefficients a and b are approximately the same for all the return periods and only c is different for different return periods. In such a case one general equation may be developed for all the return periods as given by:

$$I = \frac{KT^d}{(D+a)^b} \tag{7.3}$$

where T is the return period in years and K and d are the regression coefficients for a given location. If a and b are not same for all the return periods, then an individual equation for each return period may be used. In Figure 7.1 and 7.2 the results are given for Bhopal, as adapted from Subramanya, 1994. For Bhopal with I in mm/hr and D in hours the following parameter values in equation 7.3 hold:

K = 69.3; a = 0.50; b = 0.878 and d = 0.189

When the intensity-frequency-duration analysis is carried out for a number of locations in a region, the relationships may be given in the form of equation 7.3 with a different set of regression coefficients for each location. Alternatively, they may be presented in the form of maps (with each map depicting maximum rainfall depths for different combinations of one return period and one duration) which can be more conveniently used especially when one is dealing with large areas. Such maps are called isopluvial maps. A map showing maximum rainfall depths for the duration of one hour which can be expected with a frequency of once in 50 years over South India is given in Figure 7.3

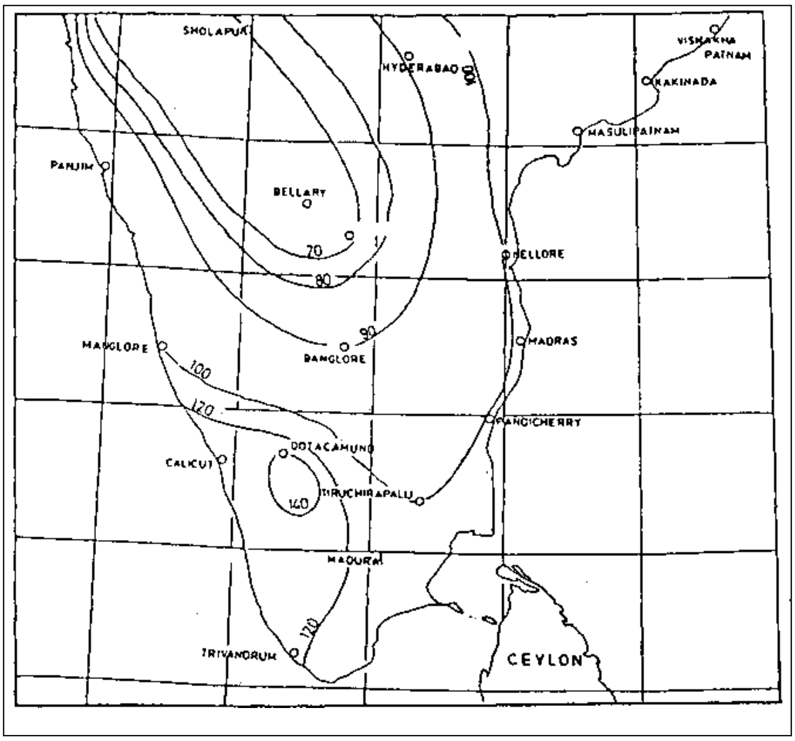

Figure 7.3: Isopluvial map of 50 years 1 hour rainfall over South India

Annual maximum and annual exceedance series

In the procedure presented above annual **maximum** series of rainfall intensities were considered. For frequency analysis a distinction is to be made between annual **maximum** and annual **exceedance** series. The latter is derived from a **partial duration series**, which is defined as a series of data above a threshold. The maximum values between each **upcrossing** and the next **downcrossing** (see Figure 7.4) are considered in a partial duration series. The threshold should be taken high enough to make successive maximums serially independent or a time horizon is to be considered around the local maximum to eliminate lower maximums exceeding the threshold but which are within the time horizon. If the threshold is taken such that the number of values in the partial duration series becomes equal to the number of years selected then the partial duration series is called **annual exceedance series**.

Since annual maximum series consider only the maximum value each year, it may happen that the annual maximum in a year is less than the second or even third largest independent maximum in another year. Hence, the values at the lower end of the annual exceedance series will be higher than those of the annual maximum series. Consequently, the return period derived for a particular I(D) based on annual maximum series will be larger than one would have obtained from annual exceedances. The following relation exists between the return period based on annual maximum and annual exceedance series (Chow, 1964):

$$T_E = \frac{1}{\ln\left(\frac{T}{T-1}\right)} \tag{7.4}$$

where: T_E = return period for annual exceedance series
T = return period for annual maximum series

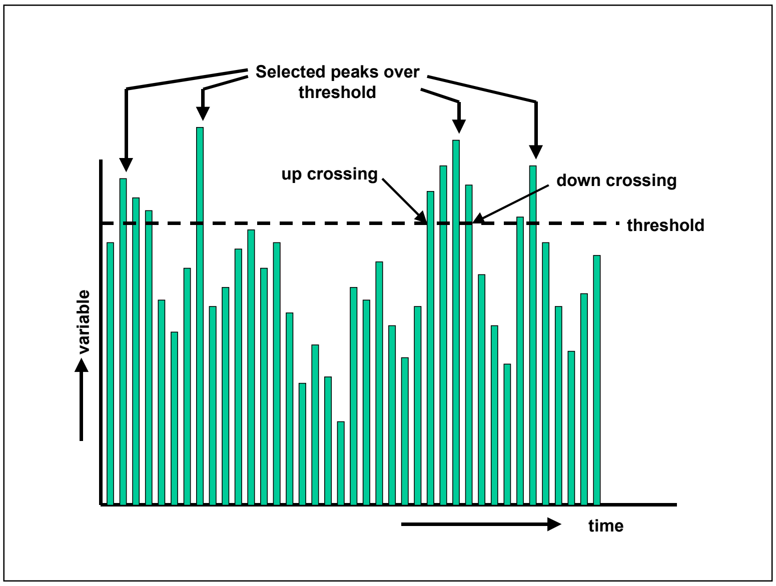

Figure 7.4: Definition of partial duration series

The ratio T_E/T is shown in Figure 7.5. It is observed that the ratio approaches 1 for large T. Generally, when T < 20 years T has to be adjusted to T_E for design purposes. Particularly for urban drainage design, where low return periods are used, this correction is of importance.

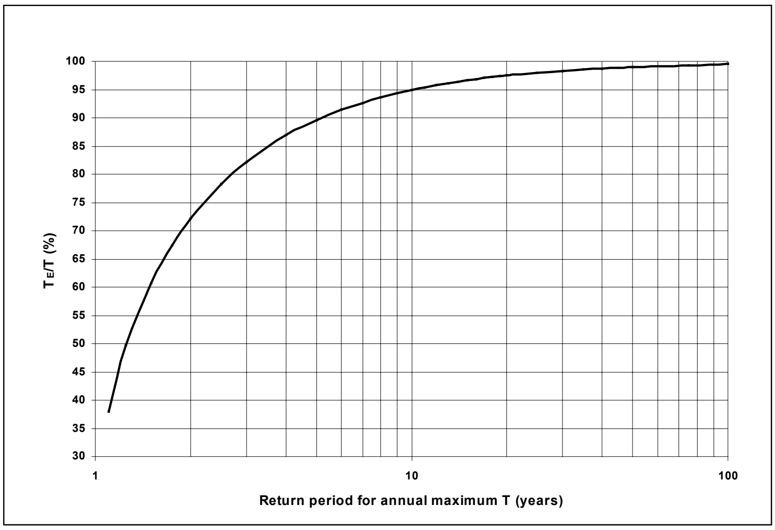

Figure 7.5: Relation between return periods annual maximum (T) and annual exceedance series (T_E)

In HYMOS annual maximum series are used in the development of intensity-duration-frequency curves, which are fitted by a Gumbel distribution. Equation (7.4) is used to transform T into T_E for T < 20 years. Results can either be presented for distinct values of T or of T_E.

Example 7.1

Analysis of hourly rainfall data of station Chaskman, period 1977-2000, monsoon season 1/6-30/9. First, from the hourly series the maximum seasonal rainfall intensities for each year are computed for rainfall durations of 1, 2, …, 48 hrs. In this way annual maximum rainfall intensity series are obtained for different rainfall durations. Next, each such series is subjected to frequency analysis using the Gumbel or EV1 distribution, as shown for single series in Figure 7.6. The IDF option in HYMOS automatically carries out this frequency analysis for all rainfall durations. The results are presented in Table 7.1. The fit to the distribution for different rainfall durations is shown in Figure 7.7. It is observed that in general the Gumbel distribution provides an acceptable fit to the observed frequency distribution.

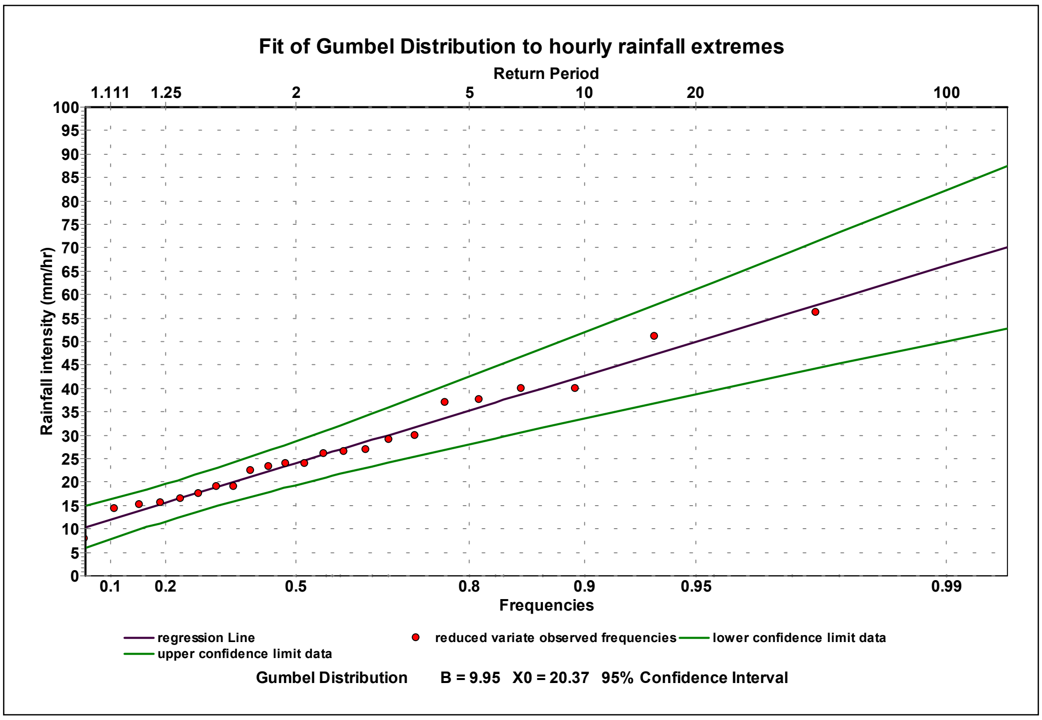

Figure 7.6: Fitting of Gumbel distribution to observed frequency distribution of hourly annual maximum series for monsoon season

Table 7.1 Example of output file of IDF option

```
Intensity - Duration - Frequency Relations

Input Timestep: 1 hour
Duration computation in: hour
Year: 1998 contains missing values, analysis may not be correct
Year: 1999 contains missing values, analysis may not be correct
Year: 2000 contains missing values, analysis may not be correct

Period from '01-06' to '01-10'
Start year: 1977
End year  : 2000


Maximum Intensities per year for selected durations
```

Year	Durations (hour) 1	2	3	4	6	9	12	18	24	48
1977	14.40	9.70	7.60	6.55	4.37	2.91	2.18	1.68	1.56	1.03
1978	19.00	14.45	10.03	7.65	5.12	3.41	2.56	1.71	1.66	1.27
1979	23.30	19.40	14.50	10.95	7.30	5.00	3.78	2.81	2.37	1.86
.										
.										
1997	37.00	29.25	27.50	25.50	23.08	18.06	15.71	11.97	9.25	5.41
1998	22.40	15.00	10.00	8.88	6.17	4.46	3.50	2.56	2.38	1.47
1999	30.00	22.60	16.17	12.38	8.30	5.53	4.15	2.77	2.07	1.08
2000	24.00	12.30	8.53	6.70	4.62	4.38	3.28	2.19	2.29	1.48

Parameters of Gumbel distribution

Duration	X0	BETA	Sd1	Sd2
1	20.372	9.955	2.140	1.584
2	13.623	7.250	1.558	1.154
3	10.322	5.745	1.235	0.914
4	8.316	4.703	1.011	0.748
6	6.308	3.567	0.767	0.568
9	4.634	2.673	0.574	0.425
12	3.605	2.166	0.465	0.345
18	2.541	1.520	0.327	0.242
24	2.192	1.211	0.260	0.193
48	1.329	0.701	0.151	0.112

IDF-data: Annual Maximum

Duration	Return Periods 1	2	4	10	25	50	100
1	11.666	24.020	32.774	42.773	52.212	59.214	66.164
2	7.282	16.280	22.656	29.938	36.813	41.912	46.975
3	5.297	12.428	17.480	23.251	28.698	32.740	36.751
4	4.203	10.040	14.175	18.899	23.358	26.666	29.949
6	3.188	7.615	10.752	14.334	17.716	20.225	22.716
9	2.296	5.614	7.964	10.648	13.183	15.063	16.929
12	1.711	4.399	6.303	8.479	10.532	12.056	13.568
18	1.212	3.098	4.435	5.961	7.403	8.472	9.533
24	1.133	2.636	3.700	4.916	6.064	6.916	7.761
48	0.716	1.586	2.202	2.906	3.571	4.064	4.553

Note that in the output table first a warning is given about series being incomplete for some years. This may affect the annual maximum series. Comparison with nearby stations will then be required to see whether extremes may have been missed. If so, the years with significant missing data are eliminated from the analysis.

Next, the table presents an overview of the annual maximum series, followed by a summary of the Gumbel distribution parameters x_0 and β, with their standard deviations (sd1, sd2) and for various rainfall durations the rainfall intensities for selected return periods. The latter values should be compared with the maximum values in the annual maximum series.

Note that Figure 7.7 gives a row-wise presentation of the last table, whereas Figure 7.8 gives a column-wise presentation of the same table. This figure is often presented on log-log scale, see Figure 7.9.

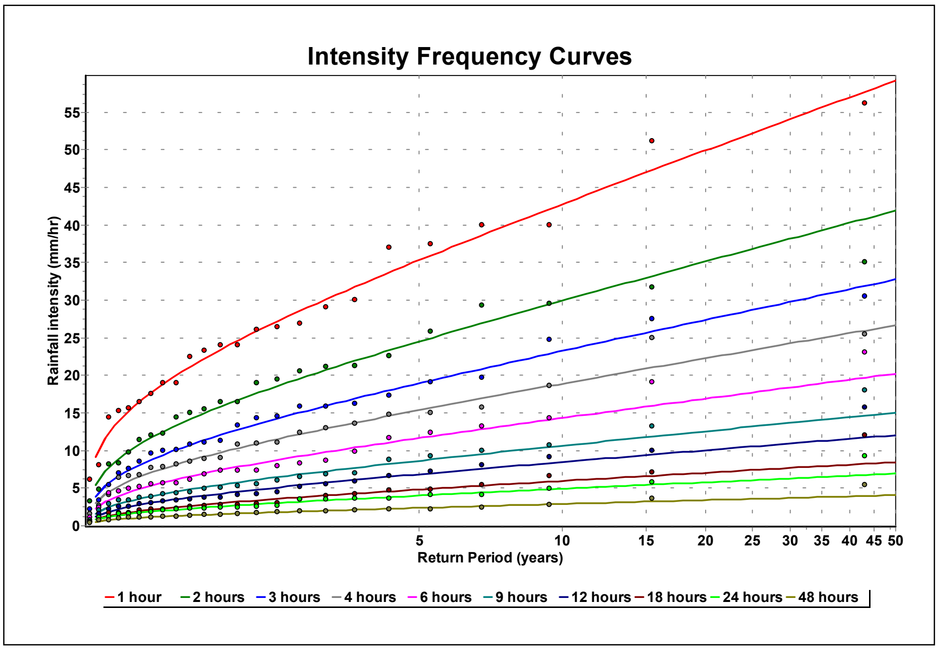

Figure 7.7: Intensity Frequency curves for different rainfall durations, with fit to Gumbel distribution

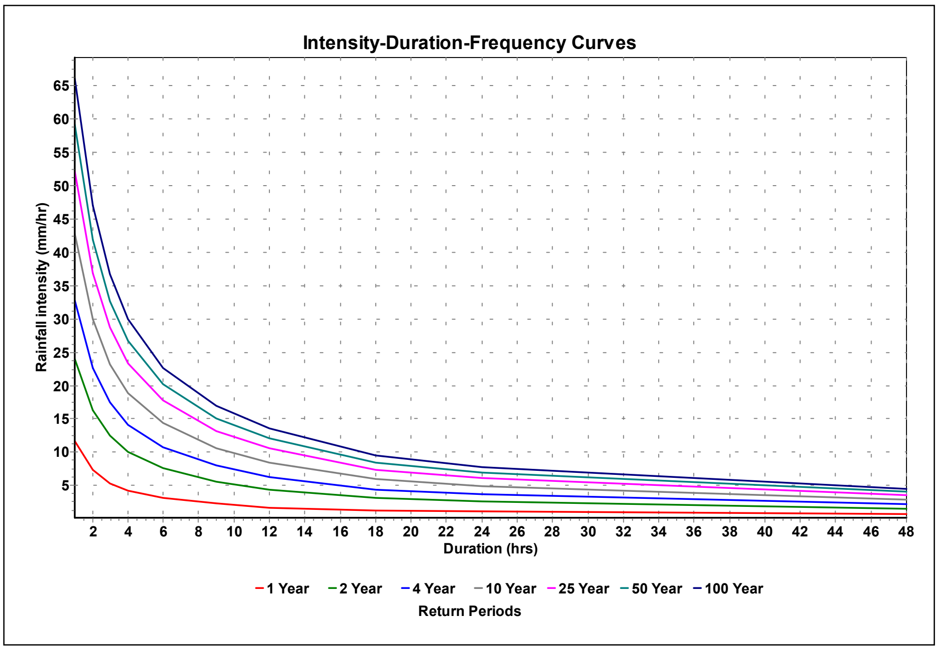

Figure 7.8: Intensity-Density-Frequency curves for Chaskman on linear scale (Annual maximum data)

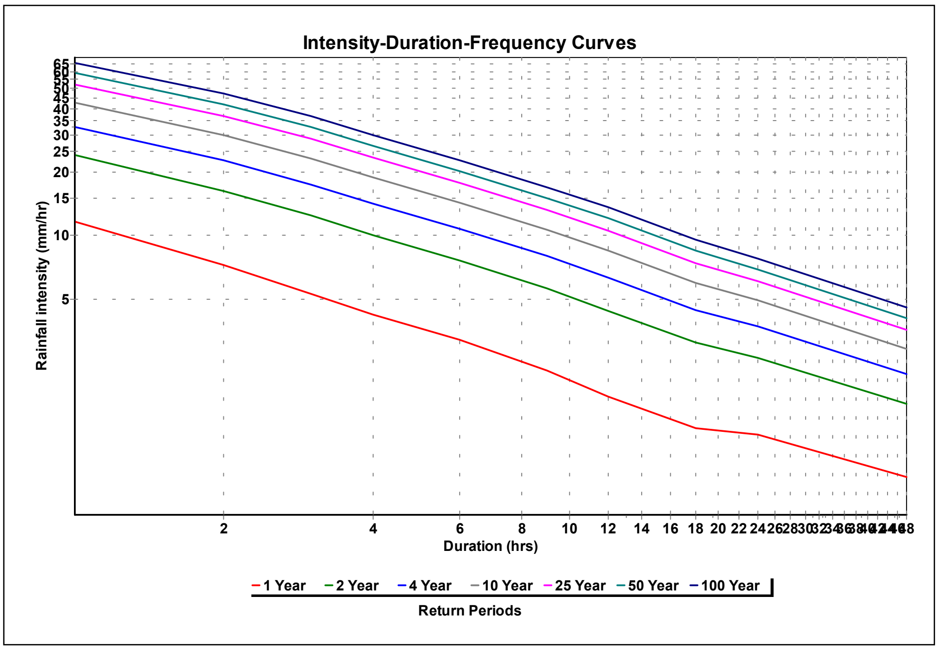

Figure 7.9: Intensity-Density-Frequency curves for Chaskman on double-log scale (Annual maximum data)

The IDF-option in HYMOS also includes a procedure to convert the annual maximum statistics into annual exceedance results by adapting the return period according to equation (7.4). Hence, rather than using the annual exceedance series in a frequency analysis the annual maximum series' result is adapted. This procedure is useful for design of structures where design conditions are based on events with a moderate return period (5 to 20 years). An example output is shown in Table 7.2 and Figures 7.10 and 7.11. Compare results with Table 7.1 and Figures 7.8 and 7.9.

Table 7.2 Example of output of IDF curves for annual exceedances.

IDF - data : Annual Exceedences

Te	Tm
1	1.581977
2	2.541494
4	4.520812
10	10.50833
25	25.50333
50	50.50167
100	100.5008

Duration	**Return Periods**						
	1	**2**	**4**	**10**	**25**	**50**	**100**
1	20.372	27.272	34.172	43.293	52.414	59.314	66.214
2	13.623	18.648	23.674	30.317	36.960	41.986	47.011
3	10.322	14.304	18.286	23.551	28.815	32.797	36.780
4	8.316	11.576	14.836	19.145	23.454	26.713	29.973
6	6.308	8.780	11.252	14.521	17.789	20.261	22.733
9	4.634	6.487	8.339	10.788	13.237	15.089	16.942
12	3.605	5.106	6.607	8.592	10.576	12.078	13.579
18	2.541	3.595	4.648	6.041	7.433	8.487	9.540
24	2.192	3.031	3.870	4.980	6.089	6.928	7.767
48	1.329	1.815	2.301	2.943	3.585	4.071	4.557

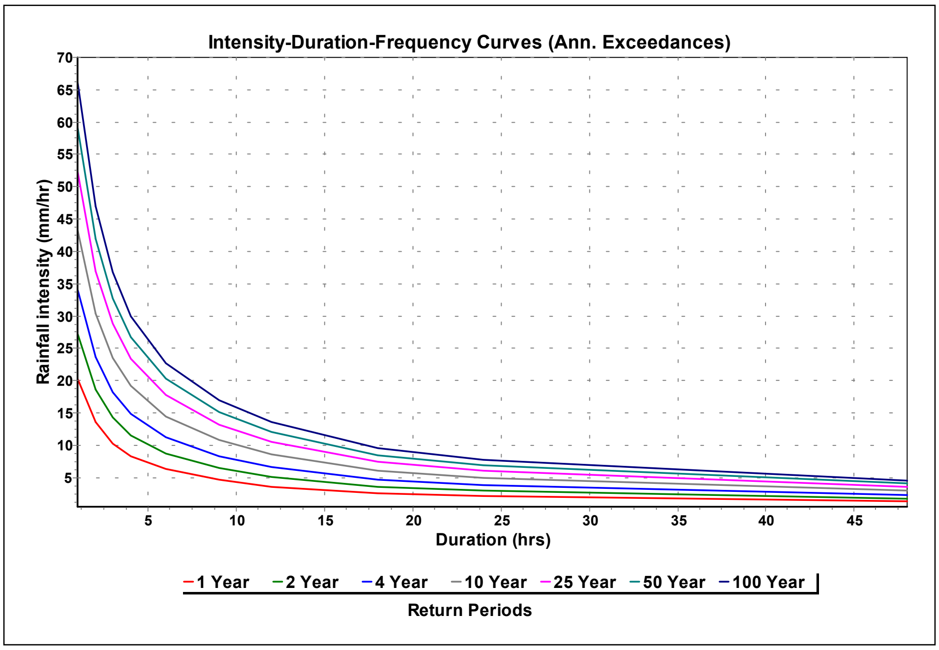

Figure 7.10: Intensity-Density-Frequency curves for Chaskman on linear scale (Annual exceedances)

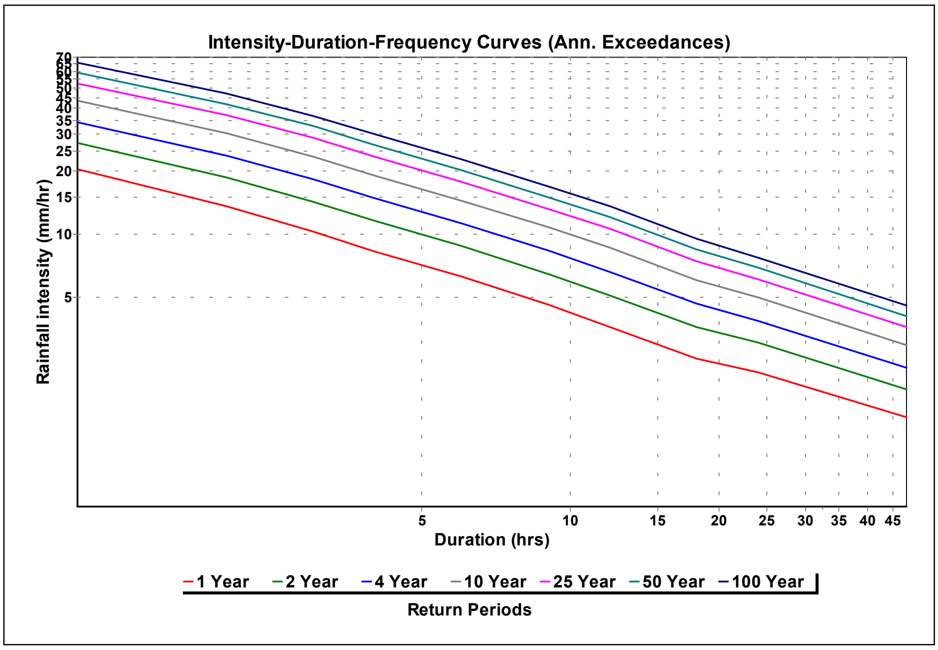

Figure 7.11: Intensity-Density-Frequency curves for Chaskman on double-log scale (Annual exceedances)

Finally, the Rainfall Intensity-Duration curves for various return periods have been fitted by a function of the type (7.1). It appeared that the optimal values for "a" and "b" varied little for different return periods. Hence a function of the type (7.3) was tried. Given a value for "a" the coefficients K, d and b can be estimated by multiple regression on the logarithmic transformation of equation (7.3):

$$\log I = \log K + d \log T - b \log(D + a) \tag{7.5}$$

By repeating the regression analysis for different values of “a” the coefficient of determination was maximised. The following equation gave a best fit (to the logaritms):

$$I = \frac{32.8T^{0.27}}{(D + 0.65)^{0.81}} \qquad R^2 = 0.993$$

Though the coefficient of determination is high, a check afterwards is always to be performed before using such a relationship!! A comparison is shown in Figure 7.12. A reasonable fit is observed.

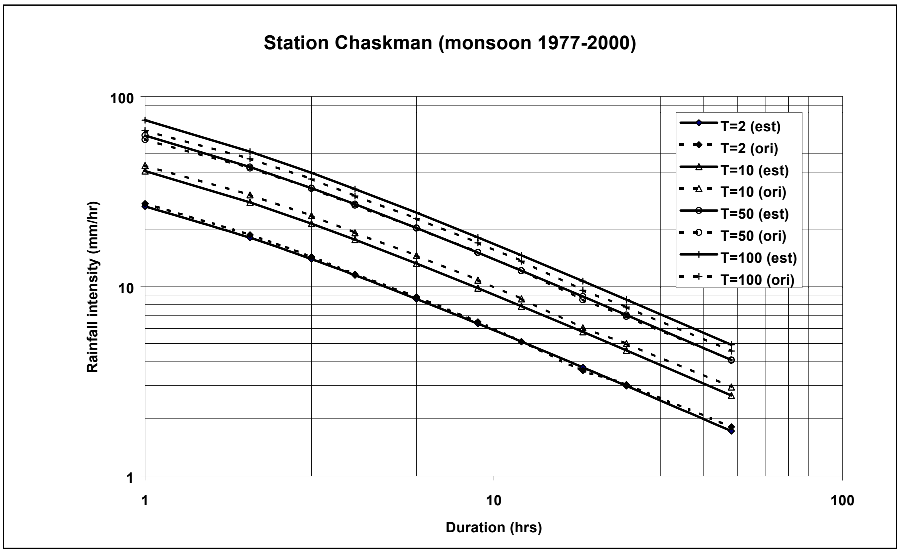

Figure 7.12: Test of goodness of fit of IDF-formula to IDF-curves from Figure 7.10 and 7.11

8 Depth-Area-Duration Analysis

In most of the design applications the maximum depth of rainfall that is likely to occur over a given area for a given duration is required. Wherever possible, the frequency of that rainfall should also be known. For example, the knowledge of maximum depth of rainfall occurring on areas of various sizes for storms of different duration is of interest in many hydrological design problems such as the design of bridges and culverts, design of irrigation structures etc.

A storm of given duration over a certain area rarely produces uniform rainfall depth over the entire area. The storm usually has a centre, where the rainfall P_o is maximum which is always larger than the average depth of rainfall P for the area as a whole. Generally, the difference between these two values, that is $(P_o – P)$, increases with increase in area and decreases with increase in the duration. Also the difference is more for convective and orographic precipitation than for cyclonic. To develop quantitative relationship between P_o and P, a number of storms with data obtained from recording raingauges have to be analysed. The analysis of a typical storm is described below (taken from Reddy, 1996).

The rainfall data is plotted on the basin map and the isohyets are drawn. These isohyets divide the area into various zones. On the same map the Thiessen polygons are also constructed for all the raingauge stations. The polygon of a raingauge station may lie in different zones. Thus each zone will be influenced by a certain number of gauges, whose polygonal areas lie either fully or partially in that zone. The gauges, which influence each zone along with their influencing areas, are noted. Next for each zone the cumulative average depth of rainfall (areal average) is computed at various time using the data of rainfall mass curve at the gauges influencing the zone and the Thiessen weighted mean method. In other words in this step the cumulative depths of rainfall at different times recorded at different parts are converted into cumulative depths of rainfall for the zonal area at the corresponding times. Then the mass curves of average depth of rainfall for accumulated areas are computed starting from the zone nearest to the storm centre and by adding one more adjacent to it each time, using the results obtained in the previous step and using the Thiessen weight in proportion to the areas of the zones. These mass curves are now examined to find the maximum average depth of rainfall for different duration and for progressively increasing accumulated areas. The results are then plotted on semi-logarithmic paper. That is, for each duration the maximum average depth of rainfall on an ordinary scale is plotted against the area on logarithmic scale. If a storm contains more than one storm centre, the above analysis is carried out for each storm centre. An enveloping curve is drawn for each duration. Alternatively, for each duration a depth area relation of the form as proposed by Horton may be established:

$$P = P_o \, e^{-kA^n} \qquad (8.1)$$

where:

P_o = highest amount of rainfall at the centre of the storm (A = 25 km^2) for any given duration
P = maximum average depth of rainfall over an area A (> 25 km^2) for the same duration
A = area considered for P
k, n = regression coefficients, which vary with storm duration and region.

Example 8.1

The following numerical example illustrates the method described above. In and around a catchment with an area of 2790 km^2 some 7 raingauges are located, see Figure 8.1. The record of a severe storm measured in the catchment as observed at the 7 raingauge stations is presented in Table 8.1 below.

Time in hours	Cumulative rainfall in mm measured at raingauge stations						
	A	B	C	D	E	F	G
4	0	0	0	0	0	0	0
6	12	0	0	0	0	0	0
8	18	15	0	0	0	6	0
10	27	24	0	0	9	15	6
12	36	36	18	6	24	24	9
14	42	45	36	18	36	33	15
16	51	51	51	36	45	36	18
18	51	63	66	51	60	39	18
20	51	72	87	66	66	42	18
22	51	72	96	81	66	42	18
24	51	72	96	81	66	42	18

Table 8.1 Cumulative rainfall record measured for a severe storm at 7 raingauges (A to G)

The total rainfall of 51, 72, 96, 81, 66, 42 and 18 mm are indicated at the respective raingauge stations A, B, C, D, E, F and G on the map. The isohyets for the values 30, 45, 60 and 75 mm are constructed. Those isohyets divide the basin area into five zones with areas as given in Table 8.2. The Thiessen polygons are then constructed for the given raingauge network [A to G] on the same map. The areas enclosed by each polygon and the zonal boundaries for each raingauge is also shown in Table 8.2.

Zone	Area	Raingauge station area of influence in each zone (km^2)						
	km^2	A	B	C	D	E	F	G
I	415	0	105	57	253	0	0	0
II	640	37	283	0	20	300	0	0
III	1015	640	20	0	0	185	170	0
IV	525	202	0	0	0	0	275	48
V	195	0	0	0	0	0	37	158

Table 8.2 Zonal areas and influencing area by rain gauges

As can be seen from figure 8.1 Zone I (affected by the rainfall stations with the highest point rainfall amounts) is the nearest to storm centre while Zone V is the farthest.

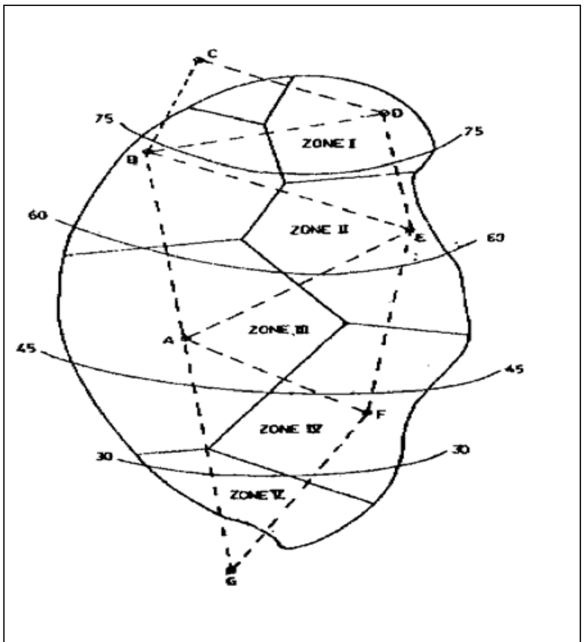

Figure 8.1: Depth-area-duration analysis

The cumulative average depth of rainfall for each zone is then computed using the data at Raingauge stations A, B, C, D, E, F and G and the corresponding Thiessen weights. For example, the average depth of rainfall in Zone I at any time, P_I is computed from the following equation.

$$P_I = \frac{105\,x\,P_B\;x\,57\,x\,P_C + 253\,x\,P_D}{(105 + 57 + 253)}$$

where P_B, P_C and P_D are the cumulative rainfalls at stations B, C and D at any given time. That is

$$P_I = 0.253\ P_B + 0.137\ P_C + 0.610\ P_D$$

Similarly for Zone II, we have:

$$P_{II} = \frac{37\,x\,P_A\;x\,283\,x\,P_B + 20\,x\,P_D + 300\,x\,P_E}{(37 + 283 + 20 + 300)}$$

or:

$P_{II} = 0.058\ P_A + 0.442\ P_B + 0.031\ P_D + 0.469\ P_E$ and so on.

These results are shown in Table 8.3 and Figure 8.2.

Time (hours)	Zone I	Zone II	Zone III	Zone IV	Zone V
4	0	0	0	0	0
6	0	0.70	7.60	4.62	0
8	3.80	7.67	12.66	10.07	1.14
10	6.07	24.07	21.66	18.80	7.71
12	15.23	29.44	31.81	27.26	11.85
14	27.30	39.77	39.47	34.83	18.42
16	41.85	46.31	46.86	40.14	21.42
18	56.09	60.53	50.87	41.71	21.99
20	70.40	64.78	52.65	43.28	22.56
22	80.78	68.25	52.65	43.28	22.56
24	80.78	68.25	52.65	43.28	22.56

Table 8.3: Cumulative average depths of rainfall in various zones in mm.

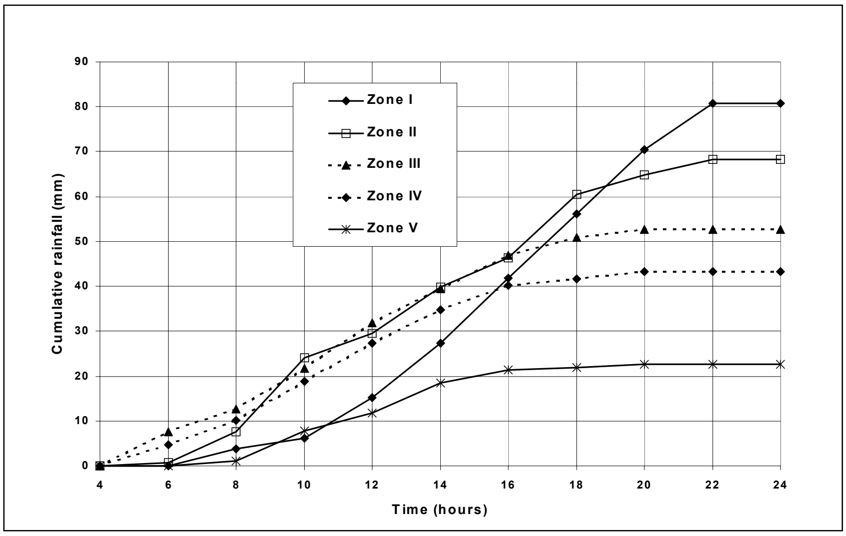

Figure 8.2: Cumulative average depths of rainfall in Zones I to V

In the next step the cumulative average rainfalls for the progressively accumulated areas are worked out. Here the weights are used in proportion to the areas of the zones. For example, the cumulative average rainfall over the first three zones is given as

$$P_{I+II+III} = \frac{415\,x\,P_I + 640\,x\,P_{II} + 1015\,x\,P_{III}}{415 + 640 + 1015}$$

$= 0.2\ P_I + 0.31\ P_{II} + 0.49\ P_{III}$

The result of this step are given in Table 8.4 and Figure 8.3.

Time hours	I 415 km^2	I + II 1055 km^2	I + II + III 2070 km^2	I + II + III + IV 2595 km^2	I + II + III+ IV + V 2790 km^2
4	0	0	0	0	0
6	0	0.43	3.94	4.08	3.79
8	3.80	6.15	9.34	9.49	8.91
10	6.07	17.00	19.28	19.18	18.38
12	15.23	23.86	27.76	27.66	26.55
14	27.30	34.87	37.12	36.66	35.38
16	41.85	44.56	45.69	44.57	42.95
18	56.09	58.79	54.91	52.24	50.12
20	70.40	66.99	59.96	56.59	54.21
22	80.78	73.17	63.12	59.11	56.55
24	80.78	73.17	63.12	59.11	56.55

Table 8.4: Cumulative average rainfalls for accumulated areas in mm

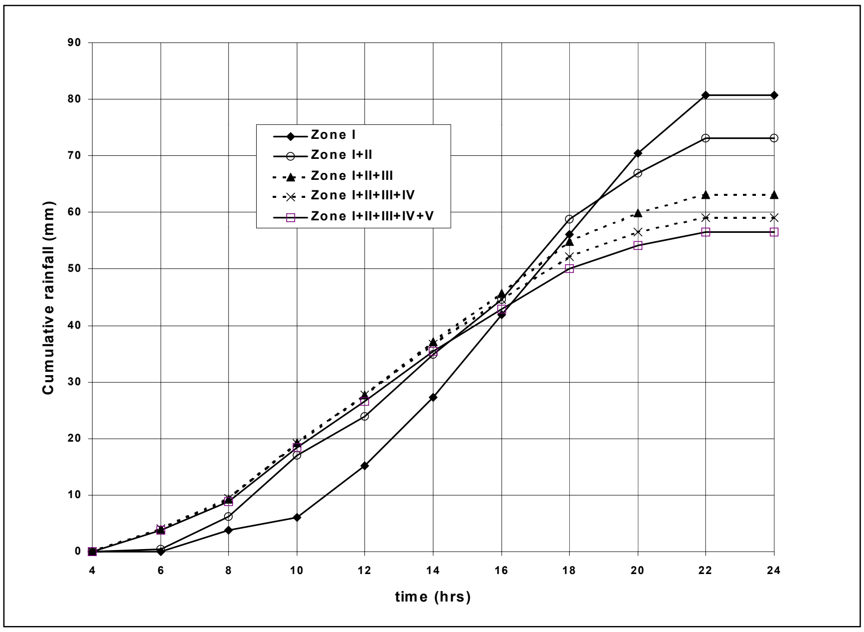

Figure 8.3: Cumulative average depths of rainfall in cumulated areas Zones I to I+II+III+IV+V

Now for any zone the maximum average depth of rainfall for various durations of 4, 8, 12, 16 and 20 h can be obtained from Table 8.4 by sliding a window of width equal to the required duration over the table columns with steps of 2 hours. The maximum value contained in the window of a particular width is presented in Table 8.5

Duration in hours	Maximum average depths of rainfall in mm				
	415 km^2	1055 km^2	2070 km^2	2595 km^2	2790 km^2
4	28.79	23.92	18.42	18.17	17.64
8	55.17	43.13	36.35	35.08	34.04
12	74.71	60.84	50.97	48.16	46.33
16	80.78	72.74	59.18	56.59	54.21
20	80.78	73.17	63.12	59.11	56.55

Table 8.5: Maximum average depths of rainfall for accumulated areas

For each duration, the maximum depths of rainfall is plotted against the area on logarithmic scale as shown in Figure 8.4

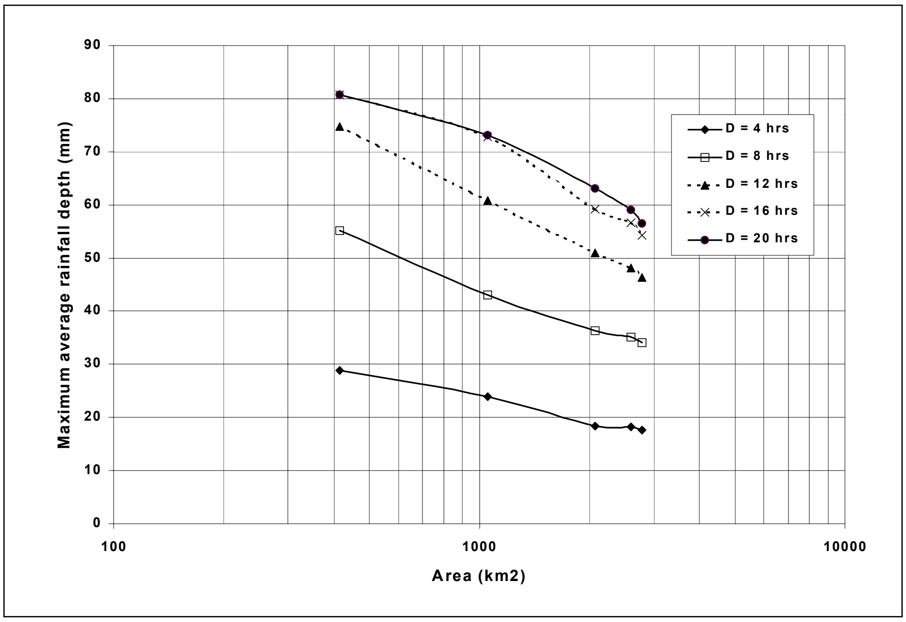

Figure 8.4: Depth-area-duration curves for a particular storm

By repeating this procedure for other severe storms and retrieving from graphs like Figure 8.4 for distinct areas the maximum rainfall depths per duration, a series of storm rainfall depths per duration and per area is obtained. The maximum value for each series is retained to constitute curves similar to Figure 8.4. Consequently, the maximum rainfall depth for a particular duration as a function of area may now be made of contributions of different storms to produce the overall maximum observed rainfall depth for a particular duration as a function of area to constitute the depth-area-duration (DAD) curve. For the catchment considered in the example these DAD curves will partly or entirely exceed the curves in Figure 8.4 unless the presented storm was depth-area wise the most extreme one ever recorded.

Areal reduction factor

If the maximum average rainfall depth as a function of area is divided by the maximum point rainfall depth the ratio is called the Areal Reduction Factor (ARF), which is used to convert point rainfall extremes into areal estimates. ARF-functions are developed for various storm durations. In practice, ARF functions are established based on average DAD's developed for some selected severe representative storms.

These ARF's which will vary from region to region, are also dependent on the season if storms of a particular predominate in a season. Though generally ignored, it would be of interest to investigate whether these ARF's are also dependent on the return period as well. To investigate this a frequency analysis would be required to be applied to annual maximum depth-durations for different values of area and subsequently comparing the curves valid for a particular duration with different return periods.

In a series of Flood Estimation Reports prepared by CWC and IMD areal reduction curves for rainfall durations of 1 to 24 hrs have been established for various zones in India (see e.g. CWC, Hydrology Division, 1994). An example is presented in Figure 8.5 (zone 1(g)).

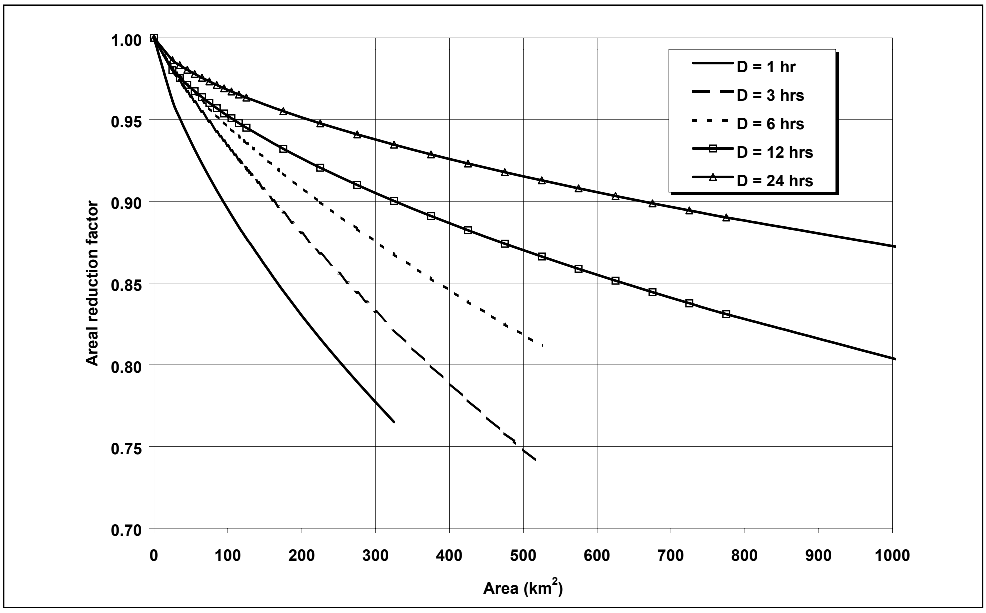

Figure 8.5: Example of areal reduction factors for different rainfall durations

Time distribution of storms
For design purposes once the point rainfall extreme has been converted to an areal extreme with a certain return period, the next step is to prepare the time distribution of the storm. The time distribution is required to provide input to hydrologic/hydraulic modelling. The required distribution can be derived from cumulative storm distributions of selected representative storms by properly normalising the horizontal and vertical scales to percentage duration and percentage cumulative rainfall compared to the total storm duration and rainfall amount respectively. An example for two storm durations is given in Figure 8.6, valid for the Lower Godavari sub-zone – 3 (f).

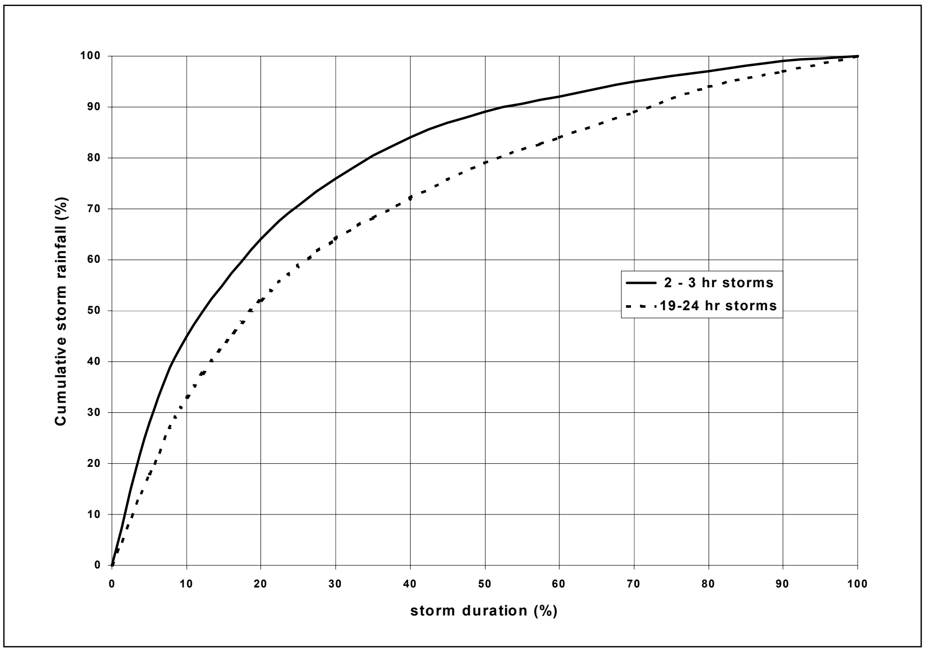

Figure 8.6: Time distributions of storms in Lower Godavari area for 2-3 and 19-24 hrs storm durations

From Figure 8.6 it is observed that the highest intensities are occurring in the first part of the storm (about 50% within 15% of the total storm duration). Though this type of storm may be characteristic for the coastal zone further inland different patterns may be determining. A problem with high intensities in the beginning of the design storm is that it may not lead to most critical situations, as the highest rainfall abstractions in a basin will be at the beginning of the storm. Therefore one should carefully select representative storms for a civil engineering design and keep in mind the objective of the design study. There may not be one design storm distribution but rather a variety, each suited for a particular use.