INSTITUTE of
HYDROLOGY

GOVERNMENT OF THE REPUBLIC OF INDONESIA

MINISTRY OF PUBLIC WORKS

DIRECTORATE GENERAL OF WATER RESOURCES DEVELOPMENT

FLOOD DESIGN MANUAL

FOR

JAVA AND SUMATRA

SEPTEMBER 1983

INSTITUTE OF HYDROLOGY

WALLINGFORD

OXON

U.K.

DPMA

BANDUNG

INDONESIA

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

During the course of this project, Dr C S Green, Mr D B Boorman and Mr F A K Farquharson of the Institute of Hydrology received help and encouragement from a large number of organisations and individuals. Without their cooperation this project could never have succeeded and we express our gratitude to all of the following.

Ir Mardjono Notodihardjo	Assistant Director General for River Development (DGWRD)
Ir Rachmat Tirtotjondro	Director, DPMA
Drs Tata Sukarta	Head, Foreign Aid Administration (DGWRD)
Ir Aziz Bocking	Assistant, Foreign Aid Administration (DGWRD)
Ir Muhadi Dipl H	Chief, Hydrology Division, DPMA
Mr Soenarso Soehardi	Chief, Foreign Aid Administration DPMA
Ir Soeharto	Chief, Hydrometry Section, DPMA
Ir Joesron Loebis	Chief, Hydrology Section, DPMA
Ir Ferry Putuhena MSc	Head, Water Resources Data Unit (P3SA)
Surin Sangsnit	Team Leader, UNDP/WMO River Forecasting Project, DPMA
Mr R Kingston	Resident Engineer, Sir Alexander Gibb and Partners, DPMA Bandung
Mr A Bonnier	SCET International, Team Leader Projek Hidrologi, DPMA
Mr A Muzet	SCET International, Engineer Projek Hidrologi, DPMA
Mr J L Irish	Regional Adviser in Hydrology UNESCO
Mr F J Kaul	Binnie and Partners, London
Mr B N Gunter	SMEC, Team Leader, West Java Design Unit Assistance Project
Mr B C Ferguson	International Engineering Co. Inc., Sederhana Irrigation and Land Development Project II
Mr E W J Crawley	First Secretary (Aid) British Embassy Jakarta

Mr I Thow	First Secretary (Aid) British Embassy Jakarta
Mr D Cottrell	Second Secretary (Aid) British Embassy Jakarta
Mr J Wilkins	Secretary (Aid) British Embassy Jakarta
Darsono Umaryadi B.E.	Staff DPU Prop. DTI Jawa Barat
Endang Sukardja	Assistant, Java Floods Study Project, DPMA
Dr A N Mandeville	Chief, ODA Section, Institute of Hydrology, Wallingford.
Miss S Smethurst	British Council, Jakarta
Dr D A Jones	Hydrologist, Institute of Hydrology, Wallingford
Mr S B Jones	Hydrologist, Institute of Hydrology, Wallingford.
Miss B J Ridler	Assistant, Institute of Hydrology, Wallingford
Mrs V Bronsdon	Assistant, Institute of Hydrology, Wallingford.

NOMENCLATURE

	river cross section area (m^2)
a	rating equation parameter
AAR	annual average rainfall (mm)
APBAR	mean annual maximum catchment 1 day rainfall (mm)
AREA	catchment area (km^2)
ARF	areal reduction factor
b	rating equation parameter
c	Chezy roughness coefficient
	rating equation parameter (m)
DGWRD	Directorate General of Water Resources Development
DPMA	Direktorat Penyelidikan Masalah Air (Directorate of Hydraulic Engineering)
DPU	Departemen Pekerjaan Umum (Department of Public Works)
EV1	extreme value distribution - Type 1
F_i	non-exceedence probability (plotting position)
FOREST =	forest index
	Darcy - Weisbach friction factor
GEOL	geology index
GF	growth factor
g	acceleration due to gravity ($m\ s^{-2}$)
h	river stage (level) (m)
h_i	series of stage peaks (m)
h_o	stage threshold (m)
JOG	Joint Operations Graphics
LAKE	lake index
M	number of floods over threshold
MAF	mean annual flood ($m^3\ s^{-1}$)

MSL		main stream length (km)
N		number of years
n	=	Manning's roughness coefficient
P		wetted perimeter (m)
PADDY		paddy index
PBAR		mean annual maximum point 1 day rainfall (mm)
PLTN		plantation index
PLN		Perusahaan Umum Listrik Negara (Electricity Authority)
POT		Peaks over a threshold
P3SA		Projek Perencanaan Pembangunan Sumber-Sumber Air (River Basin Development Planning Unit, Directorate of Planning, DGWRD)
Q		river discharge (m^3 s^{-1})
Q_{max}		maximum flood peak on record (m^3 s^{-1})
Q_{med}		median annual flood (m^3 s^{-1})
Q_T		T year return period flood (m^3 s^{-1})
q_i		series of flood peaks (m^3 s^{-1})
q_o		discharge threshold (m^3 s^{-1})
-		mean of annual maximum flood series (m^3 s^{-1})
		water surface slope (m m^{-1})
SHAPE		catchment shape index
SIMS		river slope index (m km^{-1})
SOIL		soil index
SWAMP		swamp index
S010		flood plain index (m km^{-1})
S085		river slope index (m km^{-1})
S1085		river slope index (m km^{-1})

sd	standard deviation
v	variable exponent of AREA
	velocity of river flow (m s^{-1})
,	Gumbel reduced variate
β	mean exceedence (m^3 s^{-1})
	average number of exceedences per year

PREFACE

This manual contains a rational set of flood estimation techniques applicable throughout Java and Sumatra. The techniques include a 'no data' method for use at ungauged sites and various ways of using data from the site of interest or nearby flow gauging stations. If used cautiously, the techniques can also be used on other Indonesian islands, certainly where good quality flood flow data are available. The report is the product of a two year study of Indonesian hydrology which the authors believe to be the most comprehensive assembly and analysis of Indonesian flood data undertaken to date. The 'no data' method has been developed from the project's large data base following a review of existing methods and drawing heavily on experience gained from numerous similar studies performed elsewhere in the world. A comparison of this method with previously used methods shows it to be a superior technique over a wide variety of catchment types.

Foldout maps showing the location of study catchments can be found for Java, as Figure 1.1 and for Sumatra as Figure 1.2.

The study has been carried out in two phases; Phase I in 1981 examined flooding in Java and Phase II in the following year extended the study to include Sumatra. The report produced during Phase I is superseded by this report. The study was a cooperative venture between the Institute of Hydrology, Wallingford, United Kingdom (UK) where many similar studies have been carried out for different regions of the world and the Direktorat Penyelidikan Masalah Air, Bandung, Indonesia, custodians of Indonesian flow data and the primary hydrological and hydraulic research institute of Indonesia.

The following staff worked on the project in both Indonesia and the United Kingdom.

Dr	C S	Green	IH	Phase I and II
Mr	F A K	Farquharson	IH	Phase I and II
Mr	D B	Boorman	IH	Phase II
Ir	Sunadji	Josoadiwijono	DPMA	Phase I

Drs	Trenggono	DPMA	Phase I
Ir	Moh. Arief Ilyas	DPMA	Phase II
Drs	Petrus Syariman	DPMA	Phase II

Both organisations provided support staff for the duration of the project in their respective institutes.

The project was jointly funded by the Indonesian Government and the British Overseas Development Administration.

INTRODUCTION TO FLOOD ESTIMATION

1.1 Flood statistics, return period and probability

For any flood estimation problem it is necessary to specify the return period, or probability, of the desired flood. This will vary according to the nature of the project and the consequences of the design flood being exceeded. In practice it is often useful to construct a curve relating the size of flood to its probability of occurrence. Such a curve, called a flood frequency curve, enables flood magnitudes corresponding to various design criteria to be estimated and the implementation costs and implications of failure of such criteria to be appraised. Figure 1.3 shows such a curve. The probability scale gives an exceedence probability (ie the probability of a flood level being exceeded in any one year); the scale beneath this shows return period, or average interval in years between floods exceeding this level. Return period, T, is the reciprocal of the exceedence probability and can give a more tangible appreciation of the severity of the flood.

If a very long record exists for a point on a river it is possible to construct a flood frequency curve from an examination of the record. Figure 1.4(a) illustrates one approach to this; the record is divided into hydrological years (to ensure independence of flood peaks) and the biggest flood in each year is noted. By ranking the floods and assuming a particular form for their distribution each can be assigned an exceedence probability and so a flood frequency curve can be constructed. It is interesting to note some of the properties of this annual maximum flood series. It might be expected that the mean of the annual maxima is exceeded by approximately half of the floods and so have an exceedence probability of roughly 0.5 and a return period of about two years. However, since it is possible to have floods very much bigger than the mean and because there is a limit to how much smaller they can be, the distribution of floods is skewed. In fact the mean of the annual maxima is usually taken to have an exceedence probability of 0.43 and a return period of 2.33 years. Figure 1.5(a) shows the probability density function and Fig 1.5(b) the distribution function for the annual maximum floods; this shows the skewed nature of the distribution and introduces the concept

A flood frequency curve

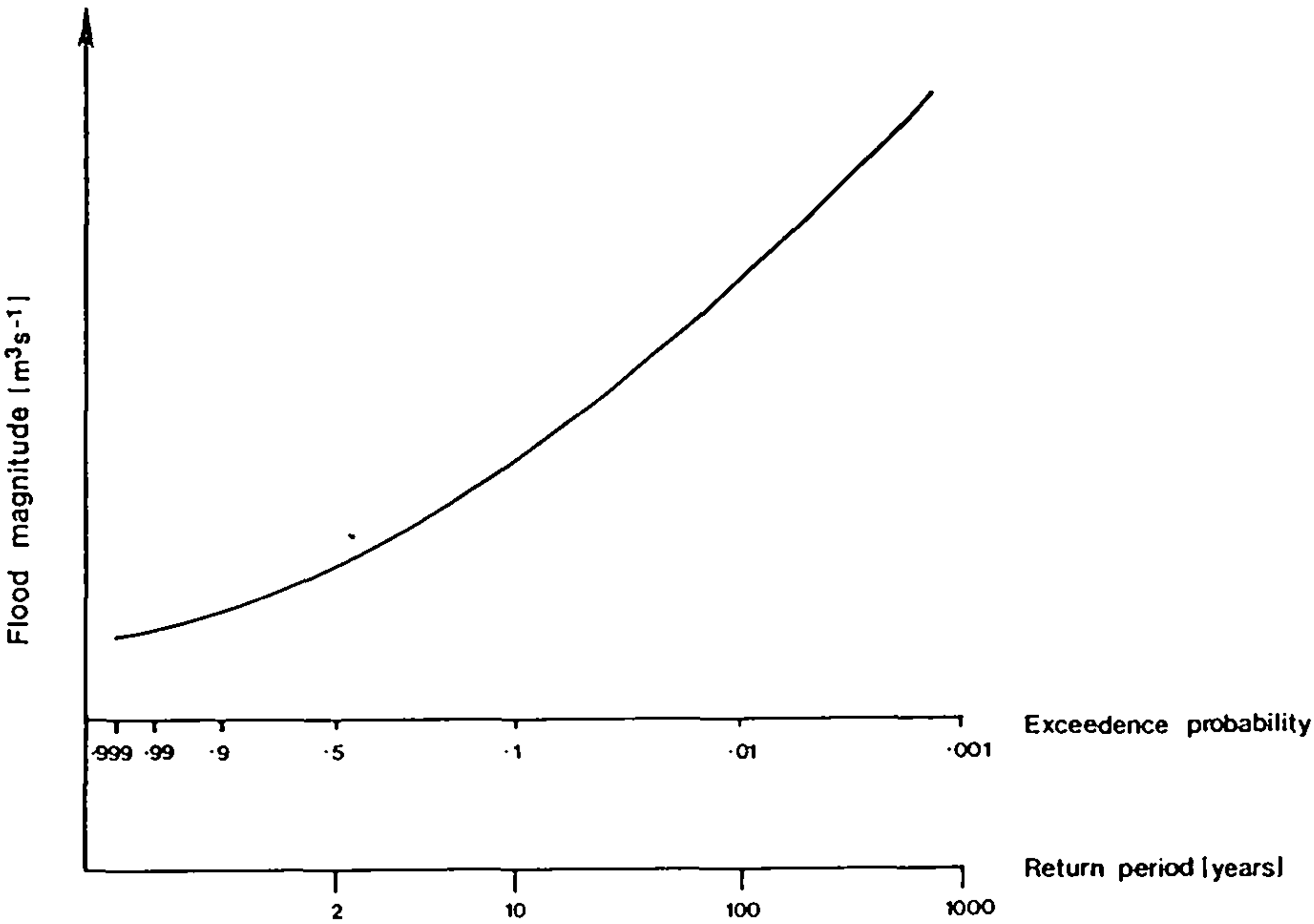


Figure 1·3

Identification of floods and their associated return periods

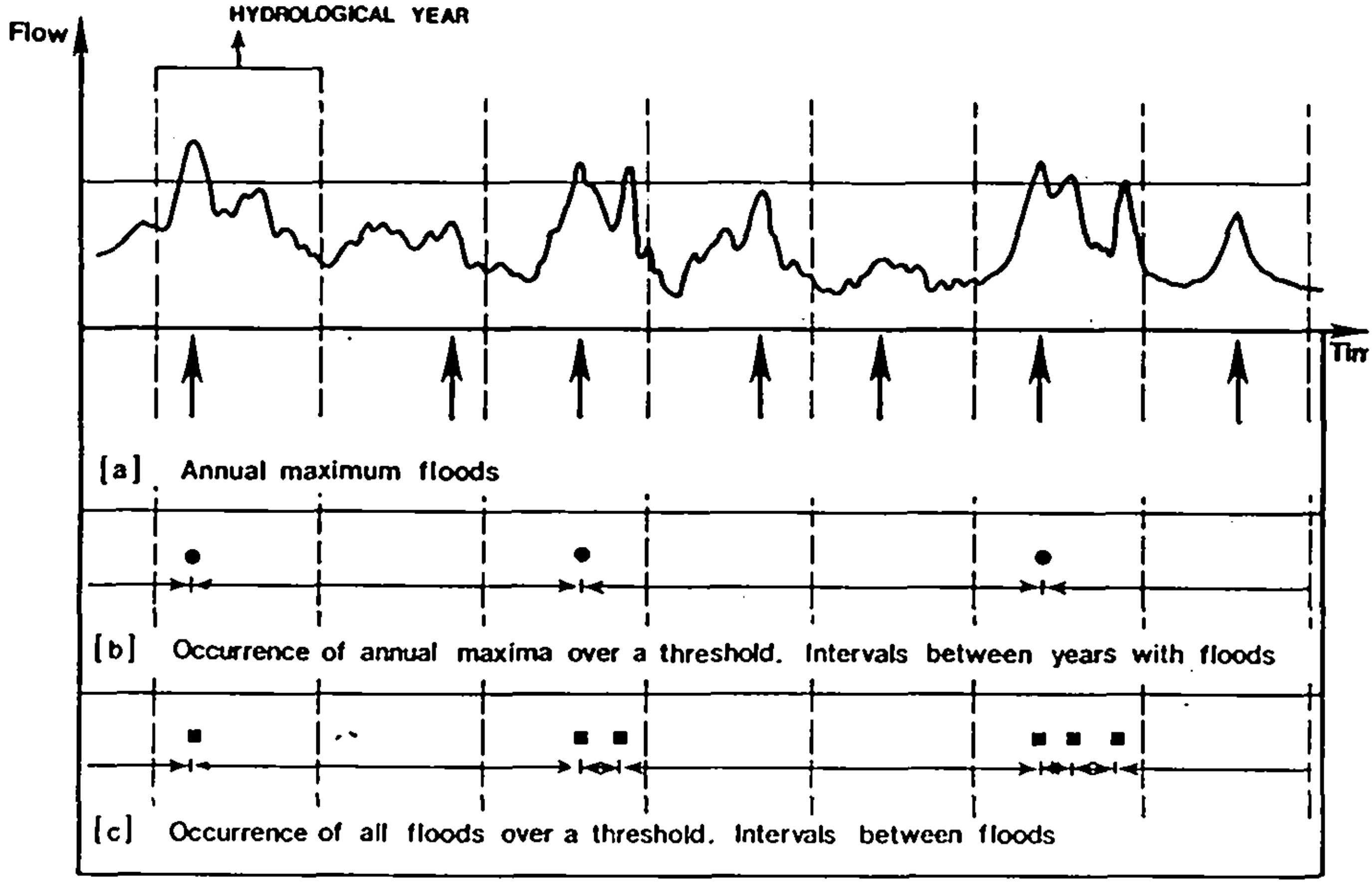


Figure 1.4

The probability density function of annual maximum floods

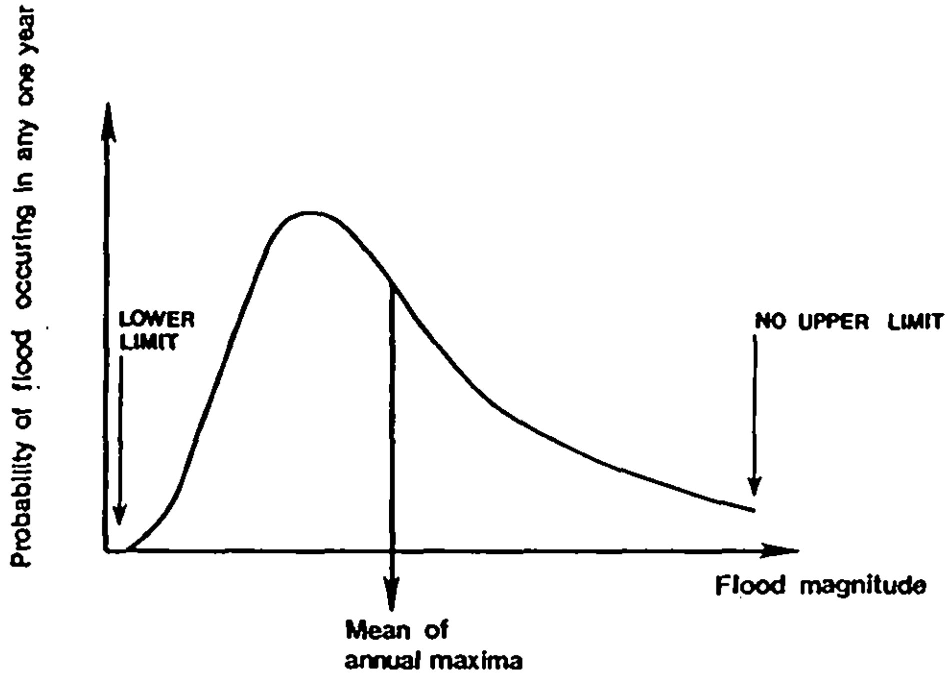


Figure 1·5 (a)

The distribution function of annual maximum floods

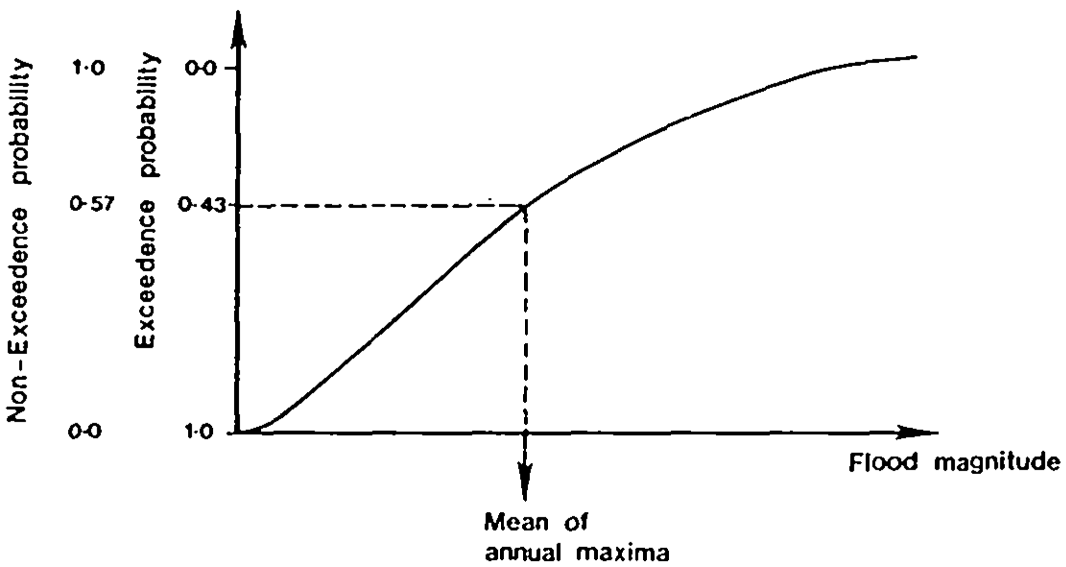


Figure 1·5 (b)

of non-exceedence probability (the probability of a flood not being exceeded in any one year) which is frequently used in preference to the exceedence probability as it leads to easy estimation of the risk of failure of any scheme (see Section 1.3).

When considering only the biggest flood in each year, the return period is not the average interval between floods of a given magnitude but the average interval between years containing floods of that size or greater. In Figure 1.4(a) it can be seen that the largest flood in some years is exceeded by the second or third largest flood in others; a second approach to flood frequency analysis that avoids this problem considers all the floods over a certain limiting size, not just the biggest in each year. Such a flood sequence is called a partial duration, or peaks over a threshold, series. In Figure 1.4(b) all years containing floods over a certain size have been marked and in Figure 1.4(c) all floods over that same size are indicated. Immediately it can be seen that the average interval between floods is less than the average interval between years with floods. The return period of the mean annual flood from the annual maximum series is about half a year greater than from the partial duration series, but the difference decreases as return period increases since as the threshold is raised the two series become identical. Although the partial duration series approach is the more fundamental one, the small difference at large return periods and the easy application of the annual maximum method makes it the more popular choice for flood frequency studies.

1.2 Flood estimation methods

When a long record is available estimation of the flood of specified return period is a straightforward task as outlined above. However, it is usually the case that only a limited period of data is available and it is either impossible to construct a flood frequency curve or to extend it to the required return period. There are three broad classes of method that can be used in such circumstances; statistical methods, rainfall-runoff methods and international empirical methods.

The international empirical methods are usually simple formulae relating flood magnitudes to physiographical properties of the drainage area. They are often based on a straightforward conceptualisation of the rainfall-runoff process and calibrated on a specific data set. The growth of flood magnitudes with return period is achieved through using rainfall frequency relationships which are generally more widely available than flood frequency curves. Because of the methods' generality and in the absence of anything better they have been adopted for use all over the world and several variations are currently used in Indonesia. Examples include the Rational method and the Creager and Franco-Rodier equations.

Rainfall-runoff methods also require rainfall frequency information although often over a variety of durations. The rainfall input is routed through a rainfall-runoff model to give the design flood. The methods have the advantage of giving a complete design hydrograph but require a considerable amount of good quality data for calibration before they can be applied to an ungauged site. This requirement makes them unsuitable at the present time for use in Java and Sumatra.

Statistical methods are based on the regional generalisation of statistical properties of flood distributions. Typically the methods involve the estimation of an index flood and the scaling of this by a factor dependent on return period to give the design flood or T year flood where T is the required return period. This method has been adopted for use in the current study as it makes best use of the available data, provides for the easy incorporation of local data in application and links in well with flood frequency concepts applicable to long records.

The index flood chosen was the mean annual flood (the mean of the annual maximum flood series) as this can be estimated at a large number of sites in Java and Sumatra from existing records. As stated earlier this can be estimated from the annual maxima for long records. However, for short records it is better to use the partial duration series (or peaks over a threshold) method as this includes information from more floods and is therefore more accurate. This method is also useful where a longer record contains breaks as the start of year is not important and incomplete years of data can be included, see Figure 1.6. For very short records even this method is not suitable. Here the no data method of an empirical locally based equation relating mean annual flood to catchment characteristics is the best. This equation is the only method available when no data at all exist.

As described above, for a very long record a flood frequency curve can be constructed from the record itself. However when a flood estimate is required for a return period much greater than the record length, averaged ratios of the magnitude of the T year flood to mean annual flood are required. These 'growth factors' have been derived in this study using local data and depend not only on return period but on drainage area as well.

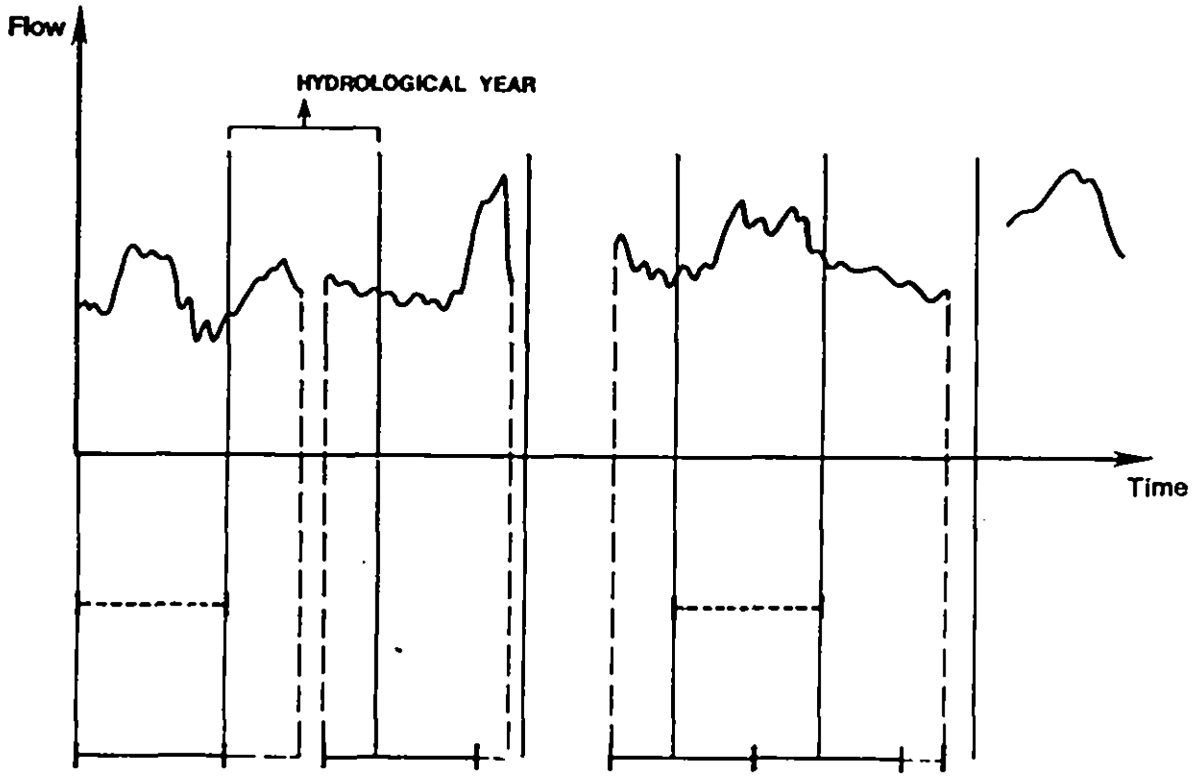
1.3 Choice of design flood

The decision of what return period is appropriate for the design of a particular project is not solely a hydrological problem. The engineer is constrained by economic, political and environmental factors in his design and so cannot improve the safety or reliability of the scheme without incurring costs elsewhere. It is however useful to consider the probability, or risk, of the design flood being exceeded during the expected life of the project. If the design flood has a return period of T years then the risk, r, of the flood being exceeded in the L year projected life of the project is given by

$$- (1 - \frac{1}{T})^{L}$$

Thus given an expected design life of 50 years for a road bridge or major irrigation offtake, there is a risk of 0.64 or 64 per cent, of the structure experiencing the 50 year flood during its lifetime. The risk of the same structure experiencing a 1000 year flood is only 5 per cent or put another way, only one such structure in 20 would be likely to experience a 1000 year flood during a designed life of 50 years.

Use of broken records: maximising use of data through use of POT method



COMPLETE HYDROLOGICAL YEARS

COMPLETE YEARS FOR P.O.T. METHOD

INCOMPLETE YEARS OF DATA ALSO USABLE FOR P.O.T. METHOD

Figure 1·6

The methods presented in this report are suitable for estimating the magnitude of floods up to the 500 year event and for tentative estimates of floods up to 1000 year return period. This is generally adequate for the design of bridges, small irrigation works and channel improvement works and to assist in the planning of urban development or the assessment of alternative large dam proposals. Many detailed studies for larger projects will require a more extreme flood to be estimated; for example the 'probable maximum flood' may be required to design the spillway of a major dam. The methods of estimation presented in this report cannot be used directly in such cases.

Users of the manual will note that with each method an estimate of likely errors is given; again it is the problem of the design engineer to decide how best to incorporate this uncertainty in estimation into his design. The accuracy of flood estimation depends greatly on the quality and quantity of available data. Although a flood estimate can be made at any site using the no data method, as little as two year's data recorded at the site will lead to a better estimate of the design flood. At a site where data has been collected for several years but the rating is good only for low flows, a flood estimate will be greatly improved by the development of a flood rating following a period of frequent flow gauging. Since the quality of rating equations is of great importance in flood hydrology, rating equations for all stations used in this study were reviewed, and frequently revised prior to use. The rating accuracy should be considered in the engineer's adoption of the design level.

HOW TO USE THIS REPORT

The flood estimation methods presented in this report use a common approach; an index flood, the 'mean annual flood' is estimated and then scaled by the appropriate 'growth factor' to give the flood of required return period. The mean annual flood (MAF) at any site is defined as the mean of all annual maximum instantaneous flood peaks. The flood of return period T years, written throughout this report Q_T, is the flood that on average will be exceeded once in a period of T years.

The flow chart in Figure 2.1 illustrates which methods should be used according the availability of data.

Where a long flood record is available the mean annual flood can be estimated by the mean of the annual maxima in the sample of available data as described in Chapter 3. For a long record the sample mean should be a good estimate of the true mean but for shorter records such an estimate becomes less good. For this reason where only a few year's data are available a better estimate of the MAF is obtained by considering all flood peaks that exceed a threshold level. This method, called the peaks over a threshold (POT) method, is detailed in Chapter 4. If no data at all are available then the estimation equation given in Chapter 5 should be used. This equation relates the size of the MAF to various physical and climatological characteristics of the catchment that are indexed by parameters obtained from maps. Wherever possible the MAF should be estimated by more than one method so that the estimates may be compared. Chapter 8 gives various methods of using data from another station situated either on the same catchment, or on a neighbouring one, in conjunction with either the estimation equation or data from the site itself.

Having obtained the best possible estimate of the mean annual flood it must be multiplied by a growth factor to give the flood of required return period. The multiplier is dependent not only on T but also on catchment area and is obtained from the table of multipliers given in Chapter 7. In the unusual case of a very long record being available then this stage of the design procedure can be replaced by the development of a flood frequency curve for the site of interest.

if the return period of the required design flood is not significantly greater than the length of the available record. The development of such a curve is described in Chapter 6

A comparison of the 'no data' methods of this report with a number of alternative flood estimation techniques currently used in Indonesia is given in Chapter 9. It is apparent that the methods of this report give consistently better results over a wide range of catchment types.

In the annexes that follow the body of the report some of the methods described only briefly in the relevant chapters are explained in more detail including variations in the basic methods and background theory. These topics include rating curve development (Annex B) development of the MAF estimation equation (Annex D), the POT method (Annex E) and growth curves (Annex F).

A data appendix (under separate cover) contains all the basic flood data collated during this study.

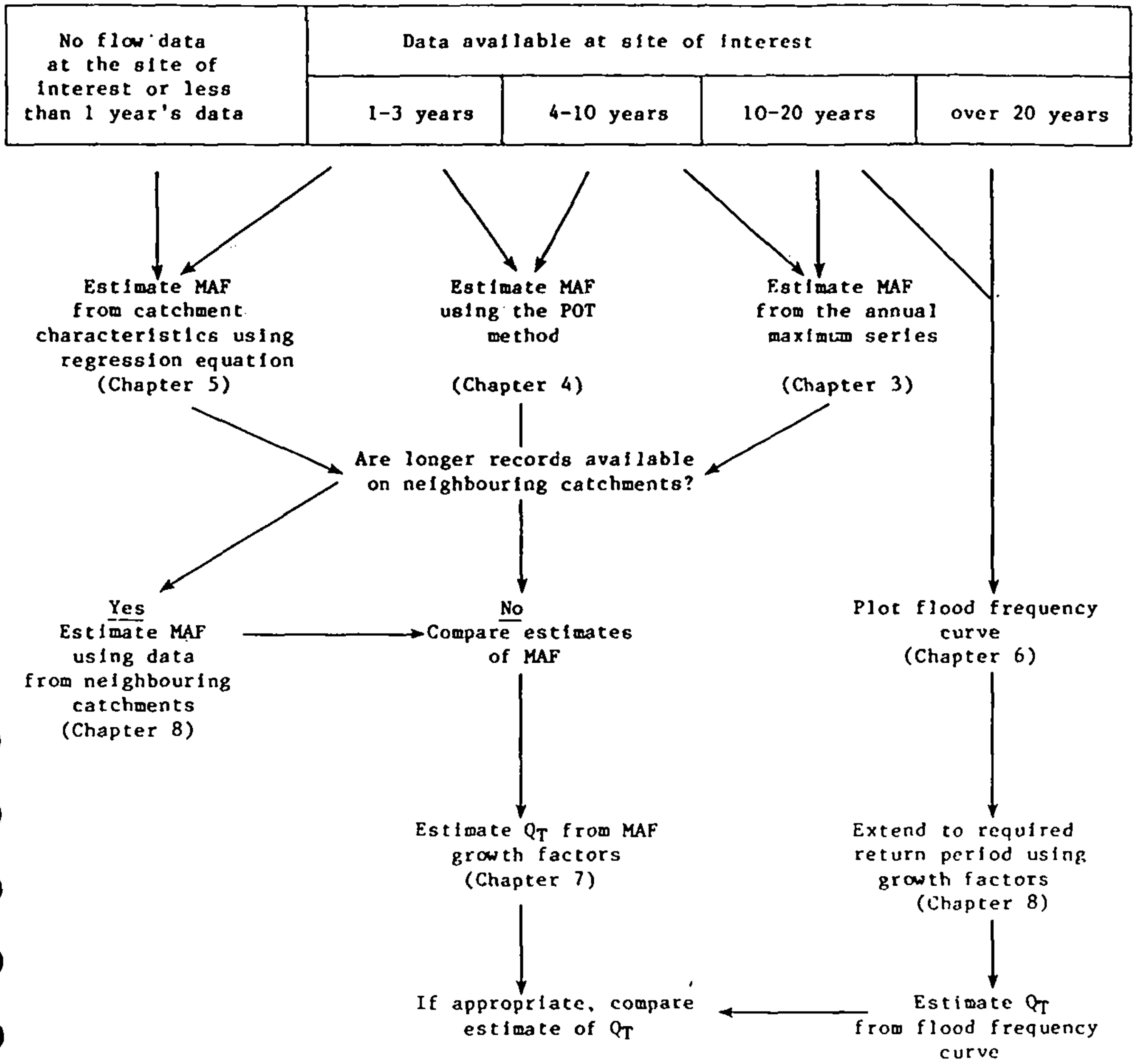


Figure 2.1 Guide to the use of this manual

ESTIMATION OF MEAN ANNUAL FLOOD FROM ANNUAL MAXIMUM SERIES

3.1 Introduction

The present chapter is concerned with estimating the mean annual flood, MAF, from the annual maximum flood series recorded at the site of interest. If the site of interest is gauged and flood peaks have been observed over a sufficiently long time, the mean of the annual maxima over the period of record may be used to provide an acceptable estimate of the MAF. It follows therefore that the MAF will be better estimated where the annual maximum series is long and its variance small. From our experience in the UK and elsewhere, we suggest that an acceptable estimate of the MAF can be derived from a minimum of 5 years' good quality data.

In order to preserve the independence of the annual maxima it is advantageous to start the hydrological year in the dry season. The chance of the same period of flooding contributing to two successive years is then least likely. Most catchments in Java and Sumatra exhibit a distinct flood season between November and April. In this study the hydrological year starts on the 1st August which, for many catchments studied, is the driest month.

3.2 Description of method

The method involves abstracting the highest flood peak in each hydrological year of record. It is important that small floods from years of incomplete record are not included in the annual maximum series and to ensure that it may be best to totally disregard such years of data. However if a particularly large flood is noted in an incomplete year its inclusion in the annual maximum series is desirable; as a guide, estimate the MAF from complete years of data and then include maxima from incomplete years greater than this and recalculate the MAF. Broken records can be used, provided only complete hydrological years are taken from within it.

If the annual maximum series contains one or more extreme floods the mean may be too high an estimate of the MAF. The UK Flood Studies Report (NERC, 1975) gives an approximate test to determine whether this is so. If Q_{max}, the maximum flood on record is greater than three times Q_{med}, the median value of the series, the record contains an outlier. It is suggested that the same test is used in Java and Sumatra as the annual maxima series for the UK and the Indonesian catchments studied exhibit a similar variability.

If the annual maximum series contains no extreme floods, the MAF is estimated as the mean of the data:

$$MAF = \frac{1}{N} \sum_{i=1}^{N} q_i$$

where,

q_i (i=1,2,...N) = flood peaks in the annual maximum series

N = number of year's data

If the annual maximum series contains one or more extreme floods, the MAF is estimated from the median of the recorded series:

$$MAF = 1.06\ Q_{med}$$

The multiplier, 1.06, in the above equation is the average ratio of mean annual flood to median annual flood for all catchments studied in Java and Sumatra. (The UK multiplier is 1.07).

3.3 Accuracy of result

The standard deviation (sd) is used here to define the accuracy of the estimation of the MAF. There is a 68% chance that the MAF estimated lies within one standard deviation of the true long term value.

The standard deviation of the annual maximum discharges is calculated thus:

$$sd(MAF) \quad \sqrt{\sum_{i=1}^{N} \frac{(q_i - \bar{q})^2}{N-1}}$$

where

N and q_i are as above and

$\bar{q}$ = mean of q_i

Although this estimate of the standard deviation strictly applies to the MAF estimated from the mean of the series, it is suggested that the standard deviation be calculated in the same way when the MAF is estimated from 1.06 x Q_{med}.

3.4 Example of application

For the Citarum at Palumbon, 31 years of flow data are available. This is sufficient to provide a good estimate of the MAF by taking the mean of the annual maximum floods.

The annual maxima are ranked in order of descending magnitude:

Year	Max. Flood $m^3\ s^{-1}$	Year	Max Flood $m^3\ s^{-1}$
39/40	2733 - Q_{max}	77/78	1328
78/79	2356	29/30	1251
68/69	2141	64/65	1241
30/31	1994	36/37	1213
35/36	1994	62/63	1158
67/68	1962	76/77	1158
66/67	1841	70/71	1121
80/81	1723	72/73	1121
37/38	1662	75/76	1009
69/70	1610	63/64	984
79/80	1610	65/66	984
73/74	1513	33/34	968
40/41	1439	71/72	943
34/35	1383	38/39	927
32/33	1348	74/75	790
31/32	1338 - Q_{med}		

The first step is to test the record for any extreme flood which could cause an overestimation of the MAF.

$$Q_{max} = 2733 \ m^3 \ s^{-1}$$
$$Q_{med} = 1338 \ m^3 \ s^{-1}$$

$$\frac{Q_{max}}{Q_{med}} = 2.04$$

The ratio of Q_{max}/Q_{med} is below the critical level of 3 indicating no extreme flood is present. The arithmetic mean of the 31 year series is therefore used to estimate the MAF.

$$MAF = 1447 \ m^3 s^{-1}$$

The measure of error associated with the estimate, the standard deviation, is calculated as 466 $m^3 \ s^{-1}$.

The accuracy of the method does not justify the implied accuracy of numbers quoted above and it is therefore more reasonable to say that the MAF for the Citarum at Palumbon is estimated as 1450 $m^3 s^{-1}$ with a standard deviation of 470 $m^3 s^{-1}$.

4. ESTIMATION OF MEAN ANNUAL FLOOD FROM PEAKS OVER A THRESHOLD SERIES

4.1 Introduction

When only a limited period of record is available, finding the mean annual flood from the annual maximum series is inappropriate as the series will be too short to estimate the mean reliably. In such circumstances the series of peaks over a threshold (POT series) may be used to estimate the mean annual flood using a larger number of flood peaks. The POT method should not be used where less than two complete years of data are available; the method is valid for long records although in practice the annual maximum method is easier to apply and equally accurate for records of over five years long. A method of applying the technique is described in the next section followed by an example. A more detailed description of the POT method is given in Annex E.

4.2 Description of Method

All available data should be assembled and the complete years of data identified (the starting date of the year is not important). Following a cursory examination of the data a flow threshold is chosen so that on average between two and five peaks per year exceed the threshold, the exact number not being critical. From the complete years of data (N years) all flow peaks exceeding the threshold, q_o, are abstracted; these M flood peaks $q_i (i = 1,2 \ldots M)$ form the peaks over a threshold series.

Where the peaks are to be taken from a stage record a stage threshold, h_o, can be chosen as the basis for peak selection. The abstracted stage peaks (h_i) are converted to flows (q_i) using an appropriate rating equation.

In selecting peaks care should be taken to ensure that they are independent. A simple test for independence is illustrated in Figure 4.1. To decide if q_2 is independent from q_1 the separation of the peaks (T_s) must be greater than three times the rise time (T_r) of the first peak and the trough (q_t) between

Test for independence of flood peaks

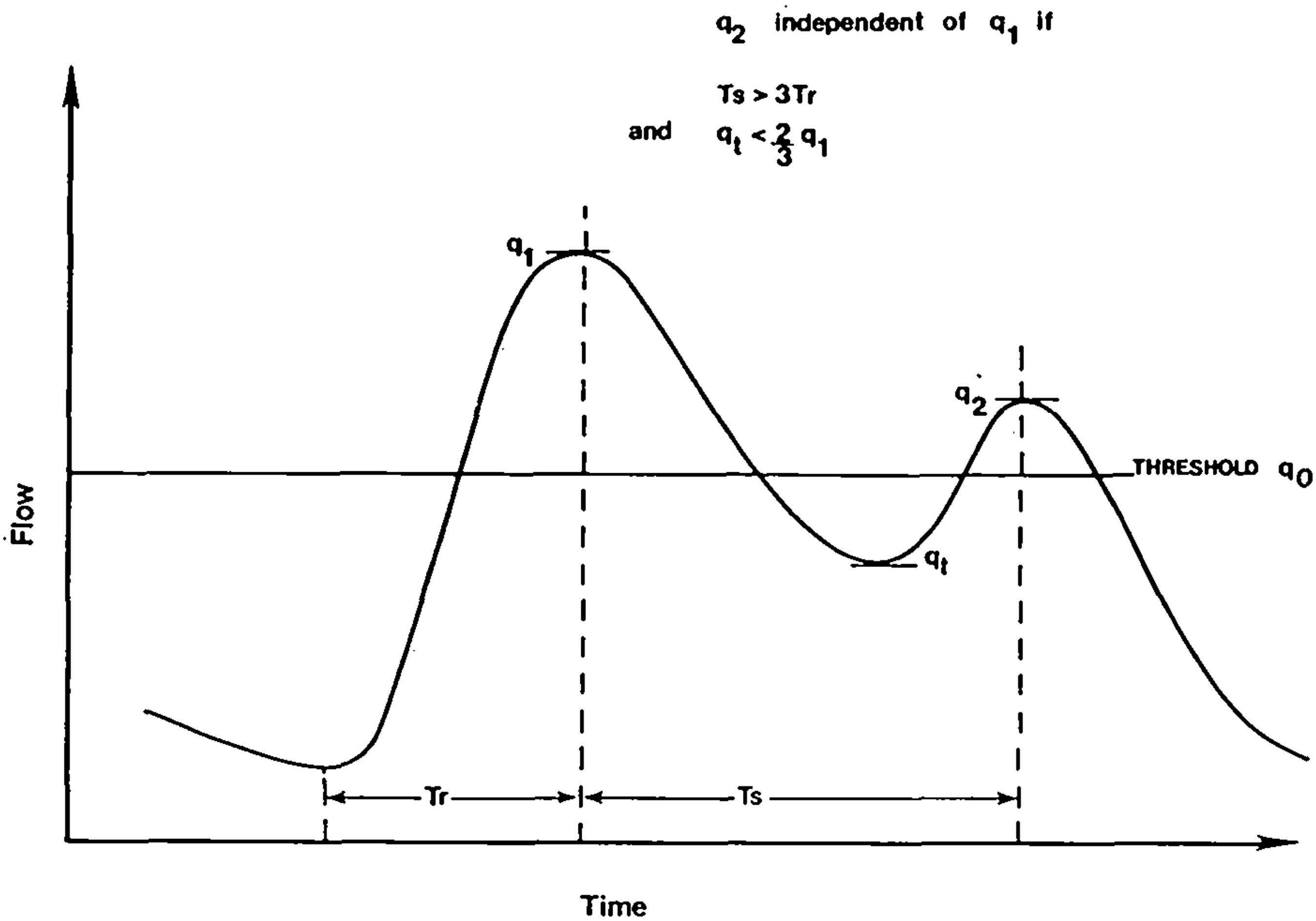


Figure 4·1

the peaks must be less than two thirds of the first peak q_1. This arbitrary rule was used in the UK Flood Studies Report (NERC, 1975) as it is objective and easy to apply. If the test suggests that the peaks are not independent, only the first peak, q_1, should be included.

From the M floods, q_i, over the threshold, q_o, the mean exceedence β is calculated from

$$\beta = \frac{1}{M} \sum_{i=1}^{M} (q_i - q_o)$$

and the average number of exceedences per year, λ, from

$$= M/N$$

The mean annual flood is then estimated from

$$MAF = q_o + \beta(0.5772 + \ln\lambda) \ m^3s^{-1}$$

where ln is the natural logarithm or $\log_e$.

4.3 Accuracy of results

The standard deviation of the estimate is given by

$$sd\ (MAF) = \frac{\beta}{\sqrt{N}} \left[\frac{1}{\lambda} + \frac{(0.5772 + \ln\lambda)^2}{\lambda}\right]^{0.5} m^3s^{-1}$$

For between three and five exceedences per year the term inside the brackets is approximately 1.1 thus

$$sd\ (MAF) \approx 1.1 \frac{\beta}{\sqrt{N}} \ m^3s^{-1}$$

To assess the accuracy of the MAF it is helpful to remember that, on average, 68 times out of 100 the estimated value of the MAF will be within one standard deviation of the 'true' value.

4.4 Example of application

For Batang Hari at Muara Tembesi just over four years of data are available. From this length of record the MAF is best estimated by the peaks over a threshold method.

From an initial scan of the record the threshold of 4000 m^3s^{-1} seemed likely to give a suitable number of floods for the POT series. Nine independent floods were abstracted as listed below

Year	Flood (m^3s^{-1})
1977	4365.6
	4032.3
	4026.1
1978	4843.4
	4340.1
	4113.3
1979	4596.2
1980	4232.6
	4461.3

Floods in the incomplete year at the end of the record were ignored.

In the notation given above

Threshold, $q_0 = 4000$ m^3s^{-1}

Number of years of data, N = 4

Number of floods over the threshold, M = 9

Therefore $\beta = 334.5$ m^3s^{-1}

$\lambda = 2.25$ floods/year

and MAF = 4464.3 m^3s^{-1}

sd (MAF) = 190.8 m^3s^{-1}

Thus, using the POT method the mean annual flood for Batang Hari at Muara Tembesi is estimated to be 4460 m^3s^{-1} with a standard deviation of 190 m^3s^{-1}.

4.5 Using incomplete years of data

It is often the case that the flow record from a station is incomplete and that much data would be wasted if only complete years of data were used. One such station is Batang Hari at Muara Kilis for which, at the time of this study data were available for the period March 1976 to October 1981, a period of over five years. Although only two complete years of data are present in this record many large floods were observed in the remainder of the record. In this situation the threshold should be chosen so that between 2 and 5 peaks are selected from complete years and then the entire record examined for exceedences. These should be listed with a note of whether or not the year is complete. For Batang Hari at Muara Kilis the threshold of 2300 m^3s^{-1} was chosen and the following peaks abstracted.

Year	Flood (m^3s^{-1})
March 1976-1977	2329.5
(complete)	2434.6
	2739.0
	2562.2
March 1977-1978	2308.6
(complete)	2661.0
	3230.8
	2609.4
	2579.3
	2337.9
March 1978-1979	2557.9
(incomplete)	2400.9
December 1980-1981	2596.5
(incomplete)	2304.4
	2583.6

The complete years of data are used to estimate the average number of exceedences per year

$$= \frac{\text{number of floods in complete years}}{\text{number of complete years, N}}$$

$$= \frac{10}{2}$$

The average exceedence, β, is estimated from all the floods, the total number in this case being $M = 15$

$$\beta = 249.04 \ m^3s^{-1}$$

To estimate the MAF the same equation is used

$$MAF = q_o + \beta(0.5772 + \ln\lambda) \ m^3s^{-1}$$

$$= 2844.56 \ m^3s^{-1}$$

The standard deviation is estimated by

$$sd\ (MAF) = \frac{\beta}{\sqrt{(\lambda.N)}} + \frac{\beta}{\sqrt{M}} (0.5772 + \ln\lambda)$$

$$= \frac{249.04}{\sqrt{(5x2)}} + \frac{249.04}{\sqrt{15}} (0.5772 + \ln 5)$$

$$219.36 \ m^3s^{-1}$$

Using data from both complete and incomplete years the mean annual flood for Batang Hari at Muara Kilis is estimated to be 2840 m^3s^{-1} with a standard deviation of 220 m^3s^{-1}. Thus for stations with incomplete years of data, only complete years should be used to calculate the average number of exceedences per year, λ, but all available data should be used to compute the average exceedence, β.

5. ESTIMATION OF MEAN ANNUAL FLOOD FROM CATCHMENT CHARACTERISTICS

5.1 Introduction

This chapter describes a method of estimating the MAF when no flow data are available at the site of interest. The method uses a regression equation relating the MAF to four readily obtainable catchment characteristics. A detailed description of how this equation was derived may be found in Annex C.

The regression equation was derived using data representing a wide range of catchment characteristics and may be applied anywhere in Java and Sumatra subject to the constraints described below. It is also recommended that the equation should not be used for flood estimation in heavily urbanised catchments as these were not considered in this study.

5.2 Description of method

It is firstly necessary to estimate the four catchment characteristics, used in the regression equation which are tabulated below; Annex D gives guidance on how these should be obtained. As well as the maps provided in this report, a topographic map of suitable scale covering the catchment area is required.

AREA = Catchment area (km^2)
APBAR = Mean annual maximum catchment 1 day rainfall (mm)
SIMS = Slope index ($m\ km^{-1}$)
LAKE = Lake index (dimensionless)

Before proceeding with this method it is necessary to check that the catchment characteristics of the basin under study are within the ranges of characteristics of the gauged catchment used in the development of the equation. Of the four catchment characteristics, AREA and APBAR are the most important in indexing the MAF. Figure 5.1 shows the spread of the AREA and APBAR data of the regression data set. It is recommended that the regression equation only be used if the AREA and APBAR combination of the catchment under study lies within the inner area shown in Figure 5.1.

Plot of APBAR against AREA for regression analysis catchments

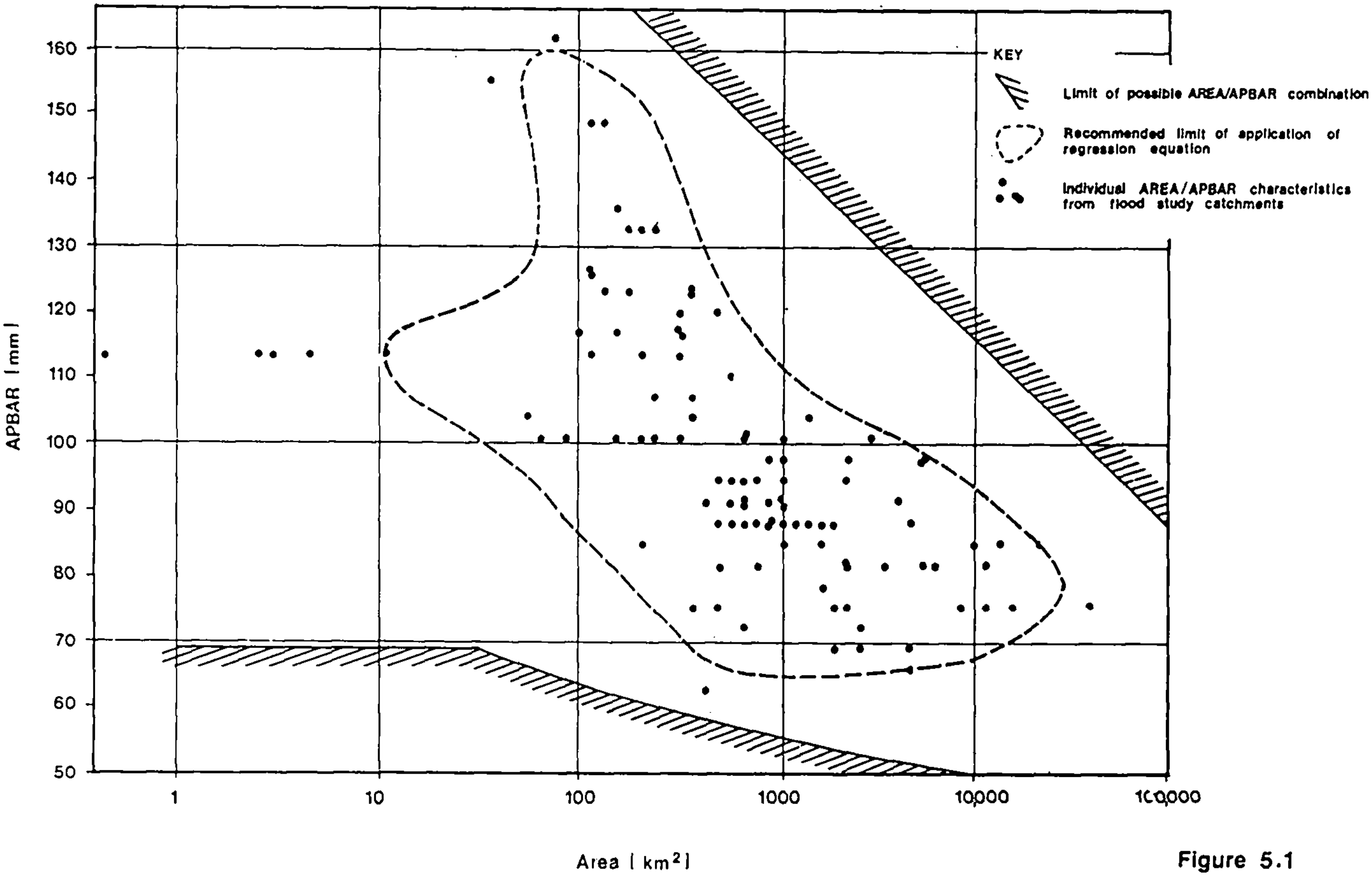


Figure 5.1

As a final check SIMS and LAKE should in the range given below:

SIMS 1 to 150 m km^{-1}

LAKE 0 to 0.25

Having obtained AREA, APBAR, SIMS and LAKE for the site of interest and checked that the values are within the acceptable range, the MAF is estimated from the following equation:

$$MAF = 8.00 \times 10^{-6} \times AREA^{V} \times APBAR^{2.445} \times SIMS^{0.117} \times (1 + LAKE)^{-0.85}\ m^3 s^{-1}$$

The exponent of AREA, V, is itself a function of catchment area and may be calculated from the formula:

$$V = 1.02 - 0.0275 \log_{10} AREA$$

The table below gives V for various catchment areas and may be used to check that the value of V calculated is in the correct range.

AREA (km^2)		AREA (km^2)	
1	1.020	500	0.946
5	1.001	1000	0.938
10	0.993	5000	0.918
50	0.973	10000	0.910
100	0.965		

5.3 Accuracy of method

The two previous methods of estimating the MAF (Chapters 3 and 4) quoted formulae for estimating the standard deviation of the MAF. With the regression equation, however, the factorial standard error of the estimate is used (Annex C). In fact this factorial standard error is analogous to the standard deviation. There is a 68% chance that any one flood estimate lies within the range MAF x 1.59 (or MAF + 59%) to MAF/1.59 (or MAF - 36%) of the 'true' MAF.

This large standard error of the prediction equation may surprise some readers. However, the factorial standard error of the estimate of the UK prediction equation (Flood Studies Report, NERC 1975) was

1.49 using data from 532 basins. These standard errors are a measure of the uncertainty of flood estimation on ungauged basins. However in many cases flood estimates may be improved by using local data; this is discussed in Chapter 8.

5.4 Example of application

The MAF is estimated for the Cimandiri at Tegal Datar using the method described in this chapter. From Table A.1 in Annex A the relevant catchment characteristics for Tegal Datar are:

AREA	=	495.1	km^2
APBAR	=	94.0	mm
SIMS	=	21.6	$m\ km^{-1}$
LAKE	=	0	

From Figure 5.1 the AREA/APBAR combination is within the acceptable range. Furthermore SIMS and LAKE are within the limits defined in Section 5.2. The regression equation may therefore be used.

First the exponent of AREA, V, in the regression equation is calculated:

$V = 1.02 - 0.0275 \log_{10} 495.1$

$V = 0.946$ (This checks with the table in section 5.2 where V for an AREA of 500 km^2 is 0.946)

MAF is then estimated:

$$MAF = 8.00 \times 10^{-6} \times 495.1^{0.946} \times 94^{2.445} \times 21.6^{0.117} \times (1 + 0.0)^{-0.85}$$

$$MAF = 271\ m^3\ s^{-1}$$

As the factorial standard error of estimate of the MAF is 1.59, the MAF may be quoted as 271 m^3s^{-1} with a 68% chance that the MAF lies between 431 m^3s^{-1} (271 x 1.59) and 170 m^3s^{-1} (271/1.59).

In fact there are 6 years of flow data available at Tegal Datar to provide a comparison within the regression method. The MAF from the mean of the 6 annual maxima is 361 m^3s^{-1} with a standard deviation of 38 m^3s^{-1}. This estimate of MAF would be used in preference to the regression equation estimate since it is based on local data and has a significantly smaller uncertainty associated with it. Although the estimation equation is seen to give a reasonable estimate of the MAF in this example it should be remembered that this will not always be the case. It should be noted that the standard deviation is unusually small for Tegal Datar, being only 10.5 per cent of the MAF. The average value from the catchments studied was 32 per cent.

6. ESTIMATION OF THE T YEAR FLOOD USING A LONG RECORD

6.1 Introduction

If more than twenty years of data are available at the site of interest a flood frequency curve can be plotted which will allow floods of return period up to the length of record to be estimated. This procedure is described in detail in this chapter.

When a more extreme flood has to be estimated the curve may be extended with reference to the flood frequency factors given in Chapter 7 by the method given in Section 8.8. A flood frequency curve can also be plotted when a shorter record is available but this curve should always be used in conjunction with the average flood frequency factors even for low return periods.

6.2 Plotting the recorded data

The annual maximum floods are abstracted from the N years of data and ordered so that the smallest flood is given rank 1 and the largest rank N. For each flood a probability of non-exceedence is assigned to it based on its position in the ranked series. This requires making an assumption about the form of the distribution from which the observed annual maxima are drawn. If the distribution is assumed to be a type 1 extreme value (EV1 or Gumbel) distribution then a good approximation to the non-exceedence probability is given by the Gringorten formula:-

$$F_i = \frac{i - 0.44}{N + 0.12}$$

where F_i is the non-exceedence probability (or plotting position) and i is the rank of the flood. In order to plot the frequency curve on linear graph paper, the EV1 reduced variate, y_i, must be calculated from the values of F_i; this can be done using the approximation

$$y_i = - \ln(- \ln F_i)$$

which is sufficiently accurate for plotting purposes.

The values of Q_i should then be plotted against the corresponding y_i on linear graph paper. The resulting plot becomes rather more useful when the reduced variate axis is rescaled in terms of return period, T. The y values corresponding to various return periods can be calculated from

$$y = -\ln\left(-\ln\left(\frac{T-1}{T}\right)\right)$$

The following table gives values of reduced variate for commonly required return periods.

Return Period,T (years)	Reduced Variate,y
2.0	0.37
2.33	0.58
5	1.50
10	2.25
20	2.97
25	3.20
40	3.68
50	3.90
75	4.31
100	4.60
200	5.30
500	6.21
1000	6.91

A smooth line should be drawn through the plotted points but need not be constrained to pass through the highest point where this lies a considerable distance from the rest of the data. If the data plot as a straight line, then the assumption of a parent EV1 distribution appears valid. However, the plot is likely to show a slight curvature suggesting the parent distribution is something other than an EV1 although in practice the sampling error is usually too large to state definitely that this is the case. Worldwide experience in plotting frequency curves suggests that steeper curves come from low rainfall areas and from smaller catchments. This trend is by no means well established (see Annex F) but should be considered when the completed curve is compared with one based on averaged flood frequency factors

For the range covered by the curve the flood corresponding to a given return period can be estimated. It should be noted that estimation of the T-year event directly by frequency curve does not require the MAF to be estimated. The upper limit of the range will depend on the variability of the plotted data about the curve; even if the data plot on a straight line it should not be extended to return periods greater than twice the length of record.

6.3 Example of application

An estimate of the 5 year return period flood is required at the Citarum at Nanjung. This station has 21 years of data which is sufficient to draw a flood frequency curve. Table 6.1 gives an ordered list of the recorded flood peaks and the corresponding values of y_i based on the rank. The flood magnitudes have been plotted against y_i in Figure 6.1. Drawing a line through these points is difficult, the best solution possibly being to draw the straight line shown. The scatter about this line for higher return periods is large and it may be best to limit flood estimation with this curve to return periods up to ten years.

From the line the five year flood can be estimated as 290 m^3s^{-1} and the ten year flood as 312 m^3s^{-1}.

Table 6.1 Ranked floods, and the corresponding values of y_i for the Citarum at Nanjung

Rank i	Flood Q_i	EV1 reduced variate y_i
21	370.0	3.62
20	303.0	2.60
19	297.0	2.07
18	293.0	1.71
17	291.0	1.43
16	288.0	1.20
15	286.0	1.01
14	284.0	.83
13	284.0	.67
12	274.0	.52
11	270.0	.38
10	270.0	.25
9	270.0	.12
8	268.0	.01
7	261.0	.14
6	253.0	.27
5	251.0	.41
4	226.0	.56
3	221.0	.73
2	208.0	.96
	205.0	-1.29

Flood frequency curve for the Citarum at Nanjung

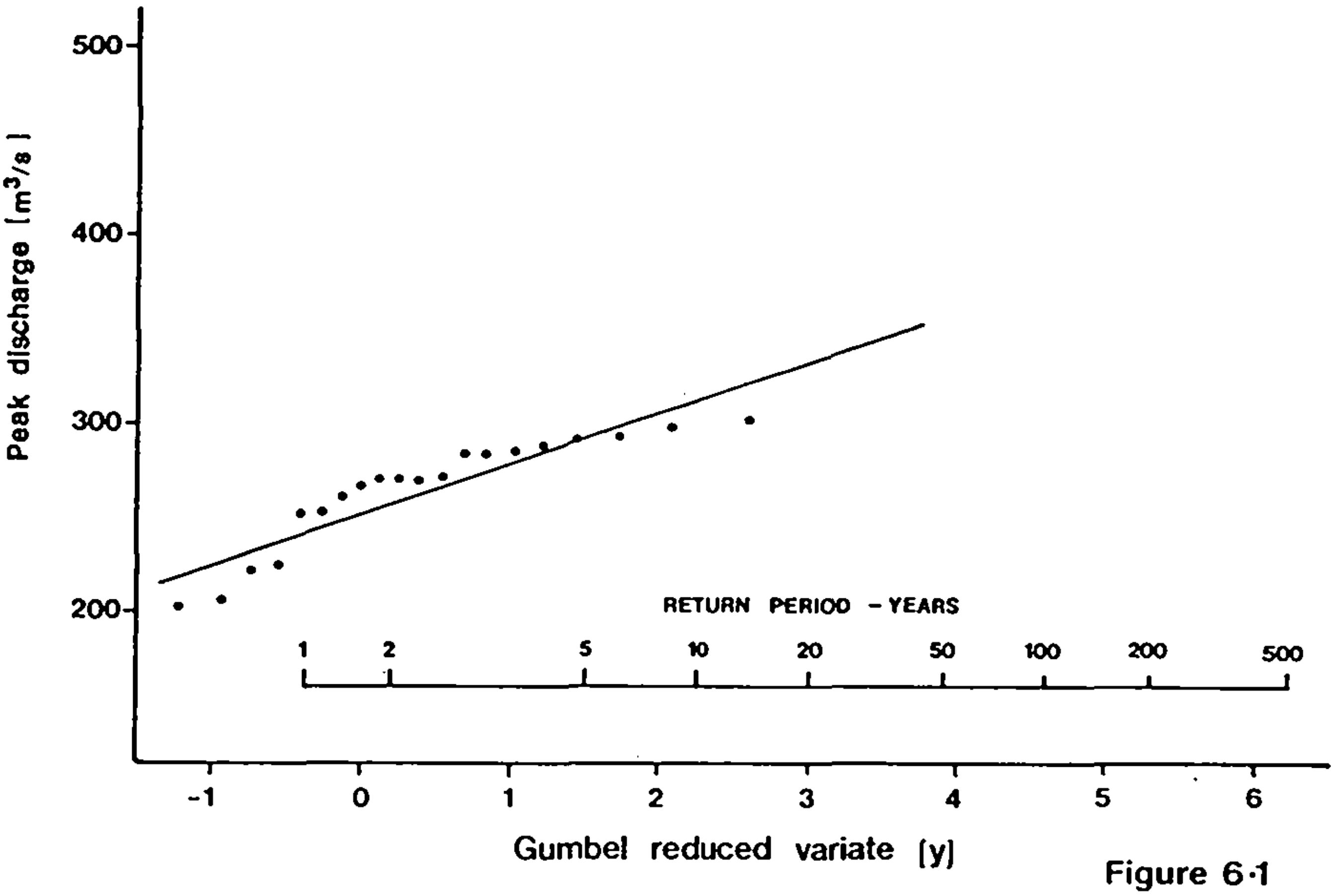


Figure 6·1

ESTIMATION OF THE T YEAR FLOOD USING REGIONAL GROWTH FACTORS

7.1 Introduction

A flood frequency curve gives a graphical representation of the relationship between the magnitude of a flood and its return period. If the graph is rescaled by dividing the flood magnitudes by the MAF, a curve of growth factors against return period is the result. For example, the all catchment average flood frequency growth curve for Java and Sumatra shown on Figure 7.1.

However, analysis in Annex F shows that growth factors in Indonesia vary not only with return period but also with the size of catchment under study. 1001 station years of data from the 92 catchments with five or more years record were used in this analysis. This allows growth factors for events up to 500 year return period to be estimated. Growth factors for the 1000 year return period are tentative and should be used with caution.

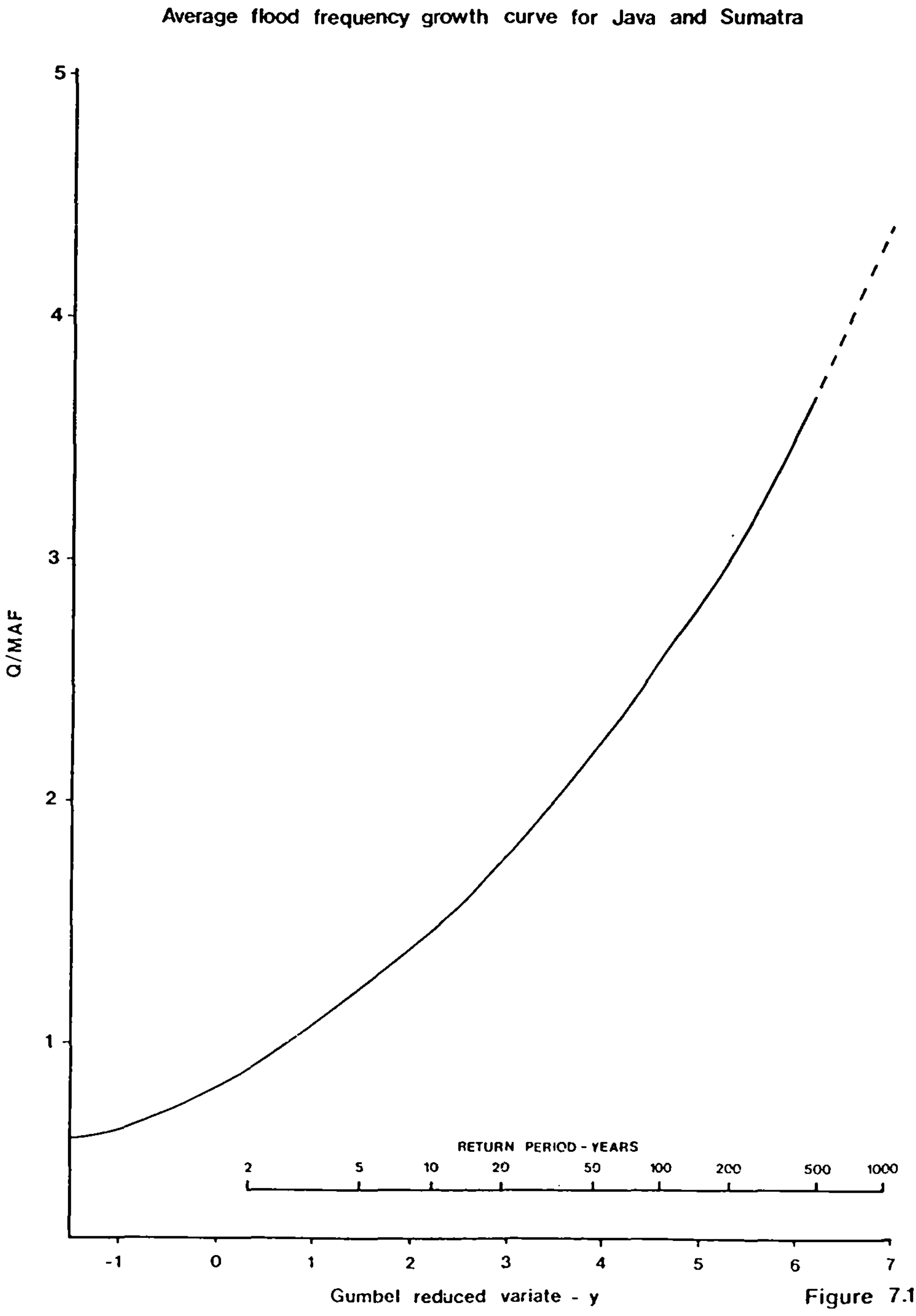
7.2 Description of method

Estimation of T year return period flood, Q_T, involves multiplying the MAF by the appropriate growth factor which is a function of T and the catchment area:

$$Q_T = GF_{(T,AREA)} \times MAF$$

It is assumed here that Chapters 3, 4, 5 and 8 have been consulted and an estimate of MAF obtained from whatever data are available.

The growth factor is obtained from Table 7.1 interpolating for both the required return period and catchment area.



Average flood frequency growth curve for Java and Sumatra
Q/MAF
5
4
3
2
1
RETURN PERIOD - YEARS
2
5
10
20
50
100
200
500
1000
-1
0
1
2
3
4
5
6
7
Gumbel reduced variate - y

Figure 7.1

Table 7.1 Table of Growth Factors GF(T,AREA)

Return Period	Reduced Variate	Catchment area (km^2)					
T	y	180 or less	300	600	900	1200	1500 or more
5	1.50	1.28	1.27	1.24	1.22	1.19	1.17
10	2.25	1.56	1.54	1.48	1.44	1.41	1.37
20	2.97	1.88	1.84	1.75	1.70	1.64	1.59
50	3.90	2.35	2.30	2.18	2.10	2.03	1.95
100	4.60	2.78	2.72	2.57	2.47	2.37	2.27
200	5.30	3.27	3.20	3.01	2.89	2.78	2.66
500	6.21	4.01	3.92	3.70	3.56	3.41	3.27
1000	6.91	4.68	4.58	4.32	4.16	4.01	3.85

Alternatively a flood frequency curve can be constructed for the required area using the MAF and growth factors for return periods given in the table; from this the flood corresponding to any return period can be read directly.

7.3 Accuracy of method

In estimating Q_T in this way errors arise from two sources: error in the MAF estimate and error in the growth factor. Chapters 3, 4 and 5 each give methods of assessing the standard deviation associated with the MAF estimate appropriate to its method of calculation. The standard error of estimate of the growth factor is hard to quantify but the UK Flood Studies Report (NERC, 1975) suggests it to be of the order of 15% at T = 10 years, 30% at T = 100 years and 50% at T = 1000 years. This approximate relationship can be summarised and used to estimate the standard deviation of the growth factor, sd(GF):

$$sd(GF) = 0.16\ (\log_{10}T) \times GF$$

where GF is the growth factor

The standard deviation of Q_T is then found from

$$sd(Q_T) = Q_T\left[\left(\frac{sd(GF)}{GF}\right)^2 + \left(\frac{sd(MAF)}{MAF}\right)^2\right]^{0.5}$$

If the regression equation is used to estimate the MAF, the term

$$\left(\frac{sd(MAF)}{MAF}\right)^2$$

may be approximated to by $(0.59)^2$ or 0.348.

7.4 Example of application

The 50 year flood is required for Batang Tembesi at Maura Inum. Twelve years of data are available at the site, sufficient to estimate the MAF by the mean of the annual maxima of the sample as described in Chapter 3.

Thus, MAF = $1164.4\ m^3s^{-1}$
sd(MAF) = $341.4\ m^3s^{-1}$

The catchment area is 1505 km^2, and from Table 7.1 the required growth factor is therefore 1.95.

$$Q_{50} = 1164.4 \times 1.95 = 2271\ m^3s^{-1}$$

To estimate the standard deviation of this estimate

$$sd(GF) = 0.53$$

$$\text{therefore } sd(Q50) \quad 2271 \sqrt{\left[\left(\frac{0.53}{1.95}\right)^2 + \left(\frac{341.4}{1164.4}\right)^2\right]}$$

$$sd(Q50) = 908\ m^3s^{-1}$$

The 50 year flood for Batang Tembesi at Maura Inum is estimated to be 2300 m^3s^{-1}, with an associated standard deviation of 900 m^3s^{-1}.

8. IMPROVING THE FLOOD ESTIMATE USING LOCAL DATA

8.1 Introduction

This chapter is concerned with making the best use of river flow data available near to the site for which a flood estimate is required. As there are a large number of river gauging stations in Sumatra and Java it is likely that some local data will be available.

If data from a station used in this study are to be used for a flood estimate then data available at the time of the study can be found in the Data Appendix. These data should then be extended to the present and checked in the light of recent developments of the rating equation.

Having obtained all relevant local data, the MAF should preferably be estimated by more than one method. The techniques which may be used, that is the annual maximum series, peaks over threshold and regression equation, have been described in Chapter 3, 4 and 5 respectively. Figure 2.1 guides users in the choice of suitable methods. A satisfactory agreement between two or more approaches would indicate that the MAF was reasonable, whereas a disagreement might indicate that special circumstances need to be considered. This point is considered in Section 8.2. Sections 8.3 to 8.6 deal with the problem of estimating the MAF where more abundant data are available at gauging stations in the vicinity of the flood estimation site. Usually it is the case that Q_T can only be estimated in one way - by using the growth factors given in Chapter 7. However, whenever possible this should be compared with an estimate using a flood frequency curve plotted from local data. A discussion on how this local data flood frequency curve may be extended to enable estimation of high return period floods is given in Section 8.8. Some general comments are made in Section 8.7 on the use of locally available flood level marks. Incorporating such historical data into the flood estimation process can only be done in a subjective way however. Each section is illustrated with one or more examples.

Flood estimation is seldom a straightforward task. Special or unusual conditions require considered judgement or even modification of the techniques described in this manual. For these reasons it is recommended that the task of flood estimation should be undertaken by an experienced engineer or hydrologist.

8.2 Using different methods at the same site

If possible the MAF should be estimated by more than one method. If there are some data at the point of interest this may be achieved using the POT method (Chapter 4) in conjunction with either the regression equation (Chapter 5) or the mean of the annual maxima series (Chapter 3). In such cases the regression equation estimate will be a poor predictor of the MAF compared to the annual maximum or POT methods based on real data.

Should there be a large disagreement between the different flood estimates, the calculations should be checked carefully for arithmetic mistakes. If the disagreement persists, the choice of an appropriate estimate of the MAF is a matter of engineering judgement. Advice on how such a choice may be made is given in general terms in later sections of this chapter.

If there is reasonable agreement between the different methods used, this may give added confidence in the estimate. How though is this added confidence reflected in the quoted value for the MAF and its error? As explained in the relevant chapters for each method of computing the MAF, an estimation error or standard deviation is associated with the calculated MAF and it has been shown how this may be found in each case. The standard deviation or standard error enables users to assess how the MAF obtained by any method might relate to the "true" long term mean annual flood at any site. At several points in the report it has been stated that using normal probability theory, the estimated MAF would be expected to lie within the range of plus or minus one standard deviation of the long term mean with a probability of 68 per cent. Similarly there is a 95 per cent probability of the long term mean annual flood being within the range of plus or minus 1.96 x the standard deviation of the estimated

MAF. The large estimation errors associated with these methods show how imprecise estimation of rare floods can be where only short periods of flow data are available or from the regression equation for the MAF of Chapter 5.

In cases where the MAF is estimated by more than one method each estimate will have a standard deviation associated with it. Normally one of the methods is most suitable for a particular application and this is the estimate that should be used. If other estimates agree with it then the various assumptions made in obtaining the estimates appear validated; if not the methods and their assumptions should be reviewed to try to explain the discrepancies. Example 2 shows how POT and regression equation estimates may be reconciled for one particular case. By either getting good agreement between methods or finding reasons for the differences confidence in the flood estimate is enhanced, albeit it in a rather intangible way.

It would normally be imprudent of an engineer to assume that because the MAF was an imprecise estimate of the long term mean annual flood one or more standard deviations should be added to the MAF estimate to allow for this imprecision. Such conservatism could increase the return period of the flood estimate dramatically without the user being aware of the fact. Given that the standard error of the regression estimate of the MAF is plus 59 per cent, additions of this error to the MAF is equivalent to the difference between the 10 year flood and the 60 year flood for a catchment of 100 km^2.

It is normally appropriate to accept the MAF estimated by the methods of this report for design purposes since this is the best available central estimate of the true mean annual flood. However, all available local flow data should be used to check and refine this central estimate as shown in later sections of this current chapter.

<u>Example 1</u>

Problem description: An estimate of the 100 year return period flood is required on the Krueng Aceh at Kampung Darang.

Three techniques are used to estimate the MAF and their results compared.

Data abstraction (regression equation)	Catchment characteristics were abstracted for this catchment as described in Annex C:

AREA = 1068 km^2

APBAR = 86 mm

SIMS = 21 m km^{-1}

LAKE = 0

Data abstraction (POT)	A threshold of 4.5 m (to the new station datum) or 266 m^3s^{-1} was chosen for the POT analysis and the following peaks abstracted:

450	350	369	282	
300	287	434	324	m^3s^{-1} (in 4 complete
309	313	427	359	years)
420	311	337		

Additional flood peaks in incomplete years are;
279 748 m^3s^{-1}

(Although the hydrological year for the annual maximum series analysis starts on 1st August, the POT method can take data from any complete 12 month period, regardless of the starting month. This explains why the largest flood on record at the station, 748 m^3s^{-1}, appears in the incomplete years here but in the complete years below).

Data abstraction (Annual maximum series)	: There are only four complete hydrological years of data at this station and the annual maxima are:

1976/1977	450 m^3s^{-1}	1978/1979	748 m^3s^{-1}*
1977/1978	434 m^3s^{-1}	1980/1981	337 m^3s^{-1}

*The original station was destroyed after this exceptional flood in August 1978. This year's data may be considered complete as it is known that the August peak was not exceeded during the same hydrological year.

MAF (catchment characteristics)

Using data from above, the AREA exponent, V, is calculated thus:

$$V = 1.02 - 0.0275 \log_{10} AREA$$

$$V = 0.937$$

MAF is estimated from the regression equation given in Chapter 5:

$$MAF = 8.00 \times 10^{-6} \times AREA^{V} \times APBAR^{2.445} \times SIMS^{0.117} \times (1 + LAKE)^{-0.85}$$

$$\underline{MAF = 422\ m^3s^{-1}}$$

The standard error of estimate of MAF is 422 x 1.59 to 422/1.59 (671 to 265 m^3s^{-1}).

MAF (POT Analysis)

Using the method described in Chapter 4, the mean exceedence, β, is calculated thus:

$$\beta = \frac{1}{M} \sum_{i=1}^{M} (q_i - q_o)$$

$$\beta = 104\ m^3\ s^{-1} \text{ (using all 17 floods)}$$

The average number of exceedences per year, λ, is calculated

$$\lambda = M/N$$

Complete years must be used in this calculation and there were 15 floods above the threshold in 4 complete years:

$$\lambda = 15/4 = 3.75$$

MAF may now be calculated

$$MAF = q_o + \beta(0.5772 + \ln \lambda)$$
$$MAF = 266 + 104(0.5772 + \ln 3.75)$$
$$\underline{MAF = 463 \ m^3 \ s^{-1}} \text{ (standard deviation } 57 \ m^3 s^{-1})$$

MAF (Annual maximum series)

The estimation of MAF from an annual maximum series of less than 5 years in length is not recommended. It is calculated here merely as a check on other methods. From the 4 years which are available.

$$\underline{MAF = 492 \ m^3 \ s^{-1}}$$

(standard deviation 178 m^3 s^{-1})

Discussion

The three estimates of MAF obtained above are in satisfactorily close agreement. However the preferred estimate of MAF must be that obtained by the POT method since this is most suitable for the short length of record at the station. This is confirmed by the low standard deviation of this method (57 m^3s^{-1}) relative to both the annual maximum series method (178 m^3s^{-1}) and the regression equation error (MAF + 249 m^3s^{-1} to MAF - 157 m^3s^{-1}). However as the MAF from the annual maximum series is slightly higher than that from the POT method, it is wise to round up the POT estimate of MAF to say 470 m^3/s. The regression equation estimate is acceptably close but it is always better to use flow data with the POT or annual maximum series whenever possible.

Solution

Q_{100} for the location is obtained by multiplying MAF by the 100 year return period growth factor of 2.41 (AREA 1068 km^2). (Table 7.1).

$$Q_{100} = 470 \times 2.41$$
$$= 1133 \ m^3s^{-1}$$

From Section 7.3 the standard deviation of the 100 year return period growth factor is 30%. This may be combined with the standard deviation of the POT estimate of the MAF to give the standard deviation of Q_{100}:

$$sd(Q_{100}) = Q_{100} \left[\left(\frac{sd(GF)}{GF}\right)^2 + \left(\frac{sd(MAF)}{MAF}\right)^2\right]^{0.5}$$

$$= 1130 \left[\left(\frac{0.723}{2.41}\right)^2 + \left(\frac{57}{470}\right)^2\right]^{0.5}$$

$$= 366 \ m^3 \ s^{-1}$$

The estimate of the 100 year return period flood is therefore 1130 m^3s^{-1} with a standard deviation of 370 m^3s^{-1}.

Example 2

Problem Description: An estimate of the 50 year return period flood is required for the Ciliwung at Kebon Baru.

Only two complete years of flow data are available for the site. This is too short to attempt an estimation of the MAF from the mean of the annual series. The MAF can, however, be estimated by both the POT method and the regression equation.

Data (catchment characteristics) The catchments characteristics for this river basin were obtained using the procedures described in Annex C:

AREA = 333 km^2
APBAR = 103 mm
SIMS = 34 m km^{-1}
LAKE = 0

Data (POT analysis):	Only two complete years of data are available and they contain the following peaks over a threshold of 90 m^3s^{-1}.

173.4	106.8	
97.7	131.9	m^3s^{-1}
92.0	93.3	
105.1	91.7	

Four additional peaks were recorded in incomplete years:

209.4	166.9	m^3s^{-1}
97.4	95.0	

MAF (regression equation)	Using data from above, the regression equation is used to estimate the MAF as described in Chapter 5.

MAF = 239 $m^3\ s^{-1}$ (standard error of estimate 380 to 150 m^3s^{-1})

MAF (POT)	Using the POT method described in Chapter 4, with the data given above the MAF is estimated as:

MAF = 152 m^3s^{-1} (standard deviation 25 m^3s^{-1})

Discussion	Two years is also really too short a record for the POT method, however it is preferable to make use of even this short record as a check on MAF estimated from the regression equation. The two methods do give noticeably different results; 239 $m^3\ s^{-1}$ from the regression and 152 $m^3\ s^{-1}$ from POT. Why should this be?

One possible explanation can be found by studying the catchment shape of the Ciliwung basin on Figure 1.1 (Catchment number 5). The catchment above Bogor may be described as typical, but thereafter it is very long and thin. As the mountainous catchment around Bogor is the main flood producing region of the catchment, the long river reach across the coastal plain towards Jakarta will signficantly attenuate floods produced in the upper catchment. Furthermore the lower catchment may not produce a large flood runoff because of its narrowness. One might expect, therefore, that the MAF produced by the regression equation to be too high - which is indeed the case.*
A tentative estimate of MAF, incorporating a factor of safety for the shortness of the record would be 200 m^3s^{-1}

From Table 7.1 the 50 year growth factor for a catchment of 323 km^2 is 2.29 giving the following estimate of Q_{50}:

$$Q_{50} = 2.29 \times 200$$

$$Q_{50} = 458 \ m^3 \ s^{-1} \text{ (standard deviation } 145 \ m^3s^{-1})$$

*Incidentally a shape factor indexing the narrowness of the catchment was consider in the regression analysis (Annex C) but not found significant when applied to the whole data set.

8.3 Using the observed MAF from an adjacent long term station

The method described in this and the following two sections uses a similar approach to adjust the MAF. If more than one of these three approaches is possible at a site, all should be tried. Engineering judgement should then be used to assess the importance of each MAF estimate in order to produce an overall balanced estimate of the MAF.

The technique described in this section is useful where a flood estimate is required at a station, A, where the period of record is short and there is a station, B, with a longer record nearby. MAF when estimated from a short period of record may be higher or lower than the true long term mean because the record may come from a time when floods are higher or lower than normal due to short term local climatic variations. This sampling error may be reduced if it assumed that the same period in the history of the station with the longer record was similarly wet or dry. For this to be true the catchments should be in close proximity and share similar climatic catchment characteristics (AAR and APBAR). This assumption may be checked by plotting the annual maxima from the common years of operation against each other to establish the degree of correlation. The adjusted MAF is calculated thus:

$$MAF_A = MAF'_A \times \frac{MAB_B}{MAF'_B}$$

where

MAF_A = Adjusted MAF for station A

MAF'_A = MAF from the record at station A (Unadjusted)

MAF_B = MAF from the entire period of operation of station B

MAF'_B = MAF from station B during period station A was operational

Example 3

Description

An estimate of the MAF is required on the Batang Pasaman at Air Gadang. Six years of data are available at Air Gadang. However the nearby station at Silaping on the Batang Batahan has a 12 year record

Data

Year	Batung Pasaman Air Gadang (342) $m^3\ s^{-1}$	Batung Batahan Silaping (343) $m^3\ s^{-1}$
39-40		139.3
40-41		247.1
41-42		388.3
72-73		317.2
73-74		303.8
74-75		466.3
75-76	898.5	170.2
76-77	1147.9	466.3
77-78	970.9	478.7
78-79	694.4	399.5
79-80	1036.1	508.0
80-81	1141.0	430.0

Catchment characteristics

AAR	3440 mm	3100 mm
APBAR	103 mm	118 mm
AREA	1267 km^2	304 km^2

Discussion

Although there is a considerable difference in catchment area, the two catchments are similar climatically, in close proximity and drain in the same direction. Although the correlation between the annual maxima of these two stations, over the common period of record (75-76 to 80-81), is poor (correlation coefficient = 0.38), the example is continued as an illustration. In the notation described above:

$$MAF'_{342} = 981.5 \quad m^3s^{-1}$$

$$MAF_{343} = 359.6 \quad m^3s^{-1}$$

$$MAF'_{343} = 408.8 \quad m^{3-1}$$

The adjusted MAF at station 342 is calculated thus:

$$MAF_{342} = 981.5 \times \frac{359.6}{408.8}$$

$$MAF_{342} = 863 \quad m^3\ s^{-1}$$

Using this technique the revised estimate of the MAF is 863 m^3s^{-1}

8.4 Using flood records from elsewhere in the catchment

Sometimes it will be necessary to make a flood estimate at a point A on a river which is some distance upstream or downstream of an established gauging station B. Provided that the differences in catchment area are relatively small, these data may be used to assist in flood estimation at the point of interest. It is suggested that this technique only be used if the difference in area between the two catchments is less than 50%.

The MAF at the point of interest, MAF_A, is calculated thus:

$$MAF_A = MAF_A^R \times \frac{MAF_B}{MAF_B^R}.$$

where

MAF_A^R = regression estimate of MAF at A

MAF_B^R = regression estimate of MAF at gauging station B

MAF_B = MAF at station B from data

Example 4

Description

The MAF is required on the Krueng Jambo Aye at Rampah. For the purposes of this example no flow data are assumed to exist there. There are, however, 8 years of flow data downstream at Lhoknibong. If the difference in catchment areas is not too great, these data may be used to refine the MAF.

	Krueng Jambo Aye	
	Rampah (117)	Lhoknibong (118)
AREA (km^2)	4061	4403
APBAR (mm)	65	67
SIMS ($m\ km^{-1}$)	10.8	8.35
LAKE	0	0

The MAF at Lhoknibong, estimated from the annual series of 8 years, is 932 $m^3\ s^{-1}$.

Analysis

Before applying the technique described in this section, we must check that the two catchment areas are within the 50% of each other. The difference in area is in fact only 8.4%. The technique may therefore be applied.

Using the catchment characteristics given above, the regression equation of Chapter 5 is used to estimate the MAF at both stations

MAF_{117}^R = 598 m^3/s (Regression estimate for station 117)

MAF_{118}^R = 672 m^3/s (Regression estimate for station 118)

From the annual maximum series at Lhoknibong

$MAF_{118} = 932 \ m^3s^{-1}$

The regression estimate of MAF at Rampah is modified thus.

$$MAF_{117} = MAF^R_{117} \times \frac{MAF_{118}}{MAF^R_{118}}$$

$$MAF_{117} = 598 \times \frac{932}{672}$$

$$MAF_{117} = 829 \ m^3 \ s^{-1}$$

The revised estimate of the MAF is therefore $830 \ m^3s^{-1}$

8.5 Using flood records from adjoining catchments

If there are no suitable gauging stations within the catchment of interest, records from a station on an adjoining river basin may be used to assist with estimation of MAF.

It is suggested that this technique only be used when the two catchments have broadly similar characteristics and in particular the two catchment areas differ by no more than 50%.

The MAF at the point of interest, MAF_A, is calculated thus:

$$MAF_A = MAF^R_A \times \frac{MAF_B}{MAF^R_B}$$

where

MAF^R_A = Regression estimate of MAF at A

MAF^R_B = Regression estimate of MAF at gauging station B

MAF_B = MAF at station B from data

Example 5

Description

Estimate MAF on the Batang Air Dingin at Lubuk Minturun for which we assume there is no flow data (for the purpose of this example).

The adjoining catchment of Batang Kuranji at Gunung Nago has 8 years of data which may, if the catchments are similar, be used to assist with the estimation of the MAF at Lubuk Minturun.

The following catchment characteristics were abstracted for the two sites using the procedures described in Annex C.

	Batang Air Dingin Lubuk Minturun (321)	Batang Kuvanji Gunung Nago (314)
AREA (km^2)	114	122
APBAR (mm)	147	147
SIMS ($m\ km^{-1}$)	75.6	70.3
LAKE	0	0

The MAF at Gunung Nago, estimated from the annual series of 8 years, is 415 $m^3\ s^{-1}$.

Analysis

The two catchments have similar catchment characteristics and are in close proximity. The difference in area is 7%. The MAF adjustment technique may therefore be applied.

The catchment characteristics given above enable MAF to be estimated at both stations by the regression equation of Chapter 5.

$$MAF^R_{314} = 266 \text{ m}^3 \text{ s}^{-1} \quad \text{(Regression estimate for station 314)}$$

$$MAF^R_{321} = 253 \text{ m}^3 \text{ s}^{-1} \quad \text{(Regression estimate for station 321)}$$

From the annual maximum series at Gunung Nago:

$$MAF_{314} = 415 \text{ m}^3 \text{ s}^{-1}$$

The regression estimate of MAF at Lubuk Minturun is adjusted thus:-

$$MAF_{321} = MAF^R_{321} \times \frac{MAF_{314}}{MAF^R_{314}}$$

$$MAF_{321} = 253 \times \frac{415}{266}$$

$$MAF_{321} = 395 \text{ m}^3 \text{ s}^{-1}$$

8.6 Using staff gauge data

It is possible that staff gauge readings have been taken over a period of time close to the site of interest and none or very few current meter gaugings exist. If it is not possible to develop a rating as suggested in Annex B, two techiques described below enable these data to be used in a flood analysis. Both techniques require five or more years data. The annual maximum series should be abstracted and will, of course, be in stages (m) not flows (m^3 s^{-1}).

The first technique involving the use of stage data requires the abstraction of the median stage and its conversion to discharge using a flow resistance formula such as the Manning or Chezy equation. The median of the annual maximum stage series when converted to a discharge is directly equivalent to the median of the annual maximum flow series Q_{med}. Analysis of data from stations used in this study showed that, on average

$$MAF = 1.06\ Q_{med}$$

This relationship may be used to derive the MAF after Q_{med} has been determined.

The second method of using data from an unrated section is useful where the station has a long record of staff gauge data. Here it is possible to plot a flood stage frequency curve and this is done in the same way as for a flood discharge frequency curve which is described in Chapter 6. There is a direct correspondence between these two curves and the T year return period flood stage in metres is equivalent to the T year return period flood flow in cubic metres per second. This stage can be converted to flow by either the Manning or Chezy formula. However in flood design it is often the maximum level of the T year return period flood which is important. In this case conversion to discharge is unnecessary.

<u>Example 6</u> Perempuan Cantik at Banyak Masalah.

Although flood peaks have been recorded by an observer for 9 years the station is unrated. A survey of the channel was undertaken to estimate the flow at the median annual stage.

The following peak stages have been recorded by the observer:

Data	Year	Peak stage m	Year	Peak stage m
	72-73	2.43	77-78	4.01
	73-74	3.07	78-79	2.44
	74-75	2.78	79-80	3.48
	75-76	3.58	80-81	2.89
	76-77	2.88		

From a survey of the river cross-section the following information was obtained for the median stage of 2.89 m:-

Cross sectional area of flow, A $= 103 \text{ m}^2$

Wetted perimeter, P $= 43$ m

Water surface slope, S $= 0.0107 \text{ m m}^{-1}$

Estimated Manning's n $= 0.04$

Firstly the hydraulic radius, R is calculated thus

$$R = A/P = 103/43 = 2.4 \text{ m}$$

The average velocity of flow, v, may now be estimated using Manning's equation:

$$v = \frac{1}{n} R^{2/3} S^{1/2}$$

$$v = \frac{2.4^{2/3} \times 0.0107^{1/2}}{0.04}$$

$$v = 4.64 \text{ m s}^{-1}$$

Median flow discharge, Q_{med}, is obtained by multiplying velocity by cross sectional area of flow at 2.89 m stage:.

$$Q_{med} = V \times A$$

$$Q_{med} = 4.64 \times 103$$

$$Q_{med} = 478 \ m^3 \ s^{-1}$$

Using the relationship given above between the MAF and Q_{med}, the MAF is then calculated:

$$MAF = 1.06 \times 478$$

$$\underline{MAF = 507 \ m^3 \ s^{-1}}$$

8.7 Using flood marks

Levels of historic floods may sometimes be found marked in the vicinity of a river. These typically occur as a line inscribed on a bridge pier or a 'tide' mark on walls close to the river. In the course of this study, isolated flood marks were found in the vicinity of several gauging stations. The existence of such flood marks is often known to the gauging station observer or local inhabitants.

Occasionally in Indonesia these exceptional floods either submerge the gauging station resulting in the float reaching its point of maximum travel, or in the station being washed away completely. In either case the peak flood stage is not recorded and may be substituted by levelling in flood marks to the station datum. It may only be possible to get a crude estimate of peak discharge at this high level due to excessive rating curve extrapolation. However it is better to include this albeit doubtful flood record in the annual maximum flood series or POT series for the station than omit it entirely.

If the extraordinary flood occurs outside the period of record of the gauging station the problem becomes more complex. Section 2.8 of the United Kingdom Flood Studies Report (NERC, 1975) discusses the problem in some detail and it is considered further in Annex E. However to make this complicated analysis worthwhile a series of flood marks above some datum and with dates are required. Such a situation is unlikely to occur in Indonesia and was not noted during the course of this study. If this exceptional flood is quoted by local people as 'the biggest flood in living memory', a very crude estimate of the return period could be obtained by assuming 'living memory' was between 30 and 60 years. This is equivalent, using the Gringorten formula given in Chapter 7, to a return period of about 50 to 100 years. At best this can be only a very crude check on the flood estimate obtained by other means. What it does show is that the river in question is capable of producing floods of that magnitude. Thus if the peak flow obtained from this extraordinary flood was greater than say Q_{500} estimated by the regression equation and growth curve it might indicate that the regression equation estimate was rather low.

8.8 Extension of a flood frequency curve

In those cases where sufficient data exist to plot a flood frequency curve it is often the case that this curve cannot be extended to the required return periods. For very large return periods of 500 years or more it is best to use a flood frequency curve based on the growth factors given in Chapter 7. However for intermediate return periods and to avoid a sudden jump from the plotted curve to that based on growth factors the following procedure is recommended.

Use the method of Chapter 6 to plot the flood frequency curve based on the available data and decide on the limiting return period up to which the curve may be used; call this L years and the corresponding flood Q_L. From the table of growth factors (Table 7.1), interpolate the values appropriate to the catchment area: denote these chosen growth factors by GF(T), where GF(T) is the growth factor corresponding to the return period T. If GF(L) represents the value of GF(T) for T=L, then the flood Q_T is given by

$$Q_T = Q_L \times \frac{GF(T)}{GF(L)}$$

for return periods greater than L up to about 10 x L years. For floods of return periods greater than 10 x L (or 500 years, whichever is smaller) estimate Q_T using the mean annual flood derived from the recorded flow data and growth factors of Table 7.1. A smooth curve should be drawn to link the three line segments. This curve can then be used to estimate the magnitude of the desired flood.

Example 7

Problem description: For the Citarum at Nanjung a flood frequency growth curve was constructed in Chapter 6. This was considered valid up to the 10 year flood but how should it be extended for use up to 500 year return periods.

Analysis: The mean annual flood can be estimated either from the mean of the annual maximum series or by reading the value from the flood frequency curve drawn from the local data corresponding to T of 2.33 years (Figure 6.1).From the annual maximum series

$$MAF = 270.1\ m^3\ s^{-1}$$

Q_{10} was estimated to be 312 $m^3\ s^{-1}$ in the example in Chapter 6. The catchment area for Nanjung is 1833 km^2.

The following table sets out the calculations for Q_T. In the first column are the return periods for which points will be plotted on the flood frequency curve. The growth factors corresponding to these obtained from Table 7.1 are given in the second column. The third column has the ratios GF(T)/GF(10) and the fourth column gives Q_T based on these ratios. The final column gives Q_T based on the growth factors of column two and the MAF.

Figure 8.1 shows the three portions of flood frequency curve from the available data, the GF(T)/GF(10) ratios and the growth factors from Table 7.1. The smoothed line drawn linking these segments should be used to estimate the required floods. Thus Q(100) is taken to be the upper limit of the extrapolated local data curve, 515 m^3s^{-1}. Q(500) and Q(1000) are read

Extrapolated flood frequency curve for the Citarum at Nanjung

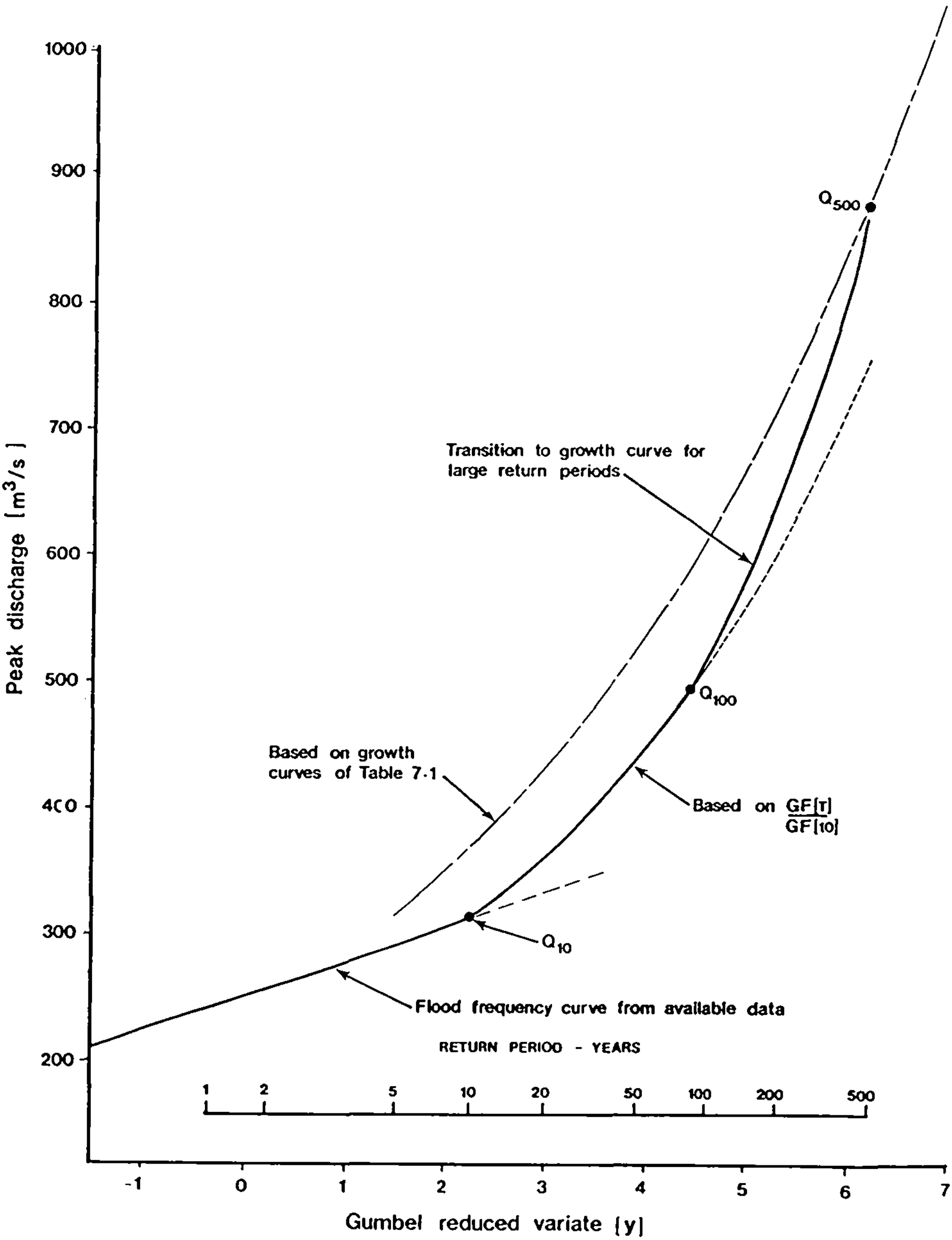


Figure 8·1

from the upper curve based on the MAF and the growth factors of Table 7.1. These are given in the final column below as 883 m^3s^{-1} and 1040 m^3s^{-1}. Floods for intermediate return periods such as Q(200) are read from the smooth transition curve and is taken to be 655 m^3s^{-1} in this case.

Return Period T	Growth Factor GF(T)	Ratio GF(T)/GF(10)	Q_T from Q_{10} and GF(T)/GF(10)	Q_T from GF(T) and MAF
2.33	1.0			
10	1.37	1.0	312.0	
20	1.59	1.16	362.0	
50	1.95	1.42	443.0	
100	2.27	1.65	515.0	(613.1)
200	2.66	(1.94)	(605.0)	(718.5)
500	3.27	(2.37)	(739.4)	883.2
1000	3.85	(2.81)	(876.7)	1040

9. COMPARISON WITH OTHER FLOOD ESTIMATION METHODS USED IN INDONESIA

9.1 Introduction

This chapter describes a comparison of different methods of flood estimation for ungauged sites currently used in Indonesia with the no data method using the regression equation described in this manual. Unfortunately no independent stations were available for a test of methods; of necessity all acceptable data were used in this study. For the first phase of the project on Java in 1981, eleven trial catchments were chosen at random before the analyses commenced to give a representative sample of catchment areas for comparison. For the second phase on Sumatra in 1982, ten catchments were similarly selected at random before starting the analyses.

The Rational method and rational type methods of Melchior, Weduwen and Hasper used in the comparison require daily rainfalls of specified return periods. For the first phase of the study on Java, the only data readily available were the maximum, the second highest and the mean annual maximum of 1 day rainfalls for all raingauges on Java (I.M.G. Met note, 1969). Data from all raingauges within each catchment were plotted to form an average rainfall growth curve using a Gumbel reduced variate and Gringorten plotting position (Chapter 6). The average of the highest and second highest daily rainfall was plotted at the position appropriate to the length of record. The mean annual maximum daily rainfall was plotted at a return period of 2.33 years. A regression line through all these data was used to estimate rainfalls of the required return periods. This procedure was not that specified by the various methods but served as the best approximation with the data available.

For the Sumatra Study in 1982, rainfall yearbooks were obtained giving details of annual maximum one day rainfalls for each of years 1951-1977 (I.M.G. Yearbooks). Raingauges on or close to each of the ten selected catchments were listed from the yearbooks, and for those with sufficient yearly data, annual maximum one day rainfalls were abstracted. Rainfall frequency growth curves were plotted for each raingauge using the same Gumbel reduced variate and Gringorten plotting position described above and average curves drawn in subjectively for each catchment. This approach should yield better estimates of the required rainfalls for flood estimation than the simplified method used for Java. Insufficient time was available to enable the earlier rainfall estimates for Java to be re-computed using the rainfall yearbook data. It is believed that any inaccuracies inherent in the simplified approach will be small.

It should be noted that most of the methods described below are variations on the Rational formula method (Section 9.2) and use the same rainfall data as input. Thus they assume that the 500 year flood is caused by the 500 year rainfall and only the rainfall areal reduction factors and runoff coefficients are changed in each method. It is interesting to note that whilst the range of growth factors for Q(500) given in Section 7.2 is from 3.27 to 4.01 with a median value of 3.7, the equivalent rainfall growth factors are rather lower. For Java the range over the eleven catchments considered was from 2.06 to 3.01 and the median is 2.48. Sumatra has a range of from 1.95 to 4.2, the latter figure being something of an outlier and the median for Sumatra is 2.23.

Flood frequency growth factors would normally be higher than the rainfall growth factors that effectively produce the floods because many of the factors controlling the conversion of rainfall to flood runoff on a catchment are relatively constant. Interception losses of rainfall on vegetation and on the soil surface are largely constant and the rate of infiltration of rainfall into the soil also varies only slightly from storm to storm. Thus the proportion of storm rainfall remaining for flood generation after these relatively constant losses have been taken off increases as storm magnitude increases for higher return periods and flood growth factors increase more rapidly than the rainfall growth factors. That seemingly small rainfalls might produce more extreme floods is not entirely suprising since many factors control the conversion of rainfall to flood runoff. The effect of antecedent catchment conditions are of great importance in this respect; the flood produced by a storm will be greatly reduced if it follows a long dry period or enhanced if it comes after a period of unusually wet weather. Such considerations illustrate the weakness of methods that assume that runoff return period equals rainfall return period and hence the advantage gained by estimation methods based on flood statistics. The following sections give a brief description of each flood estimation method. For a more detailed explanation the reader is referred to the sources quoted.

9.2 Rational method (Muhadi, 1976)

This version of the standard rational method which is used in Indonesia is adapted from Japanese practice. The principle of the method is to determine the flood peak Q_T (in $m^3 s^{-1}$) of return period T years from the formula

$$Q_T = \frac{C\ I(T)}{3.6} \text{ AREA}$$

where

C is a coefficient varying with the nature of the terrain which was taken from a table in Muhadi's paper. I(T) is a rainfall intensity corresponding to the T-year return period rainfall for a duration equal to the time of concentration of the catchment. Empirical formulae are available which relate I(T) to the stream length, slope and the 1-day rainfall of T year return period, R(T); these formulae are derived from Japanese data.

The Rational method is usually applied only to small catchments, although no information is available on the range of application for the version used here. An arbitrary upper limit of 2000 km^2 was used in this study, although this may be too large for sensible application of the Rational method.

9.3 Weduwen Method (Muhadi, 1976 and I.E.C., 1977)

This method is essentially a modification of the Rational method, and was developed for conditions near Jakarta. The flood peak Q_T (in m^3s^{-1}) of return period T-years is determined from;

$$Q_T = \alpha\beta q \text{ AREA } \frac{R(T)}{240}$$

where

R(T) is the 1-day rainfall of return period T years (mm)

and $\alpha\beta q$ is a combined areal reduction and runoff coefficient, determined graphically as a function of catchment area and slope from a figure in I.E.C. 1977.

The method is considered applicable to catchments with an area of less than 100 km^2.

9.4 Melchior method (Muhadi, 1976 and I.E.C., 1977)

This method is also developed from the Rational method, and is suitable for use on relatively large catchments. The flood peak Q_T (in m^3s^{-1}) of return period T years is determined from;

$$Q_T \quad \alpha\beta q \text{ AREA } \frac{R(T)}{200}$$

where

R(T) is the 1-day rainfall of return period T years (mm)

is an areal reduction factor dependent on the time of concentration of the catchment, t_c, (the relation is available in tabular form and t_c is determined from catchment length and slope). In fact β is usually determined graphically from catchment area using a figure given in Muhadi 1976.

is a coefficient determined from t_c and the area of an "equivalent ellipse" for the catchment by a trial-and-error graphical method.

is a runoff coefficient which is arbitrarily selected from the range $0.42 < \alpha < 0.75$. Melchior suggested an average value of 0.52 but current practice is to use higher values in the range 0.6 to 0.75.

The arbitrary nature of α means that the estimates for Q_T determined by this method must be regarded as approximate. The method is considered applicable to catchments with areas greater than 100 km^2, and the graphs required are available only for equivalent ellipses smaller than 10,000 km^2, catchment areas smaller than 7200 km^2, and times of concentration less than 20 hours. A major constraint on the method is that it can only be applied to catchments having a mainstream length of less than about 150 km. Because of this many long narrow catchments having areas of only 4000 km^2 or so may well have equivalent ellipse areas outside the range of the graphs, eg. catchments 45, 118, 201 and 707.

9.5 Hasper method (Muhadi, 1976)

This is another modified Rational method, which is very similar in concept to the Melchior method. The flood peak Q_T (in m^3s^{-1}) of return period T years is determined from;

$$Q_T = \alpha\beta q \text{ AREA } R(T)$$

where

R(T) is the 1-day rainfall of return period T years (mm)

α is a runoff coefficient determined as a function of AREA

β is an areal reduction factor determined as a function of the time of concentration of the catchment, t_c, (which in turn is determined from catchment length and slope) and the area of an "equivalent ellipse"

q is a discharge coefficient determined as a function of the time of concentration. Different functional forms are used for different ranges of t_c.

The method is applicable to catchments whose times of concentration are less than 30 hours.

9.6 Peterson method (I.E.C. 1977)

This method was developed as part of the Sederhana Irrigation Reclamation and Land Development Project for application throughout Indonesia. Multiple regression equations were obtained from which the flood peak Q_T (in m^3s^{-1}) for return periods T = 2,5,10,25 years can be estimated. These equations are:

$$Q_2 = 0.00000143\ AREA^{0.964}AAR^{1.69}$$
$$Q_5 = 0.00000174\ AREA^{0.950}AAR^{1.72}$$
$$Q_{10} = 0.00000189\ AREA^{0.942}AAR^{1.73}$$
$$Q_{25} = 0.00000159\ AREA^{0.944}AAR^{1.77}$$

where

AAR is the catchment average annual rainfall (mm). For the purposes of the comparison study, Q_2 was taken as an estimate of the mean annual flood.

These equations were derived from catchments with areas in the range 0.43 km^2 to 414 km^2 and with mean annual average precipitation in the range 1882 mm to 5226 mm. They should not be used for catchments in which the values are outside these ranges.

9.7 Comparison of results

For each of the eleven catchments on Java and ten catchments on Sumatra, flood estimates were derived using the methods just described in this chapter and also from the regression equation and flood frequency growth curve developed during this current study. Tables 9.1 and 9.2 show the results of this comparison for the mean annual flood, MAF, and for return periods of 10 and 500 years.

For each catchment, an estimate of the true MAF is available from the observed flow records at the gauging station. For all catchments except number 433 in Sumatra an estimate of Q_{10} is also available from the observed flow records although the estimate may often lack precision due to the short flow records available. A measure of the success with which the observed MAF or Q_{10} is predicted by each method is provided by the root mean square error (RMS error), where:

$$\text{RMS error } (\%) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\text{predicted MAF}_i - \text{observed MAF}_i}{\text{observed MAF}_i}\right)^2} \times 100\%$$

where n is the number of observations. A low value of RMS error indicates good agreement between the prediction method and the observed values and vice versa. It should be noted that the RMS error places greater emphasis on overprediction compared with underprediction since an underprediction can only be up to 100% less than the observed value, whereas an overprediction may be several hundred per cent greater.

Two RMS errors are given on Tables 9.1 to 9.4, the first being for all catchments to which each method was applied. Thus for Java in Table 9.1, the Weduwen method was only applicable on three catchments, numbers 25, 27 and 29 spaces. The RMS error has been computed for just these three catchments given as the first RMS error in column (a) of the table. The Indonesian Flood Studies Report (FSR) method of this report was applicable to all eleven catchments used in the comparison. Hence the RMS error given in column (a) of the table for this method is for eleven catchments and is not directly comparable with that for the Weduwen method for example. In order to compare the method of this report with other methods directly, a second RMS error was computed for the Indonesian FSR method using only those catchments common to each method in turn. Thus as a comparison with the Weduwen method, the RMS error has been computed for the Indonesian FSR method for the three common catchments, 25, 27 and 29 and this is given in column (b) of Table 9.1. These second RMS error estimates should provide the best comparison of the flood estimation methods presented in this report and others commonly used in Indonesia. Tables 9.3 and 9.4 summarise the comparisons for the MAF and Q_{10} respectively by giving the RMS errors for each method for Java and Sumatra independently and also combined.

It is apparent that the methods of this current report give consistently better estimates of MAF and Q_{10} than any other method, with the exception of the Peterson method which is a similar regression model. However, the Peterson model is only applicable for catchment areas of up to 414 km^2 and for return periods up to 25 years. This report provides a more comprehensive set of methods applicable for catchment areas up to 20,000 km^2 and for return periods up to 1000 years. Of the other methods, the Melchior approach for catchments greater than 100 km^2 produces reasonable results on the whole as does the Hasper method despite the latter's very high RMS error for Java. This RMS error is dominated by the Hasper method's gross overprediction of MAF and Q_{10} on the two very small catchments, 27 and 29. If these are excluded, the RMS error drops to only 77.2% for Java and 76.6% overall. It seems that the method should perhaps not be applied to very small catchments.

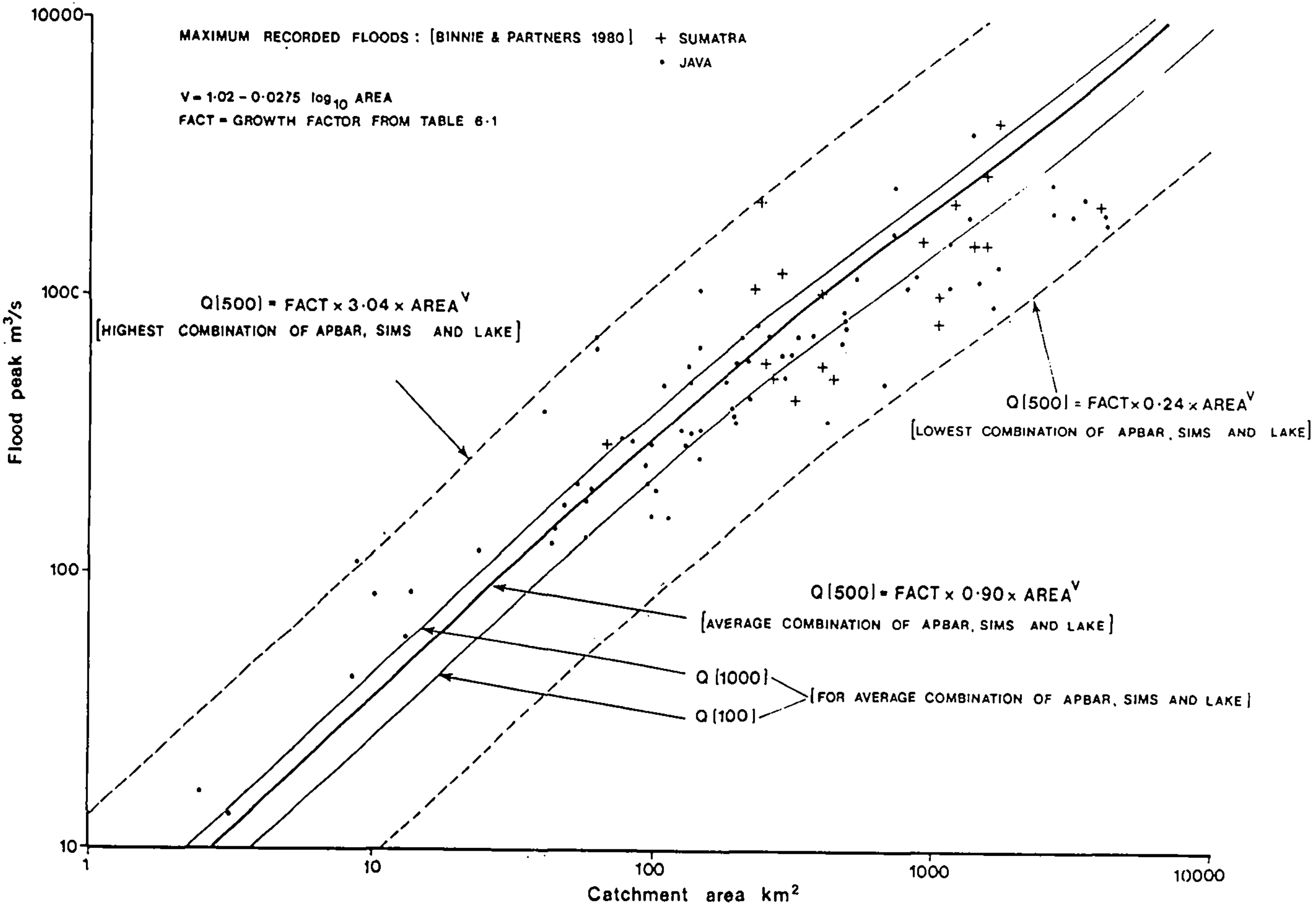
The Weduwen method does not perform particularly well and appears to consistently overestimate floods while the Rational method grossly overestimates floods. Both methods perform poorly for the very small catchments, 27 and 29. Neither of these methods seems to provide a viable alternative to the approach of this current report.

Overall, the authors believe that the range of flood estimation techniques described in this current report provide the most reliable flood estimates for Java and Sumatra. The methods may well be applicable elsewhere in Indonesia if used with care and results checked against local data as described in Chapter 8.

A second and independent assessment of the methods of this report is shown in Figure 9.1 which shows the maximum recorded floods for Java and Sumatra plotted against catchment area (Binnie and Partners, 1980). Also shown on this figure are three estimators of Q_{500} against catchment area derived from the regression equation of section 5.2 and the appropriate multiplier for Q_{500} on area given in Table 7.1. The upper line was derived assuming the highest combination of APBAR, SIMS and LAKE encountered on a catchment in this study and the lowermost line was derived assuming the lowest encountered combination of the same parameters. The central line was calculated using average values of APBAR, SIMS and LAKE.

The 500 year figure adopted here is perhaps indicative of the likely return period of the highest recorded floods shown on Figure 9.1. Many of these may in fact be much more commonplace events, having return periods of

Figure 9·1 Maximum recorded floods in Java and Sumatra and regression equation estimates



only 50 to 100 years but some at least will be very rare events. For the average combination of APBAR, SIMS and LAKE encountered in this study, the estimated Q_{100} and Q_{1000} floods are also shown on Figure 9.1. It is apparent that the variations in floods of various return periods for any particular type of catchment are significantly lower than variations of say Q_{500} between different types of catchment. The 100 year flood for a catchment having the average combination of APBAR, SIMS and LAKE may be over twice as large as the 500 or 1000 year flood for catchments with low combinations of the same characteristics.

It should be noted that there is no return period attached to the maximum floods shown in Figure 9.1 and some of the floods plotted are of doubtful accuracy. Hence some of the large scatter of points on the graph will undoubtedly be due to errors in the estimation of the magnitudes of these floods. However, as has been emphasised in this report, there are also significant errors and uncertainties in the estimation of the Q_{500} lines drawn on Figure 9.1 using the methods presented in this report. For any catchment there will be an error associated with the estimate of the MAF, whether this estimate comes from recorded flood data using the POT or annual maximum flood series, or from the regression equation as is the case for Figure 9.1. The flood frequency growth factors of Table 7.1 used to convert the MAF estimate to Q_T also have errors of estimation associated with them and consequently the plotted lines are only a best estimate of Q_{500} in each case shown. However Figure 9.1 demonstrates that the Flood Study method does produce reasonable answers over a wide range of catchment areas and types.

Table 9.1 COMPARISON OF FLOOD ESTIMATION METHODS FOR JAVA

Catchment Number: River : Name : Area (km^2):	2 Citarum Saguling 2367	7 Cilamaya Cipeundeuy 139.2	21 Cidurian Parigi 622.0	22 Kragilan Ciujung 1858.0	25 Ciletuh Cipiring 79.3	26 Cimandiri Tegal Datar 495.1	27 K Ciwadeyey Pos A 0.43	29 K Ciwadeyey Pos C 10.8	45 B. Solo Kauman 5900	46 B. Solo Bojonegoro 12429	49 Kali Asem Sancul 187.4	ROOT MEAN SQUARE ERROR (%) (a) For all catchments to which method applied	ROOT MEAN SQUARE ERROR (%) (b) FSR method; only for common catchments
(A) MAF													
Observed MAF	659.6	497.2	193.2	733.1	294.5	369.4	0.58	7.67	1477	2072	100.6		
Indonesian FSR Estimate	426.0	230.2	370.2	732.2	201.7	269.8	0.65	15.8	1264.7	1670.8	209.7	58.8	(58.8)
Weduwen	-	-	-	-	314.1	-	4.4	78.6	-	-	-	658.0	64.2
Melchior	259.5	165.4	123.2	172.1	-	229.6	-	-	-	-	287.6	92.5	64.6
Hasper	491.2	144.2	97.2	166.5	185.2	203.0	14.1	158.3	-	-	257.0	1018.3	64.5
Rational	-	387.9	542.9	805.1	392.6	845.9	21.8	236.2			923.0	1695	67.2
Peterson	-	173.1	-	-	122.8	-	0.60	13.7			191.1	66.2	73.5
(B) Q(10)													
"Observed" Q(10)	910	676	245	872	500	451	0.77	10.1	1846	2590	165		
Indonesian FSR Estimate	583.6	359.1	547.8	1003	314.7	406.7	1.01	24.6	1773	2289	327.1	68.7	(68.7)
Weduwen	-	-	-	-	463.5	-	5.8	96.7	-	-	-	622.3	87.5
Melchior	362.6	275.6	196.6	246.5	-	320.0	-	-	-	-	396.5	74.4	69.2
Hasper	686.2	216.1	155.1	238.6	273.3	282.9	16.8	198.5	-	-	354.4	933.4	75.8
Rational	-	581.5	866.1	1153	579.4	1179	28.6	310.3			1273	1658.0	79.4
Peterson	-	284.8	-	-	205.5	-	1.2	23.9			311.1	85.4	83.4
(C) Q(500)													
Indonesian FSR Estimate	830.7	541.0	807.0	1428	474.0	599.0	1.53	37.1	2466	3258	492.8		
Weduwen	-	-	-	-	817.4	-	9.0	151.2	-	-	-		
Melchior	607.1	493.1	370.2	422.5	-	534.6	-	-	-	-	654.8		
Hasper	1149	386.8	292.1	408.8	482.0	472.6	21.4	221.2	-	-	585.1		
Rational	-	1041	1631	1977	1022	1969	44.8	485.6			2101		

TABLE 9.2 COMPARISON OF FLOOD ESTIMATION METHODS FOR SUMATRA

Catchment Number:	118	201	243	316	433	522	707	803	825	827	ROOT MEAN SQUARE ERROR (%) (a) For all Catchments to which method applied	(b) FSR method; only for common Catchments
River :	Kr.Jambo Aye	S.Wampu	Bt.Ilung	Bt.Arai	Bt.Sumani	Bt.Tembesi	S.Komering	W.Sekampung	W.Pengubuan	W.Teruwan		
Name :	Lhoknibong	Stabat	Hajoran	Kandang Empat	Bd.Pundung	Muara Inum	Martapura	Pujorahayu	Gedong Harta	Gunung Batin		
Area (km^2) :	6402.7	3788.9	180.8	110.0	459.1	1506.3	4463.6	1687.9	111.3	528.5		
(A) MAF												
Observed MAF	931.6	664.6	182.4	89.1	270.0	1164	1302	463.4	67.5	108.0		
Indonesian FSR Estimate	668.8	1220	131.9	128.1	463.0	594.9	1039.0	411.6	60.0	178.4	48.7	(48.7)
Wadiman	-	-	-	338.3	-	-	-	-	215.9	-	251.6	34.9
Malchior	-	-	171.8	287.8	692.3	218.1	-	222.8	82.0	366.2	142.0	45.8
Hasper	-	985	114.6	242.9	380.1	162.9	-	187.3	68.9	97.6	76.0	52.0
Rational		-	434.5	719.7	1983.2	1602		1570	222.3	151.4	385.2	45.8
Peterson		-	77.9	126.5	-	-		-	70.3	-	41.1	30.6
(B) Q(10)												
"Observed" Q(10)	1286	750	243	160	*	1513	1927	709	108	170		
Indonesian FSR	916.2	1672	205.8	199.8	699.0	815.0	1423	563.9	90.6	267	51.1	(51.1)
Wadiman	-	-	-	451.0	-	-	-	-	297.7	-	178.8	20.0
Malchior	-	-	231.6	383.8	919.8	297.8	-	306.9	113.1	506.1	106.7	33.8
Hasper	-	1315	154.5	323.9	505.3	222.4	-	258.0	94.9	134.8	64.8	56.0
Rational		-	612.8	959.7	2635	2187		2162	306.5	209.3	242.0	33.8
Peterson		-	124.3	208.9	-	-		-	114.4	-	33.4	18.5
(C) Q(500)												
Indonesian FSR	2187	3989	528.9	513.7	1764	1945	3398	1346	240.6	669.5		
Wadiman	-	-	-	1015	-	-	-	-	443.1	-		
Malchior	-	-	369.6	863.5	2077	502.6	-	434.2	168.3	752.5		
Hasper	-	2127	246.7	728.3	1140.8	375.5		365.1	141.3	200.5		
Rational	-	-	978.0	2159.0	5950	3692		3059	456.3	311.1		

*No estimate of Q(10) available. MAF estimated using POT method from a short record.

TABLE 9.3 RMS ERROR OF MAF FOR THE FLOOD ESTIMATION METHODS USED IN THE COMPARISON

	METHOD					
	INDONESIAN FSR	WEDUWEN	MELCHIOR	HASPER	RATIONAL	PETERSON
ROOT MEAN SQUARE ERROR (%):						
(A) JAVA ONLY						
(a) For all catchments to which method applied (Number of catchments in brackets)	58.8 (11)	658 (3)	92.5 (6)	1018.3 (9)	1695 (8)	66.2 (5)
(b) Indonesian FSR RMS error for common catchments		64.2	64.6	64.5	58.8	73.5
(B) SUMATRA ONLY						
(a) For all catchments to which method applied	48.7 (10)	251.6 (2)	142.0 (7)	77.2 (8)	442.7 (5)	41.1 (3)
(b) Indonesian FSR RMS error for common catchments		34.9	45.8	52.0	45.8	30.6
(C) JAVA AND SUMATRA						
(a) For all catchments to which method applied	54.0 (21)	495.4 (5)	119.2 (13)	523.0 (19)	1213 (13)	56.8 (8)
(b) Indonesian FSR RMS error for common catchments		52.5	54.5	57.9	53.8	57.4

TABLE 9.4 RMS ERROR AS PERCENTAGES FOR THE 10 YEAR FLOOD ESTIMATES

	METHOD					
	INDONESIAN FSR	WEDUWEN	MELCHIOR	HASPER	RATIONAL	PETERSON
(A) JAVA ONLY						
(a) For all catchments to which method applied (Number of catchments in brackets)	68.7 (11)	622.3 (3)	74.4 (6)	933.4 (9)	1658 (8)	85.4 (5)
(b) Indonesian FSR RMS error for common catchments		87.5	69.2	75.8	79.4	83.4
(B) SUMATRA ONLY						
(a) For all catchments to which method applied	51.1 (9)	178.8 (2)	106.7 (6)	64.8 (7)	242.0 (4)	33.4 (3)
(b) Indonesian FSR RMS error for common catchments		20.0	33.8	56.0	33.8	18.5
(C) JAVA AND SUMATRA						
(a) For all catchments to which method applied	60.8 (20)	444.9 (5)	90.6 (12)	553.4 (16)	1186 (12)	65.9 (8)
(b) Indonesian FSR RMS error for common catchments		60.5	51.5	67.1	64.2	59.1

ANNEX A. STATION ASSESSMENT

A.1 Introduction

This annex considers the factors affecting the usefulness of data collected at gauging stations and should be considered an essential first step in starting a design flood estimation. Section A.2 describes office procedures and Section A.3, field procedures. Section A.4 gives details of how these were implemented in the current study.

A.2 Office procedure

The first stages of the flood estimation procedure should involve the compilation of a list of gauging stations at, and close to, the site of interest. The list should include not only stations near to the site within the same catchment but also those in neighbouring catchments. The primary source of information for this list should be DPMA in Bandung who are responsible for the national hydrometric network. Other bodies do operate gauging stations. Some of these are given at the foot of Table A.3.

The reason for collecting data from as many sources as possible is to ensure that all relevant information is considered in producing a balanced flood estimate.

Before site visits it is advisable to check the data which are available for each gauging station. In particular:

(1) Proximity to the site of interest.
(2) Type of station (continuously recording or staff gauge only)
(3) How many years of data available.
(4) The quality of the rating curve for flood flows.

This information may be used in assessing the relative usefulness of each station in the flood estimation procedure. A preliminary look at the data available may raise questions which can be answered during visits.

Field visits may now be undertaken as outlined in section A.3.

After field visits and the identification of useful stations peak stages should be abstracted for POT or annual maximum series analysis from the original charts if the station is automatic, or from the observer's field sheet if there is only a staff gauge. Data from a secondary source (eg year books), is prone to error. Furthermore useful information such as annotation indicating a sticking float or a large flood (benjir besar) is often only available from the original source.

Rating curves should be developed preferably with a programme of additional flood discharge measurements to reduce the degree of rating extrapolation. Guidelines for rating curve development are given in Annex B.

A.3 Station visit procedure

Preliminary site visits should involve inspection of gauging stations and discussion of gauging practice and historic floods with the observer and local people respectively. This will lead to an understanding of the relative accuracy of stations and any special local factors such as the depth and extent of flooding and the location of any historic flood. Photographing stations is generally found to be a useful aid in recalling details of station visits on return to the office.

The condition of the equipment at the station should be noted. In particular that the staff gauge is firmly fixed, its markings legible (including metre marks), and that the current reading agrees with the chart reading if the station is automatic. For non-automatic stations a check of the observer's notebook with the current staff gauge reading helps to assess the observer's diligence. It is worthwhile looking around for other staff gauges in the vicinity of the station and if found take readings on both old and new. Often old staff gauges are replaced by new ones or an automatic station installed nearby with a different datum. Noting this in the field may save problems later in the office.

An assessment of the likely behaviour of the station during flooding should be made. If the level of the maximum flood has been abstracted from the data in time for the field visit this may be visualised by reference to the staff gauge. If overbank flow occurs at this level, the depth and width of the flood plain should be estimated. The effectiveness of any overbank flow may be judged by the denseness of bankside vegetation or the presence of road or rail bridge abutments, which may confine flood flows, immediately downstram of the station. Often local people will tell of extent and frequency of overbank flow.

Special note should be made of flood marks near the station and these should be levelled into the station datum and the year of occurrence determined. When developing the rating curve for the station it is useful to know the type and shape of the hydraulic control of the station. Stable bed material, perhaps large boulders or a rock bar downstream indicate a good stable control. Bridge piers downstream usually have the same effect. Poor control is usually found in rivers with unstable bed material such as sand and gravel. This may form shoals in the river which realign after flooding. A station with a good control should have a reasonably stable rating.

A.4 Initial screening of Flood Study stations

The purpose of the station selection procedure described below was to ensure that only data from the most reliable gauging stations entered into the analysis for this project. This and the station visit procedure described in Section A.3 was one of the most important parts of the study. From sources at DPMA tables of all known gauging stations on Java and Sumatra were prepared. A total of about 1000 stations were identified, but this number was considerably reduced by eliminating all stations with short records, those affected by tides, dominated by an upstream lake or reservoir, stage only stations and those of obscure origin, and those with very poor ratings.

The procedure adopted for field visits has been discussed in Section A.3 with the necessary office work described in Section A.2.

The last stage of the station selection procedure involved studying all the available information for each station; the rating

Flood Study Number	Station / Location	Operator	Type	No. Yrs Data	MAF m^3/s	AREA km^2	MSL km	S1085 m/km	SIMS m/km	Forest km^2	Plantn km^2	Paddy km^2	Swamp km^2	Lake km^2	AAP mm	APE4P mm
1	Citarum at Nanjung	1	a	21	270.1	1832.6	65.4	4.81	13.5	576.00	0.0	549.0	4.0	71.00	2305.	69.0
2	Citarum at Saguling	1	a	7	659.6	2367.0	90.6	2.26	10.4	656.00	0.0	693.0	4.0	71.00	2256.	66.0
3	Citarum at Palumbon	1	a	31	1447.	4232.0	132.2	5.73	10.3	1226.0	0.0	1146.	4.0	71.00	2479.	70.0
4	Cianten II at Kracak	1	a	58	331.2	126.30	12.5	22.4	86.3	75.000	0.0	25.00	0.0	0.0	4950.	122.0
5	Ciliwung at Kebon Baru	9	a	2p	153.0	322.50	85.5	7.63	34.0	34.000	0.0	75.00	0.0	0.0	3469.	102.0
6	Cikarang at Cikarang	1	a	11	244.8	217.6	52.2	1.25	5.43	27.000	0.0	74.00	0.0	0.0	3482.	107.0
7	Cilamaya at Cipeundeuy	1	a	6	497.2	139.2	56.6	22.4	50.3	75.000	0.0	17.00	0.0	0.0	3631.	134.0
8	Cigulung at Maribaya	1	a	22	26.70	49.20	10.0	55.3	99.1	24.700	0.0	9.525	0.0	0.0	2709.	102.0
9	Cikapundung at Maribaya	1	a	12	31.10	75.60	11.5	43.7	85.6	38.800	0.0	10.2	0.0	0.0	2679.	98.0
10	Cimanuk at Leuwigoong	1	a	32	294.9	757.4	59.4	13.2	20.8	273.00	0.0	192.0	0.0	0.0	2560.	81.0
11	Cimanuk at Garut <Leuwidaun>	1	b	10	102.7	474.9	35.5	19.9	31.3	224.00	0.0	78.00	0.0	0.0	2715.	81.0
12	Cimanuk at Tomo	1	a	10	609.8	1996.	130.4	6.82	13.6	770.00	0.0	475.0	0.0	0.0	2631.	80.0
13	Cimanuk at Parakan Kondang	10	a	19	680.7	1514.8	114.6	7.62	14.9	385.00	0.0	366.0	0.0	0.0	2559.	79.0
15	Cilangla at Leuwineukteuk	1	a	6	460.0	179.90	32.2	18.7	21.8	18.0	0.0	53.0	0.0	0.0	3266.	132.0
16	Cisanggarung at Cilengkrang	1	b	5	391.8	622.1	81.7	7.99	10.9	74.0	0.0	180.0	93.0	23.5	2669.	88.0
17	Citanduy at Cirahong	1	a	8	588.2	634.7	54.1	6.76	18.5	114.0	0.0	230.0	0.0	0.0	3415.	101.0
18	Ciseel at Cilisung	1	b	8	143.4	178.5	30.8	17.8	36.6	30.6	0.0	6.4	0.0	0.0	3389.	122.0
19	Ciseel at Binangun	1	a	7	250.2	320.3	48.6	8.56	19.6	32.0	0.0	33.0	0.0	0.0	3279.	116.0
20	Cidurian at Kopomaja	1	b	8	302.9	300.0	68.4	13.5	26.9	83.0	0.0	69.0	0.0	0.0	3364.	112.0
21	Cidurian at Parigi	1	a	9	193.2	622.0	91.1	7.06	20.1	173.0	0.0	159.0	0.0	0.0	2970.	99.0
22	Ciujung at Kragilan	1	b	10	733.1	1858.0	111.3	1.83	6.94	639.0	0.0	364.0	2.6	0.0	3120.	94.0
23	Cikadueun at Cibogo	1	a	9	142.6	34.6	11.6	14.5	113.	14.2	0.0	14.8	0.0	0.0	3327.	153.0
24	Ciliman at Leuwikopo	1	a	6	121.5	108.5	25.0	9.74	23.7	62.7	0.0	9.4	0.0	0.0	3261.	126.0
25	Ciletuh at Cipiring	1	a	5	294.5	79.3	26.2	21.0	24.4	26.3	0.0	1.75	0.0	0.0	4086.	162.0
26	Cimandiri at Tegal Datar	1	a	6	369,4	495.1	46.3	10.4	21.6	112.0	0.0	144.0	0.0	0.0	2988.	94.0
27	Kawah Ciwidey at Pos A	1	a	10	0.579	0.43	0.4	87.0	200.	0.23	0.0	0.0	0.0	0.0	3650.	113.0
28	Cipadarum at Pos B	1	a	12	6.57	3.05	3.2	100.	150.	3.00	0.0	0.0	0.0	0.0	3490.	113.0
29	Ciwidey at Pos C	1	a	11	7.67	10.81	5.5	111.	138.	9.5	0.0	0.0	0.0	0.0	3483.	113.0
30	Cisarua at Pos D	1	a	13	5.11	4.72	3.6	111.	150.	4.72	0.0	0.0	0.0	0.0	3517.	113.0
31	Cisarua at Pos E	1	a	9	5.12	2.43	1.3	123.	188.	2.43	0.0	0.0	0.0	0.0	3480.	112.0
32	Kali Padegolan at Pajengkolan	1	a	6	627.8	212.4	36.3	15.2	25.9	35.5	0.0	62.3	0.0	0.0	4079.	132.0
33	Kali Serayu at Banyumas	1	a	11	1135.	2642.4	132.4	8.23	19.2	485.0	0.0	662.0	0.0	0.0	4050.	100.0
35	Kali Progo at Duwet	1	a	8	519.4	1749.4	86.0	10.5	37.0	60.9	0.0	331.0	0.0	0.0	2985.	87.0
36	Kali Progo at Krangan II	1	a	9	374.7	417.2	34.6	33.2	82.5	60.5	0.0	168.0	0.0	0.0	2270.	89.0
38	Kali Serang at Durungan	1	b	10	401.7	101.9	16.1	30.2	52.0	0.0	0.0	10.3	0.0	0.0	2295.	126.0
39	Kali Bogowonto at Glagah Malang	1	b	8	217.0	141.5	37.2	6.92	86.2	14.5	0.0	14.2	0.0	0.0	2423.	115.0
40	Kali Bogowonto at Bener	1	a	30	95.2	90.0	18.6	40.2	150.	0.25	0.0	3.4	0.0	0.0	2034.	116.0
41	Kali Lusi at Menduran	1	a	6	466.9	1968.1	128.0	0.4	2.17	772.0	0.0	776.0	0.0	63.0	2034.	80.0
42	Bengawan Solo at Napel	1	a	10	2150.	9578.0	213.5	0.46	4.67	1745.0	0.0	3238.0	0.0	77.0	2245.	74.0
43	Kali Serayu at Garung	1	a	48	53.2	56.4	102.9	85.5	103.	11.8	0.0	0.0	0.0	0.0	3424.	98.0
44	Bengawan Solo at Juranggempal	1	a	12	909.5	1442.0	54.1	1.60	17.0	118.0	0.0	277.0	0.0	44.0	2122.	83.0
45	Bengawan Solo at Kauman	1	b	10	1477.	5900.0	201.4	0.46	4.77	693.0	0.0	2062.	0.0	56.0	2342.	81.0
46	Bengawan Solo at Bojonegoro	1	b	11	2072.	12429.	331.9	0.33	2.72	3023.0	0.0	4094.	0.0	77.0	2189.	73.0
47	Kali Brantas at Ploso	11	a	8	832.0	8750.1	228.0	2.37	9.37	2154.0	0.0	2611.	20.0	3126.	2167.	83.0
48	Kali Brantas at Mojokerto	1	a	10	1198.	9972.9	255.2	1.84	8.86	2585.0	0.0	3207.	20.0	3126.	2176.	77.0
49	Kali Asem at Sentul	1	a	8	100.6	187.4	32.1	55.0	81.7	125.0	0.0	16.7	0.0	0.0	3250.	113.0
50	Kali Welang at Purwodadi	1	a	12	140.8	152.0	27.9	64.2	75.9	47.0	0.0	22.0	0.0	0.0	2469.	101.0
51	Kali Brantas at Kertosono	11	a	8	650.0	6902.0	202.4	2.57	11.1	1486.0	0.0	1885.	15.2	3126.	2231.	76.0
52	Kali Brantas at Cadang	11	a	4	217.7	772.2	38.3	25.5	49.2	250.0	0.0	140.0	0.0	0.0	1990.	88.0

Automatic station = a   o after station number indicates POT station
Staff gauge station = b

Note (1) Plantation not abstracted for stations in Java   (2) For key to operators see table A.3

TABLE A.1: Catchment characteristics of stations used in the analysis - Java

Flood Study Number	Station / Location	Operator	Type	No. Yrs Data	MAF m^3/s	AREA km^2	MSL Km	S1085 m/km	SIMS m/km	Forest km^2	Planta km^2	Paddy km^2	Swamp km^2	Lake km^2	AAP mm	APBAR mm
103	Kr. Tamiang at Kuala Simpang ..	1	a	6	1015.	4494.3	173.0	4.17	17.3	4173.7	236.0	17.1	0.0	0.0	2700.	95.0
105	Kr. Aceh at Kampung Darang	1	a	4p	460.0	1085.6	72.8	6.11	20.8	611.00	36.0	51.0	0.0	0.0	1850.	86.0
116	Kr. Peusangan at Beukah	1	a	3p	609.8	2214.2	137.8	11.9	17.0	1717.6	314.1	50.9	0.0	224.6	2550.	70.0
117	Kr. Jambo Aye at Rampah	1	a	4p	815.0	4060.7	187.0	4.13	10.3	3607.0	309.0	24.0	0.0	0.0	2500.	65.0
118	Kr. Jambo Aye at Lhoknibong ...	1	b	8	931.6	4402.7	231.2	2.88	5.35	3887.0	315.0	25.0	0.0	0.0	2450.	67.0
123	Kr. Suson at Kotatinggi	1	ab	3p	385.0	167.0	23.9	20.8	58.9	167.0	0.0	0.0	0.0	0.0	3200.	132.0
201	S. Wampu at Stabat	1	ab	10	664.6	3788.9	178.1	8.43	10.8	1984.7	202.8	409.5	0.0	9.8	2770.	89.0
202	S. Belawan at Asam Kumbang ...	1	ab	8	209.0	215.0	37.7	16.2	58.0	112.0	0.1	11.6	0.0	0.0	2600.	100.0
205	S. Ular at Tagor	1,5	b	10	350.0	979.0	76.3	19.2	18.7	569.6	105.2	26.4	2.2	0.0	2650.	92.0
206	S. Padang at Tebing Tinggi	1	a	5	155.0	917.6	78.9	10.6	13.1	309.2	516.6	52.7	3.2	0.0	2750.	95.0
207	Bah Bolon at Batu Gajah	1,6	ab	8	138.0	587.6	60.0	18.3	18.2	68.1	349.8	121.5	0.0	0.0	2780.	89.0
208	Bah Bolon at Nagori Bandar	1,6	ab	10	210.8	679.5	93.4	13.5	13.6	68.1	422.8	133.9	0.0	0.0	2770.	92.0
209	S. Silau at Kisaran Naga	1	b	9	257.7	1022.5	105.6	9.98	20.0	749.3	260.8	12.3	0.0	0.0	2710.	90.0
216	Bt. Pane at Gunung Tua	1	a	6	873.5	836.9	60.8	19.5	19.0	659.5	0.0	3.9	0.0	0.0	2120.	96.0
218	Bt. Gadis at Perbangunan	1,2	a	9	306.0	820.0	68.4	9.87	28.4	494.0	114.0	56.2	0.0	0.0	2290.	90.0
219	Bt. Angkola at Air Libung	1,2	ab	8	202.3	629.7	60.0	8.40	24.3	368.0	25.6	86.4	0.0	0.0	2970.	89.0
243	Bt. Ilung at Hajoran	2	a	6	182.4	180.8	45.0	24.6	40.2	151.0	0.0	1.1	0.0	0.0	1950.	98.0
246	Aek Puli at Sukamaju	2	ab	4	386.9	727.7	53.5	8.55	25.6	557.1	0.6	34.3	0.0	0.0	2040.	88.0
311	Bt.Air Haji at Kampung Danau ..	1	a	3p	475.	356.2	43.4	36.9	61.6	356.2	0.0	0.0	0.0	0.0	3190.	121.0
312	Bt.Pelepah at Tanjung Mesjid ..	1	a	4p	205.	332.1	48.5	25.9	40.8	330.4	1.2	0.5	0.0	0.0	3375.	120.0
313	Bt.Bayang at Muara Air	3	a	15	165.	209.3	21.0	63.5	68.6	195.5	4.1	2.1	2.2	0.0	3040.	145.0
314	Bt.Kuranji at Gunung Nago	4	b	8	415.	121.6	20.2	45.0	70.3	121.6	0.0	0.0	0.0	0.0	3560.	147.0
316	Bt.Anai at Kandang Empat	2	a	9	89.1	110.	23.7	53.4	106.	70.7	0.0	20.1	0.0	0.0	3450.	111.0
321	Bt.Air Dingin at Lubuk Minturun	1	a	3p	440.	114.0	20.8	62.7	75.6	114.0	0.0	0.0	0.0	0.0	4100.	147.0
331	Bt.Masang at Sipisang	1	a	11	180.	436.9	45.3	15.7	57.6	311.2	23.1	60.6	0.0	0.0	2850.	87.0
341	Bt.Tongar at Lubuk Toreh	1	a	9	263.	304.9	31.4	31.2	30.4	278.7	5.8	20.4	0.0	0.0	4050.	114.0
342	Bt.Pasaman at Air Gadang	1	a	6	981.	1267.0	67.4	12.5	19.6	1179.7	29.5	54.3	3.5	0.0	3440.	103.0
343	Bt.Batahan at Silaping	1	a	12	360.	304.0	54.6	32.6	29.9	290.5	13.5	0.0	0.0	0.0	3100.	118.0
412	Bt.Sumpur at Batu Bertindih ...	1	a	3p	132.	282.9	34.5	14.2	27.5	218.8	25.5	20.5	0.0	0.0	3970.	99.0
413	S. RokanKiri at Lubuk Bendahara	1	a	5	946.	3128.7	159.1	2.35	7.45	2875.3	120.7	68.4	13.7	0.0	2580.	80.0
422	Bt. Mahat at Sipopay	3	a	7	898.	906.1	61.5	14.6	35.0	757.6	11.2	32.5	12.6.	0.0	2790.	87.0
424	Bt.Kampar Kiri at Lipat Kain ..	1	a	2p	900.	3934.2	153.2	1.10	5.5	3846.0	83.7	4.5	0.0	0.0	3360.	87.0
431	Bt. Agam at Titi	2	a	16	96.	346.4	44.0	10.7	49.9	64.5	2.1	124.7	0.0	0.0	2340.	76.0
433	Bt.Sumani at Bd.Pundung	1	a	4p	270.	459.1	39.2	41.2	33.1	215.4	16.8	193.6	0.0	21.6	2770.	118.0
436	Bt.Kuantan at Lubuk Ambacang ..	2	ab	2p	1950.	7607.8	222.0	4.78	12.6	4534.7	356.5	1184.	0.0	1079.1	2460.	84.0
437	Bt.Kuantan at Peranap	2	a	3p	1070.	11117.	375.5	1.73	7.5	7011.1	1095.	1269.	0.0	1079.1	2550.	84.0
511	Bt.Hari at Sungai Dareh	1	a	6	2941.	4578.6	150.1	7.20	12.1	4139.0	29.7	174.8	0.0	40.3	3290.	97.0
512	Bt.Ule at Lubuk Tapus	1	a	6	351.	343.7	77.0	14.8	23.4	332.0	7.6	4.1	0.0	0.0	3220.	107.0
514	Bt.Hari at Muara Kilis	1	a	2p	2890.	18065.	429.5	0.83	4.40	15554.	1750.	326.1	0.0	40.3	2920.	83.0
515	Bt.Tabir at Muara Jernih	1	ab	8	504.3	923.	97.6	11.5	19.8	870.7	52.1	0.0	0.0	0.0	3140.	98.0
516	Bt.Hari at Muara Tembesi	1	a	4p	4490.	36370.	569.9	0.38	3.30	30774.	4394.	539.6	9.3	1029.6	2900.	75.0
521	Bt.Merangin at Lubuk Paku	3	a	15	148.	1256.	89.6	10.5	34.2	765.5	172.4	183.5	0.0	989.3	2400.	88.0
522	Bt.Tembesi at Muara Inum	1	a	12	1164.	1505.3	106.6	11.3	26.9	1241.2	264.1	0.0	0.0	0.0	3220.	87.0
523	Bt.Tembesi at Pauh	1	a	3p	1650.	10318.	339.8	2.94	11.1	8393.1	1484.	213.5	9.3	989.3	2920.	81.0
610	Air Ketaun at Tunggang	4	a	3p	400.0	946.8	73.8	13.4	23.3	782.8	7.2	76.0	1.4	0.0	3400.	89.0
701	S. Musi at Depoetah	3	a	8	217.4	649.7	50.4	15.2	26.3	555.0	49.8	10.5	4.2	0.0	3210.	93.0
707	B. Komering at Martapura	2	a	9	1302.	4463.6	166.6	3.63	8.91	3369.8	455.4	52.6	0.0	501.3	2730.	81.0
709	S. Ogan at Batu Raja	2	b	10	1409.	2046.0	105.2	7.53	18.2	1732.1	115.5	32.4	0.0	0.0	2920.	97.0
712	S. Rambang at Tanjung Rambang	2	a	9	109.5	567.0	81.4	0.8	1.63	207.0	75.7	0.0	0.0	0.0	2900.	108.0
801	W.Sekampung at Kunyir	1,8	a	9	323.4	432.1	53.4	15.7	23.3	334.4	43.8	0.0	0.0	0.0	2450.	74.0
803	W.Sekampung at Pujorahayu	1,8	a	12	463.4	1697.9	90.3	8.99	14.3	485.9	814.7	68.0	1.0	0.0	2440.	74.0
807	W.Bulok at Bulok (Jembatan) ...	8	b	6	226.2	786.0	52.4	8.36	38.4	118.8	407.9	38.3	1.0	0.0	2450.	86.0
812	W.Sekampung at Argoguruh	1,8	a	14	458.8	2102.0	122.7	5.62	10.9	490.6	1014.5	87.5	1.0	0.0	2400.	73.0
817	W.Besai at Petai	8	a	8	84.2	416.0	49.2	2.94	5.56	254.1	52.1	2.0	0.0	0.0	2460.	62.0
818	W.Besai at Banjarmasin	8	a	9	294.2	661.6	99.2	10.3	5.28	321.4	104.9	2.0	0.0	0.0	2430.	72.0
823	W.Ciham at Rantau Jangkung	8	a	9	279.6	526.7	92.3	12.8	20.2	492.4	14.6	0.0	0.0	0.0	2490.	94.0
824	W.Umpu at Negeri Batin	8	a	9	313.2	555.6	77.0	11.8	22.6	308.8	132.0	0.0	0.0	0.0	2430.	86.0
825	W.Pengubuan at Gedong Harta ...	8	a	6	67.5	111.3	44.0	20.1	39.1	106.8	4.5	0.0	0.0	0.0	2420.	85.0
827	W.Terusan at Gunung Batin	8	b	8	108.	528.5	59.0	0.31	0.97	185.5	97.9	0.0	0.0	0.0	2375.	90.0
834	W.Rarem at Kotabumi	8	a	6	278.5	903.0	71.2	7.80	29.1	645.4	223.6	0.0	0.0	0.0	2420.	84.0
839	W.Umpu at Rantau Temiang	8	a	6	88.1	205.2	30.8	27.6	48.3	200.2	5.0	0.0	0.0	0.0	2450.	84.0

Station type (a = automatic, b= Staff gauge)   p after station number implies POT station.   For key to operators see table A.3

TABLE A.2 Catchment characteristics of stations used in the analysis - Sumatra

curve, the ratio $\frac{H_{max}}{H_{obs}}$, (where H_{max} is the maximum recorded stage and H_{obs} the maximum gauged stage), the cross section when available together with the hydrologist's field assessment of the gauging station. This information on the rating was considered in conjunction with the length of record at a station to determine whether the station should be included in the analysis. Thus a station with a poor rating was more likely to be included if it had a valuable long record than if it had just a few years' data. The process was to a large extent subjective and stations were included or rejected from the analysis using the experience of the hydrologists.

The final list of 110 gauging stations given in Tables A.1 and A.2 and shown in Figures 1.1 and 1.2 includes a group of five small catchments in the Kawah Ciwidey area to the south of Bandung (Numbers 27, 28, 29, 30 and 31). These are of special interest as they have the only data on very small catchments in Java and are therefore potentially very useful in extending the range of application of the regression equation to catchments below 50 km^2 in area. The stations were abandoned in the Second World War (1939-1945) but field visits during this project discovered the existence of sharp crested weirs at four of the five sites. According to local information the fifth was constructed similarly. Although there were only low flow discharge measurements taken at these locations, the presence of the weirs gave more confidence in the rating curve extrapolation. In view of this and their potential usefulness to this study all five stations were included in the analysis.

There was some doubt about the inclusion of staff gauge only stations in the analysis because observations there are normally taken only three times a day and hence the flood peak may be missed. In some instances the observer does record peak flood levels, as is the requirement, but this is not always the case. For large catchments staff gauge readings three times daily are acceptable because the flood peak may last for many hours, but for smaller catchments the possible underestimation of peak flows may be significant. Where it appeared that these errors might be excessive, the station was removed from the analysis.

Table A.3 gives the catchment characteristics abstracted for some of the stations not used in the regression analysis. These data may be of use in the regression equation if a flood estimate is required in the vicinity of one of these stations. These stations were omitted from the analysis because of doubts about their MAF. For this reason the MAFs given in Table A.3 should not be used without first checking the basic data and rating equations.

Flood Study Number	Station / Location	Operator	Type	No. Yrs Data	MAF (m^3/s)	AREA (km^2)	MSL (km)	S1085 (m/km)	SIMS (m/km)	Forest (km^2)	Plantn (km^2)	Paddy (km^2)	Swamp (km^2)	Lake (km^2)	AAR (mm)	APBAR (mm)
14*	Cimanuk at Monjot	1	a	0	?	2617.4	138.6	6.44	12.8	845.00	0.0	684.0	0.0	0.0	?	?
34*	Kali Serayu at Mrica...........	3	a	0	?	992.3	71.7	21.6	33.9	150.0	0.0	208.0	0.0	0.0	?	?
37*	Kali Progo at Kranggan I	1	b	8	187.5	417.2	34.6	33.2	82.5	60.5	0.0	168.0	0.0	0.0	2270.	89.0
121*	Kr. Meureubo at Meunasah Rayeuk	1	a	1p	844.2	1651.5	152.7	12.1	18.4	1481.1	69.6	39.3	47.8	0.0	2720.	85.0
122*	Kr. Lambeso at Sango	1	a	4p	2000.	474.0	33.2	17.5	61.4	470.0	0.0	1.5	0.0	0.0	2450.	109.0
128*	Kr. Tripa at Cunung Kong	1	a	2p	436.4	2625.6	157.0	6.74	11.2	2271.8	268.7	60.6	0.0	0.0	2060.	69.0
204*	S. Blumai at Tanjung Morawa ...	1	b	8	133.0	248.9	50.9	7.24	29.5	152.3	10.6	5.3	0.0	0.0	2560.	105.0
223*	S. Percut at Tembung	4	b	7	70.6	173.0	61.2	6.78	30.9	97.1	0.05	4.7	0.0	0.0	2600.	105.0
224*	S. Deli at Helvetia	4	b	8	119.1	371.5	66.1	17.9	28.2	94.0	14.3	40.2	0.0	0.0	2560.	98.0
228*	Bah Tongguran at Tembakan	1,6	ab	2p	273.0	754.0	73.8	16.7	21.9	282.0	176.0	255.0	0.0	0.0	2770.	92.0
245*	Bt. Gadis at Sipotangniari	2	a	5	?	2618.0	114.6	6.23	17.2	1581.0	306.0	524.0	28.5	0.0	2520.	80.0
248*	Aek Bitumandi at Hutagalung ...	2	a	4	106.9	677.2	41.5	13.5	14.3	344.1	32.7	2.5	0.0	0.0	2100.	79.0
315*	Bt.Anai at Lubuk Bimantung	2	a	8	425.	326.1	42.6	31.7	66.6	264.3	0.0	29.0	0.0	0.0	3500.	122.0
423*	Bt.Kampar Kanan at D.Bingkuang	1	a	2p	1000.	3096.5	223.3	1.35	9.1	695.0	121.7	98.2	37.0	0.0	3250.	83.0
432*	Bt.Sinamar at Taram	4	a	2p	410.	1428.9	67.9	6.54	21.1	459.4	43.6	369.8	2.7	0.0	2220.	77.0
435*	Bt.Ombilin at Tanjung Ampalu ..	1	a	3p	1200.	2298.	119.6	6.16	12.7	616.7	134.4	507.2	0.0	1079.1	2400.	94.0
605*	Air Lais at Kuro Tidur	4	a	3p	300.0	133.2	21.0	45.4	96.3	133.2	0.0	0.0	0.0	0.0	2880.	116.0
609*	S. Manna at Bandar Agung	4	a	2p	559.0	592.0	61.0	16.9	34.9	507.1	55.5	12.6	0.0	0.0	3630.	97.0
706*	Air Enim at Dusun Lingga	3	a	6	1138.	1096.6	115.8	14.0	21.1	653.6	201.1	30.5	0.0	0.0	3360.	101.0
708*	S. Musi at Mambang	1	b	8	2103.	7891.1	289.6	1.27	6.4	4401.2	2517.	133.3	14.2	0.0	3050.	81.0
821*	W.Sekampung at Jurak	1,8	a	10	655.6	663.4	55.4	14.7	22.5	360.9	217.9	1.0	0.0	0.0	2460.	75.0
832*	W.Tahmi at Tanjung Agung	8	a	6	?	526.5	61.8	9.87	28.4	401.9	123.6	0.0	0.0	0.0	2475.	98.0
835*	W.Semangka at Liwa Road	8	a	7	?	226.4	26.6	5.4	19.1	158.9	63.0	4.5	0.0	0.0	2590.	84.0
845*	W.Ketibung at Sidomulyo	8	b	4	51.9	115.0	20.1	6.84	17.4	41.8	5.2	0.0	0.0	0.0	2420.	117.0
850*	W.Raman at Raman Hendra	8	b	4	46.1	207.4	47.7	0.63	0.77	0.0	1.9	26.1	0.0	0.0	2250.	95.0

* Station not used

Automatic station = a
Staff gauge station = b

p after station number indicates POT station.

Operators

(1) DPMA
(2) P3SA
(3) PLN
(4) Local DPU
(5) Ular Project
(6) Bah Bolon Project
(7) Wampu Project
(8) Lampung Hydrological Network Project (LHNP)
(9) Jakarta Floods Project
(10) Cimanuk Project
(11) Brantas Project

TABLE A.3. Catchment characteristics of some stations not used in the analysis.

ANNEX B. RATING EQUATION DEVELOPMENT

B.1 Introduction

A considerable amount of effort was spent during this study on the development of rating curves because of the importance of having good quality flow data in the analyses. A particular problem was the lack of flood discharge measurements resulting in large rating curve extrapolations. Because of the large number of stations being used in the project (more than 100) it was impractical to undertake additional discharge measurements in the time available. The work on rating curves therefore relied on making the best use of existing discharge measurements plus any other useful information such as channel cross sections and water surface profiles.

Section B.2 considers the development of rating curves when only discharge measurements are available; section B.3 is concerned with the use of channel cross section information and Section B.4 recommends some simple improvements in hydrometry, which would allow more flood discharge measurements to be made in the future.

B.2 Logarithmic rating curve extrapolation

The relationship between water level (stage) and discharge at a particular site is usually non-linear. A general form of rating equation, which has a sound theoretical basis (Robertson 1970), may be defined as:

$$Q = a(h + c)^b$$

where,

h	=	stage (m)
Q	=	discharge at stage h ($m^3 \; s^{-1}$)
a,b,c	=	rating curve parameters.

For any one site there can be a number of such equations covering different periods in time (when the rating is not stable) and covering different portions of the stage range.

The parameters "c" and "b" have physical interpretations

"c" is the correction applied to the head to allow for the difference between the elevation of the gauging station control and the staff gauge zero. Therefore $-c$ is the level on the staff gauge corresponding to zero flow (if there are multiple segments to the rating curve then this is only true of the lowermost segment).

"b" is the exponent of the rating curve and introduces non-linearity in the stage discharge relationship. This parameter is dependent on the shape of the river cross section at the control.

As a guide:

b = 1.5 - 1.6 for rectangular channels
b = 1.6 - 2.2 for trapezoidal or parabolic channels
b = 2.6 - 2.7 for triangular channels.

There are a number of assumptions in deriving these figures, but "b" might reasonably be expected to lie in the range 1.3 to 2.8. Where extrapolation of the rating is large "b" is the dominant parameter and as such it is important that the value chosen is reasonable.

Discharge measurements plotted on linear paper form a line of pronounced curvature often with a large cluster of points at low flows and few elsewhere. Logarithmic plotting of the same data expands the low flow range and contracts the high flow range which has the effect of spacing out the discharge measurements more evenly over the graph. Furthermore, if the parameter "c" is chosen correctly, curvature of the data when plotting $\log_{10} Q$ against $\log_{10} (h+c)$ may be substantially reduced or eliminated. This curvature is illustrated in Figure B.2 (the example is considered later in more detail). If the value of "c" is too small ($c = -0.75$ in Figure B.2) the data follow a curved path of decreasing slope. If "c" is too high ($c = -0.4$ in the example) the data plot about a curve of increasing slope. If the rating development is to be done by hand it is desirable to start initially with "c" as close as possible to the value that will give a straight line. Preferably it could be estimated on site as the level staff gauge at zero flow (in Figure B.2 it would be expected that zero flow would occur around a staff gauge reading of - 0.6 m). When this datum 'error' is unknown a sensible starting point is with 'c' of zero. If the resulting plot is strongly curved a new value of 'c' can be guessed or estimated graphically from the first graph by the method described in the Manual on Stream Gauging (WMO, 1980).

The slope of the line drawn through the logarithmic data is parameter "b":

$$\log Q = \log a + b \log (h+c)$$

Abrupt changes in cross section such as flood berms or a shift in downstream control at a particular water level may result in a two part relationship. This will be evident as a change in slope on the logarithmic plot.

Plotting discharge measurements chronologically, point by point, reveals any shift in datum. A shift could be caused by re-alignment of the channel after a major flood or by unrecorded repositioning of the staff gauge. If such a shift occurs, discharge measurement may be converted to the same datum by applying a correction to the appropriate stage values, or separate rating curves may be derived for different periods of the record.

Example

Because of the large number of rating curves which needed to be checked and re-drawn for this project, a FORTRAN computer program was written to assist this process. The procedure used is outlined below and uses the example of rating curve development for the Krueng Aceh at Kampung Darang.

(1) Plot stage discharge measurements chronologically on linear scales to determine any sudden shift in datum. If no shift go to step (4). Figure B.1 illustrates this point for the Krueng Aceh at Kampung Darang.

(2) Apply a correction to the stage values to bring all points to the same datum. If satisfactory (points plot as one line) go to step (4). Figure B1 shows such a correction for the Krueng Aceh at Kampung Darang.

Shift in station datum: Krueng Aceh at Kampung Darang

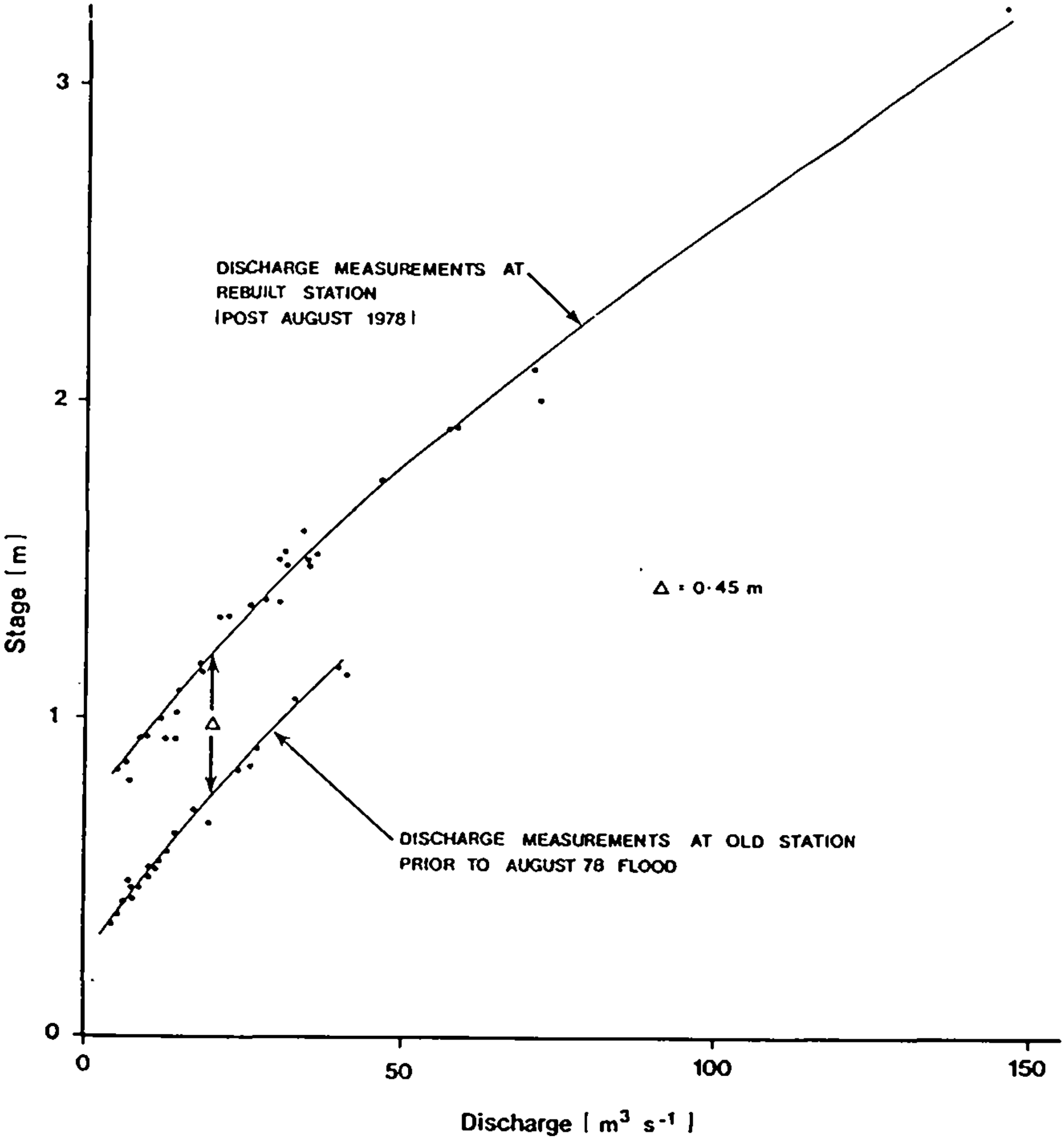


Figure B·1

(3) Divide data set and fit separate lines to each part.

(4) Compute log Q and log (h+c). Initially "c" should be set just greater than minus the minimum stage in the discharge measurement set (h_{min}). "c" cannot be less than or equal to $-h_{min}$ because it would then be impossible to calculate $\log(h_{min} + c)$.

(5) Plot log (h+c) against log Q. In our example $h_{min} = 0.79$ and c = - 0.75 (initially). Figure B.2 shows these data plotted. Note the curvature to this line which signifies an incorrect value of "c".

(6) Fit a straight line through the data using a least squares approach and calculate the error associated with the regression. In our example, with "c" = - 0.75, a relatively poor fit is obtained and the error high.

(7) Log (h+c) is recalculated several times with an increasing "c" and steps (5) and (6) repeated. Parameters "a" and "b" and associated error are recorded. Figure B.2 shows the data replotted for "c" = - 0.6 and "c" = - 0.4.

(8) Parameters "a", "b" and the regression error are plotted against parameter "c" and the point at which minimum error occurs noted. Figure B.3 shows that at Kampung Darang the optimum "c" is - .56, with a = 36.4 and b = 1.4. If a minimum is obtained go to step (10) below.

(9) In some instances, especially where a station is rated over a limited range of flows and there is considerable scatter in the discharge measurements, it may be impossible to obtain a minimum error in step (8). Here the error function initially falls rapidly, and thereafter continues to decrease slowly with an increase in "c". However the improvement in fit after a certain point is marginal. In this situation a knowledge of the river cross-section is necessary to establish a reasonable value for parameter "b".

(10) A check is then made to ascertain that the exponent "b" is reasonable for the gauging station cross-section; unless there is a special reason, "b" is only permitted to be in the range 1.3 to 2.8. The Krueng Aceh at Kampung Darang is a relatively wide river with steep banks. The exponent b = 1.4 is reasonable for this near rectangular cross-section.

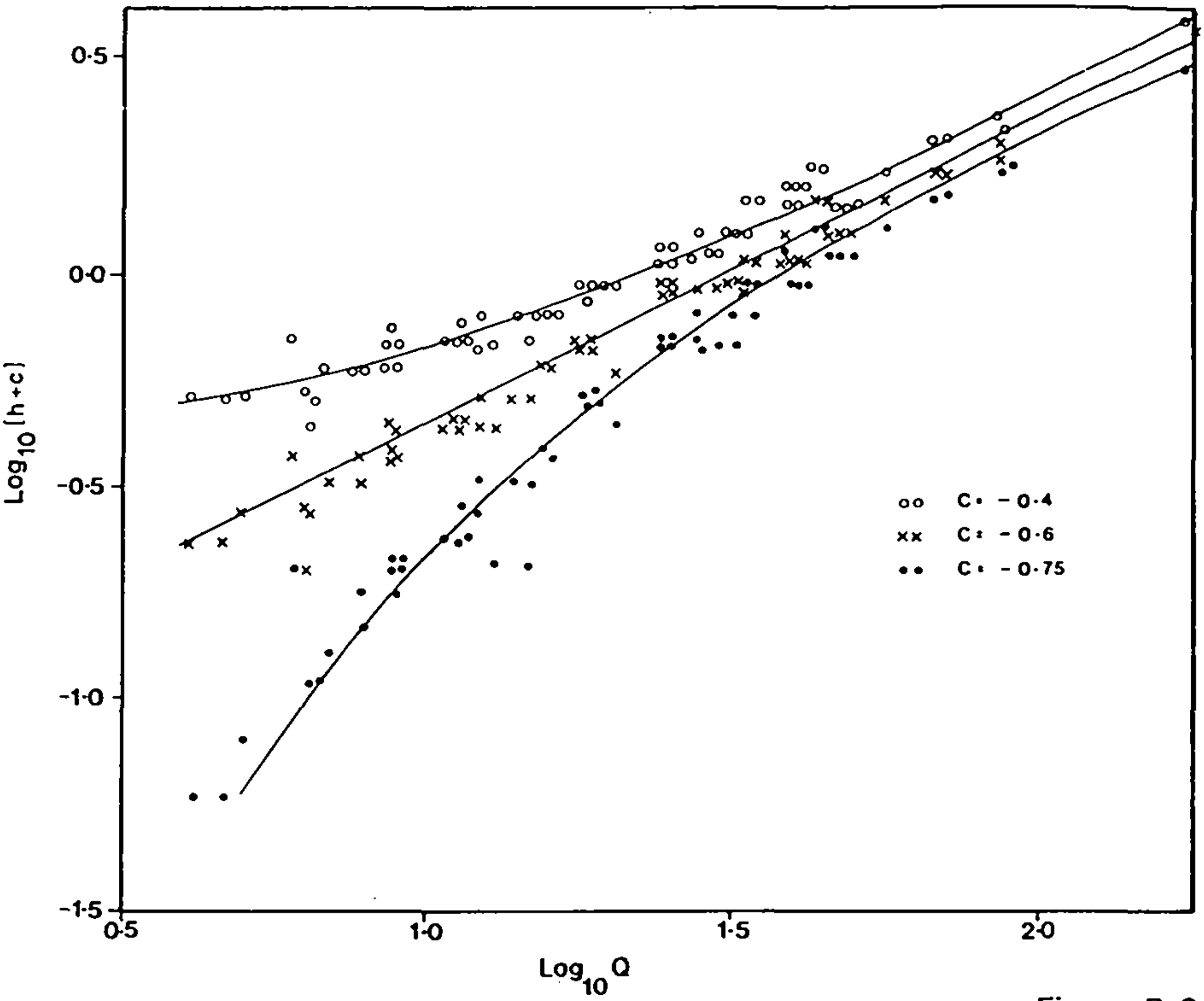


Figure B·2

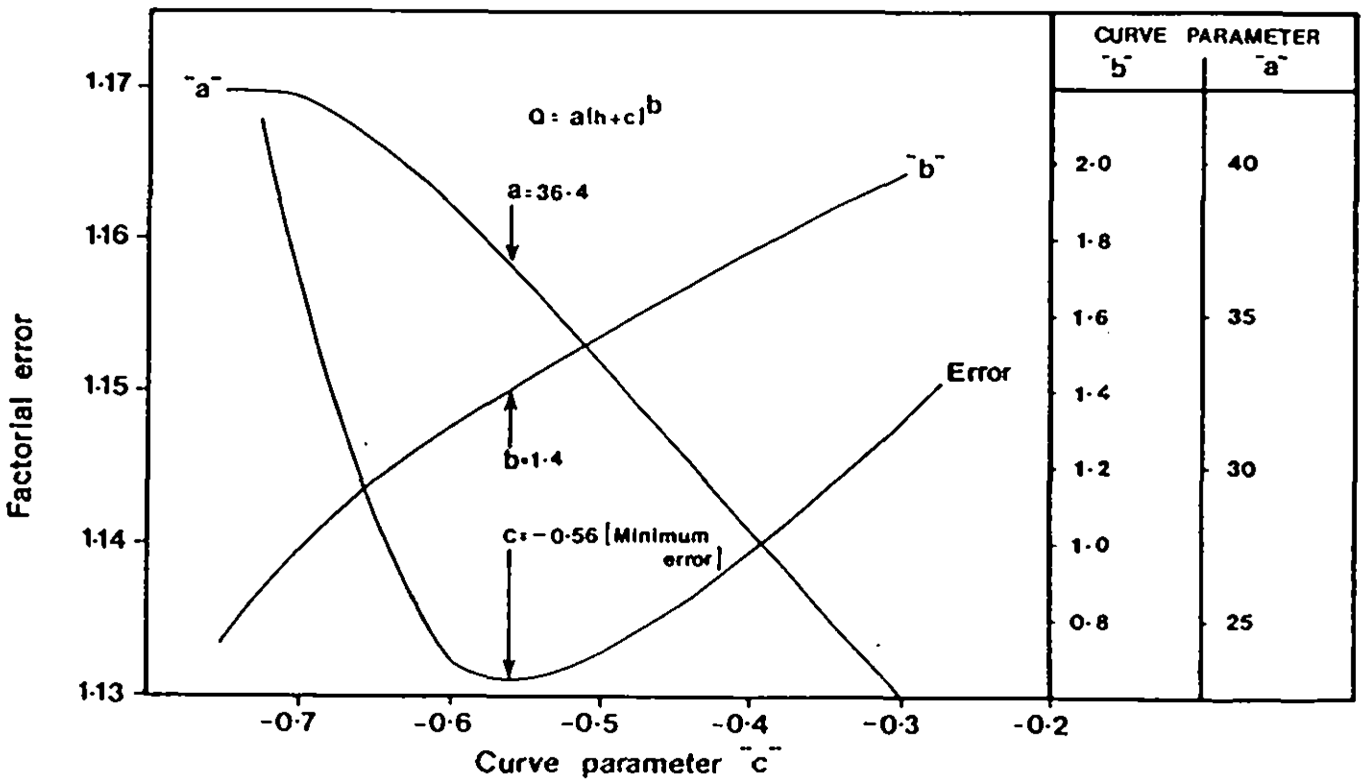


Figure B·3

(ii) If, after the best parameters have been obtained, it is not possible to obtain a single straight line on the logarithmic plot of (h+c) against q, the data may be divided into two or more straight line segments and separate rating curves developed over each range.

In the example considered above, one rating is sufficient to define the entire range of flows at Kampung Darang:

$$Q = 36.4(h - 0.56)^{1.4}$$

B.3 Slope - area rating extrapolation

There is no completely satisfactory method of extrapolating a rating curve from the highest measured discharge to the maximum flood level. Extension based on a rating curve equation fitted to low and medium flows cannot account for any marked change in the geometry of the channel at high flows. Use of a slope-area method in which the velocity of flow is calculated using a flow resistance formula, and multiplied by the flow area to give the discharge, overcomes this problem. The best known examples of such formulae are due to Manning and Chezy.

Manning's formula $= \frac{1}{n} R^{2/3} S^{1/2}$

Chezy's formula $v = C\sqrt{R S}$

where v is the mean flow velocity

R is the hydraulic radius, ie the flow area divided by the wetted perimeter of the channel

S is the longitudinal slope of the water surface

C and n are respectively Chezy's and Manning's roughness coefficients.

Example

For Batang Hari at Sungai Dareh gaugings are available in the stage range 0.84 m to 2.85 m. The logitudinal slope of the water surface, S, is 0.0016 m m^{-1}. The maximum flood level recorded at this site is over 7 m but the largest annual event is typically 5.0 m. What is the discharge at this stage?

The discharge measurements are tabulated and ordered as in Table B.1. The flow area and wetted perimeter are found from the cross section and the velocity, hydraulic radius and $1/\sqrt{f}$ calculated. Figure B.4 shows $1/\sqrt{f}$ plotted against $\log_{10}R$ and the straight line fitted to the points.

At a stage of 5.0 m the wetted perimeter is 144.2 m and the flow area 758.1 m^2

$$\text{Therefore } R = \frac{758.1}{144.2} \qquad 5.26 \text{ m}$$

From Figure B.4, $1/\sqrt{f}$ corresponding to R = 5.26 m is 4.07, which substituted in Darcy-Weisbach equation gives:

$$4.07 \times \sqrt{(8 \times 9.81 \times 5.26 \times 0.0016)}$$

$$= 3.3 \text{ ms}^{-1}$$

$$\begin{aligned} \text{Discharge} &= v \times A \\ &= 3.3 \times 758.1 \\ &= 2498 \text{ m}^3\text{s}^{-1} \end{aligned}$$

From an equation of the form $Q = a(h+c)^b$ fitted to the same flow gauging data the discharge at 5.0 m is 2862 m^3s^{-1}. The better estimate in this case must be considered to be the one from the slope-area method but the implied uncertainty in the stage-discharge relationship should be considered in making a flood estimate.

Table B.1 Calculating $1/\sqrt{f}$ from discharge measurements

Stage	Flow	Area	Wetted Perimeter	Velocity	Hydraulic Radius	$1/\sqrt{f}$
m	m^3s^{-1}	m^2	m	ms^{-1}	m	
2.85	931.0	469.5	131.7	1.98	3.57	2.95
2.73	808.0	454.0	131.1	1.78	3.46	2.69
2.31	608.0	400.4	129.1	1.52	3.10	2.43
2.30	607.0	399.1	129.1	1.52	3.09	2.43

Stage and flow from flow gauging

Area and Wetted perimeter from cross-section

Velocity = Flow/Area

Hydraulic Radius = Area/Wetted perimeter

$1/\sqrt{f}$ from Darcy-Weisbach formula

It will be seen that these formulae require the cross section to be surveyed, the longitudinal slope to measured (from scour marks after flooding) and a roughness coefficient to be estimated. The main source of error in applying such an equation is in determining the roughness coefficient, possibly based on a comparison of the channel with a table of coefficients, see, for example Chow (1959).

An alternative procedure is to estimate the roughness coefficient over the range of gauged flows and use this as the basis for extrapolation. This will be illustrated for the Darcy-Weisbach flow resistance formula, which is preferred to either the Manning or Chezy formulae as it is dimensionally correct and has a sound theoretical basis. The Darcy-Weisbach formula is

$$= \left(\frac{8gRS}{f}\right)^{0.5}$$

where R and S are as previously defined

g is the acceleration due to gravity

and f is the Darcy-Weisbach friction factor

Again the main problem in application of the formula is in estimating f. The Colebrook-White equation expresses $1/\sqrt{f}$ as a linear function of $\log_{10}R$

$$\frac{1}{\sqrt{f}} = c \log_{10}(bR)$$

where c and b are coefficients. In some applications it is possible to relate these coefficients to physically measureable properties of the bed material (Bathurst 1978, Hey 1979). In the present context it is suggested that they are estimated graphically from the available flow gaugings by plotting $1/\sqrt{f}$, calculated from the Darcy-Weisbach formula against $\log_{10}R$. To estimate the flow for a recorded flood level, the flow area and wetted perimeter are found and used to calculate R. The value of $1/\sqrt{f}$ is found from the graph of $1/\sqrt{f}$ against $\log_{10}R$ and this value is substituted into the Darcy-Weisbach formula to estimate the average flow velocity. The discharge is found from the product of the velocity and the flow area. By repeating this procedure at various stages a rating curve can be constructed.

B.4 Recommended improvements in hydrometry

Since the development of ratings is of such importance in flood hydrology the authors of this report would like to recommend a change in hydrometric practice that should lead to a rapid improvement in rating accuracy.

At present gauging teams drive each day from the local office to one or more gauging stations and return home in the afternoon or evening. Since the majority of heavy rainfalls are thunderstorms occurring late in the day the flood peaks pass the gauging stations after the gauging teams have left. If the teams could be based in the field close to a number of gauging stations and be prepared to gauge floods whenever they occur many flood gaugings would be made resulting in a marked improvement of the derived ratings. The teams should be prepared to gauge at night using a current meter from a cableway or nearby bridge but be wary of using boats especially in fast flowing or debris laden rivers. Floats are most useful in such cases. During periods of dry weather the teams can be usefully employed obtaining accurate channel cross-sections and in general station maintenance.

Graph of $\frac{1}{\sqrt{f}}$ against $\mathrm{Log}_{10} R$ for Batang Hari at Sungai Dareh

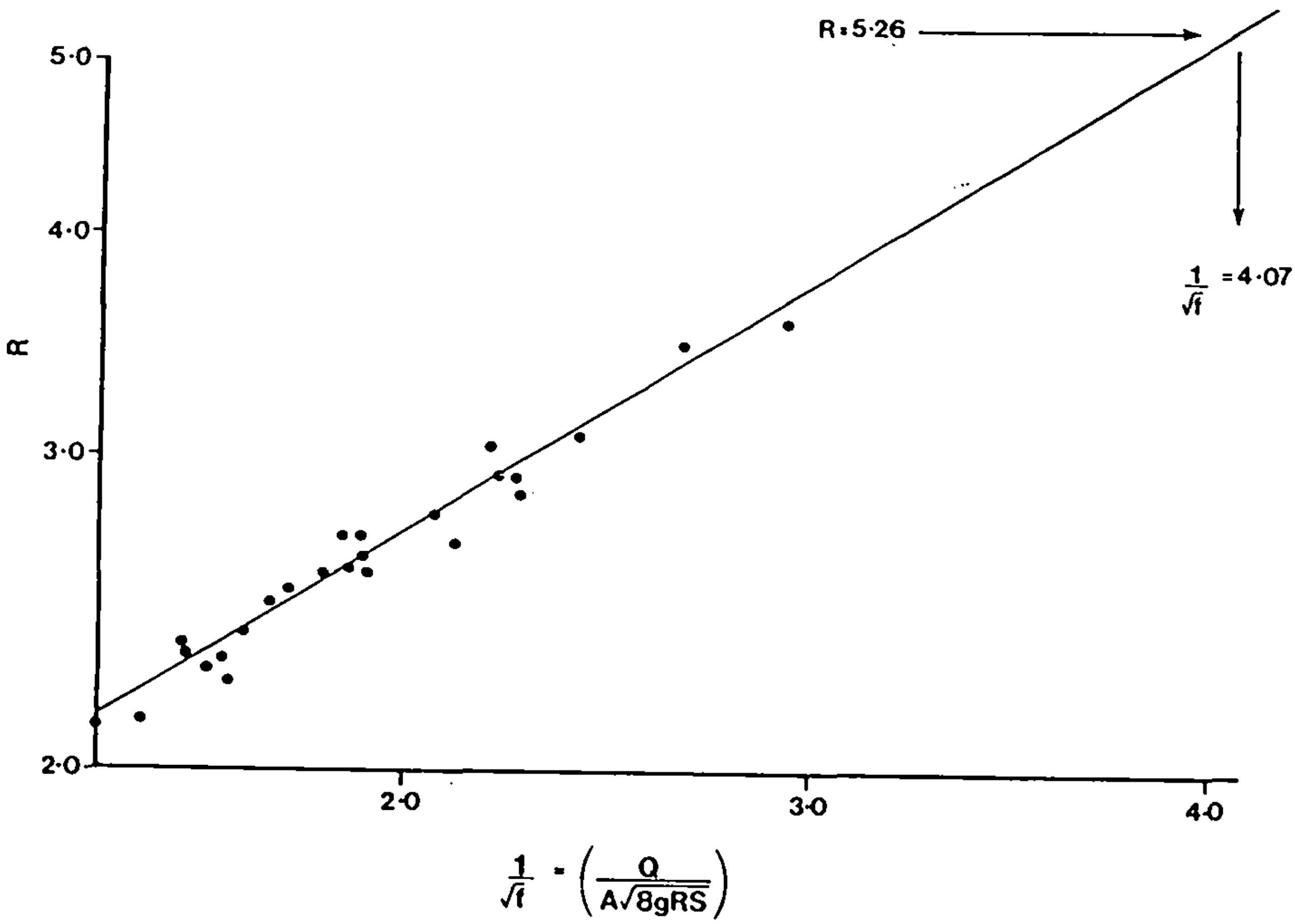


Figure B·4

B.4 Recommended improvements in hydrometry

Since the development of ratings is of such importance in flood hydrology the authors of this report would like to recommend a change in hydrometric practice that should lead to a rapid improvement in rating accuracy.

At present gauging teams drive each day from the local office to one or more gauging stations and return home in the afternoon or evening. Since the majority of heavy rainfalls are thunderstorms occurring late in the day the flood peaks pass the gauging stations after the gauging teams have left. If the teams could be based in the field close to a number of gauging stations and be prepared to gauge floods whenever they occur many flood gaugings would be made resulting in a marked improvement of the derived ratings. The teams should be prepared to gauge at night using a current meter from a cableway or nearby bridge but be wary of using boats especially in fast flowing or debris laden rivers. Floats are most useful in such cases. During periods of dry weather the teams can be usefully employed obtaining accurate channel cross-sections and in general station maintenance.

ANNEX C THE MAF ESTIMATION EQUATION

C.1 Introduction

In Chapter 5 an equation is presented enabling the MAF to be estimated from characteristics of the catchment that can be measured from maps. This annex describes the background to the formulation of this equation.

The general form of the relationship between particular catchment characteristics and the magnitude of floods is often obvious; for example, bigger catchments have bigger floods. However, to be of any use it is necessary to index both the size of flood and the characteristic of the catchment and to establish a formal relationship between the two. The index flood used in this study is the mean annual flood. The size of a catchment is given by its area although an alternative index would be main stream length. It is not possible to describe the relationship between MAF and area as a precise physical model but it is possible to develop a simple relationship based on observed values of the two indices. Values of MAF can be plotted against area and any observed relationship can be represented by a line on the figure. The subjectiveness of this can be removed by using regression analysis which provides an optimal line, in the least-squares sense. If the relationship appears non-linear, then it is necessary to transform the variables before analysis so that linear regression techniques are applicable. Regression analysis enables coefficients of the proposed relationship to be determined, the goodness of fit to be evaluated and a comparison of different relationships. Of course the magnitude of the mean annual flood is not just dependent on catchment size but also on climate, slope, geology and soils, land use, drainage density, catchment shape and storage (lakes). Regression analysis enables an equation to be developed that relates the MAF to indices of these catchment characteristics, either singly or in combination. The specific catchment characteristics used to index the various catchment features are detailed in Annex D.

This multiple regression technique is the one used to develop the MAF estimation equation and while being an empirical approach enables a straightforward flood estimation method based on local data to be established.

The quality of the resulting regression equation is greatly dependent on the data set used to derive it. The data must contain as complete a range of values for each item (MAF and catchment characteristics) as is possible and each value must be accurately determined. Tables A.1 and A.2 contain the data used in this study and which form the basis for all regression analyses. For inclusion in this set the length and quality of the record and accuracy of the rating for flood flows (as detailed in Annex A) for each catchment has been assessed and each catchment is and considered to have a reasonably well estimated mean annual flood. The mean annual floods have been calculated using the methods of Chapters 3 or 4 according to the length of available record.

A preliminary analysis of the data revealed that a transformation of the variables would be required to linearize the relationship between MAF and catchment characteristics. Thus instead of fitting a model of the form

$$MAF = a + bX_1 + cX_2 + \qquad (1)$$

where X_1, X_2, ... are the independent variables (catchment characteristics) all variables were transformed logarithmically. The model has the form

$$\log_{10} MAF = A + B\log_{10} X_1 + C\log_{10} X_2 + \qquad (2)$$

Such an equation can be expressed in terms of the original variables as

$$MAF = 10^A X_1^B X_2^C \qquad (3)$$

C.2 Regression equation for the mean annual flood.

Annex D gives a full description of the catchment characteristics described in this section, although readers should be able to read the following text without recourse to Annex D at this stage. Table C.1 gives the correlation matrix of the transformed variables from which it can be seen that size variables (AREA and MSL) are best correlated with MAF. As there is a very high correlation between the two

TABLE C.1 Correlation Matrix

	MAF	AREA	MSL	S1085	SIMS	AAR	APBAR	FOREST	PADDY	SWAMP	PLANTATION	SHAPE	LAKE
MAF	1.000												
AREA	0.885	1.000											
MSL	0.855	0.951	1.000										
S1085	-0.621	-0.759	-0.732	1.000									
SIMS	-0.621	-0.765	-0.763	0.905	1.000								
AAR	-0.219	-0.417	-0.368	0.380	0.322	1.000							
APBAR	-0.334	-0.636	-0.586	0.555	0.556	0.642	1.000						
FOREST	-0.065	-0.088	-0.098	0.236	0.149	0.234	0.078	1.000					
PADDY	0.187	0.193	0.138	-0.263	-0.152	-0.135	-0.105	-0.674	1.000				
SWAMP	0.033	0.019	0.036	-0.013	-0.070	-0.028	-0.039	-0.118	0.118	1.000			
PLANTATION	0.035	0.201	0.218	-0.117	-0.236	-0.163	-0.284	-0.070	0.197	0.028	1.000		
SHAPE	-0.344	-0.344	-0.617	0.291	0.371	0.055	0.160	0.076	0.070	-0.061	-0.149	1.000	
LAKE	0.136	0.292	0.256	0.192	-0.138	-0.243	-0.206	-0.083	-0.231	-0.047	-0.001	-0.035	1.000

Note: All variables have been transformed by taking $\log_{10}$(variable) variables only one will be useful in the regression analysis. It is interesting to note that the slope variables are all negatively correlated with MAF whereas larger floods would be expected from steeper catchments. The explanation of this is that size and slope are also negatively correlated indicating that in the data set the large catchments are flatter than the small ones, and so the slope index acts as a crude index of size too. This intercorrelation between variables makes the selection of the best sub-set of variables for inclusion in the regression equation difficult. However, by building the regression model one variable at a time it is possible to assess at each stage whether the inclusion of any extra variable is justified by a significant improvement in the equation.

Table C.2 gives details of a sequence of regression models with an extra independent variable being added at each stage. These regressions have been performed on a restricted data set; all catchments with a lake index greater than 0.1 were omitted as there were too few for the coefficient of such a term to be estimated reliably. After the inclusion of AREA and APBAR, the rainfall index, the third variable to enter the regression is a second area term, AREA2, which represents the attenuation that a large catchment can impose on a flood as it travels downstream reducing the effect of an increase in drainage area. This term is formed by squaring the log(AREA) value and is included in the regression model as a further independent variable (X_i in equation 2). In terms of the original variables (as in equation 3) this introduces a variable exponent for the AREA term which is itself dependent on AREA. While the inclusion of the slope index, SIMS, is barely significant statistically its inclusion is desirable from a hydrological viewpoint and the coefficient is consistent with values from other studies (eg UK Flood Studies Report, 1975).

To allow for the effects of lake storage on a catchment, a lake index was calibrated by adding the lake term while holding the other coefficients constant. This resulted in a coefficient of - 2.0, which is unreasonably high, although consistent with the locations of the lakes found in the data set. For this reason it was decided to adopt the lake index coefficient from the UK Flood Studies Report. This coefficient is considered to be transferable in this way as the attenuation of floods by lakes is the same process anywhere in the world unlike, for example, the nature of rainfall.

TABLE C.2 Regressions to estimate MAF

Dependent variable	Independent variables	Coefficient	Standard Error of coefficient		R^2	s.e.e.
(i) MAF	Constant	0.633	0.097	6.51	0.788	0.276
	AREA	0.671	0.035	19.40		
(ii) MAF	Constant	-5.086	0.651	-7.81	0.881	0.208
	AREA	0.852	0.033	25.74		
	APBAR	2.640	0.298	8.84		
(iii) MAF	Constant	-4.941	0.650	-7.61	0.885	0.206
	AREA	0.988	0.085	11.56		
	APBAR	2.504	0.306	8.19		
	AREA2	-0.031	0.018	-1.73		
(iv) MAF	Constant	-5.098	0.651	-7.84	0.888	0.204
	AREA	1.020	0.087	11.76		
	AREA2	-0.027	0.018	-1.54		
	APBAR	2.445	0.305	8.02		
	SIMS	0.117	0.070	1.68		
(v) MAF	as above plus					
	LAKE	-2.019	0.533	-3.78	0.889	0.196

Notes: 1. All variables were transformed by taking $\log_{10}$

2. 103 catchments were used for analysis in equations (i) to (iv), 7 extra catchments were included for equation (v).

The recommended five variable equation is

$$MAF = 8.00 \times 10^{-6} \times AREA^{V} \times APBAR^{2.445} \times SIMS^{0.117} \times (1+LAKE)^{-0.85}$$

where $V = 1.02 - 0.0275 \log_{10} AREA$

As is to be expected with any equation of this type there is considerable scatter around the regression line. Figure C.1 shows estimated against observed values of MAF for the complete data set with points labelled by catchment number. The scatter can be expressed statistically by quoting the coefficient of multiple determination, 0.89 and by giving the standard error of estimate of 0.20. This latter value is most useful in assessing the accuracy of using the equation to estimate MAF. In the $\log_{10}$ form of the equation as presented in Table C.2 the estimated $\log_{10}$(MAF) can be expected to be within ± 0.2 of the 'actual' value (68 times out of 100). In terms of the original variables this is the same as saying that in 68 times out of 100 the estimated value will lie between the actual value times 1.59 and actual value divided by 1.59. It should be noted that this large error can often be reduced by the use of local data as described in Chapter 8.

Various other regression models were considered in which the data set was divided regionally (Java and Sumatra) or according to catchment characteristic values (both AREA and APBAR were tried). However none of these alternatives produced a significant reduction in error and justified the increased complexity of such a scheme. Figure C.2 shows the geographical distribution of the factorial error of estimate from the equation and reveals no significant regional trends.

Observed MAF plotted against MAF estimated by regression equation

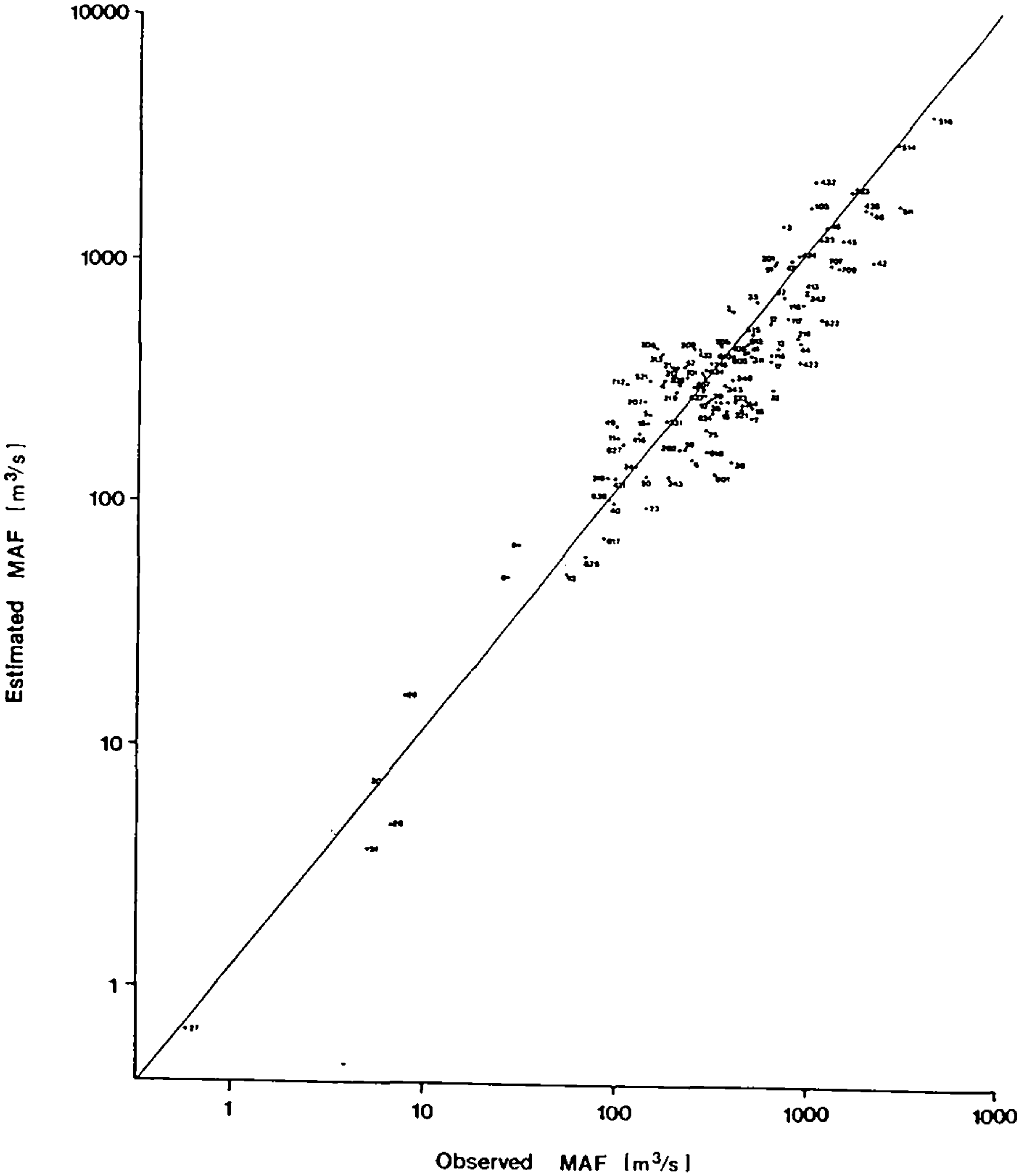


Figure C-1

Distribution of factorial standard errors of the regression equation

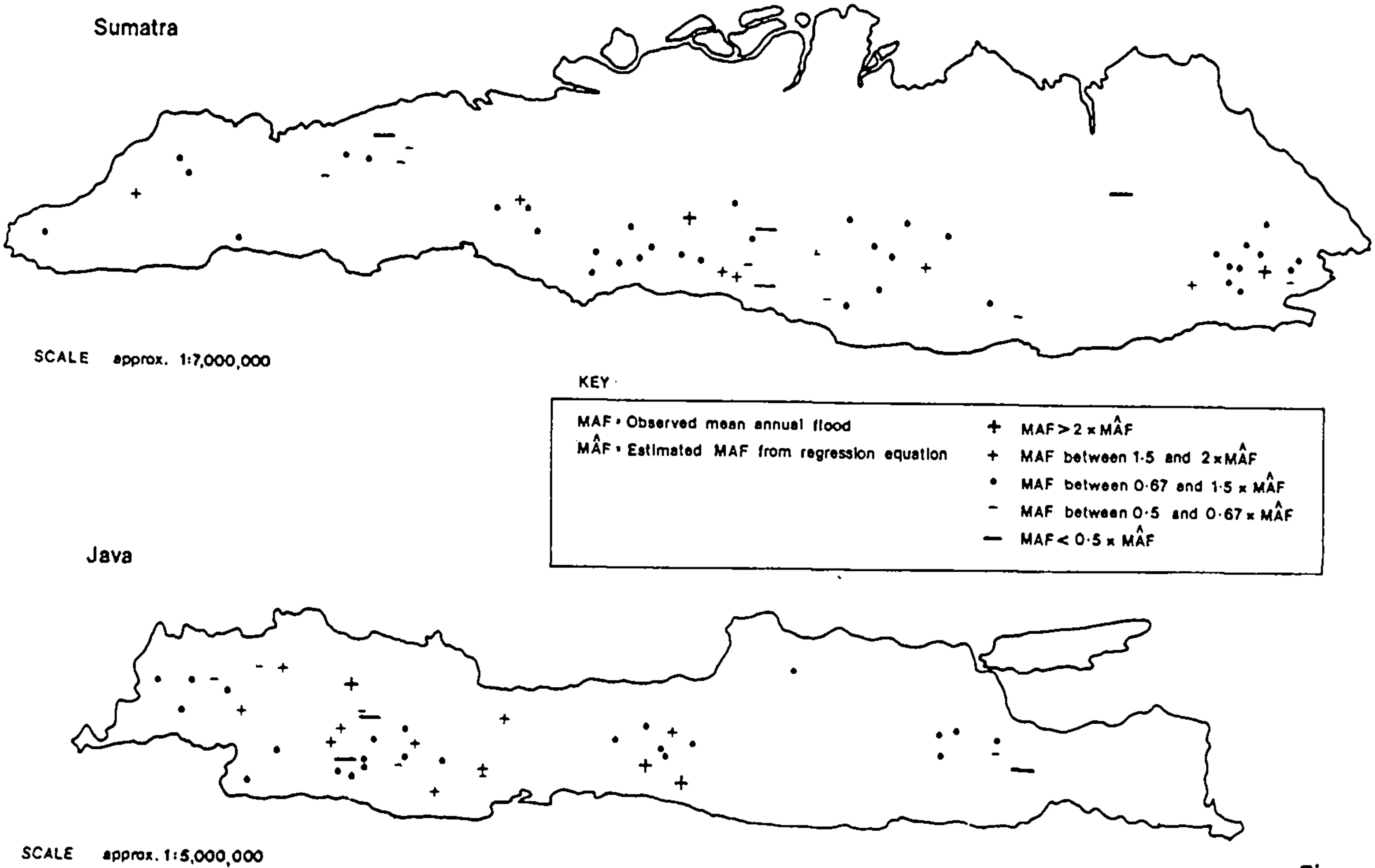


Figure C·2

ANNEX D CATCHMENT CHARACTERISTICS

D.1 Introduction

In Annex C in which the MAF estimation equation is developed several types of catchment characteristics are mentioned as being of potential usefulness in indexing the variation of flood magnitudes. This Annex gives a full description of the catchment characteristics used in this study, not just those appearing in the final MAF equation

The characteristics can be divided into seven categories as in Table D.1 which also gives the specific characteristic used to index a particular catchment feature. The categories represent aspects of catchment physiography that are known to influence flood response from either physical principles or intuitive reasoning. Although some of the characteristics represent parameters that might appear in a physics based catchment model, in the present context the variables are used as indices of catchment response only. For this reason many catchment features are represented by a simple chjaracteristic easily obtainable from maps rather than a more complicated and physically meaningful quantity.

The following sections give full descriptions of the characteristics listed in Table D.1 together with the information required for their abstraction.

D.2 Catchment Area (AREA)

Catchment area (AREA) is the most important catchment characteristic in indexing the magnitude of the flood peak. AREA is measured in km^2.

In Java, catchment area was measured from the 1:50,000 Topographic maps (US Army Mapping Service (AMS) Series) as the best available maps for this purpose. An almost complete set of these maps for the island of Java was made available by DPMA for the duration of the project. The missing maps were substituted with black and white

Table D.1 Summary of Catchment Characteristics

Category	Symbol	Section	Description	Influence on MAF
Size	*AREA	D.2	Catchment area	Larger catchments should
	MSL	D.3	Mainstream length	produce bigger floods
Climate	*AAR	D.4	Average annual rainfall	Catchments experiencing
	*APBAR	D.5	Mean annual maximum catchment 1 day rainfall	frequent heavy rainfall are more susceptible to flooding
Slope	*SIMS	D.6	Simple slope	Steep slopes lead to
	S1085	D.7	River slope over 10 - 85% MSL	faster propagation of
	S085	D.8	River slope over 0 - 85% MSL	floods
Storage	*LAKE	D.9	Lake index	Storage either by lakes,
	S010	D.10	River slope over 0 - 10% MSL (Flood plain index)	reservoirs or on the flood plain attenuates floods
Land type	GEOL	D.11	Geology index	The hydrological proper-
	SOIL	D.12	Soil index	ties of rocks and soils can influence the generation of floods
Land use	FOREST	D.13	Forest index	The land use can both
	PADDY	D.14	Paddy index	modify flood response and
	PLTN	D.15	Plantation index	index other catchment
	SWAMP	D.16	Swamp index	features
Shape	SHAPE	D.17	Catchment shape index.	The shape of the channel network influences the way flooding propagates.

An asterisk '*' indicates that the characteristic appears in the regression equation of Chapter 5. The user of this manual should consult the appropriate sections given below when using this regression equation.

prints obtainable from

Seksi Publikasi
Geological Survey of Indonesia
JL Diponegoro 57
Bandung

It is recommended that the user should, if possible, refer to the original maps of this AMS series to estimate AREA since it is sometimes very difficult to draw catchment boundaries on the black and white copies.

In Sumatra a new set of 1:50,000 topographic maps is currently being produced. At the time (April - May 1982) that maps were required for this project the only ones available from this new series were for Banda Aceh, most of Sumatera Utara (North Sumatra) and Lampung, and a few in Sumatera Selatan (South Sumatra). These maps can be obtained (with suitable authority) from

Bakosurtanal
Jl. Raya Jakarta-Bogor Km 43
Cibinong

or

Jawatan Topografi
TNI AD
Jl. Gunung Sahari
Jakarta

The remainder of Sumatra was covered with an old series of maps at 1:100,000 scale and these were obtained for this project from Jawatan Topografi (address above). It is recommended that, if possible, the new 1:50,000 series maps be used to obtain catchment area in Sumatra.

All topographic maps of Java and Sumatra used by this project are held at DPMA in Bandung.

The river basin area (AREA) should be measured with a planimeter and expressed in units of square kilometres.

D.3 Main stream length (MSL)

Main stream length (MSL) is defined as the length of the longest river channel upstream of the gauging station as defined on the 1:50,000 topographic maps. The main stream length is measured with dividers set to 4 mm. Elsewhere, on the 1:100,000 topographic maps. dividers are set to 2 mm. For both map scales MSL is calculated thus:

$$MSL = ND \times 0.2 \quad km$$

where,

ND = Number of divider steps from gauging station to the top of the longest tributary as defined on the map.

Dividers should be used to estimate MSL in preference to curvimeters as the data set used in the regressions was based on MSL (and also slope measures) obtained with dividers. Dividers should initially be set as close to 4 mm as possible and checked against a millimetre scale over at least 100 mm before and after use. A correction may then be applied to ND to allow for setting errors.

D.4 Average annual rainfall (AAR)

Catchment average annual rainfall, AAR, was obtained from 'Mean rainfall in Java and Madura 1931-1960' (Institute of Meteorology and Geophysics) which contains a 1:1,000,000 scale map of Java with isohyets of average annual rainfall and from 'Mean rainfall in the islands outside Java and Madura 1931-1960' (Institute of Meteorology and Geophysics) which contains a similar map for Sumatra at a scale of 1:3,000,000.

For Java the procedure used to obtain AAR was firstly to enlarge photographically the isohyetal map to 1:500,000 onto transparent paper. Secondly the 1:250,000 topographic maps* containing the

*1:250,000 Joint Operations Graphic (JOG) maps are available in black and white from the Geological Survey of Indonesia (address above). Use of the coloured original maps, if available, is more satisfactory.

catchment boundaries, were reduced to 1:500,000. Catchment boundaries were transferred from the topographic to the rainfall map and areas between isohyets estimated by planimeter.

In Sumatra the procedure was similar except the catchment boundary overlay was photographically reduced to the scale of the rainfall map (1:3,000,000). Although AAR was estimated by planimeter for large catchments, the small scale of the rainfall map made it impractical to use this method on other catchments. AAR was estimated in these cases either by eye or by counting squares on mm graph paper. The technique for counting squares is illustrated by the example for APBAR below.

D.5 Mean annual maximum catchment 1 day rainfall (APBAR)

Mean annual maximum catchment 1 day rainfall, APBAR, is calculated by multiplying PBAR, the mean annual maximum 1 day point rainfall for the catchment, by an areal reduction factor (ARF). PBAR is estimated as follows.

An isohyetal map of mean annual maximum 1 day rainfall (PBAR) has been reproduced from Irish (1981) and appears in this report on Figure 1.1 for Java at a scale of 1:1,000,000 and Figure 1.2 for Sumatra at 1:2,000,000. Contours of PBAR are at 20 mm intervals. The recommended procedure is to draw the catchment boundary on this map and obtain the average catchment value of PBAR as follows:

(a) Draw the catchment boundary on the appropriate 1:250,000 Joint Operation Graphic (JOG) series map. Black and white copies are available from the Geological Survey of Indonesia (address in Section D.2) but use of the originals is much more satisfactory. It is necessary to draw the catchment boundaries on these 1:250,000 scale maps since it is not practicable to reduce the 1:50,000 or 1:100,000 catchment map to 1:1,000,000 or 1:2,000,000

(b) Reduce the topographic map, the catchment boundary and, if possible, a length of coastline to 1:1,000,000 for Java or 1:2,000,000 for Sumatra either photographically or by some suitable method which will ensure accuracy.

(c) PBAR is determined by the weighted average of the map contour values where the proportion of the catchment area between each contour is used as weights. Thus given that 40 per cent of the catchment falls within the contour band of 100 and 120 mm and the remaining 60 per cent falls within the 120 and 140 mm band, PBAR is computed as:

$$\frac{40 \times 110 + 60 \times 130}{100} = 122 \text{ mm.}$$

(d) In certain areas the contours of PBAR are widely spaced and particularly for small catchments, interpolation is required to obtain the best value. For example a catchment lying completely in the area between the 120 mm and 140 mm contours (ie 130 mm band) but closer to the 140 mm line should be given a value of PBAR between 130 mm and 140 mm; the actual value depending on the position of the catchment.

If the catchment is large a planimeter may be used to estimate the catchment average PBAR. If the catchment is small it is better to count squares on millimetre graph paper.

PBAR, which refers to point rainfall, is converted to catchment areal rainfall, APBAR, by multiplication by an areal reduction factor (ARF). To date little work has been done on ARF's in Indonesia. Tabulated areal reduction factors given in the Binnie and Partners, 'Report on Hydrology' (1980) are from work by Dr Doerma during 1923-25 on a 130 km^2 area situated near Jakarta. This table covers the range 0 to 200 km^2 for durations of 30 minutes to 24 hours. Whilst this can be described as hardly satisfactory for the purposes of this study with catchments up to about 20,000 km^2 in area, it was all that was available at the time and is preferable to imported rules for ARF since ARF's are strongly dependent on the local rainfall regime. This relationship, when extrapolated for larger catchment areas, falls midway between the ARF's for Papua New Guinea and the UK, indicating that they are perhaps not unreasonable. The ARF's used in this study are therefore based on the work of Dr Doerma. In any case the effect of any inaccuracies in these ARF's will be eliminated when using the regression equation for estimation of MAF provided the same ARF's are used in design as were used in the development of the regression equation

Catchment area km^2	ARF
1 - 10	0.99
10 - 30	0.97
30 - 30,000	$1.152 - 0.1233 \log^{10}$AREA

The relationship for catchment areas between 30 km^2 and 30,000 km^2 gives a range of ARF's between 0.97 at 30 km^2 to 0.6 at 30,000 km^2.

Example

As an example of estimating APBAR consider catchment 610 shown on Figure 1.2 (Air-Ketaun at Tunggang). This is located across the 120 mm contour of PBAR. If the catchment is traced onto millimetre graph paper the following information is obtained:

Number of millimetre squares in 110 mm PBAR band = 200*

Number of millimetre squares in 130 mm PBAR band = 40*

*These figures are subject to small estimation errors due to line thickness and personal interpretation.

PBAR for the catchment is calculated thus:

$$PBAR_{610} = \frac{(200 \times 110) + (40 \times 130)}{(200 + 40)}$$

$$PBAR_{610} \quad 113 \text{ mm}$$

The ARF is calculated as follows (catchment area 946 km^2)

$$ARF_{610} \quad 1.152 - 0.1233 \log_{10} 946$$

$$ARF_{610} = 0.785$$

Hence APBAR for catchment 610 is estimated as the multiple of

$PBAR_{610}$ and ARF_{610}

$APBAR_{610} = 113 \times 0.785$

$APBAR_{610} = 89$ mm

D.6 River slope (SIMS)

Four indices of stream slope were considered. The first of these is called simple slope, (SIMS), and is the difference in height between the point of interest and the highest point above the end of the mainstream divided by the mainstream length (MSL). The 'highest point' is the highest point on the catchment divide in the vicinity of the source of the longest tributary. Linear interpolation of contours crossing the river is used to estimate the elevation of the point of interest (maps as for AREA). The units of SIMS are m km^{-1}.

D.7 River slope (S1085)

The second measure of river slope, S1085, is calculated as slope over the distance between 10% and 85% of the mainstream length measured upstream from the point of interest. S1085 may be considered to be more representative of the basin as a whole than SIMS because it excludes extremes of slope inherent in SIMS. S1085 was abstracted in a similar manner to that described for SIMS above. The units of S1085 are m km^{-1}.

D.8 River slope (S085)

The third measure of river slope, S085, is calculated as the slope over the distance between the point of interest and 85% of the mainstream length. S085 may be considered to be a measure of catchment slope which is between the extreme SIMS and the more acceptable S1085. S085 was introduced when regressions indicated SIMS was a more significant variable in influencing MAF than S1085. The units of S085 are m km^{-1}.

D.9 Lake index (LAKE)

Storage for flood waters provided by lakes and reservoirs can significantly attenuate downstream flood peaks. The degree of attenuation depends on such factors as the position of the storage within the catchment, its storage/head relationship and, for reservoirs, the operating rules. As it would be impractical to allow for all of these factors in a simple regression model of flooding, the lake index used here is simply a measure of the proportion of catchment area draining through lakes and reservoirs.

The lake index was calculated using the formula

$$LAKE = \frac{\text{Total catchment area upstream of lakes } (km^2)}{AREA}$$

For Java the total catchment area upstream of lakes was obtained from publications giving information on dams over 15 m high and are available at DPMA in Bandung (DPMA, 1980).

In Sumatra the lakes on the catchments used in this study were all natural and the total catchment area upstream of a lake was obtained from the topographic maps (Section D.2).

The regression equation should not be used if LAKE is greater than 0.25. Also, if the total surface area of the lake is less than 1% of the catchment draining through the lake, LAKE is insignificant and set to zero.

The range of LAKE is therefore 0 to 0.25. However, as logarithmic transforms of the lake index are required in the regression, zero values cannot be accepted and it is necessary to add a constant to LAKE. In this study the term which appears in the regression is (1 + LAKE).

D.10 Flood plain index (S010)

This is in fact a measure of river slope and is calculated as the slope between the point of interest and 10% of the mainstream length. S010 was introduced as an experimental variable under the hypothesis that a flood plain was more likely to occur where the river just upstream of this point of interest is very flat. The units of S010 are m km^{-1}.

D.11 Geology index (GEOL)

Information on geology in Java and Sumatra was obtained from the 1:2,000,000 scale map produced by the Direcktorat Geologi Indonesia and the United States Geological Survey.

Unfortunately the description of rock types on the maps was insufficient to categorize accurately each type according to permeability. Furthermore, at the time of the analysis in the UK. no-one was available with suitable knowledge of Indonesian geology to provide assistance. However, an attempt was made to classify the rock types into three classes of permeability and the fraction of each within each catchment estimated by eye. The geology index, GEOL, was calculated thus:

$$GEOL = (3 \times I) + (2 \times M) + (1 \times P)$$

where

I = fraction of catchment area impermeable
M = fraction of catchment area moderately permeable
P = fraction of catchment area permeable

D.12 Soil index (SOIL)

Soil maps at a scale of 1:250,000 were obtained from

Lembaga Perelitian Tanah (LPT)
Jl. IR. H. Juanda 98
BOGOR

Experience elsewhere suggests that a measure of soil type is a useful but not highly significant variable in the regression equation. Classification of the various soils into groups of runoff potential required considerable further specialised work Unfortunately this was outside the scope of the present study and therefore a soil index has been omitted.

D.13 Forest index (FOREST)

Land use information for this study was obtained as a series of 1:50,000 maps detailing Land Use throughout Java and most of Sumatra. The maps, which are in black and white. were obtained from:-

Departemen Dalam Negeri
Direktorat Jendral Agraria
Direktorat Tata Guna Tanah
Jl. Sisingamangaraja
Jakarta.

An overlay of the boundary for each catchment was prepared on transparent paper from the 1:50,000 series maps and positioned on the Land Use maps. The new series of 1:50,000 scale topographic maps for parts of Sumatra (Section D.2) also contain some land use information. This was used in preference to the Land Use maps mentioned above whenever possible as it was easier to abstract and also more uptodate.

Land use information was traced through and the total area of forest determined by planimeter. The forest index is calculated using the formula

$$\text{FOREST} = \frac{\text{Total area of forest (km}^2\text{)}}{\text{AREA}}$$

FOREST ranges from 0, for no forest cover, to 1 for complete forest cover. In order to allow logarithmic transformation of the FOREST index a constant of 1 was added in the regressions.

D.14 Paddy index (PADDY)

The paddy index was calculated using the formula:

$$\text{PADDY} = \frac{\text{Total area of paddy (km}^2\text{)}}{\text{AREA}}$$

The total area of paddy was obtained from the same source and estimated in a similar manner to the forest index, FOREST, described in section D.13.

The term used in the regression analysis was (1 + PADDY), for the reasons given above for FOREST.

D.15 Plantation index (PLTN)

The plantation index was calculated using the formula:

$$\text{PLTN} = \frac{\text{Total area of plantation (km}^2\text{)}}{\text{AREA}}$$

The total area of plantation was obtained from the same source and estimated in a similar manner to the forest index, FOREST, described in section D.13 and a constant of 1 added in the regressions to give an index (1+PLTN).

PLTN was only abstracted for stations in Sumatra and therefore could only be considered in regressions on that sub-set of stations.

D.16 Swamp index (SWAMP)

The swamp index was calculated using the formula:

$$\text{SWAMP} = \frac{\text{Total area of swamp (km}^2\text{)}}{\text{AREA}}$$

The total area of swamp was obtained from the same source and estimated on a similar manner to the forest index, FOREST, described in section D.13.

As with the other land use variables the term used in the regression had 1 added (1 + SWAMP) to avoid zero values in logarithmic transformation.

D.17 Catchment shape index (SHAPE)

It might be expected that the shape of a catchment would influence efficiency of flood generation within the catchment. All other basin parameters being equal, a long thin catchment offers more possibility of flood attenuation than one of compact shape. The shape index used in this study was.

$$\text{SHAPE} = \frac{\text{AREA}}{\text{MSL}^2}$$

SHAPE is effectively the ratio of catchment width to length. MSL is the mainstream length as defined in Section D.3.

ANNEX E. THE PEAKS OVER A THRESHOLD MODEL

E.1 Introduction

Chapter 4 described the application of the POT model to estimate the mean annual flood. The model described was one in which a threshold q_o was chosen and all peaks exceeding this threshold in the complete years of data were abstracted. The resulting POT series consisted of M floods, q_i, from N years of data. The theory behind this model is considered in this annex. It should be noted, however, that this is only one of many possible POT models several of which are described in detail in the UK Flood Studies Report (NERC 1975).

Two variations to the basic model are considered; firstly where data from incomplete years are also available and secondly where a historic series of events has been recorded.

E.2 Theory

The POT series of flood magnitudes are drawn from a conditional distribution as only floods greater than a threshold, q_o, are included. From this distribution it might be observed that 10% of floods exceed a higher value, q, but it would be wrong to state that 10% of all floods are greater than q. The conditional statement that 10% of floods greater than q_o are also greater than q, is much less useful than an unconditional statement relating to all floods. The method of deriving the unconditional statement is basically simple. Suppose that in a given POT sample selected to exceed a 4000 m^3s^{-1} threshold an average of 3 peaks per year are included and that of these floods 10% exceed the higher threshold of 4500 m^3s^{-1}. There is a probability of 0.3 that 4500 m^3s^{-1} will be exceeded in one year or that the return period of this event is 3.33 years. This concept can be expressed more formally by making assumptions about the distributions inherent in the POT model.

The distribution of flood magnitudes in the POT series is assumed to be exponential. Thus, the conditional probability statement that the probability that a flood Q exceeds q, given that q is greater than the threshold q_o, can be written

$$PR(Q > q | q > q_o) = e^{-(q-q_o)/\beta} \qquad (1)$$

where β is the scale parameter of the exponential distribution and the threshold, q_o, is the location parameter (the mean of the distribution is given by $q_o + \beta$). For convenience this probability will be written as $PR(A|B)$.

Given that i floods exceed the threshold in any year the probability that r of these exceed the value q is given by the binomial distribution:-

$$PR(r \text{ peaks} > q|i \text{ peaks}) = \binom{i}{r} (PR(A|B))^r (1 - PR(A|B))^{i-r} \qquad (2)$$

This conditional probability can be expressed in an unconditional form by assuming the probability of i floods occurring in any year is given by p_i. As $r < i$ the unconditional probability that r peaks > q occur in a year is

$$PR(r \text{ peaks} > q) = \sum_{i=r}^{\infty} PR(r \text{ peaks} > q|i \text{ peaks}) \cdot p_i \qquad (3)$$

The probabilities, p_i, of i floods occurring in a year can be assumed to come from the Poisson distribution

$$p_i = \frac{e^{-\lambda} \lambda^i}{i!} \qquad (4)$$

where λ is the mean number of exceedences per year.

Combining equations 2, 3 and 4 gives

$$PR(r \text{ peaks} > q) = \sum_{i=r}^{\infty} \binom{i}{r}(PR(A|B))^r (1-PR(A|B))^{i-r} \frac{e^{-\lambda} \lambda^i}{i!}$$

Redefining the limits of the summation

$$= \sum_{j=o}^{\infty} \binom{j+r}{r}(PR(A|B))^r (1 - PR(A|B))^j \frac{e^{-\lambda} \lambda^{j+r}}{(j+r)!}$$

Note $\binom{j+r}{r} = \frac{(j+r)!}{r!j!}$

$$= \frac{e^{-\lambda} \lambda^r}{r!} (PR(A|B))^r \sum_{j=o}^{\infty} \frac{\lambda^j(1-PR(A|B))^j}{j!}$$

Note $e^z = \sum_{j=o}^{\infty} \frac{z^j}{j!}$

$$= \frac{e^{-\lambda} \lambda^r}{r!} (PR(A|B))^r e^{\lambda(1-PR(A|B))}$$

$$= \frac{e^{-\lambda PR(A|B)} [\lambda PR(A|B)]^r}{r!} \qquad (5)$$

Comparing equations 4 and 5 shows that the distribution of peaks exceeding q is also a Poisson distribution with mean $\lambda PR(A|B)$. In a T year period $T\lambda PR(A|B)$ floods > q would be expected to occur; where only one flood > q is observed in T years q is then the T year flood

$$T\lambda PR(A|B) = 1$$

Combining with equation 1, where q now represents Q(T), the T year flood, gives

$$Q(T) = q_o + \beta \ln\lambda + \beta \ln T \qquad (6)$$

which allows the estimation of any flood Q(T) from the POT series.

Using the POT model as outlined above it has been assumed that the number of threshold exceedences per year is distributed accordingly to a Poisson distribution and that the flood magnitudes of the POT series are distributed exponentially. Neither of these assumptions is stictly true; however, the discrepancy they introduce is likely to be small for low return periods where the flows are not very much greater than the threshold value. One such flood is the mean annual flood.

The distribution of those annual maxima which exceed the threshold q_o can be deduced from the POT model and shown to be from a type 1 extreme value distribution. On the assumption that the entire annual maxima have the EV1 distribution the mean is

$$\mu = MAF = q_o + \beta \ln\lambda + 0.5772\beta$$

It will be noted that this implies the MAF to have a return period of 1.78 years whereas in the Section 1.1 it was noted that the MAF from the annual maximum series had a return period of 2.33 years. The difference arises from the fact that annual maxima approach ignores all except the biggest flood in each year but the POT method can include several floods from a single year or no flood if the annual maximum is less than the threshold. The annual maximum method therefore only considers intervals between years with floods of specified magnitudes rather than the intervals between the floods themselves. The POT approach is in fact the correct one, although in practice for large return periods the difference is slight. The two return periods TPOT and TAM are related by

$$T_{AM} = \left[1 - \exp\left(-\frac{1}{T_{POT}}\right)\right]^{-1}$$

E.3 Incomplete years of data

It is often the case that over the period of operation of a gauge there will be many breaks in the continuity of the record. In the POT model as outlined above only the complete years of data were used and the rest of the data ignored. Although by careful choice of the start date of each year the loss of data caused by interruptions in the record can be reduced, a great deal of potentially useful data is wasted. In the POT model described q_o is fixed and λ and β estimated from the series of peaks. The parameter β, the average exceedence of the threshold, is unlikely to be affected by the inclusion of peaks from incomplete years of data; in fact it should be estimated more reliably if more peaks are used. Parameter λ on the other hand could be greatly influenced if it was assumed that no peaks over a threshold occurred during a period for which there is no record. It is recommended therefore that peaks from the entire record are used to estimate β but that λ is estimated from the complete years of data only. In Chapter 4 examples using both complete years only and all available data are given.

E.4 Historic floods

Historic floods are often recorded as flood marks on a flood stone, or building. The base of the stone or building can be thought of as a threshold exceeded by all the marked floods. In such a case two flood series are available, the historic series of n' exceedences over the high threshold q_o' (corresponding to the lowest possible mark) and the recorded series of n exceedences over the lower threshold, q_o. In such a case the parameter β can best be estimated by

$$\beta = \frac{1}{n'+n} \sum_{i=1}^{n'+n} \left(q_i - \frac{n'q_o' + nq_o}{n'+n}\right)$$

λ should be estimated from the recent series only

E.5 A variation of the POT method

In the POT method outlined in the previous sections and in Chapter 4 the threshold, q_o, was fixed at a level that seemed likely to result in between two and five peaks per year being chosen. From the resulting series the parameters λ, the mean number of exceedences per year, and β, the mean exceedence are estimated. In a variation of the method the number of exceedences per year is chosen and the threshold, q_o, and mean exceedence, β, are estimated from the resulting series. This slight modification of the method allows for the restriction of the POT series originally generated by the use of a threshold that is exceeded too frequently. In the data appendix the listings of the POT analysis firstly give the results from applying the basic method to all of the abstracted peaks and then, under the heading 'POT analysis on a restricted number of peaks', this variation of the method is used in which the number of exceedences per year is reduced, in integer values, to two. While this is the correct method of restricting the POT series, in practice it makes little difference if a new higher threshold is chosen to give the required exceedences per year just by examination of the POT series; in this case λ is then considered to be estimated, and q_o is fixed as in the first case.

ANNEX F. GROWTH FACTOR ANALYSIS

F.1 Introduction

This annex considers the analysis behind the design flood frequency growth factors recommended in Chapter 7.

Individual flood frequency curves relating flood peak to return period may be drawn for any station for which a number of years' data exist. This has been described in Chapter 6. However few stations in Indonesia have a long enough record to enable individual flood frequency curves to be drawn with confidence above the 15-20 year return period. For flood design purposes, however, engineers are commonly interested in return periods in excess of 20 years. How then can estimates of high return period floods be obtained? This is achieved by pooling all the data available and obtaining a consensus on the behaviour of catchments at high return periods.

As flood frequency curves differ greatly from catchment to catchment it is desirable to scale the individual curves prior to pooling. This is achieved by using non-dimensional flood frequency curves (growth curves) in which the flood magnitude scale is divided by an index flood. The index flood is then related to floods of other return periods by dimensionless multipliers or growth factors. The index flood (the mean annual flood, MAF, in this study) is assumed to take into account catchment variables such as area, rainfall, slope etc. However, the growth factors themselves may still have some dependence on the catchment variables.

A compromise is therefore required in the pooling process such that:

(1) Sufficient catchments are grouped to enable the prediction of high return periods floods

(2) Any significant differences in growth factors due to the nature of catchments are not hidden.

The approach adopted for this study was firstly to construct a single overall dimensionless growth curve from all stations in Java and Sumatra. This curve satisfies the first criterion mentioned above where all stations are pooled to enable estimation of high return period floods, but does not permit variation of the growth curve with external factors.

F.2 Pooling of growth curves

This section describes how individual station flood frequency curves were pooled to form an 'average' growth curve for all stations used in this study.

The combined growth curve for all stations was constructed as follows:

(1) For each station a non dimensional growth curve was constructed from the flood frequency curve by dividing each flood on the record by the MAF. In each case the growth curve was stored as a series of points - reduced variate and associated Q/MAF.

(2) An average growth curve was produced by taking the mean reduced variate and mean Q/MAF from all stations within each interval of reduced variate. The intervals of reduced variate used were -1.5 to -1.0, -1.0 to -0.5, -0.5 to 0 etc.

(3) With the individual station record lengths ranging from 5 to 58 years, the smoothed average growth curve was well defined up to a return period of about 100 years. Because this is insufficient for many design purposes, the growth curve was extended by considering the five largest Q/MAF values in the data set

and plotting these as the five largest values in a supposedly independent sample*.

(4) A general extreme value (GEV) function (Flood Studies Report, 1975, Section 1.2) was fitted to the points obtained in steps (2) and (3) above such that

$$Q/MAF = u + \alpha \left(\frac{1 - e^{-ky}}{k}\right)$$

where,

y = reduced variate
u = intercept of fitted curve
α = scale parameter of the fitted curve
k = curvature of fitted curve

Parameters u, α and k were obtained by a least squares approach. The combined curve for all stations had the following parameter values

u = 0.848
α = 0.219
k = -0.2148

It should be remembered that this curve was fitted through points which contained considerable scatter, particularly at high return periods; Figure F.1 shows this curve and the points to which it was fitted. Furthermore, the individual station growth curves

---

*In fact the five largest Q/MAF values are not likely to come from a truly independent sample. Basins may be nested such that there are several gauging stations on the same river and a large flood at one almost certainly implies a large flood at all stations on that river, and possibly on adjacent rivers. However, the inaccuracy introduced by this method is small unless inter-station correlations are very high, which in Java and Sumatra they are not. The five largest floods should be plotted as the five largest in rather less than the number of station years in the group due to inter station correlations, but on the log scale used in plotting the flood frequency curve, the method gives a reasonable means of extending the curve.

showed considerable variation about the mean. The reader should bear in mind, therefore, that this and other smooth growth curves which appear in this annex in fact represent a group of points with considerable scatter.

F.3 Sub grouping of growth curves

The significance of any variation of growth curve shape with catchment characteristics (criterion (2) in Section F.1) was determined as follows:

A list was drawn up of those characteristics considered most likely to index the shape of the growth curve:

(1) Location (The two geographically convenient regions of Java and Sumatra).

(2) Catchment area (AREA)

(3) Average annual rainfall (AAR)

(4) Mean annual maximum catchment 1 day rainfall (APBAR)

For each of the above characteristics, catchments were divided into two groups (Java and Sumatra, large AREA and small AREA etc).

Using the same procedure as described above in Section F.2, pooled growth curves were produced for each of the two groups and tested to see if they were significantly different. A positive indication at this stage resulted in the catchments being divided again (into 4 groups) and the test re-applied. Thus the relationship between any characteristic and growth curve shape could be tested at its most elementary level (2 groups) and if found significant, further divisions of the data set revealed the limit to which the relationship could be adequately defined.

The first sub-grouping of catchments (according to catchment location) therefore had one pooled growth curve for Java and one pooled growth curve for Sumatra. These curves are shown together with the 'all station' curve on Figure F.1. Whether there is a statistically significant difference between growth curves in Java and Sumatra is considered in Section F.5.

Figures F.2 and F.3 show the respective effects of AAR and APBAR on growth curve shape. From these graphs it can be clearly seen that there is no significant difference from the 'all catchment' line by any subgroup.

Figure F.4 is more interesting in that it shows a trend which suggests smaller catchments have a steeper growth curve than larger catchments. The results of the significance tests in Section F.5 reveal whether the difference between growth curves on small and large catchments is statistically significant.

F.4 Significance tests

There are a number of statistical procedures which may be applied to test the significance of the difference of two distributions (Stevens and Lynn, 1978). Of those, the non-parametric χ^2 and Kolmogorov-Smirnov tests have the advantage that they are independent of any assumed plotting position. These tests, therefore, consider purely the distribution of the series of Q/MAF in each subgroup without reference to plotting position. Although the χ^2 and Kolmogorov-Smirnov tests have been shown to give similar results when applied to growth curve differentiation, (Stevens and Lynn, 1978), the Kolmogorov-Smirnov test does have advantages over the χ^2 test (Lilliefors, 1967); furthermore the Kolmogorov-Smirnov test is easy to visualise. The Kolmogorov-Smirnov test was therefore used as the basis of comparison in this study.

The Kolmogorov-Smirnov test tests the hypothesis that two distributions are not significantly different. The first step is to obtain the cumulative frequency distribution for each sample. This is achieved by dividing the Q/MAF range into intervals. In this study 28 intervals were used; 0.5, 0.6, 0.7 followed by 22 steps of 0.05 to 1.8, then 1.9, 2.0 and above 2.0. These intervals allowed roughly the same number of observations in each group. The number

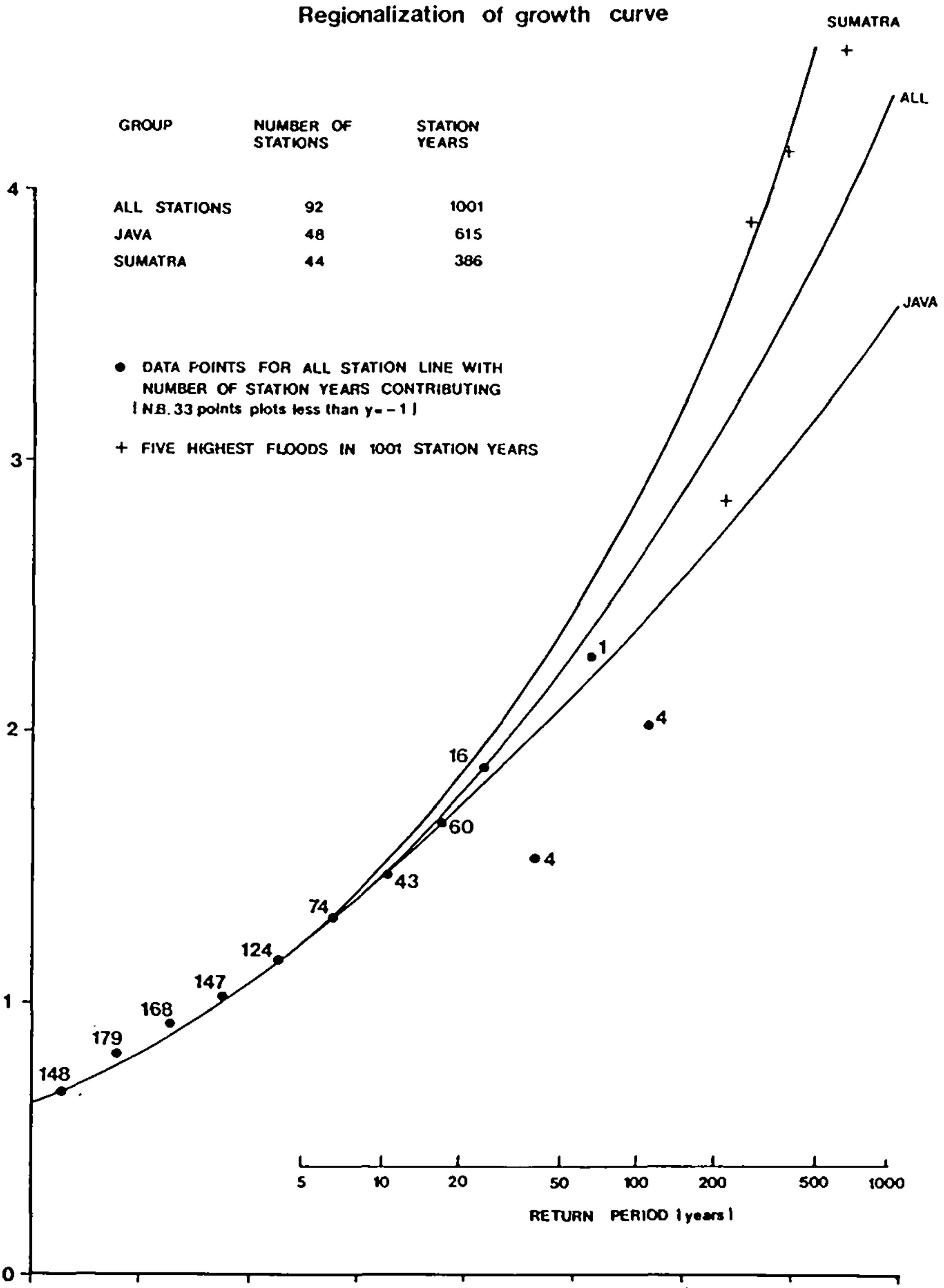

Regionalization of growth curve
SUMATRA
ALL
JAVA
GROUP
NUMBER OF STATIONS
STATION YEARS
ALL STATIONS 92 1001
JAVA 48 615
SUMATRA 44 386
● DATA POINTS FOR ALL STATION LINE WITH NUMBER OF STATION YEARS CONTRIBUTING
(N.B. 33 points plots less than y = −1)
+ FIVE HIGHEST FLOODS IN 1001 STATION YEARS
148
179
168
147
124
74
43
60
16
4
1
4
5
10
20
50
100
200
500
1000
RETURN PERIOD (years)
QT / MAF
Reduced variate y
0
1
2
3
4
-1
5
6
7

Figure F·1

Effect of "AAR" on growth curve shape

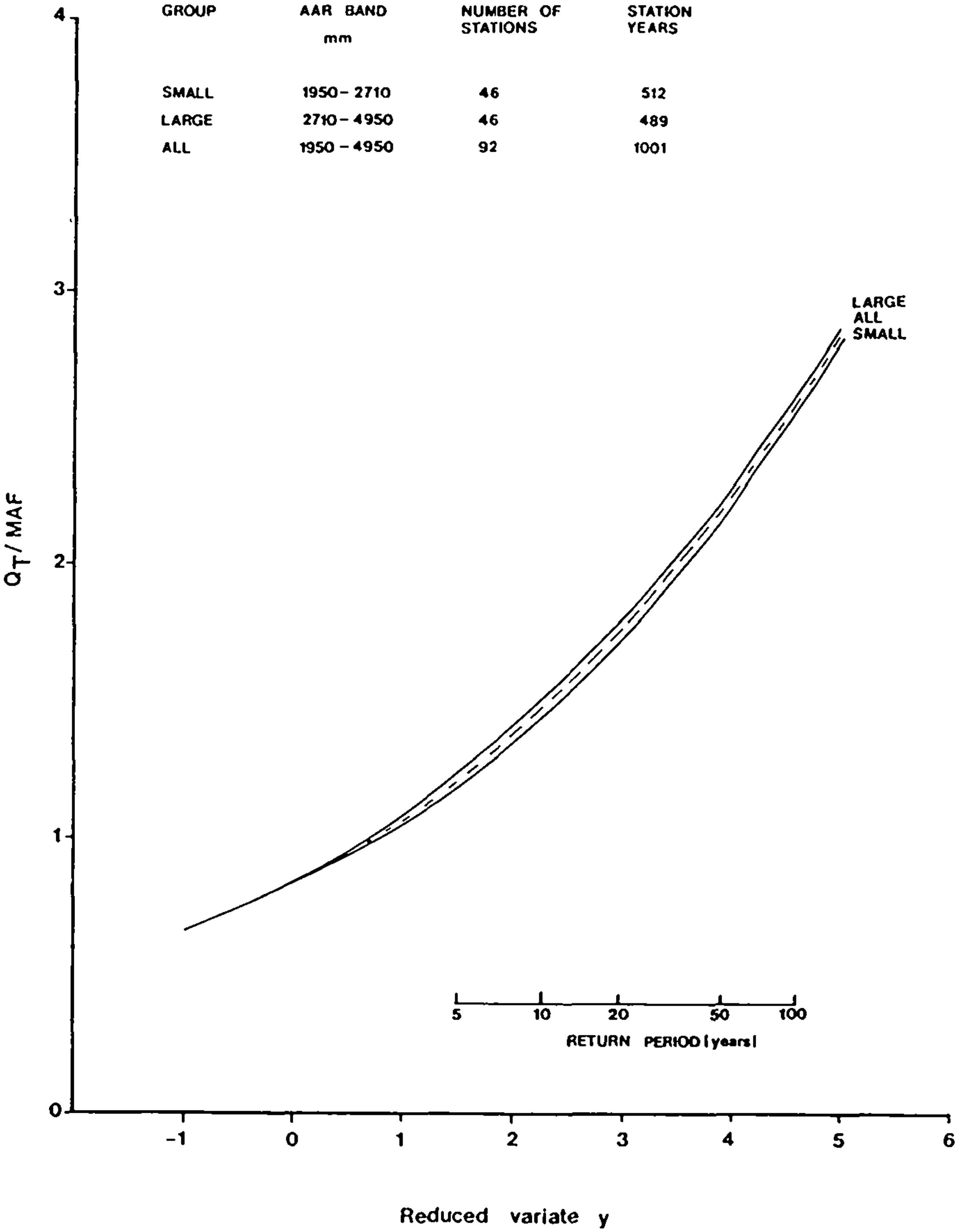


Figure F·2

Effect of "APBAR" on growth curve shape

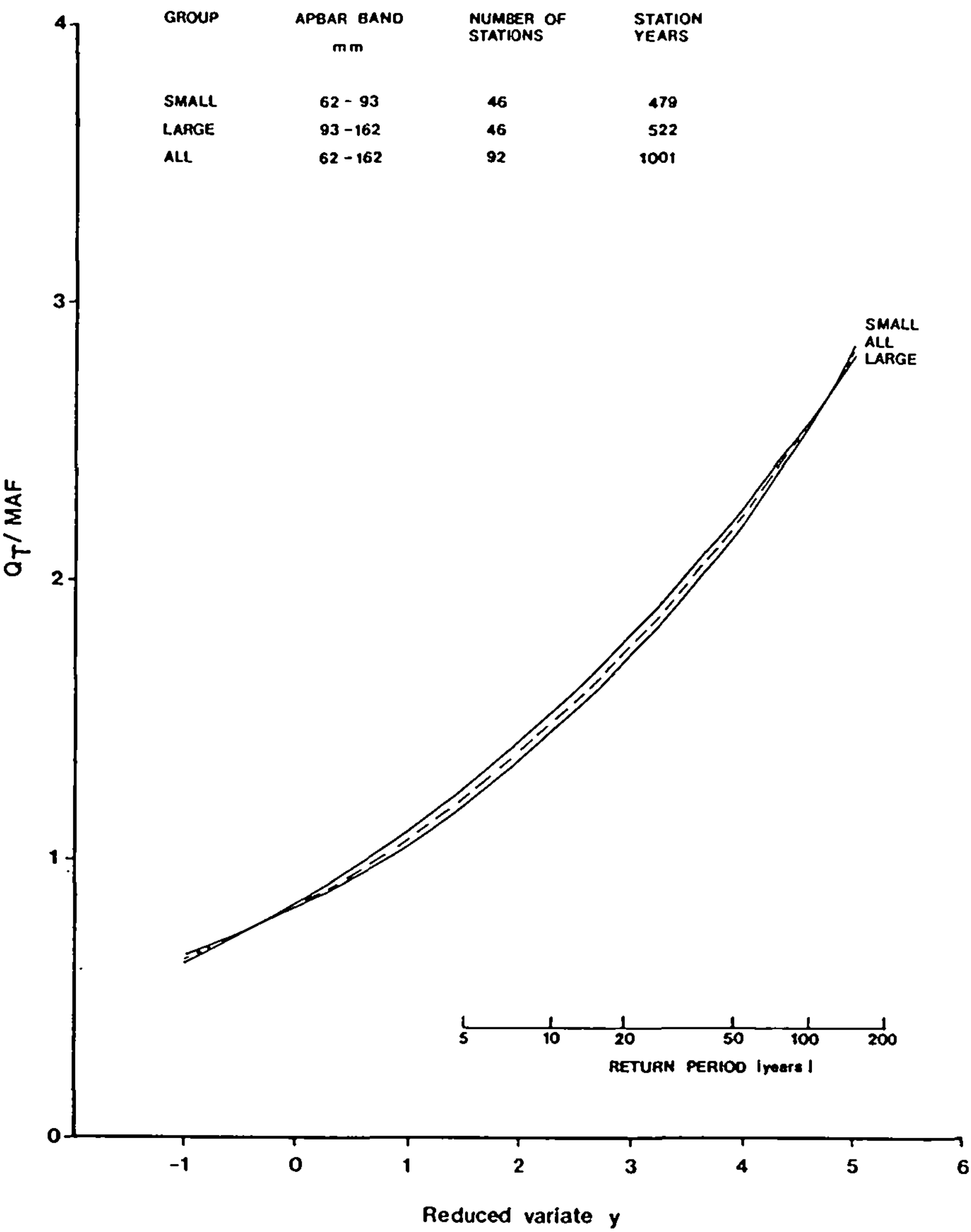


Figure F·3

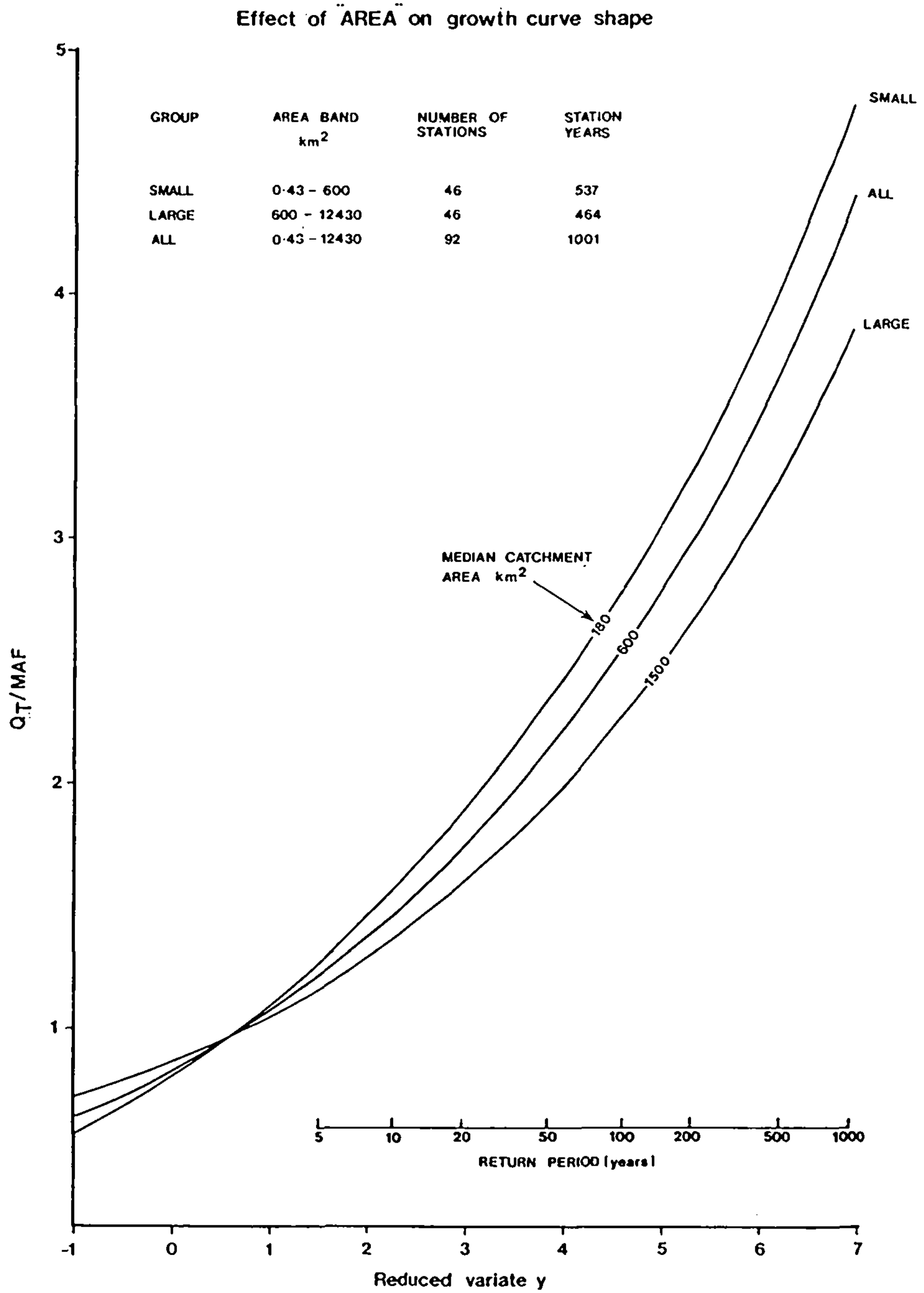

Effect of "AREA" on growth curve shape
GROUP
AREA BAND km²
NUMBER OF STATIONS
STATION YEARS
SMALL 0·43 - 600 46 537
LARGE 600 - 12430 46 464
ALL 0·43 - 12430 92 1001
SMALL
ALL
LARGE
MEDIAN CATCHMENT AREA km²
180
600
1500
QT/MAF
5
4
3
2
1
5 10 20 50 100 200 500 1000
RETURN PERIOD (years)
-1 0 1 2 3 4 5 6 7
Reduced variate y

Figure F·4

of Q/MAF values less than or equal to each interval are determined for each of the two distributions. The two cumulative frequency distributions are obtained by calculating the proportion of the total number of points in each interval for each of the two distributions.

At each Q/MAF interval, the cumulative frequency distributions are subtracted, and the absolute value of this difference obtained. The maximum of these 28 differences is the Kolmogorov-Smirnov test statistic "D".

Table F.1 gives the values of "D" for the pairs of Q/MAF distributions under comparison. The problem now is one of deciding whether the differences between these distributions are significant or not. Normally it is possible to use standard tables to obtain $d_{(0.05)}$ (the 95% confidence limit above which the distributions are dissimilar) if one set of observations is compared with an independent, completely specified, continuous distribution. In our case we are comparing two discrete sets of non-independent observations which invalidates use of tables. A simulation approach was used to obtain estimates of $d_{(0.05)}$ and is outlined below:

(1) For each comparison a simulated series of annual maximum floods was generated using the general extreme value distribution function with the parameters u, α and k obtained from fitting to all 92 stations (Section F.2). This distribution function given in the UK Flood Study Report, Section 1.2.4 is:

$$F(q) = e^{-[1-k(q-u)/\alpha]^{1/k}}$$

If this expression is inverted and F(q) replaced by U, a random number between 0 and 1, the flow generation function used in this simulation is obtained:

$$q = u + \frac{\alpha}{k}\ (1 - (-\log_e U)^k)$$

where $u = 0.848$

$\alpha = 0.219$

$k = -0.2148$

(2) The number of maxima generated for each station was the same as in the distribution under test. For example 48 values of q were generated for the Cianten II at Kracak to represent the 48 years data at that station. These values were then standardised in the same way as the basic data: division by the mean or by 1.06 x Q_{med} if Q_{max} > 3 x Q_{med}. The result was therefore a simulated set of annual maxima, generated from a function which represents how Indonesian catchments behave on average, and processed in the same way as the data under test.

(3) Having generated a series for both distributions being tested a Kolmogorov-Smirnov test was performed and "D" calculated.

(4) This procedure was repeated 100 times to give 100 estimates of "D".

(5) The 100 values of "D" were ranked and the 95th highest obtained. This then becomes our estimate of $d_{(0.05)}$. It is an estimate because only 100 samples have been taken. However the standard error of this estimated significance level at any fixed value of d can be expressed as

$$\text{s.e.} \quad (\hat{p}) = \sqrt{\left(\frac{p(1-p)}{N}\right)}$$

where,

p = the true significance level (0.95)
N = number of samples (100)

Substituting these values for p and N gives a standard error of 2.18%. In fact, in repeated trails of the same experiment 68% of the estimated significance levels at a true level of 95% would be in the interval 93% to 97%. Thus an approximate 95% confidence interval for $d_{(0.5)}$ may be obtained by referring to the values of d at the estimated 93% and 97% points obtained from the simulations.

(6) Table F.1 gives $d_{(0.05)}$ obtained from this simulation procedure for each comparison. Also given is $d_{(0.05)}$ ± one standard error.

Table F.1. <u>Kolmogorov-Smirnov test results</u>

Comparison Number	Groups		$d_{(0.05)}'$	$d_{(0.05)}$ ± s.e.	Accept Hypothesis
	Java stations	0.025	0.056	0.055	Yes
	Sumatra stations			0.059	
	Small Area	0.080	0.060	0.057	No
	Large Area			0.064	
	Small Area (1)	0.065	0.086	0.080	Yes
	Small Area (2)			0.090	
	Large Area (1)	0.072	0.086	0.086	Yes
	Large Area (2)			0.089	

Hypothesis: There is no significant difference between the two distributions.

Criterion : $D < d_{(0.05)}$

F.5 Discussion of results

Figures F.2 and F.3 show that there is little difference in growth curve grouped according to the rainfall indices AAR and APBAR. On the other hand Figures F.1 and F.4 indicate that there is a possibility that the two groupings, regionality and catchment area, may have significant differences in their Q_T/MAF distributions. Because the simulation procedure described above was time consuming, Kolmogorov-Smirnov tests were undertaken only on the regional and catchment area groupings.

Consider firstly the results of the regional groupings (Java and Sumatra) shown in Table F.1. The Kolmogorov-Smirnov "D" from the comparison of the Java and Sumatra distributions of Q_T/MAF (0.025) is well below the 95% significance level of "d" (0.056) calculated by simulation, even allowing for the margin of one standard error in $d_{(0.05)}$. The conclusion, is therefore, that there is no significant difference, according to the Kolmogorov-Smirnov test, in the Q/MAF distributions in Java and Sumatra. An explanation for this is that for return periods up to 20 years, where the bulk of the data occur, the two growth curves are very close (Figure F.1).

Above 20 years return period the Sumatra curve is steeper than that for Java. This is primarily due to the three most extreme floods in all 1001 station years data occurring in Sumatra:

Catchment number	Catchment Name	Date of Flood	Q/MAF
431	Batang Agam at Titi	January 1931	4.598
818	Way Besai at Banjar Masin	March 1981	4.476
316	Batang Anai at Kadang Empat	December 1979	4.146
43	Kali Serayu at Gurung	March 1916	3.858
23	Cikadueun at Cibogo	November 1971	2.845

These three extreme floods are important when constructing the pooled curve. Great weight is placed on these few high Q/MAF values when fitting the growth curve above 50 years return period.

However, these points form only a very small part of the total number of Q/MAF in each group and there are not enough of them to register as a significant difference in the cumulative frequency distributions between Java and Sumatra. Hence the rejection of this grouping of catchments by the Kolmogorov-Smirnov test.

Now consider the results of the Kolmogorov-Smirnov test for the Q/MAF grouped according to catchment area (Table F.1) in conjunction with the growth curves shown in Figure F.4. Unlike the regional grouping discussed above, grouping catchments into those with large AREA (greater than 600 km^2) and small AREA (less than 600 km^2) is significant. The Kolmogorov-Smirnov D for these two distributions, 0.08, is higher than the simulated 95% significant "d" of 0.06, even allowing for one standard error in "d".

From Figure F.4 it can be seen that the growth curves are dissimilar throughout the range of Q_T/MAF (except at the MAF). This is in contrast to the regional grouping, Figure F.1, where divergence only occurred at high return periods. Therefore within the body of the two cumulative frequency distributions, the Kolmogorov-Smirnov test was able to detect at least one part where the divergence of the two distributions was greater than could have been expected by chance. In other words there was a large enough difference in the number of points in one or more particular Q/MAF ranges (as defined in Section F.4) to declare the distributions dissimilar.

With a positive indication that the growth curves of the two AREA groups were significantly different, these two groups were further sub-divided according to AREA. Thus the previous group of 46 small catchments was divided into two groups of 23 catchments again according to catchment area. The original group of large catchments was similarly sub-divided. The purpose of this was to see if the trend of small catchments to have steeper growth curves than larger ones could be defined further. In other words, could the data set support four rather than two significantly different groups of catchments?

The results of this investigation are shown in Table F.1. In both cases the Kolmogorov-Smirnov difference "D" is below the simulated level of $d_{(0.05)}$. These sub-groupings by catchment area show no statistically significant difference and the hypothesis that the growth curves are essentially the same must be accepted.

It is perhaps surprising that there should be a statistical difference between the two main groupings divided at the median area of 600 km^2, but that no difference between sub-divisions of these groups can be detected. The most likely explanation for this is that insufficient data are available in the smaller sub-divisions to adequately define the pooled growth curves. These errors in the sub-divided growth curves would carry forward into the Kolmogorov-Smirnov comparisons so that no clear difference between sub-division growth curves can be detected.

On the basis of the Kolmogorov-Smirnov tests discussed earlier and the authors' observations of individual station growth curves, it is recommended that the division of the data set into large and small catchments be accepted as a sound basis for flood estimation.

F.6 Recommended growth factors

The above analysis has shown that there is a statistically significant difference in growth factor if the catchments are divided into two groups according to catchment area. This section considers how these results were incorporated into the design recommendations given in chapter 7.

In order to simplify the application of the recommended growth curves for users of this report we have replaced the curves by a tabulated set of growth factors in Chapter 7.

Thus we have given the growth factor, or ratio Q_T/MAF, for a range of useful return periods, which we feel is easier for users to apply. The remainder of this section discusses the choice of these recommended growth factors rather than considering the growth curves discussed so far.

It is important that the design recommendations are both easy to apply and credible. With this in mind three options were considered:

(1) Separate design growth factors for catchments greater than 600 km^2 and those less than 600 km^2 (600 km^2 being the median catchment area of all stations)

(2) A continuous relationship between growth factor, return period and catchment area over the whole range of catchment area.

(3) As option (2) but over only part of the range of catchment area.

Option 1 is the easiest to apply. The user simply decides which of two curves is appropriate to the catchment in question. The problem comes around the transition catchment area of 600 km^2. In reality, there is unlikely to be a discrete jump in growth factor at any one catchment area. Some form of continuous relationship is more likely. Option 1, therefore, does not satisfy the criterion of credibility for average size catchments at least.

Options 2 and 3 require the development of a continuous relationship between growth factor and catchment area. With only two groupings of area being significant, the only reasonable relationship would be a linear transition between sets of growth factors. Although not as easy to apply as option 1, since a linear interpolation is required, these two options do recognise that there is some form of continuing decrease in growth factor with catchment area. Option 2 assumes this trend to be continuous throughout the range of catchment areas studied. Considering the looseness of the relationship and the relatively few catchments at the extremes of catchment area, and the fact that the four sub-groupings of AREA failed to produce a significant different growth curves, option 2 was rejected.

Option 3, which permits a continuous change in growth factor over a limited range of catchment areas, was considered to be the most credible and developed as follows into a set of design recommendations:

(1) The median catchment area was found in each of the following three grouping of catchments:

(a)	Small catchments	180 km^2
(b)	All catchments	600 km^2
(c)	Large catchments	1500 km^2

(2) Table 7.1 was constructed by linearly interpolating between the growth factors associated with the three catchment groups in (1) above.

(3) When the catchment area is 180 km^2 or less, the 'SMALL' growth curve (or the first column of growth factors in Table 7.1) is used.

(4) When the catchment area is 1500 km^2 or more, the 'LARGE' growth curve (or the last column of growth factors in Table 7.1) is used.

(5) If the catchment area is between 180 km^2 and 1500 km^2 linearly interpolate between two adjacent columns. For example the 1000 year return period growth factor for a 425 km^2 catchment is calculated thus:

$$Q_{1000}/MAF\ (300\ km^2) = 4.58$$

$$Q_{1000}/MAF\ (600\ km^2) = 4.32$$

$$Q_{1000}/MAF\ (425\ km^2) = 4.32 + \frac{(600-425)}{(600-300)} \times (4.58 - 4.32)$$

$$Q_{1000}\ MAF\ (425\ km^2) = 4.47$$

REFERENCES

Bathurst, J.C. 'Flow resistance of large scale roughness' Journ. Hydraulics Div., A.S.C.E., HY1, December 1978.

Binnie and Partners et al. 'Report on Hydrology'. For the Central Planning Consultancy, Directorate of Planning and Programming, River Basin Planning Division, D.G.W.R.D., D.P.U., Republic of Indonesia, August 1980.

Chow, V.T. 'Open Channel Hydraulics', McGraw-Hill, 1959.

D.P.M.A. 'Penyusunan Inventarisasi Danau dan waduk di Pulau Jawa Tahap I - Jawa Tengah. Seksi Banganan Air, 1980.

D.P.M.A. 'Penyusunan Inventarisasi Danau dan waduk di Pulau Jawa Tahap II - Jawa Barat-Jawa Timur'. Seksi Banganan Air, In press.

Hey, R.D. 'Flow resistance in gravel bed rivers'. Journ. Hydraulics Div., A.S.C.E., HY4, April 1979.

International Engineering Company et al. 'Sederhana Irrigation, Reclamation and Land Development Project - Preliminary Hydrological Study Report on Estimating Low and Flood flows from small watersheds'. For DGWRD, Republic of Indonesia, 1977.

Institute of Hydrology and D.P.M.A. 'Flood Design Manual for Java', November 1981.

Institute of Meteorology and Geophysics (IMG). 'Mean Rainfall and Mean Number of Raindays in Java and Madura', Met. Note No.8, Part I, 1931-60 and Volume 83A 1961-70.

Institute of Meteorology and Geophysics. 'Monthly Rainfall and Monthly Number of Raindays'. Vol. 67 for 1951-52 to Vol.90 for 1977.

Irish, J.L. 'Report on the estimation of short duration design rainfall intensities for Indonesia', Part I, 'One day design rainfall amounts', Draft, 1981

Lilliefors, W.L. 'On the Kolmogorov-Smirnov test for normality with mean and variance unknown'. American Statistical Journal, June 1967, pp 399-402.

Ir Muhadi, Dipl H. 'Menghitung Debit Maximum Dgn. Cara Weduwen, Melchior dan Rasional'. Seminar Cara Perhitungan Design Flood. D.P.M.A., Bandung, 1976.

N.E.R.C. 'Flood Studies Report' In 5 Volumes, available from the Institute of Hydrology, Wallingford, U.K., 1975.

Robertson, A.I.G.S. 'Logarithmic plotting of stage discharge observations'. Water Resources Board, Reading, U.K., Reprinted October 1970.

Stevens, M.J. and Lynn, P.P. 'Regional growth curves'. Institute of Hydrology Report No.52, October 1978.

Sutcliffe, J.V. 'Use of the Flood Studies Report Overseas'. Flood Studies Report-5 years on. Thomas Telford Ltd., London, U.K., 1981.

World Meteorological Organisation. 'Manual on Stream Gauging'. Operational Hydrology Report No.13, Two Volumes, 1980.