Institute
of
Hydrology

A regional analysis of river floods and low flows in Malawi

by
R S Drayton, C H R Kidd,
A N Mandeville & J B Miller

Wallingford
Oxon

Report No 72
July 1980

INSTITUTE OF HYDROLOGY

A REGIONAL ANALYSIS OF RIVER FLOODS AND LOW FLOWS IN MALAWI

by

R S DRAYTON, C H R KIDD,
A N MANDEVILLE and J B MILLER

ABSTRACT

River flow data for Malawi have been examined in a regional analysis to provide methods of estimating floods and low flows for engineering and agricultural applications. Floods for 28 stations have been analysed to produce (a) a regression model for the Mean Annual Flood on catchment characteristics and (b) a regional flood frequency curve. An archive of daily flows for 53 stations was used to produce estimation methods for (a) the average daily flow, (b) the flow duration curve, (c) the low flow frequency curve, (d) storage-yield analysis, and (e) recession forecasting. This study should permit the estimation of a variety of high and low flow measures to be undertaken in Malawi irrespective of the amount of hydrological data available at the site of interest.

REPORT No 72
Malawi Water Resources Division
Report No TP 8

July 1980

CONTENTS

1. INTRODUCTION

Background

Malaŵi's economy has expanded rapidly during the fourteen years since independence; this has brought about a corresponding increase in water development, and pressure to provide more information on the water resources of the country. The government department responsible for the hydrometric network throughout the country has in the past provided information about the behaviour of the rivers to individual requests from outside enquirers on an *ad hoc* basis, but this service has suffered from a number of deficiencies. Many of these enquiries concerned the extreme behaviour of the rivers, namely flood discharges or low flows, which are unreliably measured on many rivers. In addition, many of the rivers of interest to users did not necessarily possess river gauging stations.

The most satisfactory solution to this problem is obtained by a regional analysis of existing hydrological data. In very general terms, regionalisation means the development of procedures for the estimation of hydrological variables irrespective of the amount of data, if any, available at the site of interest. A regional analysis allows the transfer of information from gauged sites to another site where sufficiently detailed information is not available. This report describes a regional study where the type of information required relates to extreme high and low river flows.

This study was proposed by the Water Resources Division of the Ministry of Agriculture and Natural Resources, Malaŵi Government; and was undertaken in cooperation with the Institute of Hydrology, a research station under the Natural Environment Research Council, of the United Kingdom. The part of the studies conducted in the U.K. was financed by the Overseas Development Administration of the U.K. Foreign and Commonwealth Office, and that part conducted in Malaŵi by the Malaŵi Government. The study commenced in March 1978, and was completed in May 1980.

Aims

The aims of the studies were:-

(a) To provide methods of estimation of the reliability of low flows in (i) gauged and (ii) ungauged rivers in Malaŵi.

(b) To provide methods of estimation of the frequency of floods in (i) gauged and (ii) ungauged rivers in Malaŵi.

The results of the low flow study can be used in the design of intakes, reservoirs, and treatment works for irrigation, water supply, and hydro-electric power-generation schemes. The results of the flood study will be useful for the design of bridges, irrigation headworks, river training embankments, and reservoir or barrage spillways.

Analysis methods

For the network of gauged stations the recorded flows were analysed and the various measures of flow variability were presented in a form suitable for design. The studies also provided other techniques by which these measures of flow can be estimated for ungauged catchments, or for catchments with only short records. In the low flow study these measures included (i) the flow duration curve, (ii) the low flow frequency curve, (iii) the storage-yield curve, and (iv) the recession during the dry season. The flood study provided estimates of (i) the mean annual flood and (ii) the flood frequency curve for the estimation of floods of other return periods.

Although the measures of interest are different for the two studies, the basic method of analysis was similar:-

(a) Selection of suitable gauged catchments;

(b) Detailed scrutiny of data to assess reliability of the records;

(c) Extraction of low flow and flood measures from data records for these catchments;

(d) Extraction of catchment characteristics from suitable maps;

(e) Development of relationships between extracted flow measures and catchment characteristics, or countrywide mapping of flow measures.

Items (a) and (b) were undertaken by the Water Resources Division, items (d) and (e) by the Institute of Hydrology, and item (c) jointly.

The studies of Malaŵi records were modelled on the techniques developed by the Institute of Hydrology for previous studies of United Kingdom data (Natural Environment Research Council, 1975; Institute of Hydrology, 1980). Certain modifications were introduced, some in order to analyse data from a different climatic regime, others because certain relationships developed in U.K. did not hold for Malaŵi.

Topography

Malaŵi is located at the southern end of the East African rift valley which dominates the topography of the country (Pike and Rimmington, 1975). Lake Malaŵi, the third largest lake in Africa occupies the northern two thirds of this section of the rift. The lake is at an altitude of 470 m and its only outlet, the River Shire, drains southwards to the lower rift valley at 90 m altitude, before joining the Zambezi River (World Meteorological Organisation, 1976). All rivers in Malaŵi drain eventually into the Shire River, except those

in the eastern part of the country. Here there is an extensive catchment area draining into Lake Chilwa, from which there is no outlet, and a smaller area draining into the lakes and rivers of Mozambique.

The topography of Malaŵi is particularly varied, but the country may be divided into four broad hydrological zones (see figure 1.1) as follows:-

(a) plateau

(b) highland

(c) escarpment

(d) rift valley

The plateau is at an altitude of between 900 and 1,200 m and features broad undulating plains. The underlying rocks are those of the Precambrian Malaŵi Basement Complex. The climate is temperate, and the original vegetation is Brachystegia-Julbernadia woodland. However, these are the most densely cultivated regions of the country, and little of the original woodland remains. An important feature of the plateau is the dambos, broad grass-covered swampy valleys which overlie impervious strata and which become saturated with water during the rainy season. These dambo areas exhibit a peculiar hydrological behaviour, which sets them apart from all other types of catchment within the country.

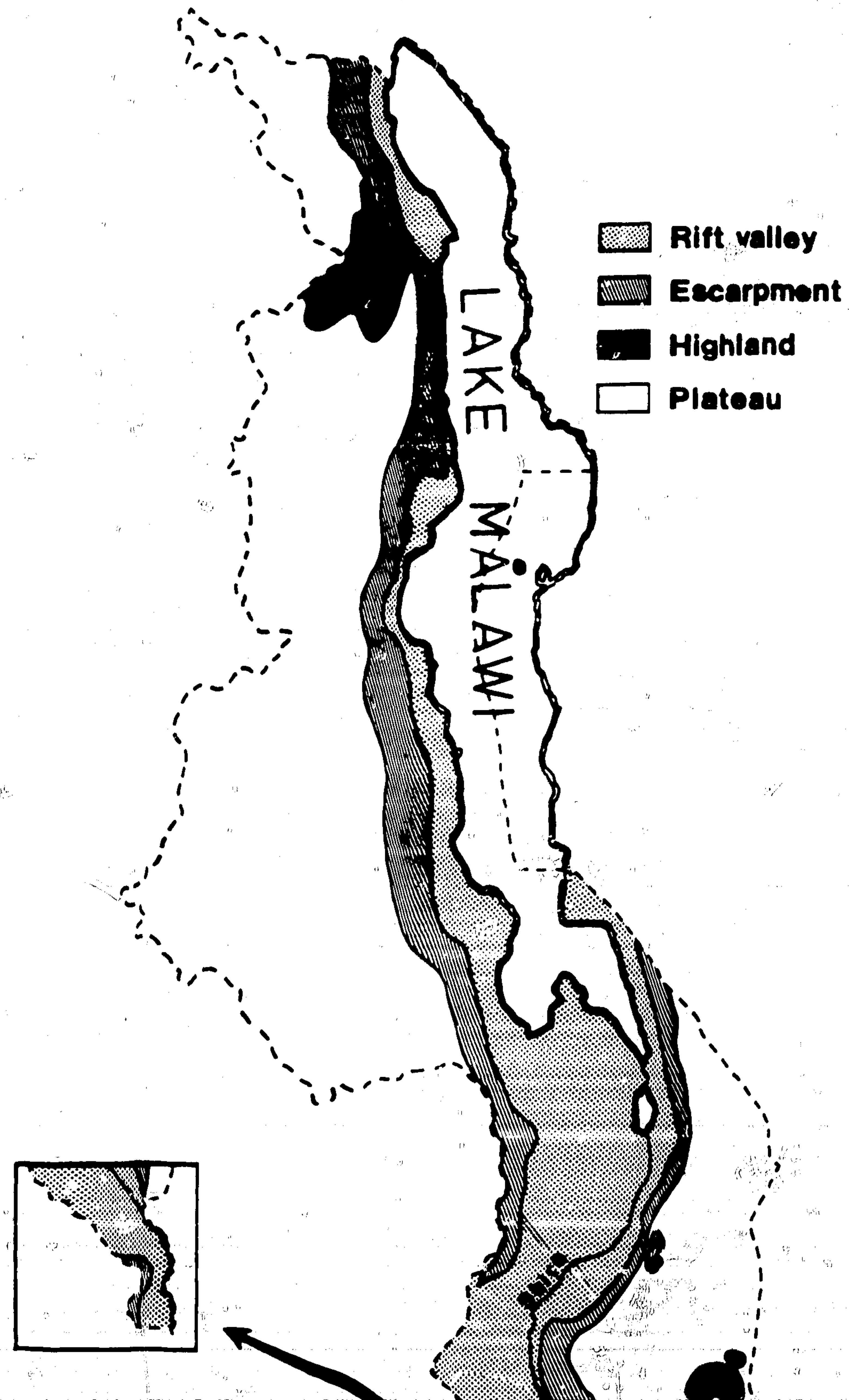
The highland rises abruptly from the plateau reaching altitudes of between 2,100 and 3,000 m. The underlying rocks are granites, phyllonites or syeno-granites. The climate is cool, and the natural vegetation is forest relics and open grassland. These areas are now either forest reserves or game reserves, partly covered with exotic trees.

The rift valley is mainly covered by alluvial deposits of the Quaternary Age. The climate is semi-tropical, and the original vegetation is mixed savannah woodland. The most favourable soils in this zone have been developed into irrigated rice and sugar schemes.

The escarpment marks the boundary between the plateau and the rift valley; it drops down in a series of shelves, and is an area of major faulting. Considerable portions of the escarpment are protected by forest or game reserves, but pressure to secure arable land results in some steep land-forms being cultivated outside these protected areas.

Hydrology

The variation in the average annual values of the primary hydrological variables measured on the catchments used in the studies (Hill and Kidd, 1980) is summarised below:-


Rift valley
Escarpment
Highland
Plateau
LAKE MALAWI

Variable	Mean (mm)	Minimum (mm)	Maximum (mm)
Rainfall	1130	710	2100
Runoff	310	37	980
Actual evaporation	820	630	1100

However the rugged topography, range of altitude, and the temperature reservoir of Lake Malaŵi all ensure that climatic conditions are complex (Agnew and Stubbs, 1972).

In most areas the rainy season extends from November to March, with the maximum in January, resulting from the general migration of the Inter-Tropical Convergence Zone. Convective rainfall over the land occurs normally in the afternoon in the form of local thunderstorms. However, along the lakeshore, storms are commonest in the early morning due to convection over the relatively warmer waters of the lake. Rainfall is also strongly influenced by orography, and the highlands and escarpments directly exposed to the prevailing south-easterly winds receive an annual rainfall which can be three times that of adjacent areas. This effect is particularly marked in April, with the northwards retreat of the Inter-Tropical Convergence Zone, when these areas receive their maximum monthly rainfall, while all other areas receive only light showers.

Two other types of rainfall also occur. Firstly, occasional torrential downpours are associated with the movement inland of tropical cyclones from the Indian Ocean. Secondly, inflow of cool maritime air causes orographic rain, known locally as "chiperoni", over south-east facing highlands and escarpments; this effect is particularly apparent when contrasted with the otherwise fine weather of the dry season, which extends from May to October.

The average annual runoff expressed as a percentage of the average annual rainfall varies between 4% and 54%. The lower runoff occurs in the drier parts of the plateau, where the streams dry up or remain as stagnant pools for between one and six months every year. Higher runoff occurs in the highlands, where most streams are perennial.

The values of average annual actual evaporation may be contrasted with average annual pan evaporation which ranges from 1100 mm in the highlands to 2300 mm along the rift valley (Van der Velden, 197[illegible]). The average annual evaporation from Lake Malaŵi has been estimated as 1610 mm (World Meteorological Organisation, 1976).

2. FLOW DATA

Hydrometric records

In November 1979 there were 159 river gauging stations operating in Malawi. The network covers most districts of the country, with the density of stations highest in the south. The earliest station was opened in 1947, and some of the most important stations have more than 20 years of record.

There is a great variety of controls at these stations, ranging from a V-notch or compound weir, through a solid rock bar or boulders infilled with concrete, to shifting shingle bar, sand bar, channel control, or even no visible control. The gauge boards are normally read twice a day by a villager living close by. At 54 stations these readings are supplemented by charts, mainly monthly, from automatic recorders.

The records from 55 stations had been selected previously for computer processing under a contract with the Southern Water Authority in the U.K. These stations had been chosen either because of their importance as monitors of the country's water resources or because of the high quality of their record. The processed records up to a common end-date of October 1975, which are stored as daily discharges on magnetic tape, were used as basic data for the low flow study.

Selection of catchments

A selection of station records was made for the low flow and flood studies; details are given in Table 2.1 and Figure 2.1. Fifty three stations were used for the low flow study. These correspond to the 55 stations whose records were processed by the Southern Water Authority, less those for stations 1.C.2 Lilongwe and 9.A.7 Lufira. Three stations, namely 1.D.11 (Ruo), 1.G.1 (Shire) and 1.B.1 (Shire), had low flow measures abstracted from their records, but were not included in the regional analysis. The Ruo station was excluded because half the catchment lies in Mozambique, so preventing extraction of catchment characteristics. The Shire stations were excluded because their low flows are dominated by the effects of Lake Malawi.

From the distribution of low flow stations shown in Figure 2.1, it may be seen that there is a need for more low flow stations in:-

Station Ref. No. pre June 1980	Station Ref. No. post Jun '80	River	Low Flow Study Years	Low Flow Study Grade	Flood Study Years	Flood Study Grade
1.B.1	1.B.1	Shire	7	C		
			10	D		
1.D.3	14.B.2	Tuchila	24	D	26	C
1.D.10	14.C.3	Lichenya	4	B		
1.D.11	14.D.2	Ruo	17	C		
1.D.14	14.C.4	Chapaluka	5	B		
1.D.23	14.A.2	Luchenza	17	C		
1.D.24	1.D.24	Kwakwazi			23	C
1.F.2	1.F.2	Tangadzi East	13	B		
1.G.1	1.G.1	Shire	21	C		
1.K.1	1.K.1	Mwanza	23	B		
1.R.3	1.R.3	Rivi-rivi			18	C
1.R.18	1.R.18	Mpamadzi	5	C		
2.B.8	2.B.8	Mulunguzi	5	A		
2.B.21	2.B.21	Likangala	16	D		
2.B.22	2.B.22	Thondwe	15	C	12	C
2.C.8	2.C.8	Naisi	16	B		
3.E.3	3.E.3	Livulezi	7	B		
3.F.3	3.F.3	Nadzipulu	13	A		
4.A.3	4.A.3	Lifidzi	8	B		
4.B.1	4.B.1	Linthipe	12	B	22	E
			10	C		
4.B.3	4.B.3	Linthipe	18	A	12	C
4.B.4	4.B.4	Diamphwe	18	A	20	A
4.D.4	4.D.4	Lilongwe	10	A	23	A
			10	C		
4.D.6	4.D.6	Lilongwe	15	C		
4.E.1	4.E.1	Lingadzi	21	B	14	C
					10	B
4.E.2	4.E.2	Lingadzi	15	A		
4.F.5	4.F.5	Lumbadzi	8	D		
4.F.6	4.F.6	Lumbadzi			4	B-
5.A.8	15.A.8	Lingadzi	13	A	16	E
5.B.13	15.B.13	Kaombe			8	E
5.C.1	5.C.1	Bua			21	B
5.D.1	5.D.1	Bua	16	C	20	B
5.D.2	5.D.2	Bua	21	B	26	C
5.D.3	5.D.3	Mtiti	17	C	21	C
5.E.1	5.E.1	Namitete	18	C		
5.E.2	5.E.2	Bua	13	C		

Table 2.1: <u>List of stations used in low flow and flood studies</u>

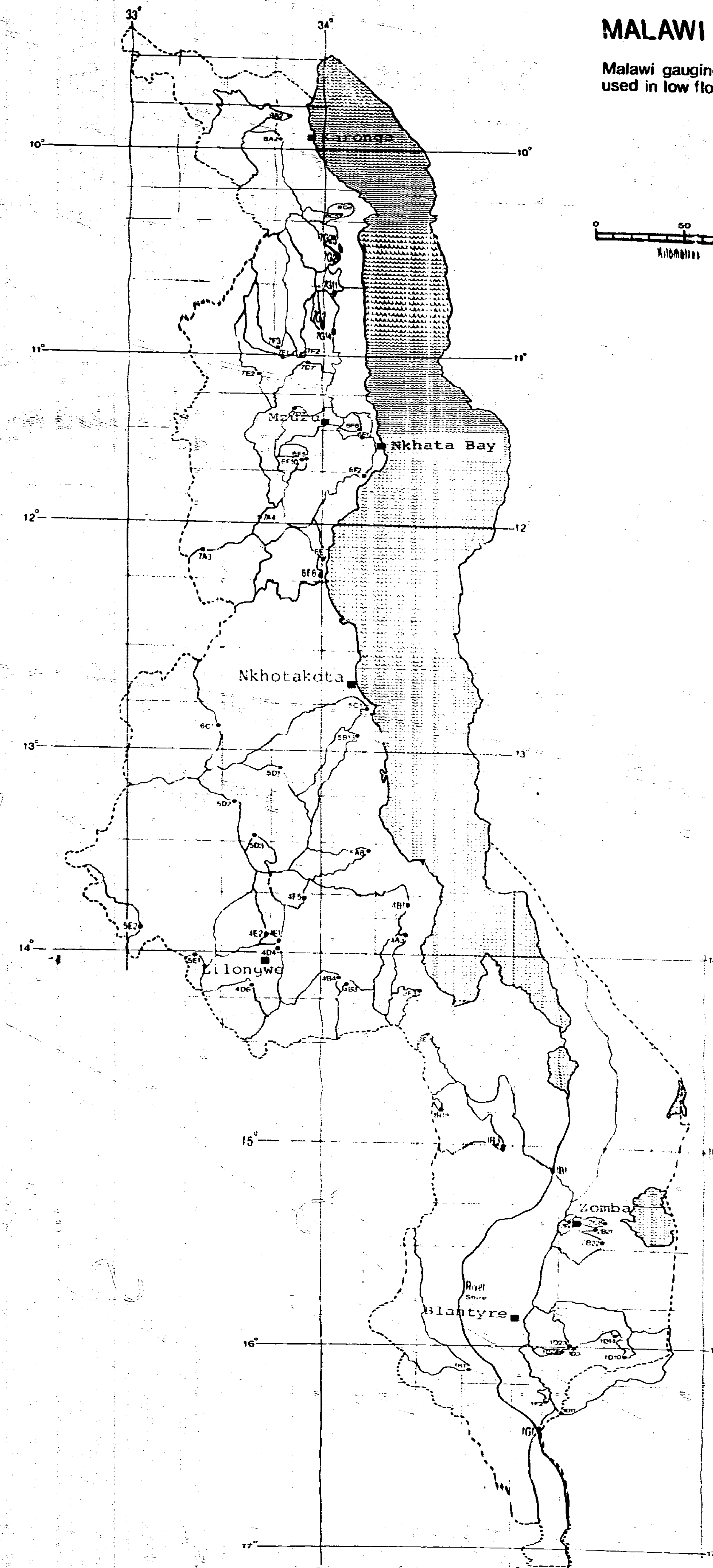
Station Ref. No pre June 1980	Station Ref. No post Jun '80	River	Low Flow Study Years	Low Flow Study Grade	Flood Study Years	Flood Study Grade
6.C.1	6.C.1	Dwangwa	14	B	19	B
6.E.6	16.E.6	Dwambadzi			10	A
6.E.7	16.E.7	Mlowe			7	D
6.F.1	16.F.1	Limphasa	22	B	10	D
6.F.2	16.F.2	Luweya	14	B	14	B
6.F.5		Luchelemu	7	A	20	D
6.F.6	16.F.6	Luwawa	16	A	18	E
6.F.10	16.F.10	Luchelemu	16	B		
7.A.3	7.A.3	South Rukuru	18	B	18	C
7.A.4	7.A.4	Mzimba	11	A		
			7	C		
7.D.3	7.D.3	Lunyangwa	22	C	24	E
7.D.7	7.D.7	Kasitu			23	D
7.E.2	7.E.2	South Rukuru	15	D		
7.F.1	7.F.1	Runyina	20	B		
7.F.2	7.F.2	Chelinda	16	B	17	B
7.F.3	7.F.3	Runyina			9	A
7.G.2	7.H.1	North Rumphi	19	B		
7.G.3	7.G.3	Muhuju	10	B		
7.G.11	7.G.11	Kambwiya	20	A		
7.G.14	7.G.14	South Rukuru	18	C		
7.G.15	-	Kaziwiziwi *	11	A		
7.G.25	7.H.2	Kaziwiziwi *	4	B		
8.A.2	8.A.2	North Rukuru	22	B	23	D
8.A.5	8.A.5	North Rukuru			9	B
8.C.5	8.C.5	Wovwe	8	B		
8.C.6	17.C.6	Wovwe	5	E		
9.A.2	9.A.2	Lufira	17	B		

* 7.G.15 and 7.G.25 were combined and treated as one continuous station.

Table 2.1 (continued): List of stations used in low flow and flood studies.

MALAWI

Malawi gauging stations
used in low flow and flood studies



STATION NUMBER	RIVER	STATION NAME	LOW FLOW DATA SET	FLOOD DATA SET
1.B.1	SHIRE	LIWONDE	•	
1.D.3	TUCHILA	LUCHENZA	•	•
1.D.10	LICHENYA	MINI-MINI	•	
1.D.11	RUO	SANKHULANI	•	
1.D.14	CHAPALUKA	CHAMBE PLATEAU	•	
1.D.23	LUCHENZA	STEVENS ESTATE	•	
1.D.24	KWAKWAZI			•
1.F.2	THANGADZIEAST	MLOLO	•	
1.G.1	SHIRE			
1.K.1	MWANZA	TOMALI	•	
1.R.3	RIVI-RIVI			•
1.R.19	MPAMADZI	NCHEU	•	
2.B.8	MULUNGUZI	ZOMBA PLATEAU	•	
2.B.21	LIKANGALA	KOKANGUO ESTATE	•	
2.B.22	NTHONDWE	JALI	•	•
2.C.[illegible]	NAISI	MWANDAMA	•	
[illegible].E.[illegible]	LIVULEZI	KWEKWELELI	•	
[illegible].F.[illegible]	NADZIPULU	NTAKATAKA	•	
4.A.[illegible]	LIFIDZI	KADYAPULA	•	
4.B.1	LINTHIPE	SALIMA	•	•
4.B.[illegible]	LINTHIPE	LINTHIPE	•	•
4.B.[illegible]	DIAMPHWE	DEDZA-LILONGWE RD.	•	•
4.D.[illegible]	LILONGWE	LILONGWE	•	•
4.D.[illegible]	LILONGWE	MALIGUNDE	•	
4.E.[illegible]	LINGADZI	LILONGWE	•	•
4.E.[illegible]	LINGADZI	PHELENI	•	
4.F.[illegible]	LUMBADZI	KAPULULA	•	
5.A.[illegible]	LINGADZI	KAMPHUNGU	•	•
5.B.[illegible]	KAOMBE			•
[illegible]	BUA			•
[illegible]	BUA	BUA DRIFT	•	•
[illegible]	BUA	KASESE	•	•
[illegible]	MTITI	MPONELA	•	•
[illegible]	NAMITETE	NAMITETE TR. CENTRE	•	
[illegible]	BUA	TEMBWE	•	
[illegible]	DWANGWA	KASUNGU/MAZIMBA	•	•
[illegible]	DWAMBADZI			•
[illegible]	[illegible]			•
[illegible]	LIMPHASA	NKHATA BAY RD.	•	•
[illegible]	LUWEYA	CHINTECHE	•	•
[illegible]	LUCHELEMU	MAZAMBE	•	•
[illegible]	LUWAWA	TIMBIRI RD.	•	•
[illegible]	LUCHELEMU	MAZAMBA	•	
[illegible]	SOUTH RUKURU	LUNDAZI	•	•
[illegible]	MZIMBA	CHIBUTHI	•	
[illegible]	LUNYANGWA	EKWENDENI	•	
[illegible]	KASITU			•
[illegible]	SOUTH RUKURU	LAKE KAZUNI	•	
[illegible]	RUNYINA	CHIKWABA	•	
[illegible]	CHELINDA	RUMPHI	•	•
[illegible]	RUNYINA			•
[illegible]	NORTH RUMPHI	[illegible]	•	
[illegible]	[illegible]	[illegible] TR. CENTRE	•	
[illegible]	KAMBWIYA	NCHENACHENA	•	
[illegible]	SOUTH RUKURU	PHWEZI	•	
[illegible]	KAZIWIZIWI	CHOKA	•	
[illegible]	NORTH RUKURU	BORDER ROAD	•	
[illegible]	WOVWE	NJALAYANKUNDA	•	
[illegible]	WOVWE	KAFIYIRA	•	
[illegible]	LUFIRA	NGERENGE	•	

Figure 2.1

(a) the Karonga and Chitipa areas north of the Nyika highlands

(b) the lakeshore area between Chintheche and Chia lagoon

(c) much of the area south of Lake Malaŵi

All 159 current stations were reviewed for possible inclusion in the flood study. The majority were immediately eliminated, principally because at high flows either (i) there were insufficient data to construct stage-discharge ratings or (ii) stage readings were missing or unreliable. The remaining 54 were reviewed in detail and a further 24 found to be of inadequate quality. Data from the remaining stations were thoroughly reviewed as described below. Of this total of 30 stations, 20 were common to both the low flow and flood study.

For a country with such diverse topography and climate, a total of 30 stations is barely sufficient for a regional study, but the 30 stations do include representatives of all of the major hydrological zones, with the exception of the low rainfall areas of the southern escarpment and rift valley in the South. For a good multiple regression analysis, it is necessary to have as wide a variation as possible in each variable, and the 30 stations do show considerable variation.

From Figure 2.1, it may be seen that the number of flood stations is inadequate in:

(a) the Karonga and Chitipa areas north of the Nyika highlands

(b) the streams originating from the Nyika highlands

(c) the whole region south of a line between Dedza and Salima.

<u>Low flow information</u>

The computer tape containing the daily discharges that had been returned from the Southern Water Authority was sent to the Institute of Hydrology. In addition, a large quantity of corrections to the daily discharges were discovered on checking of the computer listings, and these corrections were also provided. The rating curves used for the computer processing of twice - daily gauge readings were accepted without further amendment for use in the low flow study. The quality of the rating curve was assessed as (i) well-defined, (ii) mediumly - defined, or (iii) poorly - defined according to the criteria shown in Table 2.2. If a particular rating curve was extremely poor, that period of record was recommended to be omitted from the low flow analysis.

Upstream influence affecting the record were grouped into the four major categories: (a) surface abstractions, (b) surface effluents, (c) ground-water abstractions, and (d) reservoirs. The combined effect of these influences was labelled (i) negligible, (ii) minor, or (iii) major, relative to the normal low flows in the rivers. No attempt was made to correct the daily discharges recorded at the

Rating curve quality	Comments
Well - defined:	Discharge measurements at all gauge heights at lower end of rating curve. Scatter confined to ± 10% of variables. Theoretical weir formula with several check discharge measurements.
Mediumly - defined:	Discontinuous discharge measurements along lower end of rating curve, scatter more than ± 10%, but curve still fairly obvious. Theoretical weir formula with few or no check discharge measurements.
Poorly - defined:	Few discharge measurements, widely scattered, and position of rating curve open to subjective decision.

Table 2.2: Criteria used to judge quality of rating curve at a low flow station.

	Effect of upstream influences		
Rating curve quality	Negligible	Minor	Major
Well-defined:	A	B	C
Mediumly-defined:	B	C	D
Poorly-defined:	C	D	E

Table 2.3: Criteria used to allocate grade to records from a low flow station.

gauging station by adding the abstraction flows or by subtracting the effluent flows. Fortunately, Malaŵi river records generally commenced before these upstream effects were of much consequence, and the majority of records represent the natural state of the river. However, these upstream influences are now starting to make an appreciable change to the low flows in parts of the country.

The grade letter allocated to a low flow record is based both on the quality of the rating curve and the effect of upstream influences, as shown in Table 2.3. The grade may be adjusted subjectively by up to one letter taking into account the quality and continuity of the gauge readings themselves.

Flood information

The stage-discharge rating for high flows at each of the 30 flood stations was thoroughly reviewed. For just a small number out of those stations common to both studies, the established rating curves as used in the computer processed daily data for the low flow study were also used in the flood study. Recent flood measurements were used to improve earlier ratings wherever the hydraulic control could be considered to be sufficiently stable.

The twice-daily records were carefully scrutinised, and the stage reading giving the maximum flood in each water-year (November to October) was extracted. Since so much reliance is placed on single observations by gauge readers, these were checked very carefully, and cross-checked with both rainfall records and with stage records from neighbouring stream-flow stations. Very few reliable chart records were available, but, where possible, these were used to confirm the maxima, and extracted as above to provide stage readings for instantaneous maxima.

The maximum flows were calculated and tabulated for each available year of record. Where the gauge-reader's notes, or chart records or other information such as wrack-marks provided a gauge height of an instantaneous maximum this was tabulated too. These were finally combined to give the highest estimated flood in each water year.

These highest floods were tabulated, together with comments relevant to their interpretation, the grid reference of the gauge and its catchment area. Each station was then given a reliability grading, based principally on the quality of the flood rating, but modified by a subjective assessment of the quality of the gauge readings, as shown in Table 2.4. Dates of the maxima were also given so that a check could be made for independence of events when combining them in a regional analysis.

Details of historic floods from outside the period of record were also included, but these were generally of very low reliability.

Grade	Comments
A:	Stable, well-defined rating. Good control. High discharge measurements taken by current meter.
B:	Stable rating defined over most of its range. Good control. High discharge measurements taken by direct or indirect measurement.
C:	Shifts in rating at low flows only. Rating defined at upper limits
D:	Shifts in rating at low and medium flows. Rating defined at top end.
E:	Shifts in rating. Rating extrapolated to high flows.

Table 2.4: Criteria used to allocate grade to records from a flood station.

Data quality

The full list of catchments used in the low flow and flood studies is shown in Table 2.1. Against each station is shown the length of record and grading allocated for each study.

The length of records used in the low flow study was found by taking the difference between the start and finish year of the record on the computer tape, less any complete years recommended to be omitted from the analysis, less any complete years of missing data on the computer tape. Subsequently data corrections were made to the computer tape, but some other single months still remained missing. Calculations of certain low flow measures require additional years of record to be omitted due to critical missing months, so the length of period should in this case be taken as only an indication of the maximum period of usable record available. The length of records shown for the flood study is the number of annual maxima for those years during which the station was given a grading.

Lengths of record for the two studies are summarized below:-

Type of Study	Length of record (years)						
	26-30	21-25	16-20	11-15	6-10	1-5	Total
Low flow	-	17%	40%	23%	9%	11%	100%
Flood	7%	30%	30%	10%	20%	3%	100%

The low flow records have a mean length of 15.0 years with nearly two thirds of the records falling within the 11 to 20 year bracket. There is a minor peak in the 1 to 5 year bracket due to some of the processed gauge board records at the chart stations being arbitrarily confined to a 5 year length. The flood records have a mean length of 17.2 years. The percentage of records with lengths below 10 years is similar to that of the low flows. Although the percentage of flood stations with records over 20 years is more than twice as much as that for low flow stations, the actual number of stations is very similar at 11 and 9 respectively, of which 6 are common stations. The secondary peak in the 6 to 10 year bracket is partly attributable to the construction of bridges within the road building programme in recent years.

Grades allocated to each station are summarised below:

Type of Study	Grade					
	A	B	C	D	E	Total
Low flow	19%	42%	28%	9%	2%	100%
Flood	13%	30%	20%	20%	17%	100%

It is noticeable that few of the stations common to both studies share good grades, and there is only one station, 4.B.4, Diamphwe,,which features an A grade for both. The grade of the low flow stations varies from A to E; the grade varies enormously, grade A stations being of very good quality, but grade E being of very poor quality. Some of the reasons for poor quality low flow records are summarised below:-

(a) silting up of gauges, particularly at the foot of the escarpment in recent years;

(b) unstable controls such as gravel or sand bars, leading to annual change in rating curve;

(c) gauging stations being visited infrequently by field teams;

(d) lack of information sheets showing changes of gauge zero or location of gauges in past years;

(e) data processed up to 20 years after observation;

(f) theoretical rating formulae for weirs not checked by discharge measurements.

Although the estimates of flood maxima represent the best information available it should be recognised that the quality is generally not very high, with only 43% of stations being Grade A or B. The principal reasons for poor flood records are:-

(a) Some gauging stations inaccessible to vehicles during the wettest months;

(b) doubtful slope-area estimates of peak discharges;

(c) flood peaks on flashy escarpment and other small catchments passing the staff gauges without being recorded;

(d) gauge readings often fabricated or taken only once per day;

(e) outflanking of control section;

(f) gauge washed away or overtopped;

(g) data processed up to 25 years after observation;

(h) lack of high flow measuring structures, such as cableways, bridges or traverse gears;

(i) frequent occurrence of floods at night-time on the lakeshore, when survey teams are not working;

(j) silting up of stilling wells during the flood season.

3. CATCHMENT CHARACTERISTICS

Introduction

The hydrological response of a catchment may depend on several factors including its size, the slopes of the stream and the land surface, the density of the drainage network, the soils covering the catchment and the underlying geology. Though there is no hydrological theory available to predict how a catchment will behave from a knowledge of the above factors many investigators have been successful in deriving empirical equations relating flow statistics to numerical measures of catchment characteristics taken from maps. It is often the case when a project is being designed that no flow data are available for the site. In this case, the catchment characteristics may be measured off maps and the empirical equations used to provide estimates of flow statistics required for project design.

The Water Resources Division provided (i) a complete set of 1:50,000 series and 1:250,000 series topographic maps, and (ii) maps showing climate, geology, soils and agriculture, from which the Institute of Hydrology measured the catchment characteristics. Detailed description of the procedures to be used in measuring catchment characteristics are given below. It is important that those using the results of this study should understand these procedures and follow them carefully.

Characteristics describing catchment size, slope, climate, density of the stream network and the amount of dambo and swamp were used in the study. No indices of geology or of soil, other than the dambo measure, were used. While it is recognised that such indices might be important in estimation of flow statistics they are not suitable as catchment characteristics because they do not easily lend themselves to simple estimation from maps. This does not mean that users of the methods described in this report should ignore such information, which can often be used subjectively. A separate detailed description of the various characteristics is given below and Table 3.1 lists the values for each of the stations.

Catchment area

Catchment area is a fundamentally important characteristic for the simple reason that large catchments have in general higher values of flow statistics than small ones. For most catchments, the area was supplied by the Water Resources Division. The Institute of Hydrology checked the area for some catchments, and in these cases the area was measured from maps of the 1:50,000 series using a D-mac digitiser. Catchment areas are quoted in square kilometres, and this variable is called **AREA**. Under normal circumstances, the catchment divide should be sketched on 1:50,000 series map(s), and the area measured by planimeter or square counting.

Station	AREA (km^2)	STMFRQ (km^{-2})	S1085 (m/km)	DAMBO	AAR (mm)	ESCARP
1.B.1	130 200	-	-	-	-	-
1.D.3	1400	1.65	2.31	.004	1110	0
1.D.10	62.3	3.72	78.6	.000	2800	-
1.D.11	4700	-	-	-	-	-
1.D.14	10.6	7.45	47.5	.000	2090	-
1.D.23	500	3.22	8.61	.000	1070	-
1.D.24	6[illegible].5	3.68	15.3	.000	1240	0
1.F.2	45.5	3.50	3.61	.000	1300	-
1.G.1	148 900	-	-	-	-	-
1.K.1	1680	2.05	5.97	.010	990	-
1.R.3	748	2.87	12.4	.000	970	1
1.R.18	7.0	1.71	83.0	.000	1080	-
2.B.8	18.1	1.99	50.0	.000	2060	-
2.B.21	144	1.59	13.3	.000	1430	-
2.B.22	302	1.05	11.1	.000	880	0
2.C.8	75.0	1.07	14.1	.000	1280	-
3.E.3	452	2.80	5.90	.000	1000	-
3.F.3	224	2.07	33.9	.000	1000	-
4.A.3	434	1.97	1[illegible].2	.003	990	-
4.B.1	[illegible]180	[illegible]	[illegible].84	.082	880	1
4.B.3	[illegible]	[illegible]	3.85	.066	910	0
4.B.4	[illegible]0	[illegible]	1.05	.238	830	0
4.D.4	1870	[illegible]	2.06	.037	930	0
4.D.6	763	.48[illegible]	4.43	.017	940	-
4.E.1	928	.19[illegible]	2.97	.169	810	0
4.E.2	585	.144	3.25	.228	840	-
4.F.5	359	.585	6.88	.095	900	-
5.A.8	3[illegible]	2.12	14.2	.000	870	1
5.B.13	430	1.93	7.14	.000	1120	1
5.C.1	10 500	.324	1.59	.284	1100	1
5.D.1	3410	.164	0.52	.215	900	0
5.D.2	6790	.076	0.92	.272	900	0
5.D.3	233	.159	7.60	.100	710	0
5.E.1	147	.333	6.11	.032	930	-
5.E.2	348	.178	5.46	.126	910	-
6.C.1	2980	.304	2.02	.117	745	0
6.E.6	778	1.34	20.7	.006	1080	1
6.E.7	113	.91	14.2	.000	1300	1
6.F.1	261	.762	11.3	.031	1300	1
6.F.2	2420	.895	10.6	.065	1480	1
6.F.5	297	.963	21.7	.024	1090	0
6.F.6	114	.588	9.56	.000	1240	1
6.F.10	138	.957	23.6	.014	1130	-
7.A.3	953	.361	4.78	.069	810	0
7.A.4	269	1.49	10.5	.000	1250	-
7.D.3	513	1.23	3.74	.023	1210	0
7.D.7	2280	.951	3.71	.020	980	0
7.E.2	6640	.451	0.61	.099	990	-
7.F.1	880	1.24	18.7	.007	920	-
7.F.2	476	1.50	20.0	.007	1260	0
7.F.3	602	1.79	22.1	.002	1320	-
7.G.3	39.4	1.47	103	.000	1200	-
7.G.11	119	1.24	13.8	.000	1370	-
7.G.14	11 800	.600	0.51	.065	880	-
7.G.15	117	2.03	40.3	.000	1360	-
8.A.2	1760	1.62	15.4	.000	910	1
8.C.5	165	1.76	56.2	.000	1350	-
8.C.6	313	1.29	43.5	.024	1060	-
9.A.2	1410	2.06	9.55	.009	1180	-

Table 3.1: Catchment characteristics

Stream frequency

Drainage density is a measure of the degree of development of the stream network, but it is difficult to estimate simply. Use is therefore made of the strong correlation between drainage density and stream frequency (Melton, 1958), since stream frequency is much simpler to measure. All other things being equal, one would expect catchments with high values of stream frequency to exhibit higher floods and lower low flows than catchments with low values. It is measured by counting stream junctions on a 1:50,000 scale map and dividing by catchment area in square kilometres. This stream frequency variable is referred to as STMFRQ. The method to be followed in measuring stream frequency is as follows:

Select the latest available 1:50,000 scale maps of the catchment. Count the natural stream junctions going upstream from the site of interest, which is also counted as a junction. It is best to work progressively up each tributary, marking the running total at each major junction. Where the stream passes through a dambo, swamp or lake it is often not marked; however the arrangement of stream channels can be sketched in and the junctions counted. Where the headwaters of a stream are in a dambo, sketch a stream up the centre of the dambo and up any side lobes and count the junctions so produced. Divide the number of junctions counted by the catchment area in square kilometres.

Slope

The slope of catchments was indexed by the slope of the main stream. The overland slope of the basin can only be measured by a tedious grid sampling technique. Benson (1949) has shown that only mainstream slope need be measured because it is closely correlated with tributary slopes, and Strahler (1950) has demonstrated the correlation of channel slope and valley side slopes. The slope measure used is that defined by the US Geological Survey and is the slope between the 10 and 85 percentiles of the mainstream length, measured upstream from the site of interest to the end of the blue line on the map. This definition excludes the highest and lowest gradients at either end of the stream. All other things being equal, one would expect steep catchments to have higher floods and lower low flows than flat catchments.

In this study, the Institute of Hydrology measured these slopes from 1:50,000 maps by means of a D-mac digitiser and a computer program simulating the divider method described below. Users should follow the following procedure:

> Use the latest available series of the 1:50,000 scale map(s). Choose the mainstream from the maps. Usually this is the longest stream in the basin but, in cases of difficulty, work upstream and at each junction follow the stream draining the larger area. Distances are measured upstream from the gauging station with a pair of precision dividers set to 4 mm (0.2 km on the 1:50,000 map) and the distance to each contour is tabulated against its elevation

Once the total length of the stream, to the end of the blue line, is known the lengths and elevations of the 10% and 85% points are used to calculate the slope which is expressed in parts per thousand or metres per kilometre.

This slope measure is referred to as S1085.

Dambo

The proportion of the catchment area falling within a dambo or swamp was measured off the 1:250,000 series maps. Dambos are broad grass-covered swampy valleys which form a dominant topographic feature of the plateau. Balek and Perry (1973) have investigated dambos in Zambia. Originally the dambo and swamp proportions (the latter is very small) were measured separately, but it was eventually found more convenient to add these proportions together. In general terms, one would expect dambo to damp floods and augment low season flows, although it was observed that only the first of these effects was significantly apparent. This combined measure is referred to as DAMBO, and is dimensionless.

Average annual rainfall

The average annual yield of a catchment depends heavily on the average annual rainfall. It was originally proposed to measure the average annual rainfall off the 1:1,000,000 scale map "Climatic Regions of Nyasaland" dated 1959. However, this map proved to be unsuitable as it was based on too few data and lacked detail. Instead the average annual rainfall for each catchment was calculated from available daily raingauge records using Thiessen polygons. Two or three gauges were used for each catchment; comprehensive details of this analysis are given by Hill and Kidd (1980).

At the time of going to press, a map of average annual rainfall is being prepared by the Malaŵi Meteorological Office and should be used in applying the results of the work described in this report. The method of estimating catchment average annual rainfall is as follows:-

Trace a map of the catchment boundary and overlay it on the rainfall map. Measure the area between each pair of rainfall isohyets on the catchment using a planimeter or by counting squares on graph paper. Associate with each area the average rainfall of the two bounding isohyets and calculate the catchment average rainfall according to the formula:

$$R = \frac{1}{A} \sum_{i=1}^{n} A_i R_i$$

where R is catchment average annual rainfall,
A is catchment area,
R_i is the rainfall on the ith segment of the catchment,
A_i is the area of the ith segment,
and n is the number of segments.

Until a suitable map is available, it is recommended that average annual rainfall be measured from available raingauge records using Thiessen polygons. The average annual rainfall is called AAR, and the units are millimetres.

Escarpment

This is a so-called dummy variable: stations on or immediately below the escarpment as defined in Figure 1.1 were given the value unity and all others the value zero. No account was taken of the proportions of the catchment area on the escarpment or on the plateau above, the criterion being simply the position of the gauge. However, all the stations studied had a large part of their catchment area on the escarpment. This variable was only used in the flood study as a means of studying the possibility of further regionalisation. It is referred to as ESCARP, and is dimensionless.

Use of the catchment characteristics

As will be seen later, only AREA, STMFRQ and AAR are in fact used in the recommended estimation procedures described in chapters 4, 5 and 6. The other characteristics are described for completeness. As was discussed in the introduction to this chapter with respect of geology and soil type, an appreciation of catchment characteristics not directly used in estimation techniques can lead to useful subjective insights in a particular application. This is particularly true where comparison is being made with flow data collected nearby.

4: AVERAGE DAILY FLOW

Definition of Average Daily Flow (ADF)

All the low flow measures described in chapter 5 are expressed as a percentage of the average daily flow from the catchment. The average daily flow (ADF) is defined as the long-term average rate of runoff from the catchment and is expressed in cubic metres per second. It is normally estimated by averaging the daily mean flows over a sufficiently long period. A related variable is the average annual yield (AAY) which is the average annual runoff volume expressed as a depth in millimetres per year over the catchment. These two variables are related by

$$\text{ADF} = \frac{\text{AAY} \times \text{AREA}}{31500} \qquad 4.1$$

where AREA is the catchment area in km^2.

Evaluation of ADF from runoff data

The runoff for any particular year from a catchment depends strongly on the catchment rainfall for that year, and, because the annual rainfall is extremely variable, then the annual runoff is correspondingly variable. For a reliable calculation of ADF from flow data alone, a minimum of 5 years of data is desirable. An examination of the rainfall over this period in relation to its long-term mean could suggest whether this estimate of ADF is high or low in comparison to its long-term mean. If it is considered necessary, an appropriate adjustment can be made accordingly. The ADF was estimated from runoff data for the 52 stations in this study, and is tabulated in table 5.1. No adjustment was made for variation in annual rainfall over the period of record.

Evaluation of ADF from rainfall data

When insufficient flow data are available, ADF can be derived from rainfall data by an empirical equation. An examination has been made of the relationship between annual yield and annual rainfall for 47 stations. In this analysis, a number of raingauges, usually two or three, were chosen on or near to the catchment in question. Annual rainfall totals at these gauges were used in Thiessen polygon analysis to estimate the annual catchment rainfall corresponding to the annual yield calculated from the daily flow data. Where either rainfall or flow data were missing for a given year, this year was rejected, and some stations were rejected because there were no suitable raingauges. Catchment averages of annual rainfall and annual yield over a concurrent period were thus obtained, and these values for the 38 acceptable stations are shown in table 4.1. More details of these analyses are given by Hill and Kidd (1980).

As might be expected, there is a strong relationship between rainfall and yield, as shown in figure 4.1. A regression of average annual yield on average annual rainfall yields the following equation:

$$AAY = 0.71AAR - 490 = 0.71\ (AAR - 690) \qquad 4.2$$

This equation has a correlation coefficient of .93 and a standard error of estimate of 90 mm.

The above linear relationship is naturally unrealistic as rainfall approaches 690 mm. In practice, one would expect a nonlinear relationship between AAY and AAR, as demonstrated by Midgley and Pitman (1969). In this case, this trend tends to be masked by the presence of dambo in catchments having low average annual rainfall. Equation 4.2 should therefore be used with great caution for AAR below 800 mm. Hill and Kidd (1980) have derived a nonlinear equation wherein allowance is made for loss of yield due to dambo. Equation 4.2 will be improved as more data become available.

The estimate given by equation 4.2 is not as accurate as might be desirable, especially in view of the fact that it is at the basis of all the low flow estimation procedures. The equation should only be used in conjunction with a subjective appraisal of adjacent catchments having

Station	annual average rainfall	annual average yield	potential evaportation E_o	percentage of area covered by dambo	observed actual evaporation	E_a/E_o
1.D.3	1110	193	1710	0	917	.54
1.D.11	1280	395	1800	0	885	.49
1.D.14	2090	978	1500	0	1112	.74
1.D.23	1070	270	1800	0	800	.44
1.D.24	1240	332	1800	0	908	.50
1.F.2	1300	370	1930	0	930	.48
1.R.18	1080	199	1650	0	881	.53
2.B.8	2060	983	1500	0	1077	.72
2.B.21	1430	499	1520	0	931	.61
2.B.22	880	208	1700	0	672	.39
2.C.3	1730	882	1500	0	848	.57
2.C.8	1280	312	1520	0	968	.64
3.F.3	1000	337	1730	0	663	.38
4.B.1	880	133	1790	8	747	.42
4.B.3	910	192	1630	7	718	.44
4.D.4	930	155	1700	4	775	.46
4.D.6	940	182	1710	2	758	.44
4.E.1	810	82	1730	17	728	.42
5.A.8	870	298	1960	0	572	.29
5.D.1	900	83	1890	21	817	.43
5.D.2	900	73	1930	27	827	.43
5.D.3	710	53	1780	10	657	.37
5.E.1	930	228	1790	3	702	.39
5.E.2	910	119	1800	13	791	.44
6.C.1	740	37	2030	12	703	.35
6.F.1	1300	403	1850	3	897	.48
6.F.2	1480	500	1750	6	980	.56
6.F.5	1090	260	1700	2	830	.49
6.F.6	1240	340	1780	0	900	.50
7.D.3	1210	232	1800	2	978	.54
7.E.2	990	40	2080	10	950	.46
7.F.1	920	233	1850	1	687	.37
7.F.2	1260	348	1750	1	912	.52
7.G.2	1320	661	1590	0	659	.41
7.G.11	1370	625	1650	0	745	.45
7.G.14	880	81	1900	6	799	.42
8.A.2	910	227	1730	0	683	.39
8.C.6	1060	428	1600	2	632	.39
mean	1130	309	1750		817	
s.d.	320	247	140		128	

Table 4.1: <u>Catchment average annual rainfalls and yields</u>

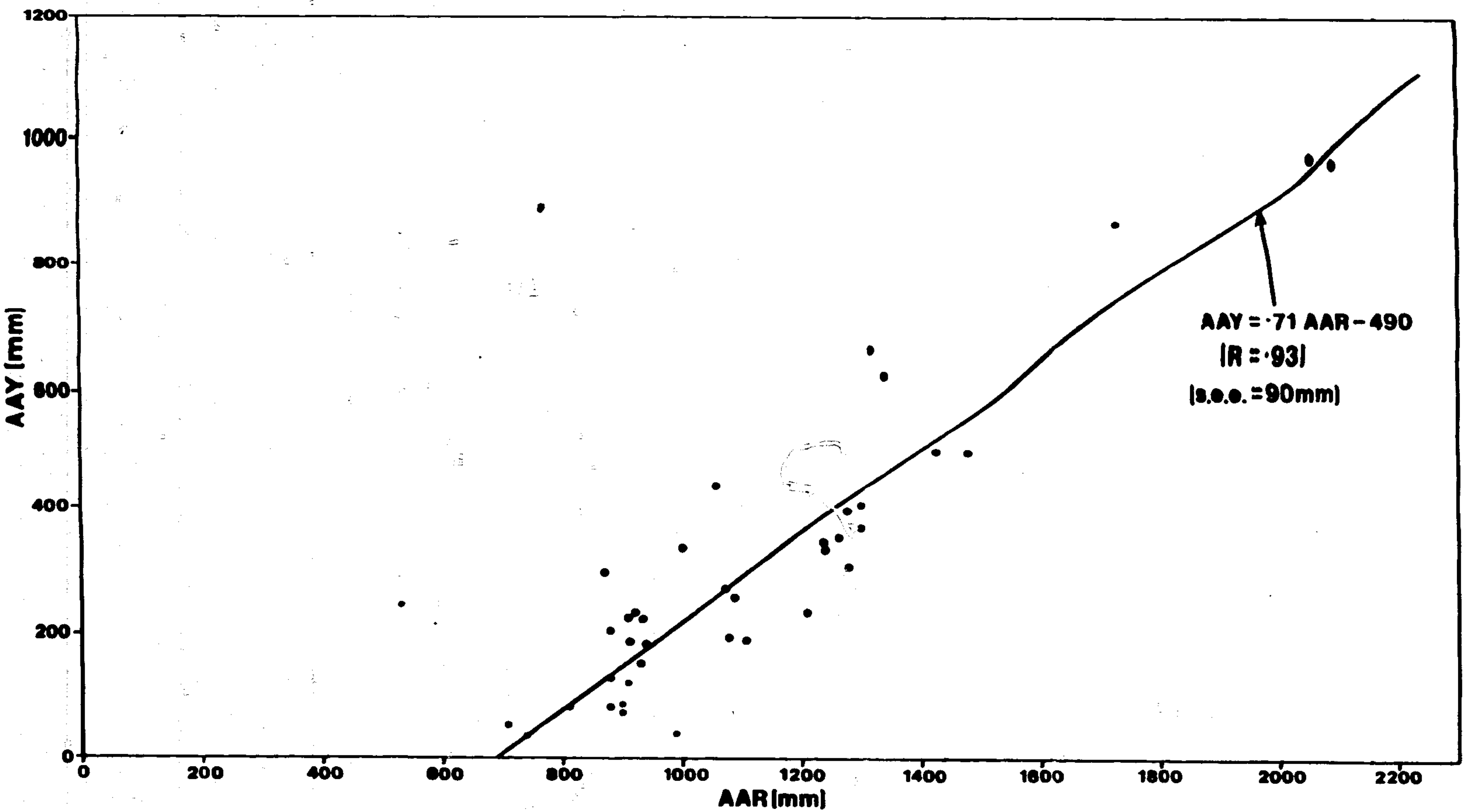


FIGURE 4.1 : Relationship between average annual yield and rainfall

similar topographical and climatic characteristics. Particular note should be taken of information from gauged stations upstream or downstream of the site of interest.

Great assistance can be given to the estimation of ADF by the acquisition of as little as one year's data and figure 4.2 illustrates this point. This figure shows a plot of the annual runoff in millimetres at 4.E.1 and 4.E.2 which are 2 stations on the Lingadzi river. A single year's data at one of these sites would provide a good estimate of its long-term ADF from a knowledge of how the same annual runoff at the other site related to its own long-term mean. In this respect, spot gauging at the point of interest can also be useful - this gauging can be used in conjunction with measured flows at nearby long-term stations.

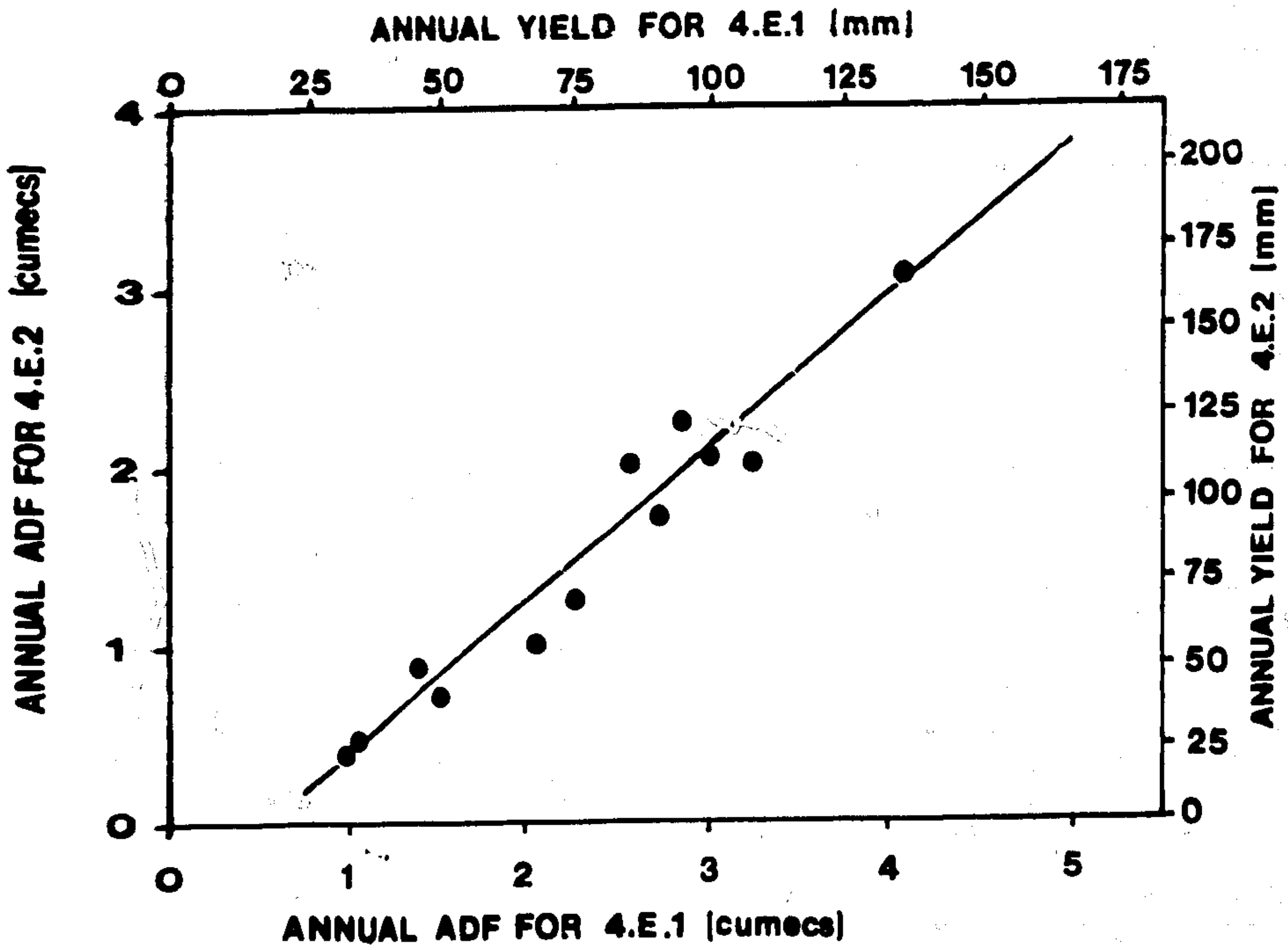


Figure 4.2 Comparison of annual ADF's at two sites on the Lingadzi river

5. LOW FLOW MEASURES

Introduction

The term 'low flow measures' is used to describe the many ways in which the low flow regime of a river may be summarised, often in diagrammatic form. The four major measures, described in this chapter, are the flow duration curve, the low flow frequency curve, the storage-yield diagram and recessions. These low flow measures are relevant to a variety of different applications such as licencing of abstractions, sewage works design, reservoir operation and design, and forecasting. The term 'low flow index' is used to describe a single numerical value specially adopted from a particular low flow measure.

Low flow measures were derived for the low flow study stations listed in Table 2.1. The records for stations 7.G.15 and 7.G.25 on the Kaziwiziwi were combined and treated as a single record - these two stations are in close proximity, and 7.G.25 was built to replace 7.G.15.

Two methods of estimation are described for each measure: (i) from data at a gauged site, and (ii) from catchment characteristics at an ungauged site. The methods are very similar to those adopted in the UK Low Flow Study (Institute of Hydrology, 1980).

Definition of the flow duration curve

This curve shows the relationship between any given discharge and the percentage of time during which this discharge is exceeded. Its main uses are in the licensing of abstractions and in sewage works design. Figure 5.1 is an example of such a curve for station 6.F.1 Limphasa for the period 1952-1975. To allow a comparison between catchments, all the flows are expressed as a percentage of the average daily flow (ADF).

It will be noticed that there are a number of curves, each one relating to a different duration. For all practical purposes, the 1-day curve can be assumed to be the instantaneous flow duration curve; and the interpretation of a point on the 10-day curve, say, is that it gives the proportion of 10-day periods during which the average discharge is above a given value. It is appropriate to use curves with duration greater than 1-day when there is an element of storage involved in the proposed scheme. In practice, there is little or no difference between the curves for durations less than about 10 days.

In general, it has been found that when the flows are plotted on a logarithmic scale along the y-axis and the percentage of time exceeded is plotted on a normal probability scale along the x-axis, the flow duration curve often approximates a straight line, which greatly facilitates extrapolation beyond the range of the data. This is to say that the logarithms of the daily discharges are normally distributed.

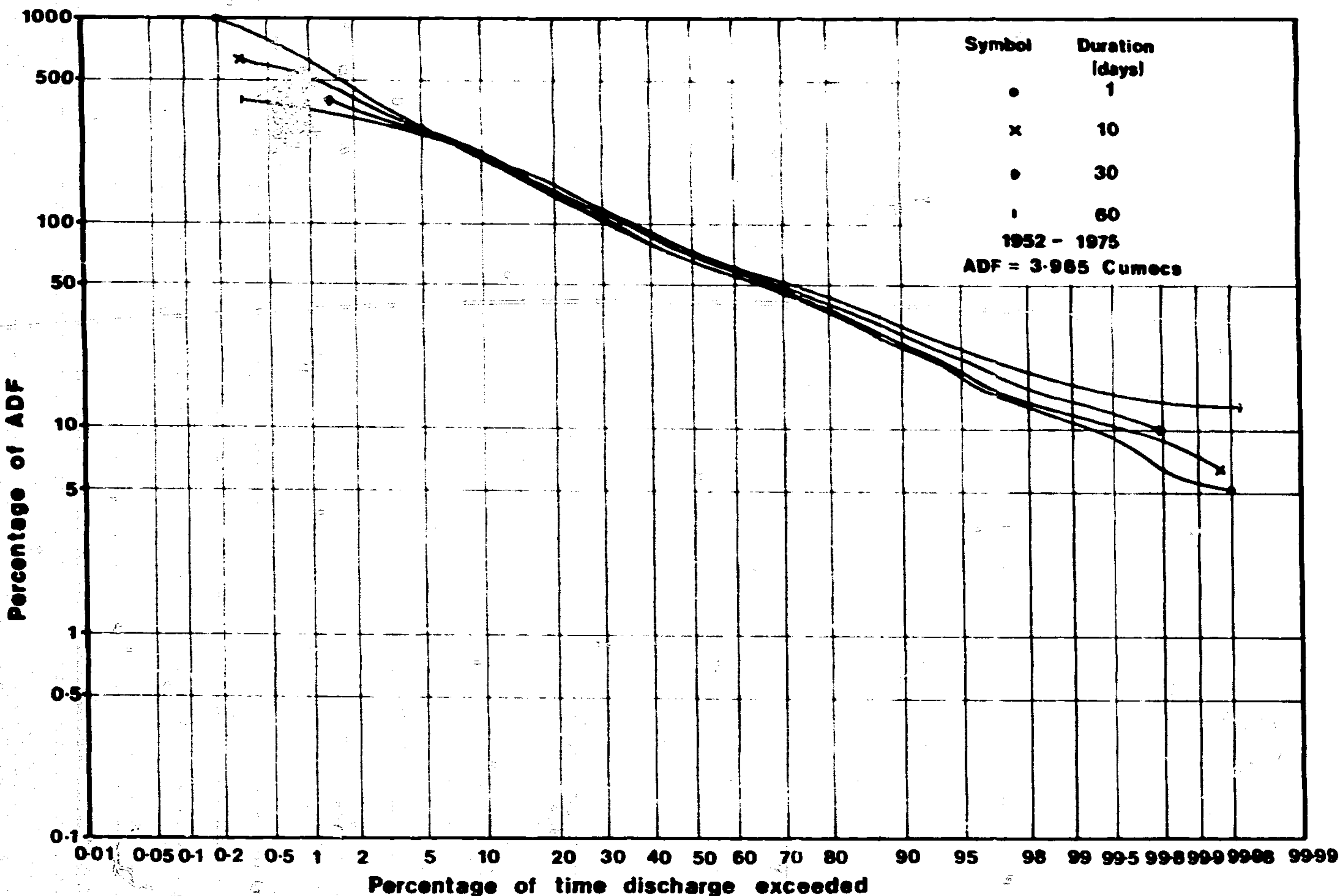


Figure 5.1 Flow duration curves for station 6.F.1

This has been found to be the case with most Malawi rivers, as reflected in the example in Figure 5.1. An exception to this general rule applies in rivers in which zero flows occur - in these circumstances the flow duration curve drops sharply at approximately the probability of occurrence of zero flows.

The point on the flow duration curve to be used as a low-flow index is the 75 percentile 10 day discharge, Q75(10), expressed as a percentage of ADF. This is the average flow over 10 days which is exceeded 75% of the time, and is 42.9% of ADF for station 6.F.1 as can be seen from Figure 5.1. Q75(10) is more suitable as an index than Q75(1) because the series of 10 day average discharges is less sensitive to data error than the 1 day series. Catchments with relatively high Q75(10)'s and consequently flat flow duration curves are those that have characteristically persistent dry season recessions whereas 'flashy' catchments have low Q75(10)'s and steep flow duration curves.

Derivation of the flow duration curve from data

The 1-day flow duration curve can be derived from data by hand. The method involves assigning daily flows for the complete record to class intervals and counting the number of occurrences within each interval. The total number of occurrences above the lower limit of each class interval is expressed as a percentage of the total number of days in the record and this percentage is then plotted against the lower limit of the interval to give the flow duration curve. The averaging process required to determine the flow duration curve for other durations makes it impractical for derivation by hand but is easily performed by computer.

Even when standardised by ADF, an estimate of the flow duration curve from a number of year's data will, on average, be flatter than the true curve (derived from an infinite number of years). In practice, this bias is small and 3 years of data is considered adequate for deriving the flow duration curve from data - and even 1 or 2 complete years will provide information which can be considered valuable.

Flow duration curves have been derived for 52 stations in this study. The values of Q75(10) for these stations can be found in Table 5.1.

Derivation of the flow duration curve without flow data

The flow duration curves from 49 catchments were studied to derive a method of estimating a synthetic flow duration curve. The proposed method is in two parts: (i) estimation of the 75 percentile flow for a 10-day duration, called Q75(10), and (ii) the use of Q75(10) to fix the slope and position of the flow duration curve for any other duration.

The first attempt at estimating Q75(10) from the gauged stations used a model based on regression analysis involving catchment characteristics. Included in the catchment characteristics was a geology-related Base Flow Index - itself derived as far as possible from data - developed in the U.K. Low Flow Study (Institute of Hydrology, 1980). In practice, the complex interaction of various geomorphological characteristics in this Malawi data set yielded problems which were not easily solved

Station	ADF (cumecs)	Q75 (10) (%ADF)	MAM (10) (%ADF)	KREC (months)
1.B.1	340.	68.1	58.6	11.6
1.D.3	9.45	8.6	3.3	3.0
1.D.10	5.09	34.6	18.5	6.1
1.D.11	56.0	25.0	10.4	3.9
1.D.14	.345	41.6	24.2	4.8
1.D.23	5.67	11.8	5.4	4.2
1.F.2	.519	47.6	26.6	7.8
1.G.1	410.	64.4	49.9	9.3
1.K.1	3.18	0.0	0.0	0.5
1.R.18	.0463	13.0	4.4	2.2
2.B.[illegible]	.562	31.3	16.0	2.9
2.F.21	2.66	15.2	5.5	2.9
2.G.[illegible]	2.46	8.7	2.4	3.4
2.C.8	.696	3.0	0.5	1.5
3.E.3	4.30	10.9	3.4	3.3
3.F.3	2.55	17.4	5.0	3.0
4.A.3	2.92	14.0	7.5	3.3
4.B.1	36.7	5.1	1.1	2.1
4.B.3	3.78	3.5	0.32	1.7
4.B.4	8.68	4.2	0.59	2.4
4.D.4	8.25	7.7	2.1	2.2
4.D.6	4.21	10.9	2.7	2.2
4.E.1	2.76	3.3	0.78	3.4
4.E.2	1.92	1.8	0.25	2.6
4.F.5	3.16	0.17	0.01	2.2
5.A.B	3.63	12.3	3.7	2.9
5.D.1	22.1	2.5	0.40	[illegible]
5.D.2	16.7	1.6	0.21	1.2
5.D.3	.807	1.1	0.12	[illegible]
5.E.1	1.21	4.4	0.89	2.3
5.E.2	1.74	8.6	2.4	2.5
6.C.1	3.43	0.0	0.0	0.4
6.F.1	3.96	42.9	22.4	3.5
6.F.2	37.9	46.1	28.8	3.3
6.F.5	2.93	50.5	27.8	4.9
6.F.6	1.76	46.8	29.[illegible]	4.3
6.F.10	1.30	50.0	28.1	4.4
7.A.3	1.90	6.2	1.5	2.1
7.A.4	1.82	29.5	18.0	4.9
7.D.3	3.42	22.8	8.8	2.8
7.E.2	9.26	2.3	0.35	1.5
7.F.1	6.44	44.0	22.3	3.4
7.F.2	5.15	53.3	30.4	4.7
7.G.2	9.06	51.7	34.1	5.3
7.G.11	2.59	50.7	34.5	4.2
7.G.14	31.5	24.0	9.0	2.8
7.G.15	3.21	58.0	46.3	5.2
8.A.2	12.7	25.2	11.7	3.1
8.C.5	3.41	53.6	41.3	5.1
8.C.6	4.05	36.2	20.8	2.8
9.A.2	19.4	8.6	4.1	3.4

Table 5.1: Low flow indices

by traditional regression methods. As an alternative, it was felt that mapping of Q75(10) might give an improved method of estimation, and this approach was finally adopted.

Figure 5.2 is a contour map of Q75(10) expressed as a percentage of ADF. It should be noted that the map does not give point values of the variable. Estimation of Q75(10) for a point on a given river involves averaging Q75(10) over the catchment area down to the point under consideration, as is done for catchment average annual rainfall (see Chapter 3). The map evolved out of a trial-and-error process - the strongest control on the position of the contours was the 49 data points. Over this information were overlaid (i) average annual rainfall, (ii) geology and (iii) topography. Where data were common (such as the Central Region), the positions of the contours were dictated by the data alone. In other areas of the country (particularly the Southern Region), data is somewhat sparse and the contours are drawn substantially according to intuition based upon annual rainfall, geology and topography. A first sketch of the contours was made, from which estimates of Q75(10) were calculated for the gauged catchments. Errors in these estimates were noted, and the contours adjusted. This process was repeated until a satisfactory fit was obtained. As such, this map is not a unique solution to the problem and another person could arrive at a different map which still fulfils the criteria by which the one in Figure 5.2 was drawn. Furthermore, no estimate of the accuracy of the map is possible, since it provides near enough an exact solution to the data points at all but 2 stations (1.F.2 and 7.G.14). Despite these drawbacks, this map provides what is considered to be the best method available for estimating Q75(10).

In order to produce a flow duration curve for any other duration D, it is first necessary to establish the relationship between Q75(D) and Q75(10). Data from 44 stations have been examined, and the most appropriate relationship, for D > 10 days is:

$$\frac{Q75(D)}{Q75(10)} - 1 = \text{CONST}\ (D - 10)^{\text{EXP}} \qquad \ldots\ 5.1$$

Optimum estimates for CONST and EXP can be obtained for a given station by linear regression in log space. The 2 parameters are inevitably strongly intercorrelated, so EXP was fixed at the mean value of 1.33 for the 44 stations. Best estimates of CONST for EXP = 1.33 were generated, and Figure 5.3 indicates the strong correlation between CONST and Q75(10). This yields by regression:

$$\text{CONST} = .0089\ \{Q75(10)\}^{-.77} \qquad \ldots\ 5.2$$

This gives an equation for predicting Q75(D) from Q75(10):

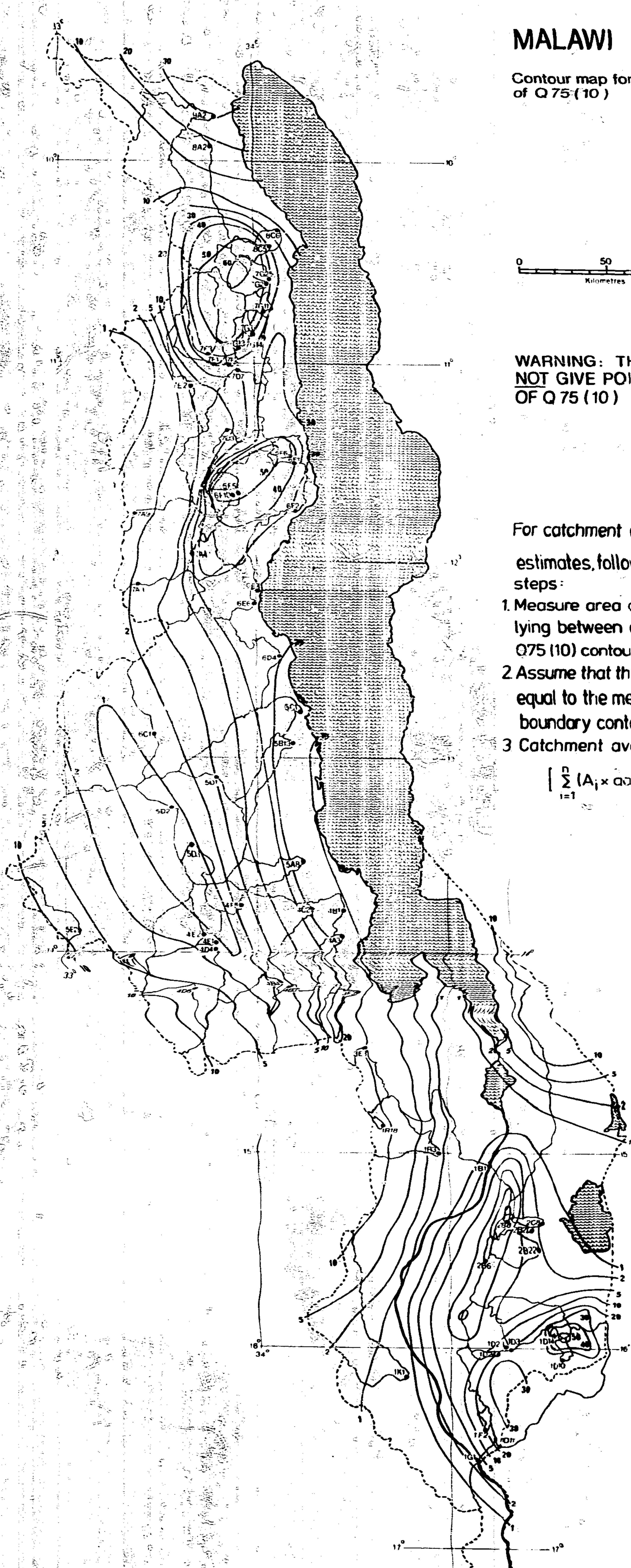
$$Q75(D) = Q75(10) + .0089\{Q75(10)\}^{.23}\ (D-10)^{1.33} \quad \text{for } D > 10 \qquad \ldots\ 5.3$$

MALAWI

Contour map for estimation of Q 75 (10)

0 50 100
Kilometres

WARNING: THIS MAP DOES <u>NOT</u> GIVE POINT ESTIMATES OF Q 75 (10)



For catchment average Q75(10) estimates, follow the following steps:

1. Measure area of catchment (A_i) lying between any pair (i) of Q75 (10) contours.
2. Assume that this area has Q75(10) equal to the mean of its two boundary contour values (qav_i)
3. Catchment average Q75(10) =

$$\left[\sum_{i=1}^{n} (A_i \times qav_i) \right] \div \text{total catchment area}$$

Figure 5.2

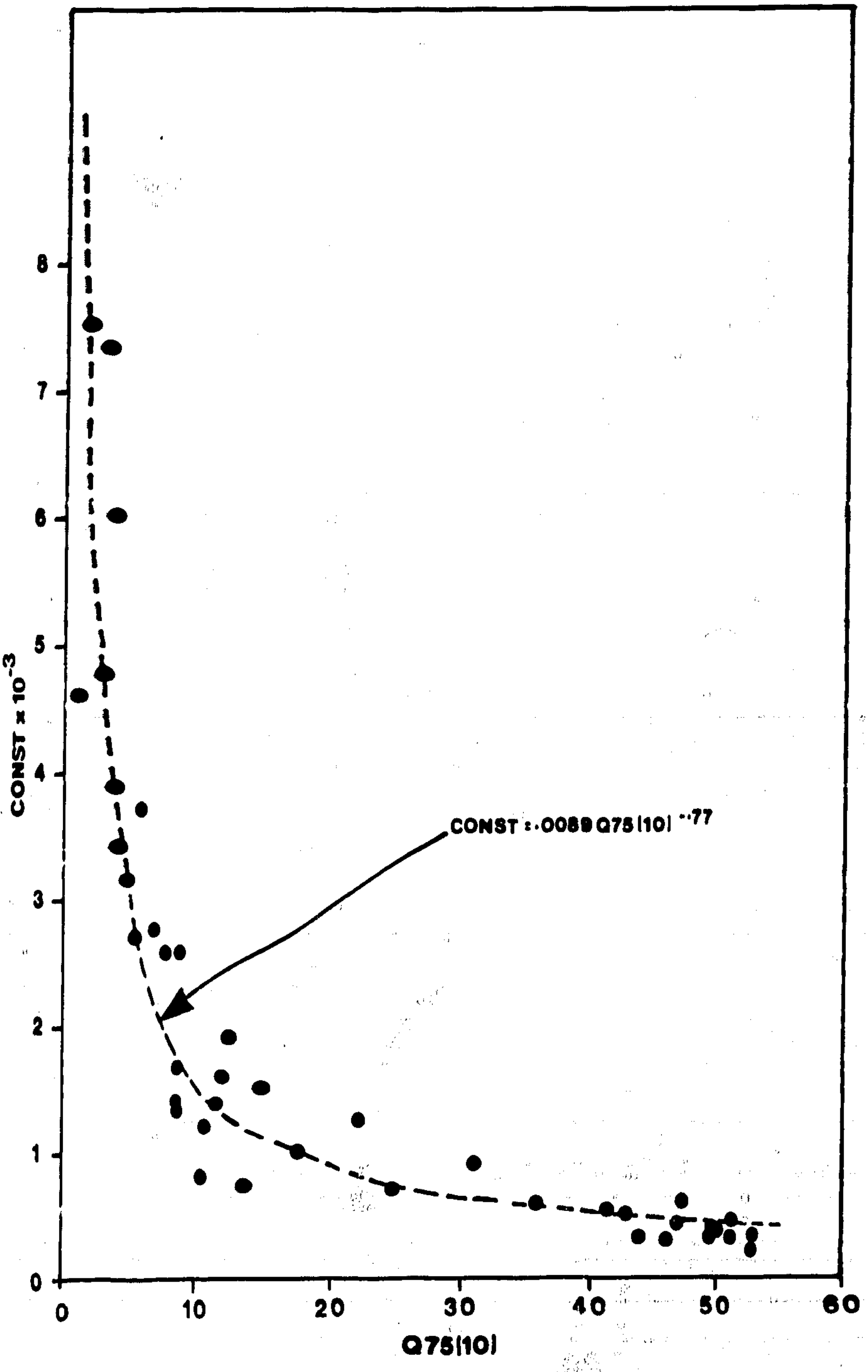


Figure 5.3 Relationship between CONST and Q75[10]

Equation 5.3 is derived from data for durations in excess of 10 days and cannot be used for D < 10. Differences in Q75(D) are small for D < 10, and a symmetrical relationship to equation 5.3 is assumed:

$$Q75(D) = Q75(10) - .0089\{Q75(10)\}^{.23}\ (10-D)^{\ 1.33} \quad \text{for } D < 10 \qquad \ldots.\ 5.4$$

Having determined Q75(D), it remains to determine the variation of flow with frequency to complete the flow-duration curve. As described above, the flow duration curves can be satisfactorily approximated by a straight line, so that only one other point is required to define the curve. The 25 percentile point Q25(D), or that flow which is exceeded 25% of the time, is a suitable point, and it was found that Q25(D) could be related to Q75(D) alone. Analysis of the data from 44 catchments led to the family of curves in Figure 5.4, which enables Q25(D) to be estimated from Q75(D) and duration D. The result in Figure 5.4 is reproduced in Table 5.2 which may be easier to use.

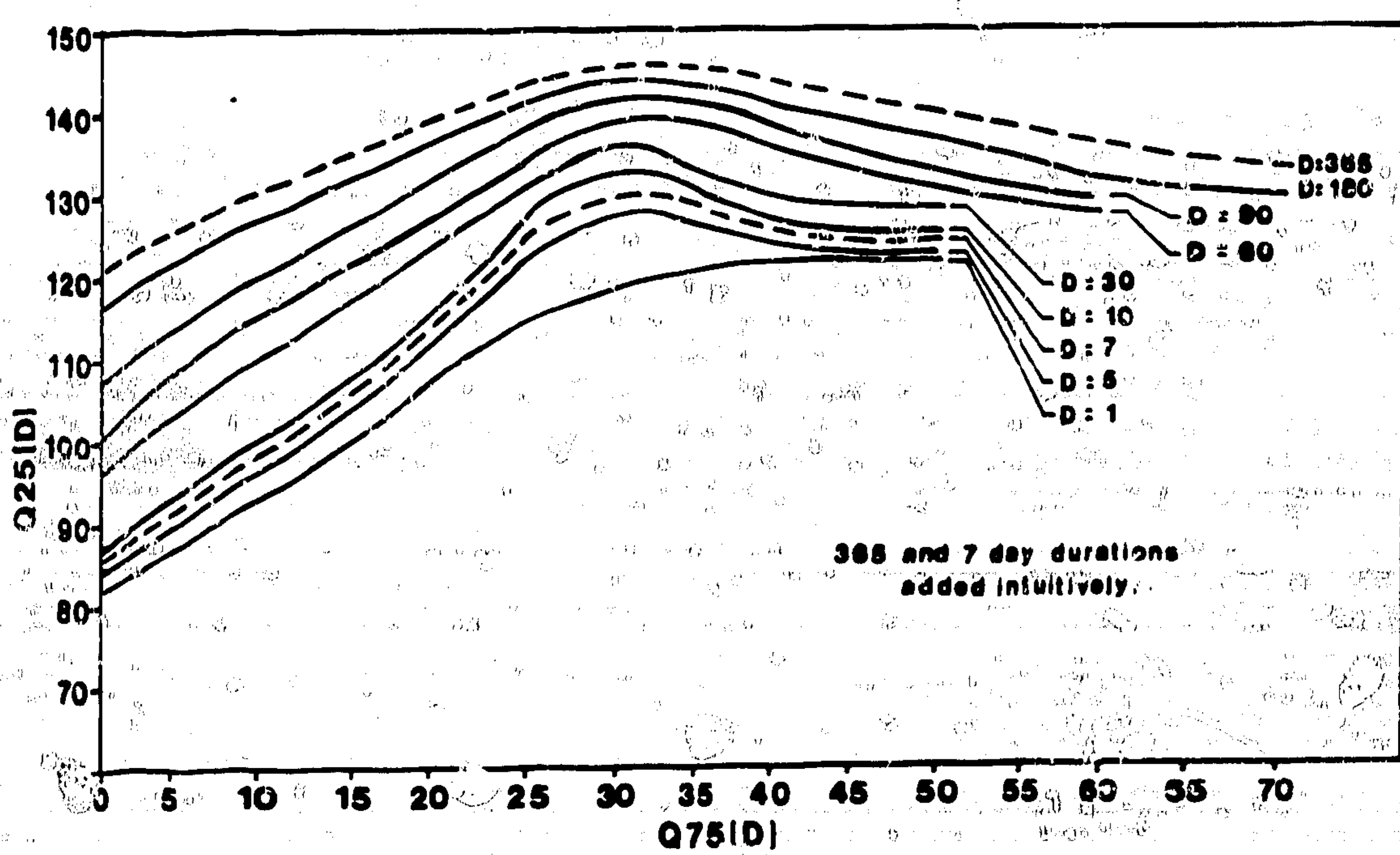


Figure 5.4 Estimation of Q25(D) from Q75(D) and duration D

DURATION	Q75 (%ADF)									
	0	5	10	15	20	30	40	50	60	70
1	82	87	93	100	108	120	122	123	123	123
5	85	90	96	104	112	127	124	124	124	124
7	86	92	98	106	114	130	126	125	125	125
10	88	94	100	108	117	133	127	125	125	125
30	96	103	110	118	125	136	130	129	127	126
60	101	108	115	122	128	140	136	131	128	127
90	108	114	120	126	132	142	139	133	130	129
180	116	121	126	131	136	144	142	137	132	130
365	121	126	131	135	140	146	144	140	136	133

‥e 5.2: Q25(D) for values of Q75(D) and duration D

Deiinition of the low flow frequency cur‥

While the flow duration curve is concerned with the proportion of time that a given flow is exceeded, the low flow frequency curve indicates the proportion of years in which the flow falls below a given value. It is analogous to the annual maximum flood frequency curve. Its uses are basically the same as for the flow duration curve, but it is better equipped to deal with more extreme events. Figure 5.5 is an example for station 7.F.1 Runyina at Chikwaba.

As for the flow duration curve, this curve can be drawn for any duration. The flow can be expressed in cumecs, as percentage of ADF, or as a fraction of the mean annual minimum (called MAM(D) and analogous to the mean annual flood).

The curves are drawn using, for convenience, the Weibull distribution, which can be shown to be equivalent to the Extreme Value Type III distribution. If the data were indeed distributed according to the Weibull distribution (with appropriate parameters), the curve would plot as a straight line. In practice, the curves are not straight, particularly at low return periods and Figure 5.5 demonstrates the observed general shape of the curve. The curve is generally well-behaved enough to permit limited and judicious extrapolation beyond the range of the data.

The index related to the low flow frequency curve is the mean annual 10 day minimum, MAM(10), expressed as a percentage of ADF. For station 7.F.1 in Figure 5.5, MAM(10) is 22.3% ADF.

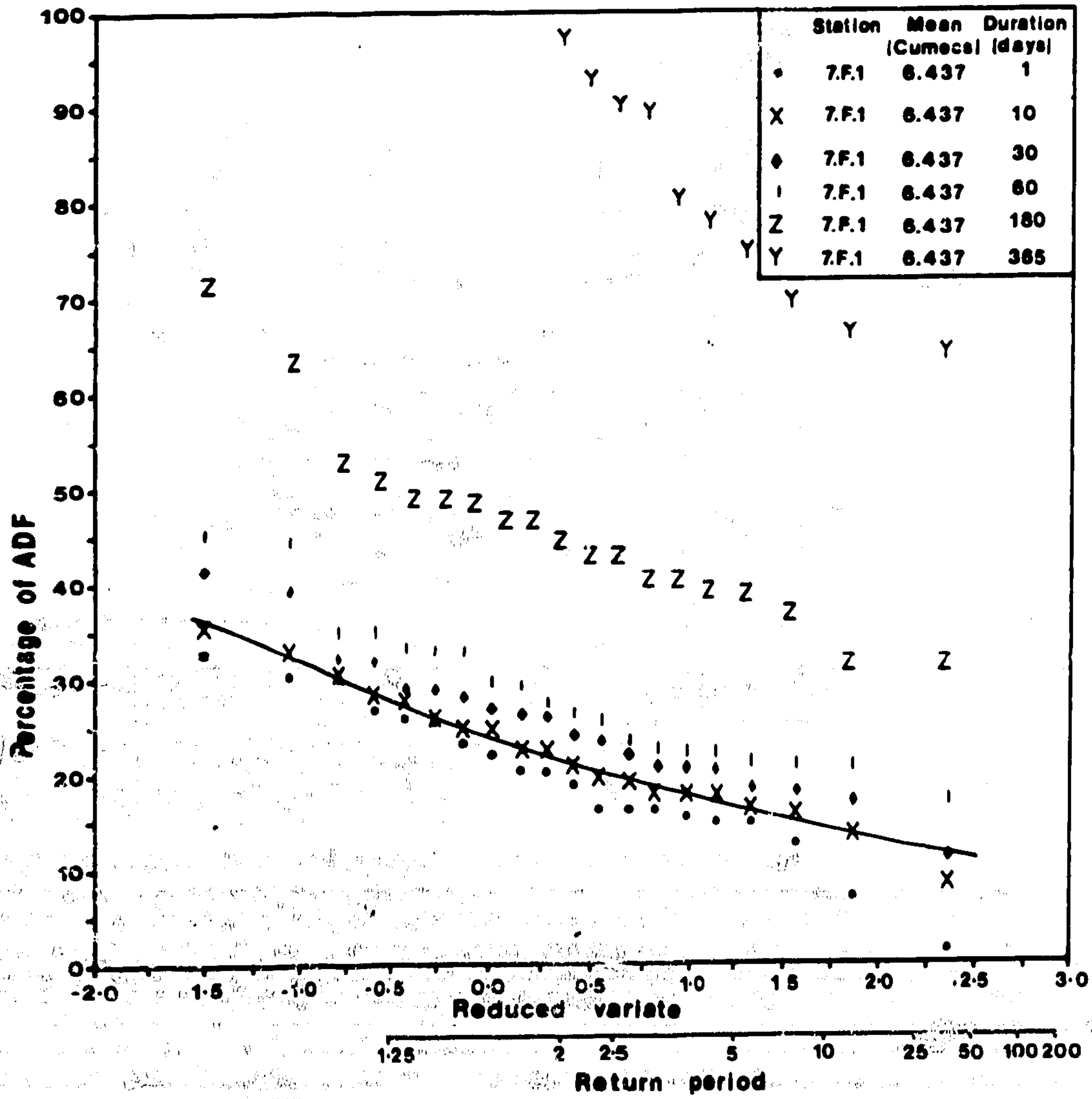


Figure 5.5 Low flow frequency curves for station 7.F.1

Derivation of low flow frequencies curve from data

The method for deriving the low flow frequency curve from data involves the following steps: (i) find the lowest average flow over D days in each calendar year, (ii) express these flows as a percentage of ADF, (iii) rank these flows and assign a plotting position to each, (iv) plot the flow against the plotting position, and (v) draw by eye a smooth curve through the data. The plotting position formula is given in Appendix A.

A minimum of 15 years of data is needed for a complete and adequate description of the low flow frequency curve. Where there is less than 15 years of data, a compromise between this and the methods described in the next section is required. The low flow frequency curve has been derived for all stations, and MAM(10) for each catchment is included with the low flow indices in Table 5.1.

Derivation of low flow frequency curve without flow data

Where there are little or no data available on a catchment, the low flow frequency curve can be obtained from the generalised regional method described below. As for the flow duration curve, the appropriate low flow index, MAM(10), is obtained externally; the curve for this or any other duration is then obtained, as before, by internal duration and frequency relationships.

Deriving a low flow frequency curve from data is equivalent to fitting a probability distribution to the annual minimum data. Some Malaŵi rivers dry completely in some dry years, and in these circumstances the annual minimum series will have several zeros. If the proportion of dry years is too high (1 year in 5 is the threshold used here), then the data will not fit the distribution at the left hand end because the data are truncated at zero. In this case, the arithmetic mean of the data will not be a satisfactory estimator of the distribution mean. Of the 44 stations used, 16 had zero 10-day minima in more than 1 year in 5. A regression between MAM(10) and Q75(10), both expressed as a percentage of ADF, for the remaining 28 stations produced:

$$\mathrm{MAM}(10) = .101\{Q75(10)\}^{1.47} \qquad \ldots\ 5.5$$

This equation has a correlation coefficient of 0.98. This implies that, as might be expected, the geomorphological and climatic factors which determine the value of Q75(10) for a given catchment control the value of MAM(10) in the same fashion. The two indices are therefore very strongly related. No catchment characteristics were significant after inclusion of Q75(10) as an independent variable. Instead of using the arithmetic mean of the annual minima, equation 5.5 was used to provide an estimate of MAM(10) for the 16 drying catchments. This estimate will be needed in the next stage of the analysis.

A similar analysis using 44 stations was performed for the internal duration relationship between MAM(D) and MAM(10), similar to that described earlier for the flow duration curve. This yielded the following equations:

$$\mathrm{MAM}(D) = \mathrm{MAM}(10) + .0186\{\mathrm{MAM}(10)\}^{.27}\ (D-10)^{1.28} \text{ for } D > 10 \qquad \ldots\ 5.6$$

$$\mathrm{MAM}(D) = \mathrm{MAM}(10) - .0186\{\mathrm{MAM}(10)\}^{.27}\ (10-D)^{1.28} \text{ for } D < 10 \qquad \ldots\ 5.7$$

For defining the frequency relationship (the shape of the low flow frequency curve through a given MAM(D)), the 44 stations were grouped into 8 categories according to their MAM(10) value. The 10-day frequency curves for all stations in each group were plotted on a common base with annual minima normalised by their MAM(10). These curves were averaged and then smoothed. This yielded 8 curves, each representative of a different MAM(10) band, which could be plotted on a common base. The family of curves in Figure 5.6 is derived from this last graph after further smoothing and interpolating. This provides an estimate of the frequency curve which is more reliable for high values of MAM(10) than for low values. The same analysis repeated for other durations suggested that the relationships in Figure 5.6 could be used for any duration.

Where no data exist, MAM(10) can be estimated using Equation 5.5 from Q75(10) which itself can be estimated from the map in Figure 5.2. Equations 5.6 and 5.7 can be used to obtain MAM(D) if required, and Figure 5.6 gives the remainder of the low flow frequency curve. Where less than 15 years of data are available the low flow frequency curve should be estimated using both above methods and interpolating between the resulting curves.

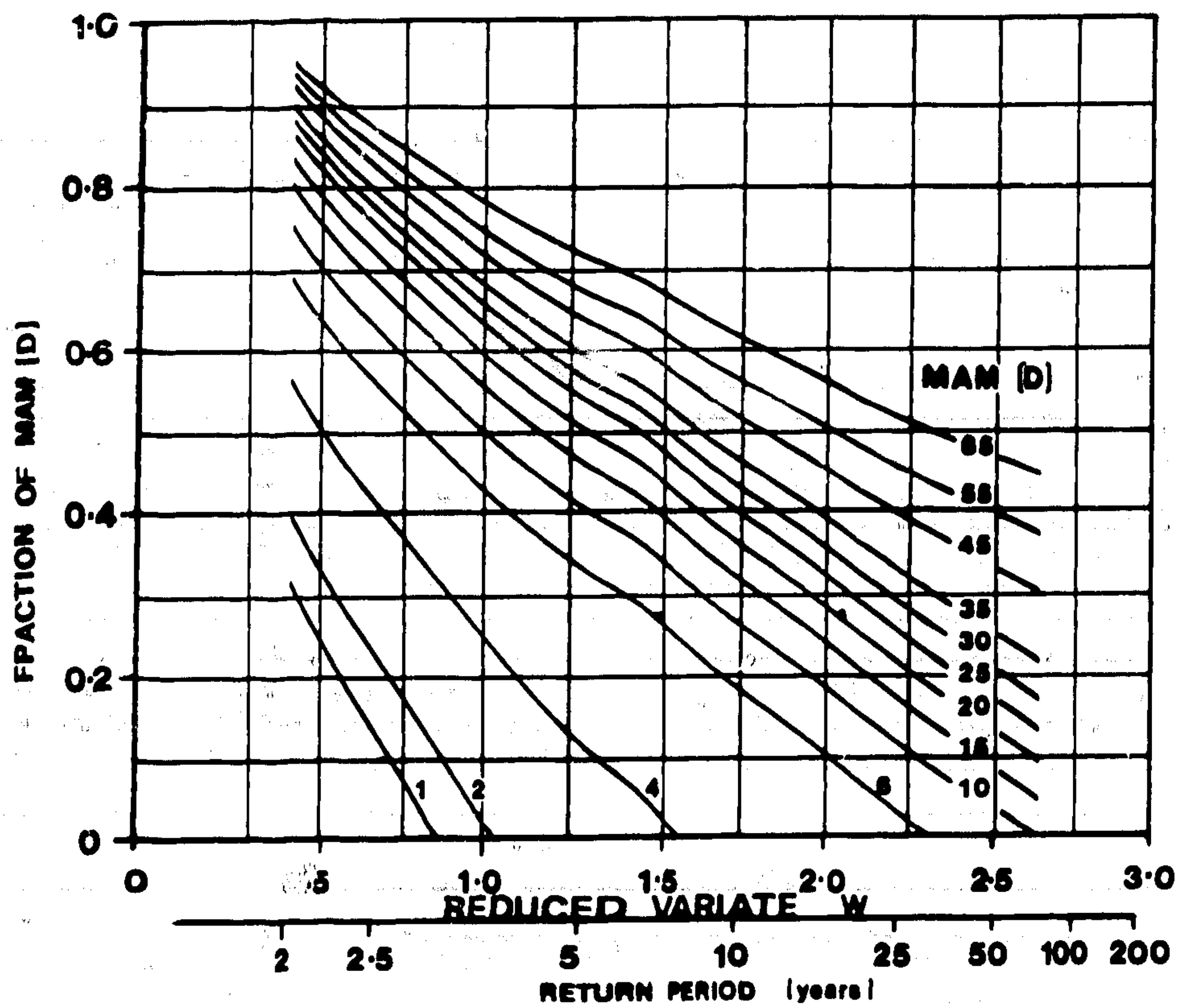


Figure 5.6 Estimation of low flow frequency curve from MAM (D)

Storage-yield analysis

Storage-yield analysis is used in reservoir design. For a yield of a given uniform supply rate, the storage-yield diagram is used to determine the volume of storage required for a given probability of failure. The analysis is time-consuming and therefore normally done by computer.

Storage-yield analysis using flow data

Figure 5.7 illustrates the form of analysis used. A daily flow hydrograph for a prospective reservoir site is first prepared, and a horizontal line corresponding to a given yield can be drawn through the hydrograph. The reservoir is assumed to be full at the start of the analysis. When the hydrograph is below the yield line (e.g. T_1 to T_2, and T_3 to T_4) demand exceeds supply and the reservoir is emptying. When the hydrograph is above the yield line (e.g. T_2 to T_3, T_4 onwards) the reservoir is filling and possibly spilling (e.g. T_5 onwards). For the purposes of this analysis, a semi-infinite reservoir is assumed - this is a reservoir which can spill but never empties. A depletion event is defined as occurring between the two times T_1 and T_5 when the

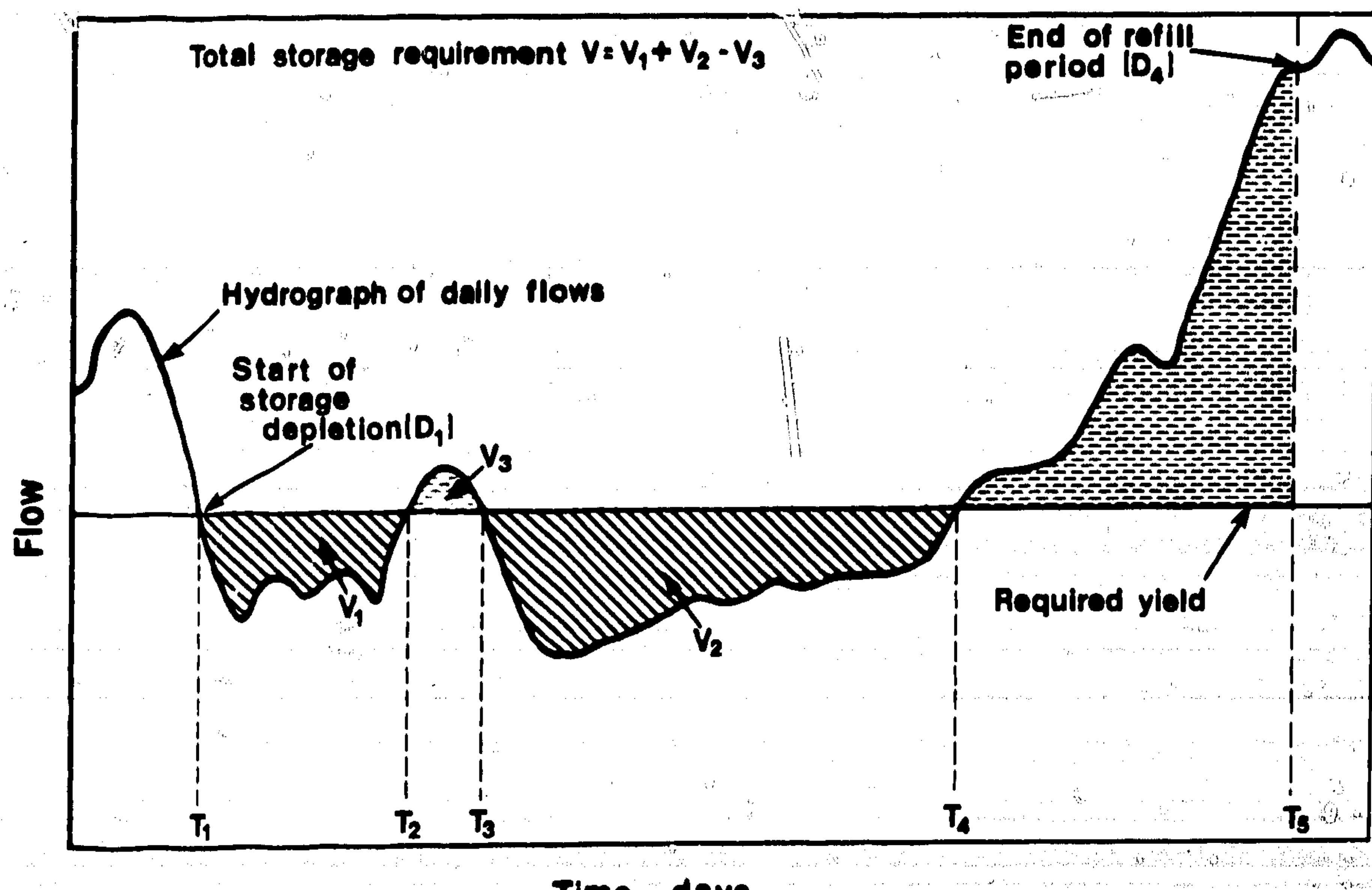


Figure 5.7: Definition of a depletion event for storage yield analysis

reservoir is full. The storage requirement of an event is the maximum drawdown, which occurs at time T_4. A new depletion event will start at a time after T_5 when the hydrograph again falls below the yield line. For a given yield, the depletion events can be ranked in order of the storage requirement. A diagram can be produced relating the storage requirement, the yield and the probability of this storage requirement being exceeded; Figure 5.8 is an example of such a diagram for station 5.D.2 Bua at Kasese. This analysis has been performed for 49 stations. A minimum of 15 years data is desirable for this analysis. Although daily data have been used in the analysis performed in this study,it is normally adequate to use monthly data for this purpose.

This form of analysis has 3 major drawbacks. Firstly, the effect of evaporation from, and direct rainfall on, the reservoir surface has been ignored - these effects can be significant in the Malaŵi climate (chapter 1). Although this phenomenon is a time varying one, a rough estimate of the yield required to satisfy the evaporation component can be made if the storage-area relationship of the reservoir are known. The second drawback is that, in practice, reservoir operating rules would be implemented to reduce the yield as the reservoir become seriously depleted. Thirdly, the analysis uses a semi-infinite reservoir in which the reservoir state at the outset of a depletion event depends on the performance of a reservoir which just satisfied the previous deficit rather than some fixed reservoir storage. These three drawbacks mean that the storage yield analysis described here is only suitable for preliminary screening of reservoir sites. Detailed design is best achieved by a more rigorous approach.

<u>Storage-yield analysis with little or no data</u>

The example shown in Figure 5.8 has a number of yield lines, the position of which will depend upon the variability of flows at the station in question. In general terms, for two catchments with the same ADF, the more flashy catchment will require a greater storage to give a specified yield. But this flashy catchment has a lower Q70 (or flow exceeded 70% of the time) than its less flashy counterpart, and so the two lines corresponding to a yield of Q70 on each catchment may be close together. It was decided therefore to test the hypothesis that all flows corresponding to a particular percentile on the flow duration curve would have the same storage-yield curve. This hypothesis assumes that the flow-duration curve is an appropriate index of the storage-yield characteristics of a river - if this hypothesis is true, then a single yield line corresponding to, say, Q70 should be applicable to all Malaŵi rivers.

The storage-yield analysis described above was performed for 44 stations and for a range of yield corresponding to Q40, Q50, Q60, Q70, Q80 and Q90 (the last only for stations with non-zero Q90's). The stations were divided into 4 groups according to their Q75 value (or their flashiness). The storage-yield diagrams for a given percentile yield for the stations within each group were plotted on a common base. Figure 5.9 is an example for Q70 for catchments with Q75 > 30% ADF. An average line can be drawn through the scatter of points, and if this average line is not significantly different from those of the other 3 groups for the same percentile yield then the hypothesis may be accepted.

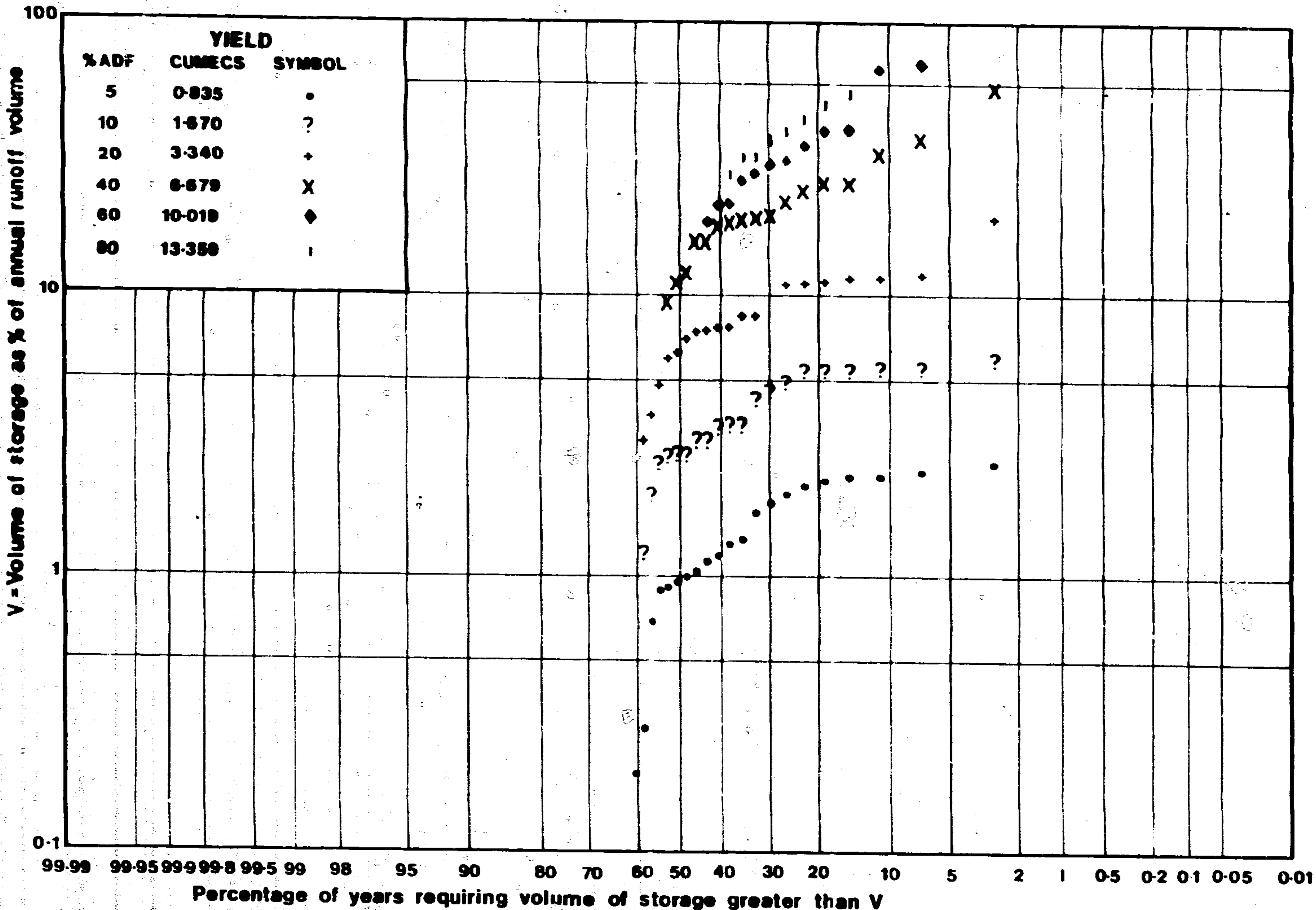


Figure 5.8 Storage yield analysis for station 5 D.2

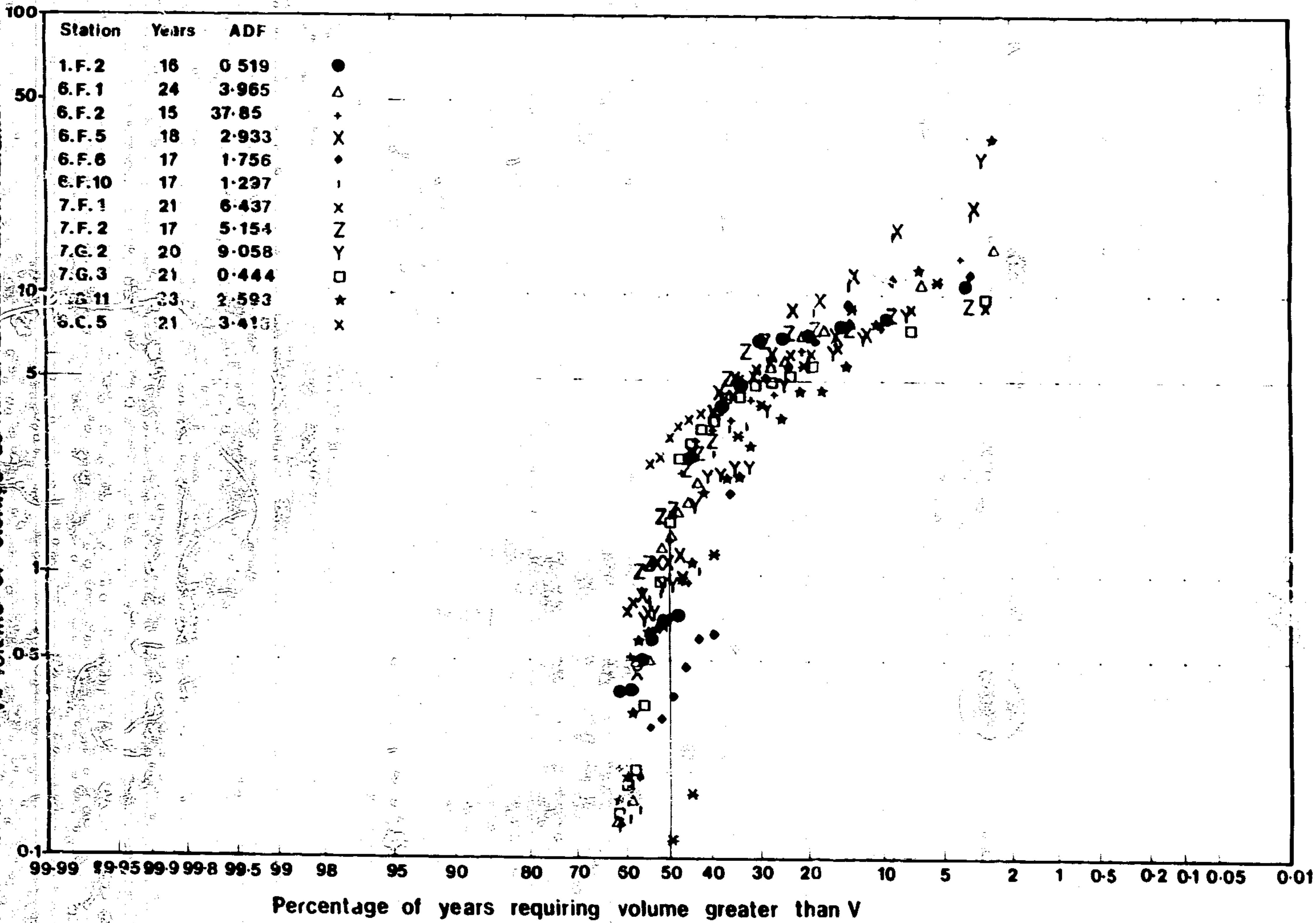


Figure 5.9: Storage yield analysis for a yield of Q70(10) on all catchments with Q75(10) > 40% ADF

The set of standardised storage-yield curves shown in Figure 5.10 was derived by the analysis described above. Within the context of preliminary screening, the hypothesis was generally found to be reasonable, although there was a slight tendency for the slope, but not the average position, of the curve to increase with Q75. The Q40 to Q70 lines are applicable to all the stations, but the Q80 and Q90 lines will over-estimate the storage requirement for stations with Q75 less than 10% ADF. The Q80 and Q90 percentiles relate to very small yields in these catchments, and so this shortcoming is unlikely to be important. The procedure for using these curves is given in Appendix A.

A correction for reservoir evaporation can be made (see Appendix A). The nature of this correction, and of the analysis method described, is the subject of a further study (Kidd, 1980).

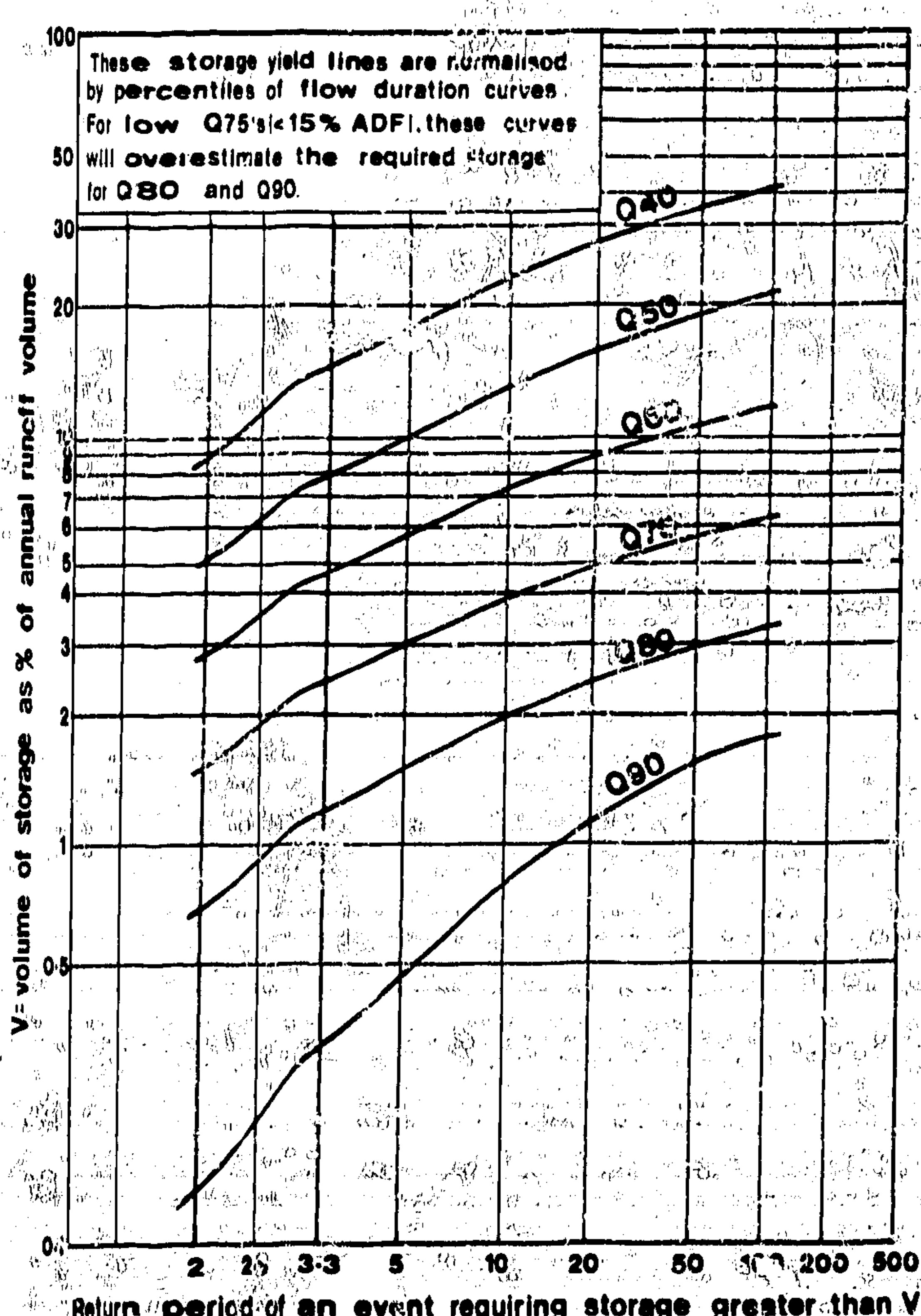


Figure 5.10: Standardised storage yield curves

Recessions and forecasting

In Malawi, all rivers are predominantly in recession between the months of May and October. For irrigation and water supply schemes, operational decisions can be facilitated by the ability to forecast future flows during the recession. Using mean monthly flows, two approaches to the forecasting problem have been tried:

(i) fitting a characteristic recession to the data;

(ii) predicting October flows directly from May flows.

These two approaches should be regarded as demonstrations of possible methods; where the importance of the application warrants, a more intensive investigation along these lines can be undertaken to provide a more accurate forecasting method.

The first approach leads to a generalised model which allows the user, knowing the present flow, to predict the flow at any time in the future. An exponential decay (or linear reservoir) has been fitted to each dry season between May and October, as follows:

$$Q_t = Q_o e^{-t/K} \qquad \ldots\ldots 5.8$$

where Q_t is the flow at some time t months in the future;

Q_o is the flow at the present time;

t is the intervening period (months) - or forecast lead time;

K is the recession constant (months)

Each year yielded a different value of K, and Figure 5.11 is an example for station 1.D.3 (the Tuchila at Muchenza) of the scatter of recessions. Although dry season flows here are augmented by occasional rainfall on Mulanje Mountain, this example is not extreme in respect of the degree of scatter. As shown on Figure 5.11, an average straight recession line can be drawn through average recession constant, called KREC, takes the value of 3.0 months in this case. The values of the recession constant KREC for each of the stations analysed can be found in Table 5.1.

Figure 5.12 shows a plot of KREC against Q75(10), and the following relationship represe[illegible] b[illegible]t linear equation for prediction of KREC:

$$\text{KREC} = 2.17 [illegible] \qquad \ldots\ldots 5.9$$

This equation has [illegible] coefficient of 0.80. No other catchment characteristics were si[illegible]nt after the inclusion of Q75(10). There is a great deal o[illegible]r both in the variation of KREC from year to year at a given [illegible]on and also from station to station, so

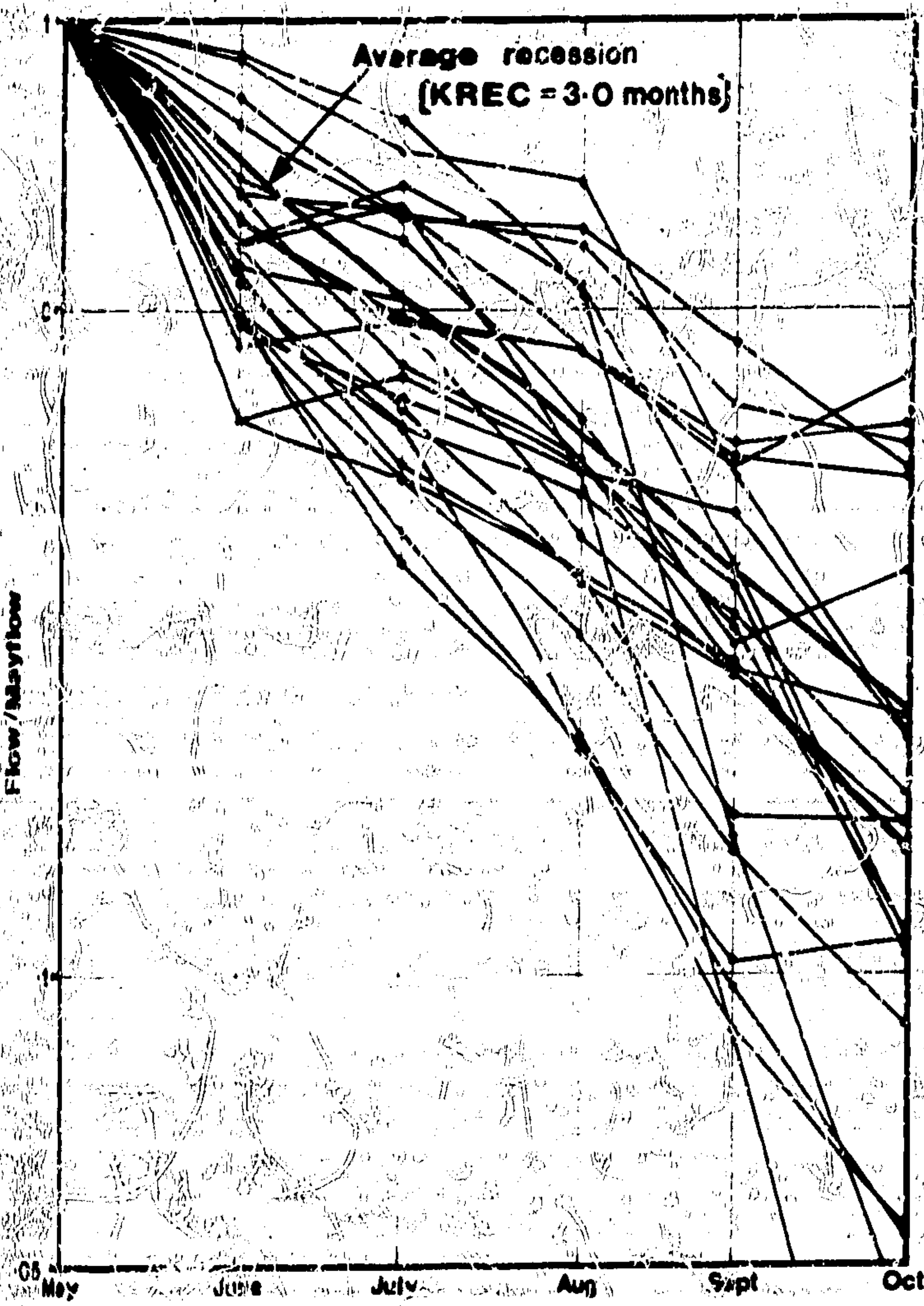


Figure 5.11 Annual recessions for station 1.D.3

No great confidence can be placed in this method of forecasting. This approach has the merit of providing a best average estimate of the flow at any future time at any site where the present flow is known. Little is known, however, about the reliability of this estimate.

The second approach is based on the assumption that the flow at some future date is a function of 2 components: (i) a systematic component depending on the storage at the present time, and indexed by the present flow, and the manner in which this storage is depleted, and (ii) a random component depending on elements (such as rainfall) over which the present conditions have no influence. For a given station, for instance, the October flow (Q_{oct}) and the May flow (Q_{may}) can be related according to the following relationship:

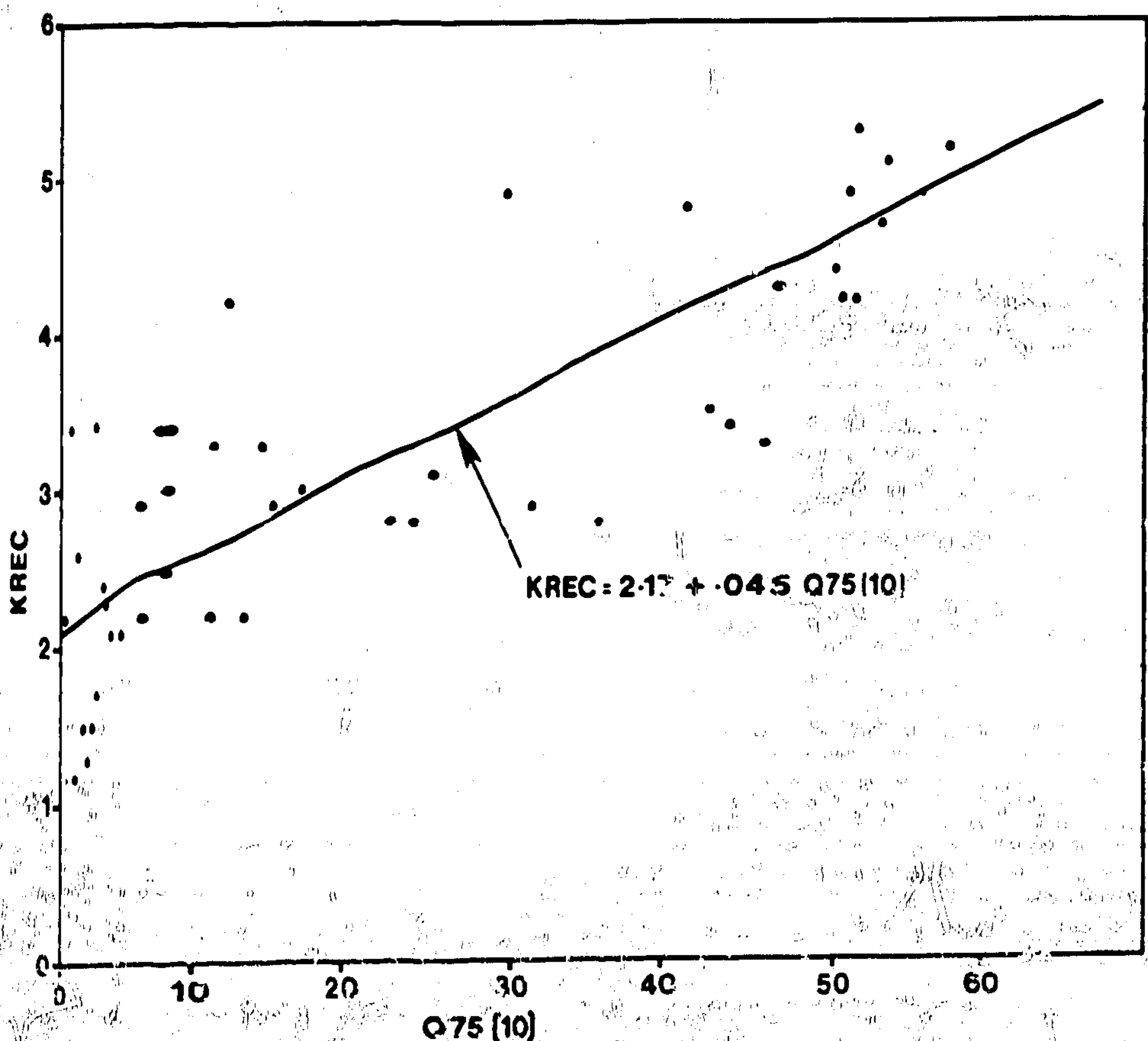


Figure 5.12: Relationship between recession constant KREC and Q75 (10)

$$Q_{oct} = K_1 Q_{may} + K_2 \qquad \ldots . \quad 5.10$$

where K_1 and K_2 can be found by regression analysis. An example is shown in Figure 5.13 for station C.F.1 Limphasa. The systematic component is represented by the straight line, and the random component by the scatter of points about this line. In this case the standard deviation of October flows was .61 and of the residuals about the regression line .41, so that it can be said that the systematic and random components of the variability are in an approximate ratio of 1:2. The analysis described has been performed for a total of 40 stations, the results of which are given in Table 5.2. In about 2/3 of the stations analysed, the random exceeded the systematic component of the variability in October flows. The parameters K_1 and K_2 allow an estimate of the Q_{oct} for a given May flow Q_{may}. A knowledge of the standard error of this estimate, assuming the residuals to be normally distributed, permits a study to be made of the reliability of this estimate.

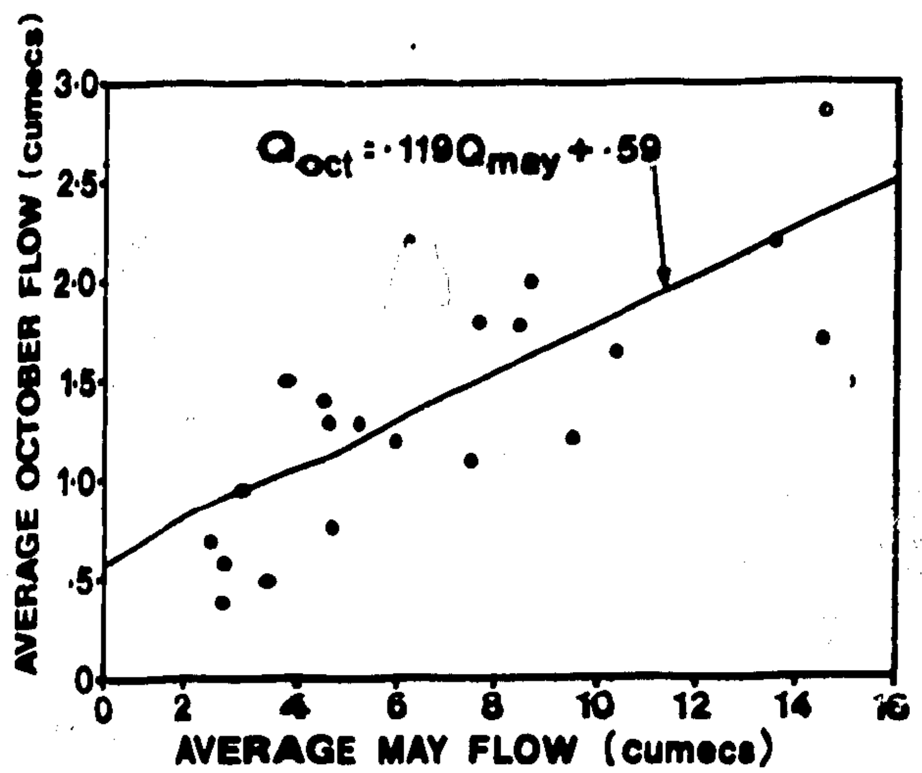


Figure 5.13 Relationship between October flow and May flow for station 6.F.1

Table 5.3 gives the proportion of years in which a given flow can expect to be exceeded given the estimate (Q_{oct}) and its standard error (SE). As an example of this calculation, if the average May flow Q_{may} = 12.0 cumecs for station 6.F.1, then equation 5.11 gives Q_{oct} = 2.0 cumecs. Table 5.3 yields the information that, given Q_{may} = 12.0 cumecs, an October flow of 2.0 cumecs will be exceeded in 50% of years, a flow of 1.59 cumecs in 84% of years, a flow of 1.39 cumecs in 93% of years, etc. The analogous calculation can be made for flows greater than Q_{oct}. While this approach has the added feature of being able to give a measure of the reliability of a forecast flow, its use is limited to these stations where this analysis has been performed. Similar forecasting methods could be developed for other periods (e.g. September flow forecast from May flow). Climatic data, such as total wet season rainfall, could also be used as predictors if appropriate.

$$Q_{oct} = \text{SLOPE } Q_{may} + \text{CONST}$$

Station	Mean Q_{oct} (cumecs)	Standard Deviation (cumecs)	SLOPE	CONST	Correl. Coeffnt.	Standard error (cumecs)	Ratio of systematic to random components
1.D.3	0.54	.39	.097	.20	.64	.29	.3
1.D.10	2.34	.75		insufficient information			
1.D.14	.120	.021	.203	.061	.77	.015	.4
1.D.23	.49	.34	.220	.03	.83	.20	.7
1.F.2	.204	.104	.408	.003	.94	.037	1.8
1.K.1				Q_{oct} usually zero			
1.R.18	.004	.003		insufficient information			
2.B.8	.219	.020	.066	.074	.70	.016	.2
2.B.21	.27	.33	.125	-.01	.95	.10	2.3
2.B.22	.24	.33	.156	.06	.37	.32	.0
2.C.8	.008	.007	.010	.000	.93	.003	1.3
3.E.3	.29	.24	.251	.00	.92	.10	1.4
3.F.3	.24	.09	.132	.03	.85	.04	1.0
4.A.3	.32	.26	.177	.11	.49	.25	.0
4.B.1	1.26	2.11	.087	-.22	.48	1.90	.1
4.B.3	.037	.032	.022	.014	.55	.027	.2
4.B.4	.16	.16		no systematic component			
4.D.4	.34	.12	.016	.26	.48	.11	.1
4.D.6	.29	.18		no systematic component			
4.E.1	.052	.051	.009	.044	.19	.051	.0
4.E.2				insufficient information			
4.F.5	.012	.011	.098	0	.97	.003	2.7
5.A.8	.28	.19	.174	-.02	.80	.12	.6
5.D.1	.52	.94		insufficient information			
5.D.2	.36	.38	.020	.02	.65	.29	.3
5.D.3	.035	.090	.225	-.010	.86	.048	.9
5.E.1	.034	.030	.047	.010	.57	.025	.2
5.E.2	.10	.13	.088	.01	.64	.10	.3
6.C.1				Q_{oct} usually zero			
6.F.1	1.38	.60	.119	.59	.75	.41	.5
6.F.2	12.91	3.83	.110	6.34	.73	2.79	.4
6.F.5	1.18	.59	.299	.01	.91	.25	1.4
6.F.8	.69	.26	.172	.23	.87	.13	1.0
6.F.10	.52	.26	.275	.02	.88	.13	1.0
7.A.3	.15	.16	.163	-.11	.90	.07	1.3
7.A.4	.48	.31	.397	-.09	.92	.12	1.5
7.D.3	.49	.27	.106	.10	.73	.19	.4
7.E.2	.26	.35	.054	-.18	.76	.24	.5
7.F.1	1.95	.55	.194	.35	.73	.35	.6
7.F.2	2.07	.50	.232	.65	.79	[illegible]	.6
6.G.2	3.85	.95	.218	1.17	.94	[illegible]	1.9
7.G.3	.21	.06	.080	.16	.37	.03	.2
7.G.11	1.17	.35	.088	.8	.49	[illegible]	.1
7.G.14	4.86	1.91	.152	.53	[illegible]	[illegible]	1.1
7.G.15	1.64	.57		no systematic component			
8.A.2	2.25	.95	.168	.47	.65	.75	.3
8.C.5	1.72	.52	.261	.57	.94	.19	1.7
8.C.6	1.06	.32	.105	.43	.88	.17	.9
9.A.2	1.01	.48	.078	.53	.73	.35	.4

Table 5.[illegible] Relationships between October flows and May flows

Average monthly flow Q	Percentage of years when flow > Q	Return period of Q
Q_{oct}	50%	
Q_{oct} - .1SE	54%	
Q_{oct} - .2SE	58%	
Q_{oct} - .5SE	69%	
Q_{oct} - 1.0SE	84%	
Q_{oct} - 1.5SE	93%	
Q_{oct} - 2.0SE	98%	0

Table 5.4: Reliability of the forecast of the flow in October

6. FLOOD FREQUENCY

Introduction

This study aimed to provide a flood frequency relationship for rivers in Malawi which could be used to make estimates of the flood of a given return period for any site in the country irrespective of the availability of river flow data. The methods used to derive these relationships are based on those developed during the UK flood studies (Natural Environment Research Council, 1975). Because of lack of autographic raingauge data it was not possible to study the floods resulting from specific rain storms and thus, for example, no methods are given for estimating the probable maximum flood for a site.

An annual flood is defined as the largest instantaneous flow in any given year. An important feature of flood hydrology is the high variability of annual floods which makes it very difficult to make reliable estimates of high return period floods even with quite long periods of record. In practice, hydrological records of 20 to 25 years, which occur in Malawi, are not long enough to give reliable flood estimates. To overcome this difficulty, the flood data from a more or less geographically homogeneous region are pooled to provide a more stable estimate of the flood frequency curve. Experience has shown that this regional flood frequency curve, as it is known, is a reliable means of predicting flood frequencies, and is to be preferred to estimates made from the data of a single station. The flood data for analysis are standardized to the same scale by dividing by the mean annual flood, which is the arithmetic mean of the annual floods. When the regional curve is used to estimate the magnitude of floods at a particular site, the mean annual flood is required to scale the results from the curve. If a few years of flow data are available at the site, then the mean annual flood can be calculated directly. However, in many cases, flood estimates will be required for sites where no flow data have been collected. To meet this requirement, an empirical equation has been derived enabling the mean annual flood to be estimated from catchment characteristics. The reliability of this equation is much less than that of the regional curve, and, when flow data are available, these should be used in preference for estimating the mean annual flood.

Flood data processing

The Water Resources Division provided estimates of the annual maximum instantaneous flows and the dates on which they occurred for 30 gauging stations (Chapter 2). Station 4.F.6. Lumbadzi was discarded from the analysis as it had only 4 annual maxima, that for 1975 being about ten times the other three. Station 8.A.2, North Rukuru, has now been abandoned and 8.A.5 is used in its place. For the purposes of this study the records of the two stations have been combined to give a single record. This combined record is referred to below as 8.A.2 with the floods transferred from 8.A.5 being reduced by 5%, which is the difference in catchment area of the two stations. These changes left 28 stations for analysis with 509 station years of record (average length 18.1 years).

Computer processing took part in two stages: individual station analysis and regional combination. In the first stage the floods for each station were ranked in order and plotted against the appropriate plotting position. Statistics of the flood record such as mean, median and coefficient of variation were calculated. At the second stage a regional curve was produced by averaging the individual station curves.

Individual station statistics

Flood statistics for the 28 stations analysed are shown in Table 6.1;

Station	Mean annual flood (cumecs)	CV annual flood (%)	Median annual flood (cumecs)	No. of floods
1.D.3	306	96	214	26
1.D.24	43.6	90	33.0	23
1.R.3	326	70	225	18
2.B.22	169	100	92.0	16
4.B.1	452	36	468	22
4.B.3	185	71	134	12
4.B.4	191	65	172	20
4.D.4	142	75	125	23
4.E.1	68.0	52	59.0	24
5.A.8	163	97	104	16
5.B.13	577	85	340	8
5.C.1	775	58	629	21
5.D.1	142	54	130	20
5.D.2	120	74	106	26
5.D.3	65.9	149	25.0	21
6.C.1	77.5	56	69.0	19
6.E.6	108	42	104	10
6.E.7	55.1	30	47.0	7
6.F.1.	31.9	29	36.1	10
6.F.2	312	55	272	14
6.F.5	14.2	35	13.0	20
6.F.6	22.4	60	19.5	18
7.A.3	30.0	67	25.4	18
7.D.3	166	113	84.0	24
7.D.7	76.4	55	65.0	23
7.F.2	22.7	55	19.0	17
7.F.3	24.6	25	24.0	9
8.A.2	254	85	142	24

Table 6.1. Annual flood statistics

two measures of the general size of floods at a station are given, the mean and the median. The use of the mean annual flood (MAF) to scale the regional frequency curve has been mentioned above, and this is the preferred estimator of the general size of floods. However, in the case when the record contains one very large flood much greater than the rest, the mean can be distorted, and in this case the median, which does not depend upon the extremes of the record, can be used as an alternative measure of the scale of floods. For the 28 stations studied the average ratio of the mean to the median was 1.3:1. In situations where there is a single very large flood, the mean could be estimated by multiplying the median by 1.3. This technique should only be used when the highest flood in the record is more than, say, seven times the median, and considerably higher than the second largest.

The coefficient of variation (CV) is a dimensionless measure of the variability of the annual floods at a station and is defined as the standard deviation of the annual floods divided by the mean expressed as a percentage. The average value of CV is 67% with values ranging from 25% to 149%. This average CV is not unusually high by world standards, though it is much greater than the value for Great Britain (35%). The wide range of CV is not unusual and is probably due more to sampling error than to any inherent difference in variability between stations. Figure 6.1 shows a map with point values of CV, no obvious pattern of variation being discernible.

Regional frequency curve

Figure 6.2 shows the regional flood frequency curve together with the points from which it was derived. Table 6.2 gives the ordinates of this curve for various return periods. Return period is defined as the average interval (in years) between occurrences of a flood greater than or equal to a given value - it is also the reciprocal of the exceedance probability of a flood in any given year. The ratio of the annual flood divided by the mean annual flood is plotted against the reduced variate, y. The reduced variate is related to the return period, T, by the following equations

$$F = 1 - \frac{1}{T}$$

$$y = -\ln(-\ln F)$$

If T is greater than 5 this relationship may be approximated by

$$y = \ln(T-\tfrac{1}{2}).$$

The graph is then entered with this value of y to give the flood of the required return period. The curve drawn assumes that the floods follow a general extreme value distribution, that is to say

$$\frac{Q(T)}{MAF} = u + \frac{\alpha}{k}(1 - e^{-ky}),$$

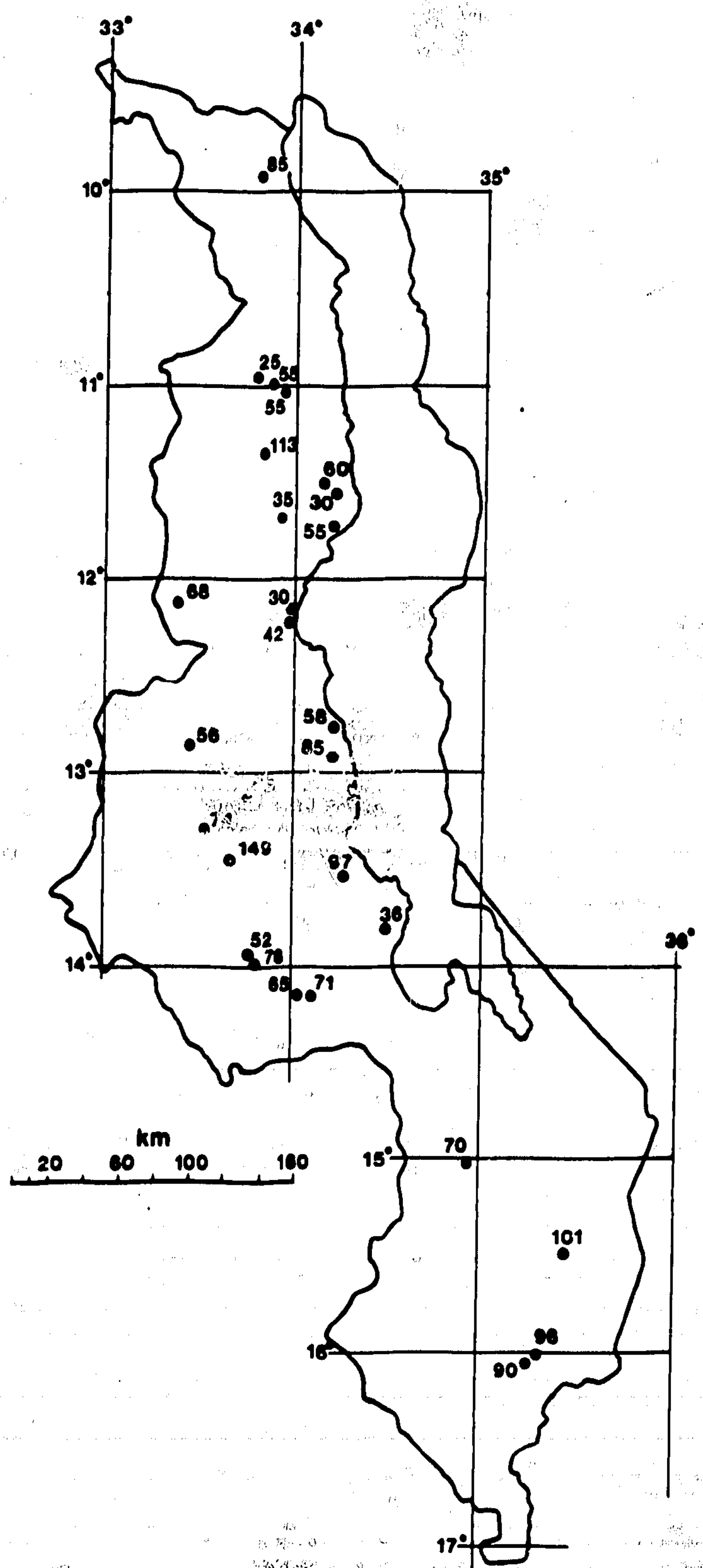


Figure 6.1 Coefficients of variation of annual flood %

Return period T (years)	$\frac{Q(T)}{MAF}$	Reduced variate y
2	.83	.37
4	1.27	1.25
5	1.40	1.50
10	1.85	2.25
20	2.33	2.97
40	2.87	3.68
50	3.05	3.90
100	3.67	4.60
200	4.36	5.30
400	5.14	5.99
500	5.41	6.21
1000	6.32	6.91

Table 6.2. Ordinates of regional flood frequency curve

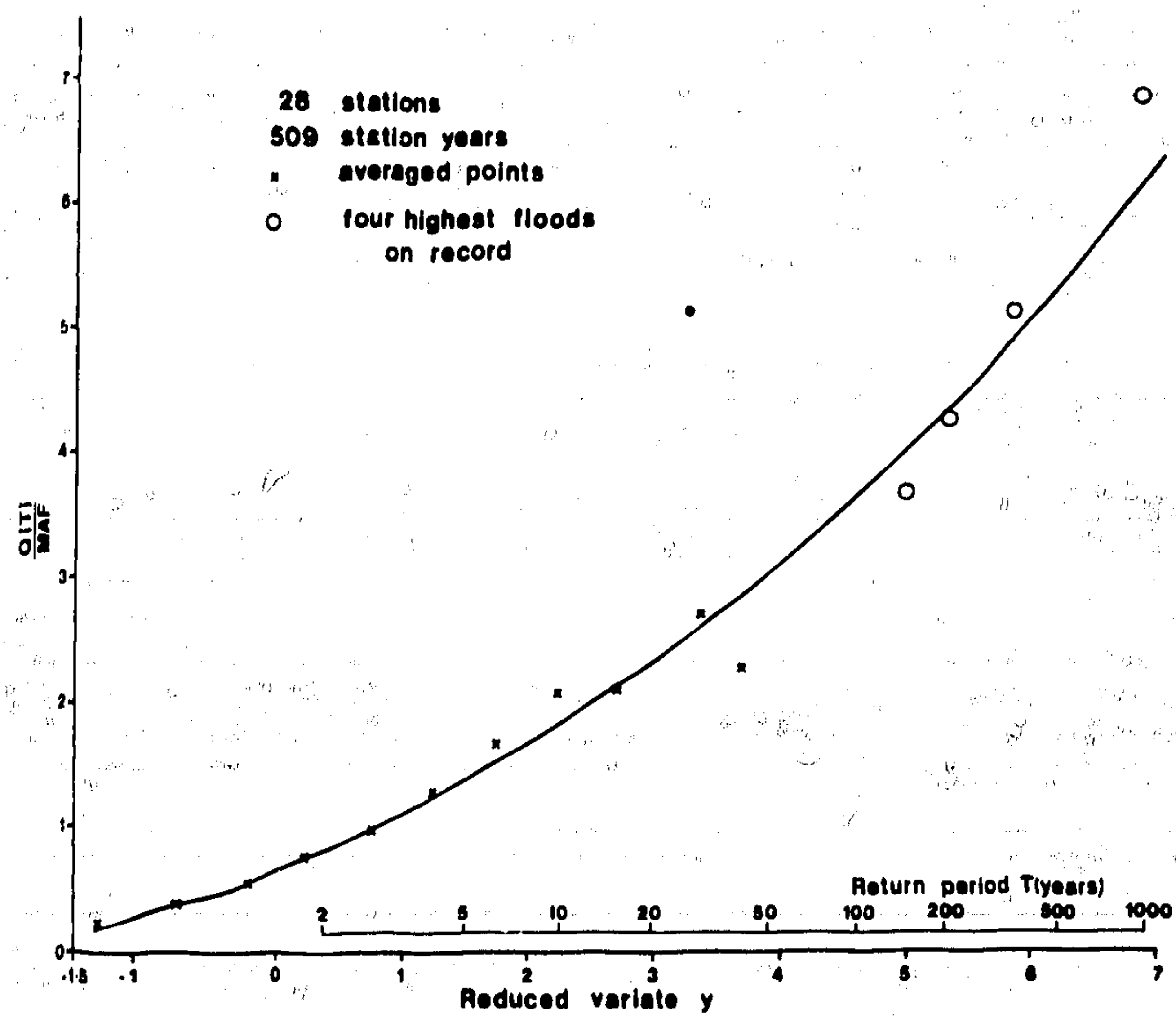


Figure 6.2: Regional flood frequency curve

where $Q(T)$ is the flood of return period T
MAF is the mean annual flood
y is the reduced variate for the return period T

and u, α and k are parameters of the distribution.

In this case $u = 0.67$, $\alpha = 0.43$, $k = -0.17$ and the above relationship reduces to

$$\frac{Q(T)}{MAF} = -1.92 + 2.58 \exp(0.17y)$$

or on substituting the approximate relationship for reduced variate in terms of return period

$$\frac{Q(T)}{MAF} = -1.92 + 2.58(T - 0.5)^{0.17}.$$

Derivation of the regional frequency curve

The regional frequency curve is drawn by averaging the individual frequency curves for each station. The annual floods for each station were first standardised by dividing by the mean annual flood for the station. The plotting positions for plotting the curves are the expected value of the order statistics for the extreme value type 1 or Gumbel distribution. A table of these plotting positions is given in the Flood Studies Report (Natural Environment Research Council) and they can be approximated by Gringorten's formula

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

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

where n is the number of years of record,
and i is the rank (from the smallest upwards).

To average the curves of individual stations the plotting position axis was divided into steps of length 0.5 (ie - 1.5 to - 1.0, - 1.0 to - 0.5, - 0.5 to 0 etc.) and all the points for all the stations lying in each step were averaged and plotted as a single point. These are shown by crosses on Figure 6.2. The four highest values of Q(T)/MAF were then extracted as though they were the four largest floods in a record of 509 years and plotted as circles on Figure 6.2. The floods concerned are as follows:

Station	Date	Flood	Q(T)/MAF	y_i
5.D.3	16.12.75	450	6.83	6.81
7.D.3	27. 1.71	368	5.13	5.81
1.D.3	15.3.67	1302	4.26	5.29
4.D.4	13.3.78	425	3.67	4.96

It will be seen that they are all from different stations and on widely differing dates so the assumption that they are statistically independent events is seen to be justified. The general extreme value curve was fitted by least squares and can be seen to be a good representation of the points.

Averaging the station curves enables the regional frequency curve to be drawn with confidence up to a return period of about 50 years. Beyond here the individual points for the four highest floods allow an extension to higher return periods, but with reduced accuracy. The highest point is plotted at a return period of 907 years and thus 1000 years should be regarded as the limit for reliable extrapolation. Beyond this there are great uncertainties in the flood frequency relationship and the curve should only be used with caution.

Prediction of mean annual flood from catchment characteristics

For each catchment the following characteristics were measured: AREA, STMFRQ, S1085, MAR, DAMBO, ESCARP. Definitions of these characteristics are given in chapter 3. In addition Q75(10) was measured from flow data or, when data were not available, taken off the map given on Figure 5.2. Q75(10) is defined as the 10 day flow exceeded 75% of the time, expressed as a percentage of the Average Daily Flow, ADF (see chapter 5). For just one station, namely 6.C.1 Dwangwa, Q75(10) was zero, so the value was arbitrarily set to 0.1 to allow the use of a logarithmic transformation. The mean annual flood (MAF) and the catchment characteristics were transformed by taking logarithms to the base 10, except for ESCARP (which was not transformed) and DAMBO which was transformed by $\log_{10}(1 + \text{DAMBO})$ - this is appropriate for proportions between 0 and 1. The correlation matrix of the transformed data is shown in Table 6.3.

The equation to be used in estimating the mean annual flood, when no flow data are available, is:

$$\text{MAF} = 2.89\,\text{AREA}^{0.55}\,\text{STMFRQ}^{0.36} \qquad \text{.... } 6.1$$

where MAF is the mean annual flood in m^3/s
AREA is the catchment area in km^2
and STMFRQ is the stream frequency in junctions/km^2.

The corresponding equation in the log variables,

$$\log(\text{MAF}) = 0.461 + 0.55 \log(\text{AREA}) + 0.36 \log(\text{STMFRQ}),$$

has a standard error of estimate (s.e.e.) of 0.378. This means that in 95% of estimates made using this equation the predicted log(MAF) will be within $\pm$ 0.74 (1.96 times s.e.e.) of the correct value. By taking antilogs of these figures we get the factorial error of estimate as 2.39 and 95% of estimates of MAF in the range 5.5 MAF to MAF/5.5. Thus it will be seen that the answer given by this equation is not very precise. There are a number of reasons for this lack of precision, one

	MAF	AREA	STMFRQ	Q75(10)	AAR	S1085	DAMBO	ESCARP
MAF	1.0	0.542	0.004	-0.278	-0.175	-0.380	0.260	0.213
AREA	0.542	1.0	-0.473	-0.457	-0.347	-0.706	0.659	-0.091
STMFRQ	0.004	-0.473	1.0	0.627	0.543	0.720	-0.833	0.376
Q75(10)	-0.278	-0.457	0.627	1.0	0.771	0.650	-0.621	0.433
AAR	-0.175	-0.347	0.543	0.771	1.0	0.477	-0.429	0.413
S1085	-0.380	-0.706	0.720	0.650	0.477	1.0	-0.791	0.426
DAMBO	0.260	0.659	-0.833	-0.621	-0.429	-0.791	1.0	-0.251
ESCARP	0.213	-0.091	0.376	0.433	0.413	0.426	-0.251	1.0

Table 6.3. Correlation matrix of transformed characteristics and mean annual flood.

Variable	Coefficient	std error	t	p
Constant	0.461	0.408	1.13	0.2
AREA	0.553	0.141	3.92	0.0006
STMFRQ	0.360	0.191	1.88	0.072

Analysis of Variance

Source	SS	df	MS
Due to regression	2.1950	2	1.0975
About regression	3.5645	25	0.1426
Total	5.7595	27	

F ratio = 7.70 with 2 and 25 degrees of freedom,

probability of exceedance 0.0025

Multiple determination coefficient (R^2) 0.38

Multiple correlation coefficient (R) 0.62

Standard error of estimate 0.378

Table 6.4: Statistical description of the recommended equation for predicting mean annual flood from catchment characteristics

of which is the small number of stations available for analysis. Another cause arises from the relatively high variability of annual floods in Malawi with average coefficient of variation equal to 67%; this means, for example, that the mean annual flood estimated from a 20 year record has a standard error of $\pm$ 15%. A fuller statistical description of this regression equation is given in Table 6.4.

Derivation of the regression equation

Several other combinations of variables were tried in deriving the equation recommended above. Table 6.5 shows a selection of the better equations. AREA is by far the most important explanatory variable and regression on AREA alone gives a standard error of estimate of 0.40, and adding STMFRQ to give the best two variable equation reduces this to 0.38. Other two variable combinations such as AREA and Q75(10) showed no reduction in standard error. After AREA and STMFRQ are in the regression adding the slope variable S1085 gives the greatest reduction in standard error. However, the reduction is small from 0.38 to 0.37 and the coefficient of S1085 is negative, whereas one would expect it to be positive (that is to say greater slopes would be expected to be associated with larger floods). Replacing S1085 with Q75(10) gave [illegible] the same standard error but the coefficient of Q75(10) [illegible] of S1085, was not significantly different from zero. [illegible]ntly, it was decided to recommend the best two variable equation using AREA and STMFRQ.

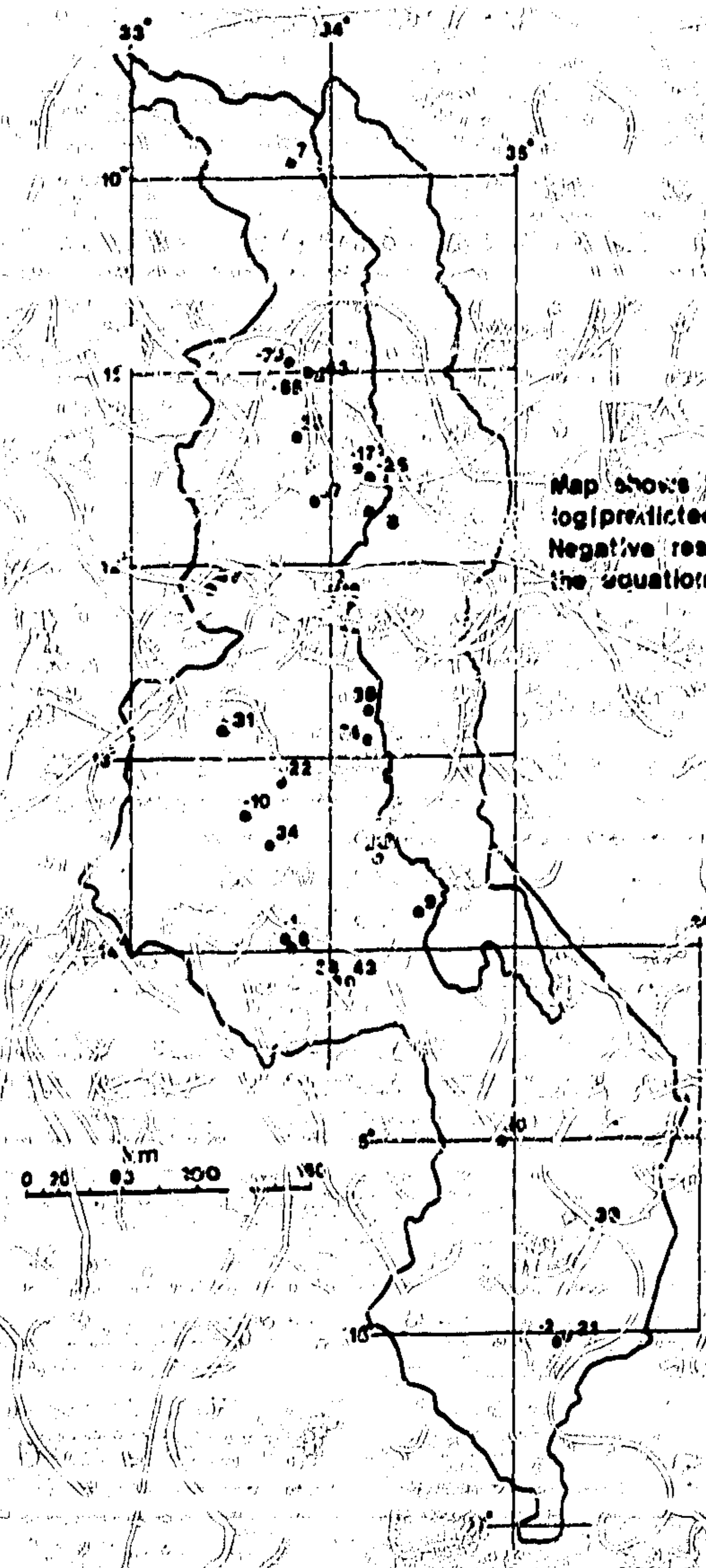
A map of residuals from the recommended equation is shown in Figure 6.[illegible] The residuals are defined as log (observed MAF) - log (predicted MAF), so that a negative residual means that the equation overestimates the mean annual flood. It will be seen that there is some tendency for stations on the escarpment to be under-estimated.

An attempt was made to improve estimation by splitting Malawi into two separate regions using the variable ESCARP. If it is included in the regression it effectively gives a different constant term in the regression equation for each region. The regression on AREA, STMFRQ and ESCARP has much the same standard error as the regression without ESCARP, though the value of R^2 is slightly higher. The use of different regressions in the two regions was also tested but this proved to be of no advantage. An analysis of variance was undertaken of the three options: (i) single equation, (ii) different constant terms but the same coefficients for the variables, and (iii) completely different coefficients. No differences were identified which were significant at less than the 5% level.

Finally the best three variable equation is shown in Table 6.5. It uses STMFRQ, S1085 and ESCARP. It was rejected for the following reasons: (i) its improvement on the recommended equation is small, (ii) AREA is an important variable and was omitted, and (iii) S1085 has a negative coefficient.

Estimation of the mean annual flood from data

When more than five annual floods are available these should be used, in preference to the equation, to estimate the mean annual flood. If less than five are available then both the equation and the data can be combined to give an estimate. The weight to be given to the estimate from the equation must be decided individually in each case and will depend upon a number of factors such as, the reliability of the flood records from the station in question, the performance of the equation on nearby or similular catchments, and the number of floods in the data.



Map shows (log(observed) - log(predicted)) x 100. Negative residual means the equation overestimates.

Variables	Coefficient	Standard error	t	p	R^2	R	s.e.e.
Const	0.77	0.39	1.98	0.06	0.293	0.541	0.396
AREA	0.43	0.13	3.29	0.003			
Const	0.46	0.41	1.13	0.27	0.381	0.617	0.378
AREA	0.55	0.14	3.92	0.0006			
STMFRQ	0.36	0.19	1.88	0.07			
Const	0.84	0.52	1.62	0.12	0.295	0.543	0.40[illegible]
AREA	0.41	0.15	2.77	0.01			
Q75(10)	- 0.029	0.14	0.21	0.84			
Const	1.27	0.67	1.89	0.07	0.434	0.659	[illegible]
AREA	0.40	0.17	2.33	[illegible]			
STMFRQ	0.53	0.24	2.44	0.02			
S1085	- 0.41	0.28	1.50	0.15			
Const	0.84	0.48	1.76	0.09	0.431	0.656	0.370
AREA	0.51	0.14	3.56	0.002			
STMFRQ	0.53	0.22	2.40	0.02			
Q75(10)	- 0.22	0.15	1.45	0.16			
Const	3.11	3.07	1.01	0.32	0.400	0.632	0.379
AREA	0.54	0.14	3.77	0.0009			
STMFRQ	0.45	0.22	2.06	0.05			
AAR	- 0.86	0.99	0.87	0.39			
Const	0.61	0.39	1.57	0.13	0.363	0.602	0.383
AREA	0.45	0.13	3.53	0.002			
ESCARP	0.24	0.15	1.65	0.11			
Const	0.42	0.41	1.04	0.31	0.407	0.638	0.377
AREA	0.54	0.14	3.80	0.0009			
STMFRQ	0.28	0.21	1.35	0.19			
ESCARP	0.16	0.16	1.04	0.31			
Const	2.68	0.21	12.6	$<10^{-4}$	0.444	0.666	0.365
STMFRQ	0.56	0.24	2.35	0.03			
S1085	- 0.93	0.23	- 4.12	0.0004			
ESCARP	0.38	0.16	2.45	0.02			

Table 6.5: Statistical summary of some rejected equations for predicting mean annual flood

7. CONCLUSIONS AND RECOMMENDATIONS

Conclusions

The results of a regional low flow study and a regional flood study have been presented. These studies should permit the estimation of a variety of low and high flow measures to be made in Malaŵi irrespective of the amount of hydrological data available at the site of interest. On the low flow side, the flow measures considered are the average daily flow, flow duration curve, low flow frequency curve, storage-yield characteristics and recession characteristics. For the flood study, estimation procedures have been developed for the mean annual flood and the flood frequency curve.

In all these studies, the whole of Malaŵi has been considered as one hydrological region. It has been suggested that particular parts of the country might behave in a fundamentally different way from other areas, and that these should be treated separately. Some of these reputed differences can be discerned in the results of the study but statistical analysis does not confirm their existence. In general, the study is based on too few data to detect regional differences and consequently it has been necessary to consider Malaŵi as a single region.

One exception to the above is the river Shire, the lake outflow. The existence of the lake upstream of stations 1.B Shire at Liwonde, and 1.G.1., Shire at Chiromo, has such a funda[illegible] effect on the flow characteristics that no attempt was made to include them in the analyses. Likewise, the results of this study should not be applied to the Shire, which is treated more fully elsewhere (World Meteorological Organisation, 1976; Drayton, 1980).

A familiar theme in the preceding chapters has been the need to combine estimates made from data with the regionalised empirical estimates described here. It cannot be emphasised too strongly that the methods in this report should not be used in isolation. Whenever possible any flow data available at the point of interest <u>or</u> at nearby sites should be examined and compared with the empirical estimates. It is neither possible nor prudent to quantify the relative weights to be applied to the data and to the empirical estimates. This must be left to the judgement of the hydrologist. The factors he should consider include (i) the quality of the available data, (ii) whether the period when the data were collected was abnormally wet or dry, and (iii) the performance of the empirical method on nearby gauged catchments having similar characteristics.

A regional study of this nature can often be of greatest value when dealing with small catchments. Projects using the water from these small catchments will themselves be small and the cost of a lengthy hydrological investigation would not be justified. In the low flow study there are seven catchments under 100 km^2 in area with the smallest being 7.0 km^2 (Table 3.1). However, the distribution over the country is very uneven, with five in the areas termed highland (two on Mulanje, two on Zomba and one on Nyika) and the remaining two catchments under 100 km^2 on tributaries of the lower Shire. The position is less satisfactory with respect to flood catchments. The smallest catchment, and the only one under 100 km^2, has an area of 63.5 km^2. The behaviour of small catchments is usually more variable than larger ones and there may be difficulties in extrapolating these results to small catchment areas. However, it should be possible to use the results of both studies in areas as small as 20 to 30 km^2. More data from small catchments are needed if this is to be improved. The Water Supply Branch and the Water Resources Branch have both made measurements on very small persistent streams in October, to assess their use for rural gravity piped water supplies. Though these data are obviously biased, in that only persistent streams are measured, they should be examined to see whether general guidelines can be developed.

The flood analysis methods proposed here are not intended to be applied to urban drainage problems. Firstly, catchments being considered in urban drainage are usually very small, below the range where these methods are considered applicable. Secondly, although the density of settlement in towns in Malaŵi is low, one effect of urban development is to fundamentally alter the hydrological catchment response. The Water Resources Branch has instrumented an urban catchment in Capital City Centre, Lilongwe and the results from this should provide a means of assessing the applicability of standard urban drainage design methods in Malaŵi.

The behaviour of a river is often characterised within a spectrum of conditions ranging from "flashy" catchments through to "damped" catchments. A flashy river is one whose catchment response is dominated by direct runoff; the annual hydrograph is generally characterised by great variability in wet season flows with large floods and a dry season flow which recedes rapidly. A "damped" river is one whose catchment response is dominated by subsurface runoff - the annual hydrograph has a very damped wet season response and a very persistent dry season recession. An exception to this behaviour pattern was observed in Malaŵi for catchments dominated by dambo along the main stream. In this situation, the dambo provides a lot of flood plain storage resulting in a sluggish wet season response, characteristic of the "damped" river. However, the dry season recession is more characteristic of a "flashy" river, and the late dry season flows are comparable to neighbouring unaffected catchments. This is not to say that there is no water being retained in the dambo through the dry season - it seems that its low transmissivity provides a source of water for vegetation during this period without significantly releasing water to the river. These dambo-affected catchments are not significantly distinguishable from their unaffected neighbours as far as their values of Q75(10) is concerned. On the other hand, it has been suggested in Chapter 4 that the presence of dambo increases evaporation with a corresponding decrease in average annual yield. The nature and behaviour of dambos is not yet well understood, and is an area which could be a useful subject of further study.

The study has shown that the most persistent flows occur on the rivers draining the highland areas of Mulanje, Zomba and Nyika. The high values of Q75(10) in these regions are a striking feature of the contour map shown in Figure 5.2. This is at first surprising because the geology of these massifs consists of impervious intrusive rock, mostly granite. In practice, such geology provides a good deal of groundwater storage, although a hydrogeologist would not normally view these rocks as being good aquifers due to the problems of extraction. This feature of granites has also been noted in the UK low flow study (Institute of Hydrology, 1980). A secondary factor in these persistent flows off the massifs could be the occurrence of rainfall, usually light, during the dry season.

A number of catchment characteristics which may be measured off maps are described in Chapter 3. Only catchment area, average annual rainfall, stream frequency and proportion of dambo were found to be useful in predicting the flow measures used in this report. The other characteristics, S1085 and ESCARP, proved to have little predictive power. Land use is often suggested to be an important determinand of catchment response. In a developing country like Malawi, the land use patterns will be changing rapidly as more land is brought under cultivation and cultivation methods themselves change. This makes it difficult to apply a land use index in a retrospective study of this nature and for the user of the results of the study to determine the land use pattern of his catchment. In addition it has been found very difficult to isolate the effects of land use in catchment experiments conducted for this purpose in many parts of the world. It thus seems unlikely that land use will have a very significant effect on the rather crude measures of flow variability used in this report.

There is considerable variability between the flood records taken from different catchments, as can be seen from the coefficient of variation values given in Table 6.1 and Figure 6.1. Most of this variability can be ascribed to sampling variation only, rather than to inherent differences between catchments. Because of this it is possible to derive the regional flood frequency curve shown in Figure 6.2, which is more reliable for estimating floods of all return periods than individual station curves. Figure 6.2 gives the flood of a particular return period as a multiple of the mean annual flood. This can be estimated from annual flood data or, where these data are not available, from a regression equation using catchment area and stream frequency. This equation is not very reliable and should not be regarded as a substitute for using data, when several years data are available. When less than five years of annual flood data are available it is suggested in Chapter 6 that the result of the equation could be combined with the available flood data.

An attempt was made to improve the equation for estimating the mean annual flood by dividing the country into two regions along the top of the escarpment. The improvement obtained is neither statistically significant nor practically useful and this division is not recommended. There are probably not enough stations in the analysis to establish, unambiguously, any regional differences.

The residual map, Figure 6.3, shows that in the south of the country (south of Lilongwe, say) most residuals are positive. In other words, the equation underestimates the mean annual flood. This is also the part of the country which is open to cyclones coming in from the Indian Ocean and this may be the cause of the floods being higher than predicted. Against this view it should be noted that there are very few data points in this region, and it is quite possible that the result may have arisen by chance. Also the stations in this region do not show evidence of large outliers which might show the effects of occasional cyclones. Because of the few stations involved and the difficulty of defining the region affected by cyclones no formal statistical analysis of possible differences was made.

Further work on floods

This report has only dealt with the frequency analysis of floods. Many engineers dislike using this approach for the design of major structures such as reservoir spillways, particularly when loss of life could result from failure. The design return period to be adopted in these cases would involve extrapolation well beyond the range of the data used in this analysis. The preferred alternative uses the probable maximum flood (PMF). Two steps are required for estimation of the probable maximum flood: (i) estimation of the probable maximum precipitation (PMP), and (ii) conversion of the PMP into a PMF.

A regional estimate of PMP would require a national study of severe storm records by the Meteorological Service. It is recommended that such a study should be an important priority. The conversion of this PMP into a flood flow would require the development of a rainfall-runoff model suitable for storm events; a unit hydrograph is often used. At present the data do not exist for deriving unit hydrographs on a wide scale in Malaŵi, but the installation of level recorders in recent years means that suitable flow data are progressively becoming available. However, there is a lack of autographic raingauges on catchments with level recorders, and this would have to be remedied before a storm rainfall-runoff study could be started.

A rainfall-runoff model can also be used with smaller storms than the PMP to give useful flood estimates. If sufficient flood data are not available for estimating a design flood, then an estimate can be made from rainfall statistics using the model. This approach requires that the available autographic raingauge data be analysed to provide rainfall statistics. As an example, the rainfall-runoff model could be calibrated from one or two years data, less than are needed to give a reliable estimate of the mean annual flood. Alternatively, a regional study of, say, unit hydrographs would give the necessary parameters.

Thus to extend the flood analysis described here, more short duration rainfall data should be collected and analyzed. In addition, concurrent rainfall and runoff data over durations down to one hour should be collected for rainfall-runoff model development.

Further work on low flows

The average daily flow (ADF) is an important parameter in all the estimation techniques described here. Its estimation from rainfall was described in Chapter 4 and the urgent requirement for a reliable map of annual average rainfall to assist in this has already been discussed. The method described for estimating ADF from average annual rainfall is not as accurate as would be considered desirable. It was derived from data which were readily available in computer-based form at the time of the study. However, rainfall and river flow have been measured at many more sites than were considered here, and some of these extra sites could be analysed to provide more data for the analysis of ADF. When more data are available, it may be possible to develop a contour map of annual average yield along the same lines as the Q75(10) map in Figure 5.2. The coverage of stations used in the analysis described here was inadequate to justify such an approach.

Q75(10) is the other key parameter in the analysis methods described in this report. The map in Figure 5.2 has a number of shortcomings due to lack of data in certain areas and the coverage of stations was discussed in Chapter 2. At some time in the future (say, 5-10 years), the map should be updated using data from new stations and revised estimates at existing stations using data since 1975.

ACKNOWLEDGEMENTS

Dr Mandeville is seconded by the Overseas Development Aministration of the U.K. Foreign and Commonwealth Office as Principal Hydrologist to the Department of Lands, Valuation and Water of the Malaŵi Government. Mr Drayton formerly held the post of Water Resources Engineer in the same Department, and this post was funded by the World Meteorological Organisation under project number MLW/77/012 (Advancement of Hydrological Services in Malaŵi). The U.K. component of the study was funded by the Overseas Development Administration under its research subvention to the Institute of Hydrology.

REFERENCES

Agnew, S. and Stubbs, M., 1972. Malaŵi in Maps. University of London Press.

Balek, J. and Perry, J. E., 1973. Hydrology of seasonally-inundated African headwater swamps. Jour. Hydrology, 19, 227-250.

Benson, M. A., 1959. Channel slope factor in flood frequency analysis Jour. Hyd. Div., A.S.C.E., 85, HY4.

Drayton, R. S., 1980. A study of the causes of the abnormally high levels of Lake Malaŵi in 1979. Water Resources Division Report No. T.P.5, Lilongwe 3, Malawi.

Institute of Hydrology, 1980. Low Flow Report. Institute of Hydrology, Wallingford, Oxon, UK.

Hill, J. L. and Kidd, C. H. R., 1980. Rainfall-runoff relationships for 47 Malaŵi catchments. Water Resources Technical Report in press. Lilongwe, Malaŵi.

Kidd, C. H. R., 1980. Evaluation of a storage-yield analysis method for Malaŵi. Institute of Hydrology Report, in press.

Melton, M. A., 1958. Correlation structure of morphometric properties of drainage systems and their controlling agents. J. Geol., 66, 442-460.

Midgley, D. C. and Pitman, W. V., 1969. Surface water resources of South Africa. Hydrological Research Unit Report No. 2/69, Dept. of Civil Engineering, University of the Witwatersand, South Africa.

Natural Environment Research Council, 1975. Flood Studies Report.

Pike, J. G., 1964. The estimation of annual runoff from meteorological data in a tropical climate. Jour. Hydrology, 2, 116-123.

Pike, J. G. and Rimmington, G. T., 1975. Malaŵi: a geographical study. Oxford University Press.

Strahler, A. N., 1950. Equilibrium theory of slopes approached by frequency distribution analysis. Am. Jour. Science, 248, 800-814.

Van der Velden, J., 1979. A summary of rainfall, pan evaporation and temperature at pan evaporation stations in Malaŵi. Water Resources Division, Lilongwe 3, Malaŵi.

World Meteorological Organisation, 1976. A water resources assessment of Lake Malaŵi. Project 71/15B Report.

APPENDIX A: LOW FLOW ESTIMATION PROCEDURES

Maps at a scale of 1:50,000 will be required for these procedures. The boundary of the catchment to the point of concern on the river should be sketched in pencil. More than one 1:50,000 map may be necessary for this purpose. A sketch of this boundary may also have to be transposed on to a 1:1,000,000 map. Estimation of flow measures are dealt with under separate headings, the first two of which (the average daily flow and the flow duration curve) are mandatory.

1. AVERAGE DAILY FLOW (ADF):

Five years of data could be adequate for estimation of ADF solely from data, although a check against the regional estimate will be valuable. Otherwise use the following steps for sites with no data:

1.1 Measure catchment area (AREA) in km^2 from map(s) by planimetering or counting squares.

1.2 Estimate average annual rainfall (AAR) in mm from a map by measuring areas between adjacent isohyets (see Chapter 3).

1.3 Estimate average annual yield (AAY) from:

$$AAY = .71AAR - 490 \quad (mm)$$

1.4 Calculate ADF from:

$$ADF = \frac{AAY \times AREA}{31500} \quad (cumecs)$$

1.5 Compare result with adjacent sites or limited data for site of interest as appropriate.

2. FLOW DURATION CURVE (F.D.C.)

2.1 Three years of data are adequate for estimation of F.D.C. as (%ADF). For manual derivation of F.D.C. from data for D = 1 day duration, follow the following procedure (where computer is available for other durations, repeat procedure after one pass of a D-day moving average filter):

2.1.1 Choose about 15-20 class intervals of flow from the highest flows down (it is best to use smaller intervals at low flow rates).

2.1.2 Assign all daily flows to a class interval, and fill out a table such as that in Table A.1.

Class interval (cumecs)	Tally of days in each interval	Total in each interval	Cumulative total	%> than btm of interval	Btm of interval (% ADF)

Table A.1 Calculation table for flow duration curve

2.1.3 Plot the discharge at the bottom of each class interval against the percentage of days exceeded (as calculated in Table A.1) on log-normal probability paper (log flows on y-axis, probability on x-axis). Where no normal probability paper is available, the x-axis can be linearised by:

$$x = \text{signum}\ (p-\tfrac{1}{2})\ \{1.238t\ (1 + .0262t)\}$$

where b = exceedance probability in proportional terms,
$\text{signum}\ (p-\tfrac{1}{2}) = +1$ for $p > \tfrac{1}{2}$; $= -1$ for $p < \tfrac{1}{2}$,
and $t = \{-\ln 4p\ (1-p)\}^{\frac{1}{2}}$

Up to 4 cycles may be needed on the y-axis for "flashy" rivers.

2.2 Where no data are available, the flow duration curve can be reproduced by the following procedure:

2.2.1 Estimate Q75(10) from map (as%ADF) - note that map does not give point estimates of Q75(10); see map for instructions.

2.2.2 Estimate Q75(D) from:

$$Q75(D) = Q75(10) + .0089\ \{Q75(10)\}^{.23}\ (D-10)^{1.33} \text{ for } D > 10$$
$$Q75(D) = Q75(10) - .0089\ \{Q75(10)\}^{.23}\ (10-D)^{1.33} \text{ for } D < 10$$

2.2.3 Estimate Q25(D) from Q75(D) and D by interpolation in Table 5.1

2.2.4 Convert Q25(D) and Q75(D) to cumecs by multiplying by ADF

2.2.5 Plot Q25(D) and Q75(D) on log-normal probability paper - see 2.1.3 for construction if special paper not available.

2.2.6 Draw straight line through these two points, extrapolating as far as Q98 if possible.

2.3 Appropriate compromise between 2.1 and 2.2 can be made when some data are available. Spot gaugings will be of particular value.

3. LOW FLOW FREQUENCY CURVE

3.1 15 years of data are required for a complete description of the low flow frequency curve from data alone, although comparison with a regional estimate will always be useful. To generate the low flow frequency curve for a D-day duration, follow these steps:

3.1.1 Find the lowest average flow over D-days in each year.

3.1.2 Rank these flows (from the highest value downwards) and assign a plotting position to each; the plotting position w is obtained from:

$$P = (i - .44)/(N + .12)$$

$$w = 4\left[1 - \{-\ln P\}^{\frac{1}{4}}\right]$$

where w is the plotting position

P is the exceedance probability = $\frac{T-1}{T}$

i is the rank of the flow

N is the number of years of data

T is the return period (years)

3.1.3 Plot the flow v. the plotting position w for each data point (it is useful at this stage to draw a return period T scale as well as w; T will be non-linear).

3.1.4 Draw (by eye) a smooth curve through the data points, extrapolating to w = 2.5 (T=50 years).

3.2 Where no data are available, the low flow frequency curve can be constructed in the following way:

3.2.1 Estimate the 10-day mean annual minimum - MAM(10) - as %ADF from:

$$MAM(10) = .101\ \{Q75(10)\}^{1.47}$$

where Q75(10) is estimated as in step 2.

3.2.2 Calculate MAM(D) from:

$$MAM(D) = MAM(10) + .0186\{MAM(10)\}^{.27}\ (D-10)^{1.28} \quad \text{for } D > 10$$

$$MAM(D) = MAM(10) - .0186\{MAM(10)\}^{.27}\ (10-D)^{1.28} \quad \text{for } D < 10$$

3.2.3 Using MAM(D), read fraction of MAM(D) off curves in Figure 5.6 for w = .5, 1.0, 1.5, 2.0, 2.5

3.2.4 Convert ordinates at w = .5, 1.0, 1.5, 2.0, 2.5 to cumecs

3.2.5 Construct curve as in 3.1.3 and 3.1.4 above

3.3 Where limited data are available (<15 years), construct the low flow frequency curve by both 3.1 and 3.2; compromise between 2 curves according to quantity of data available.

4. STORAGE-YIELD ANALYSIS

4.1 The analysis from daily data is complex and usually done by computer. From monthly data, the analysis is just about tractable by hand. The basis of the analysis method is outlined in Chapter 5.

4.2 The analysis uses the standardised type curves of Figure 5.10, and requires an estimate of ADF (Appendix B.1) and the F.D.C. (Appendix B.2). For an allowance for evaporation loss, approximate estimates are required for (i) the annual average open water evaporation in mm, and (ii) the relationship between the surface area of the proposed reservoir and the associated storage volumes under consideration.

4.2.1 Determine the gross yields Q40, Q50, Q60, Q70, Q80, Q90 from the F.D.C. (step B.2)

4.2.2 Calculate the gross yields Q40, Q50, Q60, Q70, Q80, Q90 in cumecs by multiplying by ADF/100.

4.2.3 Select an acceptable return period (T) of failure

4.2.4 For each gross yield Q40, Q50, Q60, Q70, Q80, Q90, read the required storages from Figure 5.10. Convert these storages to m^3 by multiplying by ADF x .316 x 10^6.

4.2.5 Determine net yield for each storage by allowance for exploration loss:

(a) For each storage determined in 4.2.4, determine the equivalent surface area (A) in km^2.

(b) For open-water evaporation E mm/year, loss = $\frac{AE}{31500}$ cumecs (assuming reservoir always full).

(c) Correct for error of assumption in (b) by multiplying loss by 2/3.

(d) Derive net yield by subtracting loss from gross yield.

4.2.6 Plot net yield v. storage for all 6 yields on plain graph-paper. Sketch a line through these 6 points.

4.2.7 Repeat procedure from 4.2.3 for other return periods if required.

5. RECESSION CHARACTERISTICS

5.1 For stations with complete records through the dry season (assumed here to be May to October inclusive), the following procedure can used to estimate a dry-season recession constant.

5.1.1 Calculate the monthly average flow for May, June, July, August, September, October for each complete year of record.

5.1.2 Plot log (monthly flow) against the month for each year separately.

5.1.3 Fit a straight line (by eye) through the data for each year.

5.1.4 Calculate recession constant K (in months) for each year.

5.1.5 Catchment average recession constant (KREC) is arithmetic mean of the yearly values.

5.1.6 Future flows can be forecast from:

$$Q_t = Q_0 \exp(-t/KREC)$$

where Q_t is monthly average flow to be forecast in the future (cumecs)

Q_0 is monthly average flow now (cumecs)

t is the forecast lead time (months)

5.2 Where no records are available, the following procedure may be used:

5.2.1 Calculate KREC from:

$$KREC = 2.17 + .045\,Q75(10)$$

5.2.4 Repeat 5.1.6 as appropriate

5.3 A satisfactory compromise between 5.1 and 5.2 can be found where a limited quantity of data is available.

APPENDIX B: FLOOD ESTIMATION PROCEDURES

The procedures described here are intended to provide an estimate of the flood which has a return period of T years (known as the T-year flood). There are two separate parts to the procedure, relating (i) to the mean annual flood, and (ii) to the flood frequency curve, which are combined to give the required flood.

1. ESTIMATION OF THE MEAN ANNUAL FLOOD (MAF):

1.1 Five years of data are considered adequate for estimation of MAF from data alone, although a check with the regional estimate will be useful. However, even one complete year's data is useful, and the method is as follows:

1.1.1 Select the highest flood (cumecs) in each water-year (November to October)

1.1.2 MAF = the arithmetic mean of the annual floods

1.2 The MAF can be estimated from catchment characteristics as follows:

1.2.1 Measure the catchment area (AREA) according to the instructions given in Chapter 3

1.2.2 Measure the stream frequency (STMFRQ) according to the instructions given in Chapter 3

1.2.3 Calculate the MAF from:

$$MAF = 2.89\ AREA^{.55}\ STMFRQ^{.36}$$

1.3 Where a few years of data are available, estimate MAF by both methods and choose some intermediate value. Guidance in selecting this value is given in Chapter 6.

2. ESTIMATION OF THE FLOOD FREQUENCY CURVE:

It is recommended that the regional flood frequency curve given in Figure 6.2 is more reliable than a curve derived from even 20 years of data at a particular station. The procedure for obtaining a T-year flood is as follows:-

2.1 Choose the design return period T.

2.2 Estimate the ratio Q(T)/MAF from Table 6.1. Alternatively, this ratio can be obtained for any value of T from:

$$y = -\ln(-\ln(1-{}^{1}/T))$$

$$\frac{Q(T)}{MAF} = -1.92 + 2.58 \exp(0.17y)$$

If T > 5 years, $\frac{Q(T)}{MAF} = -1.92 + 2.58\,(T - \tfrac{1}{2})^{.17}$ is accurate enough.

2.3 Multiply $\frac{Q(T)}{MAF}$ determined in 2.2 by MAF determined in 1 to give Q(T) in cumecs, the flood of the required T-year return period.

APPENDIX C: WORKED EXAMPLE OF ESTIMATION PROCEDURES

We will undertake an example analysis for an imaginary site of interest - this is where the Lilongwe-Salima road crosses the Lumbadzi river. Figure C1 is a plan of the catchment at a scale of 1:250 000, and the area (AREA) is nominally 430 km^2. We require to derive the average daily flow, the flow duration curve, the low flow frequency curve and the recession characteristics of the Lumbadzi river at this point. In addition we will estimate the 100 year return period flood. The procedural steps relate to the estimation procedure given in appendices A and B.

AVERAGE DAILY FLOW

What is the average daily flow (ADF) at the site in question? For convenience we have reproduced a map of average annual rainfall for the catchment in Figure C.2.

1.1 Catchment area (AREA) is given as 430 km^2.

1.2 Estimate AAR from Figure C.2. AAR = 953 mm

1.3 Estimate average annual yield AAY = 187 mm

1.4 Calculate ADF ADF = 2.55 cumecs

FLOW DURATION CURVE

Derive the 10-day and 1-day flow duration curves for the point of interest. For convenience, the map of Q75(10) - the 10-day 75 percentile flow as % ADF - has been reproduced for the catchment in Figure C.3.

2.2.1 Estimate Q75(10) Q75(10) = 2.15 %ADF

2.2.3 Estimate Q25(10) Q25(10) = 91 %ADF

2.2.4 Convert to cumecs Q75(10) = .055 cumecs

Q25(10) = 2.32 cumecs

2.2.5 Plot Q75(10) and Q25(10) on Figure C.4

2.2.6 Draw straight line through 2 points.

2.2.2 Estimate Q75(1) Q75(1) = 2.05 %ADF

Repeat 2.2.3 - 2.2.6 for D = 1 day

Q75(1) = .052 cumecs

Q25(1) = 84 %ADF

Q25(1) = 2.14 cumecs

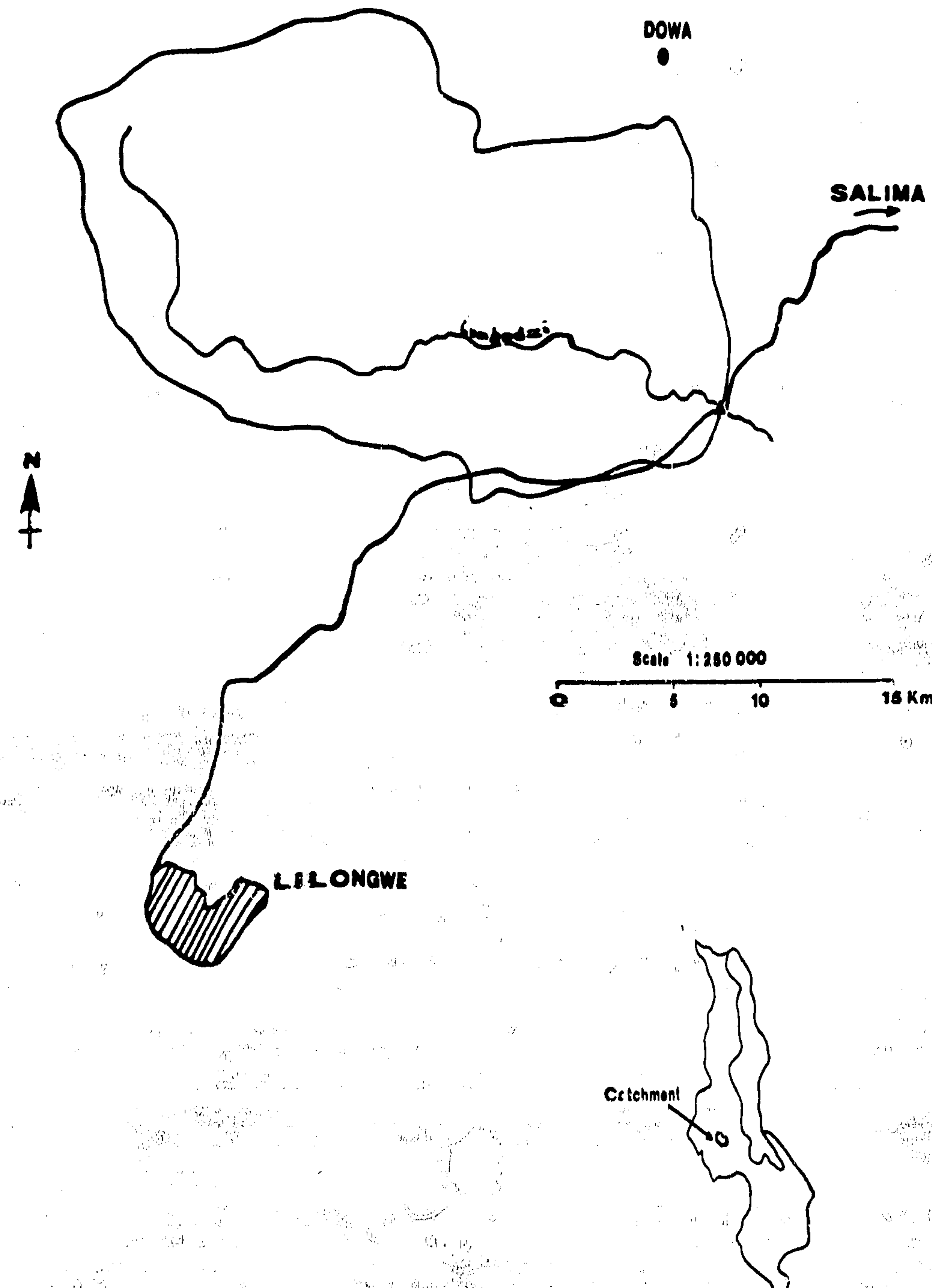


FIGURE C1: The Lumbadzi river on the Lilongwe-Salima road

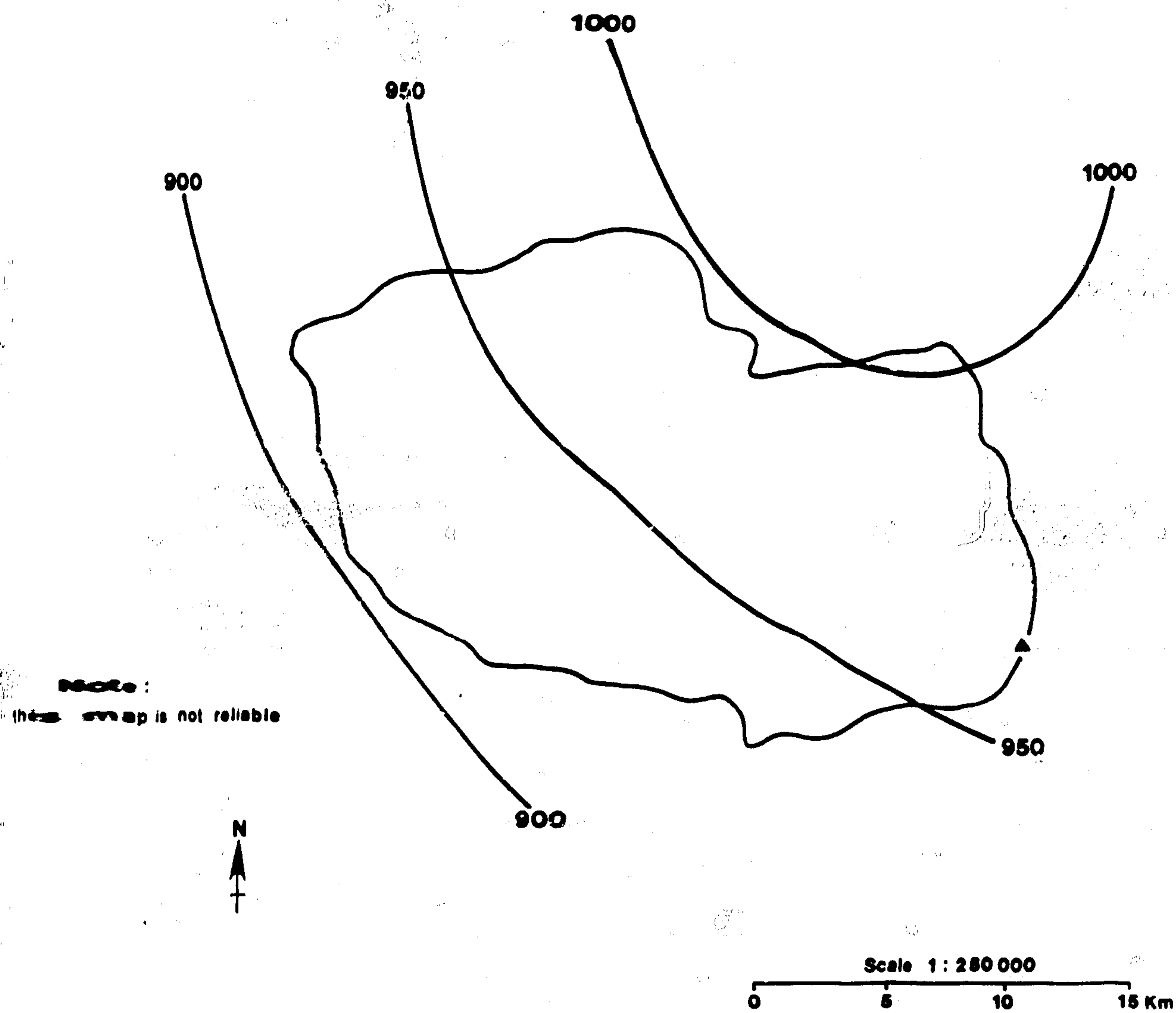


FIGURE C2: Isohyetal average annual rainfall map

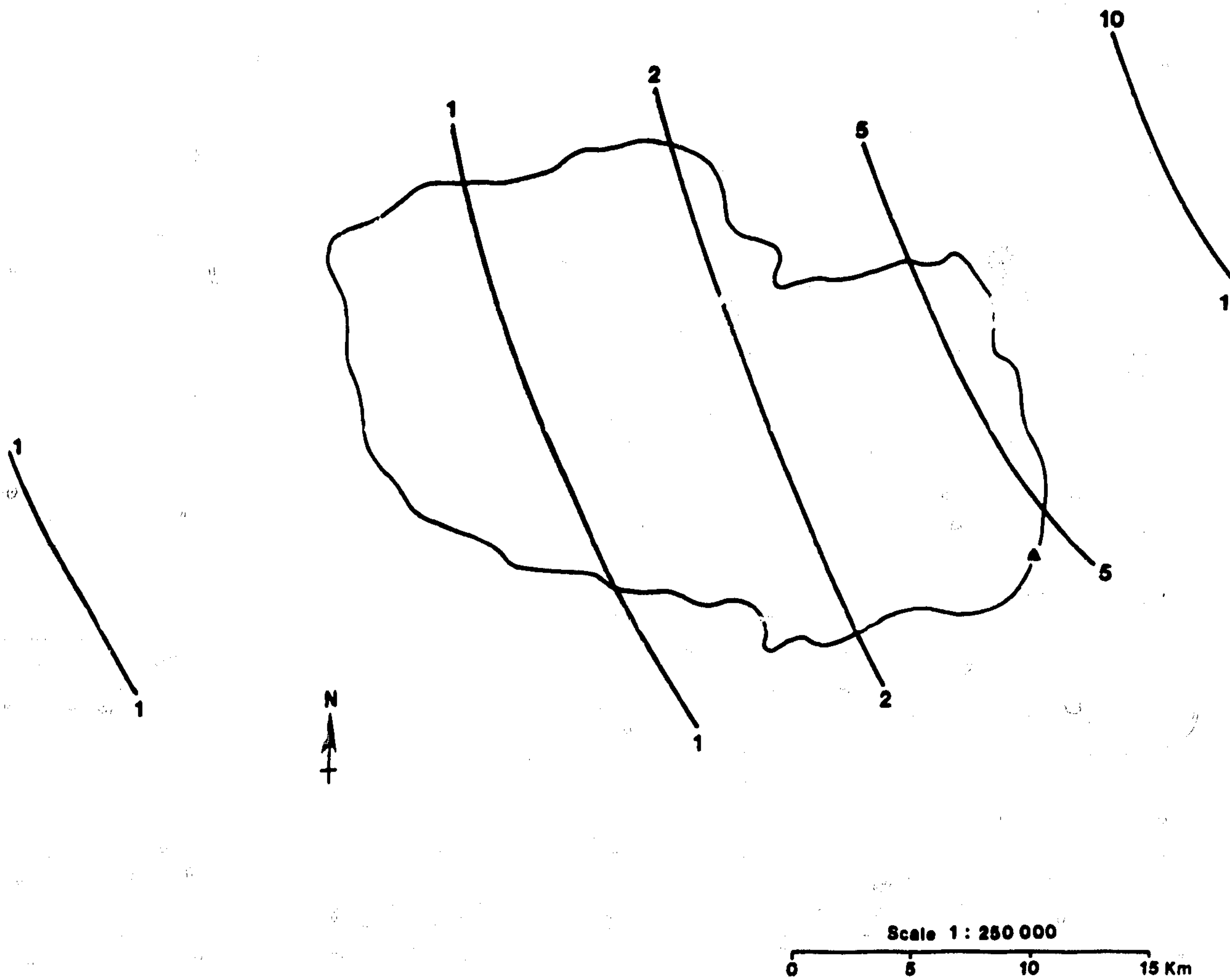


FIGURE C3 : Contour map of Q75 (10)

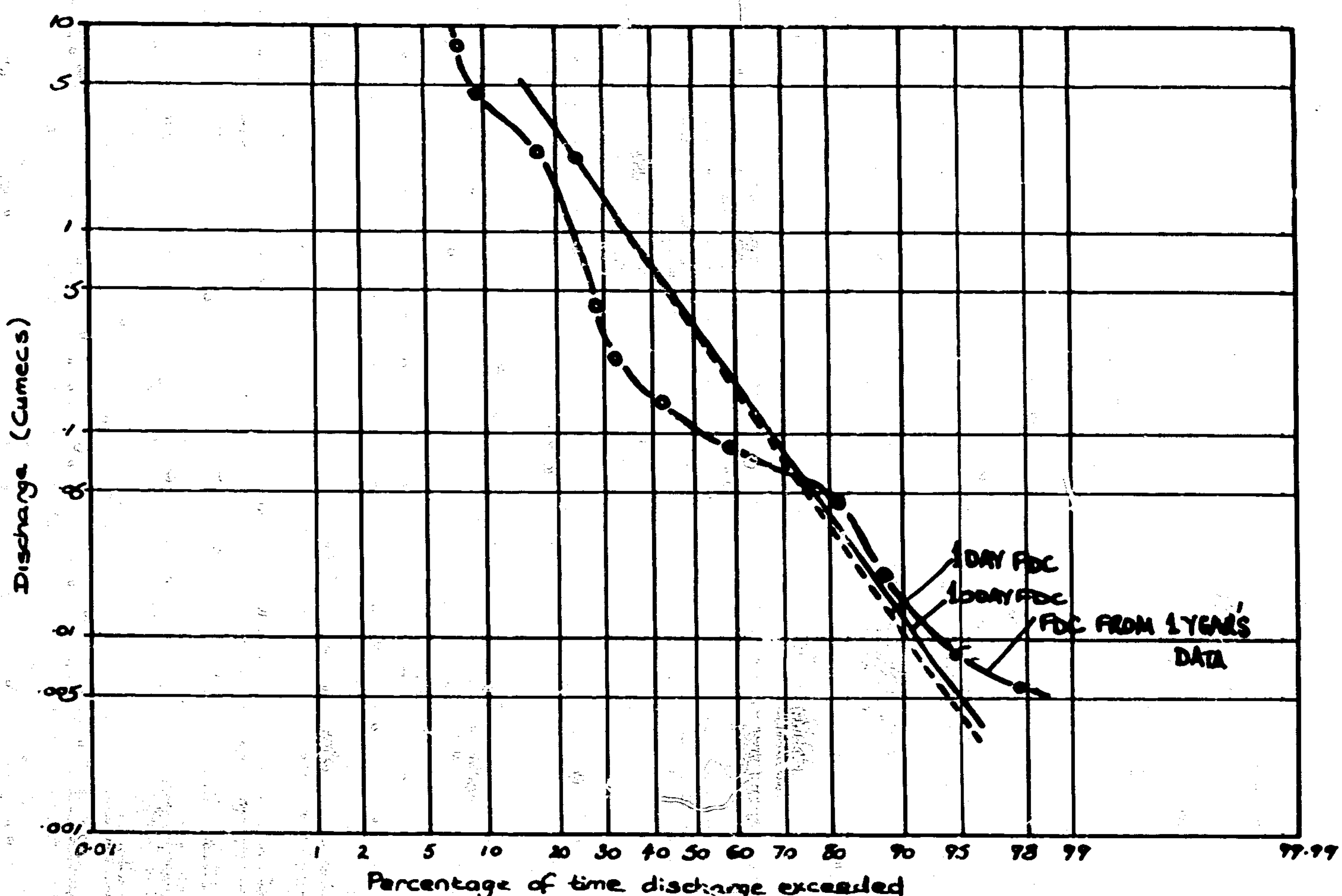


Figure C4: Flow duration curve

We have one year's flow data at this site, which is given in Table C.1. Using Table C.2 and the procedure given in steps 2.1.1 and 2.1.3, derive the flow duration curve for this year's data (in Table C.2, the last column is produced using the year's ADF - 5.99 cumecs). Using the long term ADF derived above, reproduce this flow duration curve on Figure C.4 for comparison

* *

What is the 10-day average flow which we could expect to be exceeded 95% of the time? Q95(10) = .005 cumecs

LOW FLOW FREQUENCY CURVE

Derive the 90-day duration low flow frequency curve.

3.2.1 Estimate MAM(10) - the 10-day mean annual minimum MAM(10) = 0.31 % ADF

3.2.2 Estimate MAM(90) MAM(90) = 4.00 % ADF

3.2.3 Derive ordinates of low flow frequency curve:-

W	= 0.5	1.0	1.5	2.0	2.5	
Ordinate	= .49	.24	.02	0	0	(fractions of MAM(D))
Ordinate	= 1.96	0.96	.08	0	0	(% ADF)

3.2.4 Ordinate = .050 .024 .002 0 0 (cumecs)

3.2.5 Construct low flow frequency curve on Figure C.5 as in steps 3.1.3 and 3.1.4.

What is the approximate return period of a 90-day continuous period of zero flow? T = 8 years

RECESSION CHARACTERISTICS

Derive the characteristic recession constant (KREC).

5.2.1 Estimate KREC KREC = 2.3 months

Given that the average flow in June was 0.15 cumecs at the site of interest, what is the average flow in October likely to be? Use the equation given in step 5.

Forecast lead time t = 4 months

Forecast October flow = .026 cumecs

FLOOD ESTIMATION

The stream frequency (STMFRQ) has been measured at 1.15 junctions per square kilometre.

DAY	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
1	4.04	19.00	5.30	2.22	.365	.238	.237	.170	.158	.041	.020	.069
2	3.35	8.15	7.24	1.91	.365	.230	.230	.164	.149	.038	.021	6.12
3	8.86	5.01	8.52	1.47	.364	.230	.230	.164	.144	.036	.020	64.93
4	31.06	3.84	8.79	0.93	.349	.230	.230	.164	.136	.036	.020	34.00
5	121.90	4.08	8.58	0.76	.336	.230	.233	.164	.132	.036	.019	25.98
6	74.85	8.74	8.40	0.66	.336	.230	.230	.164	.128	.036	.016	3.92
7	21.57	7.52	8.19	0.64	.336	.230	.223	.164	.124	.036	.015	0.83
8	[illegible]	21.41	7.68	0.62	.328	.230	.229	.164	.124	.036	.013	1.58
9	[illegible]	8.64	7.42	0.62	.323	.230	.220	.164	.124	.036	.013	1.64
10	[illegible]	6.77	9.40	0.64	.314	.237	.223	.164	.124	.036	.012	1.59
11	3.36	6.47	9.15	0.63	.307	.230	.223	.166	.124	.036	.011	1.53
12	[illegible]	[illegible]	7.82	0.61	.308	.230	.229	.164	.124	.035	.011	1.49
13	3.06	[illegible]	5.76	0.61	.320	.230	.230	.159	.124	.036	.012	1.44
14	3.08	[illegible]	3.69	[illegible]	.314	.230	.230	.164	.124	.035	.012	1.40
15	3.17	[illegible]	3.52	0.48	.306	.230	.230	.164	.124	.035	.011	1.35
16	3.86	54.78	3.43	0.48	.286	.229	.230	.164	.120	.034	.041	[illegible]
17	14.72	46.18	[illegible]	0.45	.270	.230	.230	.164	.123	.033	.406	[illegible]
18	208.33	[illegible]	3.25	0.46	.253	.230	.222	.166	.124	.033	.052	.059
19	277.20	42.39	3.14	0.44	.248	.225	.209	.164	.124	.034	.033	.054
20	114.30	41.48	3.06	0.42	.232	[illegible]	.196	.163	.120	.034	.041	.053
21	18.70	40.08	2.97	0.42	.229	.222	.183	.164	.120	.031	.068	.137
22	8.07	39.00	2.90	0.42	.230	.229	.170	.164	.119	.028	.131	.128
23	4.86	[illegible]	2.97	0.40	.238	.219	.159	.166	.112	.025	.155	.087
24	9.76	38.06	3.01	0.39	.238	.215	.159	.164	.110	.022	.150	.062
25	11.81	37.82	[illegible]	0.39	.230	.222	.164	.166	[illegible]	.022	.132	.057
26	35.22	33.87	2.65	0.39	.237	.230	.164	.164	[illegible]	.022	.125	.052
27	34.20	5.44	2.61	0.39	.238	.230	.166	.164	[illegible]	.020	.053	.050
28	33.35	4.61	2.55	0.38	.230	.223	.164	.166	.093	.020	.040	.050
29	32.54	-	2.50	0.38	.229	.229	.164	.164	.084	.020	.036	.049
30	31.74	-	[illegible]	0.38	.230	.223	.164	.164	.077	.019	[illegible]	.048
31	30.60	-	[illegible]	-	.230	-	.164	.134	-	.019	-	[illegible]

Yearly average daily flow = 5.99 cu[illegible]

Note: These data are entirely fictional

Table C.1: <u>One year's data for Lumbadzi at Lilongwe-Salima Road</u>

Class interval (cumecs)	Tally of days in each interval	Total in each interval	Cumulative total	%> than btm of interval	Btm of interval (% ADF)
>50.00		7	7	1.9	885
20.01 - 50.00		22	29	7.9	334
10.01 - 20.00		7	36	9.9	167
5.01 - 10.00		27	63	17.2	83.6
1.01 - 5.00		44	107	29.3	16.9
.51 - 1.00		12	119	32.6	8.51
.31 - .50		32	151	41.4	5.18
.21 - .30		65	216	59.2	3.51
.11 - .20		78	294	80.5	1.84
.051 - .109		16	310	84.9	.851
.021 - .050		36	346	94.8	.351
.015 - .020		11	357	97.8	.250
<.015		8	365	100.0	-

Table C.2: Calculation for flow duration curve

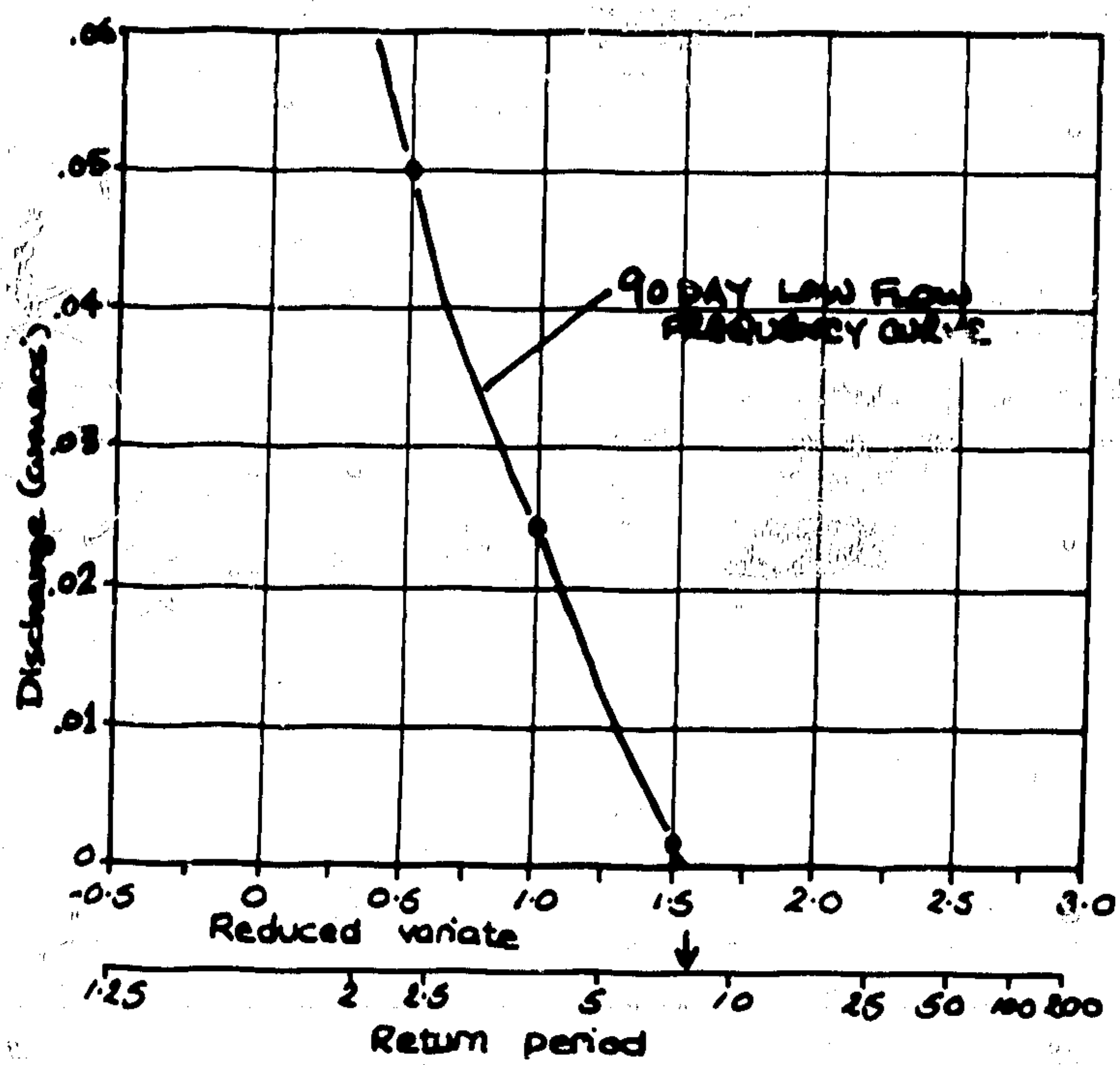


Figure C5: Low flow frequency curve

1.2	Estimate the mean annual flood	MAF = 85 cumecs
2.1	Estimate Q(T)/MAF for T=100 years	y = 4.60
		$\frac{Q(T)}{MAF}$ = 3.72
2.2	Estimate the 100-year flood	Q(T) = 316 cumecs