73I7-1001

INSTITUTE OF HYDROLOGY

Report No. 25 August 1976

MAPWORK FOR FLOOD STUDIES

PART II : ANALYSIS OF INDICES AND RE-MAPPING

by

Malcolm Newson

ABSTRACT

Part I of this Report covered the background to the selection of morphometric indices for use as independent variables in the U.K. Flood Study (N.E.R.C., 1975) and the derivation of values for those indices on 755 catchments in the British Isles. Part II now describes the interrelationships between variables, including those extra ones derived for unit hydrograph shape prediction, their re-mapping on national maps and their usefulness for flood prediction at ungauged sites. The results support the systematic organization of the drainage basin and its flood response (even though the relationships are statistical) which first guided the choice of variables.

CONTENTS

FIGURES

TABLES

I INTRODUCTION

Scheme and variables

The production of the largest-ever archive of catchment characteristics in the British Isles allows a general assessment of morphometric interrelationships, hitherto only achieved on such a scale in the United States. Although not so many variables were available as in some smaller regional studies, the basic features of size, slope and channel network were indexed (catchment area, two indices of main channel slope and stream frequency) and, in addition, there were indices of climate, soils, and land-use (see Table 1). Interrelationships are of interest in marshalling variables prior to regression; they also point to causative links. Certain of the latter can be developed as linear regression models. The analysis can be used to indicate the effectiveness of the variables before their utility is tested in regression of flood parameters. Residuals from flood prediction may be tested against morphometric irregularities. As a result of all aspects of the investigation further research topics are suggested.

Table 1 gives a definition of catchment characteristics as used in this Report.

Data summaries

Lack of space prevents duplication of the basic catchment data here; Volume 4 of the Flood Study (N.E.R.C., 1975), pages 120-135, gives catchment data for 755 catchments in the British Isles. These are a subset of 1294 catchments examined. Further subsets of the 755 referred to here are the 533 finally used for predicting mean annual flood and 410 (Great Britain only) on which certain regression and principal components analyses were performed. Subsetting has been necessary for reasons of data quality, completeness and format. Finally 130-150 catchments have been used for unit hydrograph analysis.

Instead of listing basic data, parameters of the distribution of values on each variable used in the analysis of 533 catchments are given here (Table 2)

For all subsequent analyses the data have been transformed to their logarithm of the base 10. This has a distinct normalizing effect as can be gathered from the skewness and kurtosis values when reading Table 3. However, the resulting distributions are not completely normal; with 533 values the sample skewness for normal data should be in the range $\pm$ 0.17 and kurtosis in the range 2.68 to 3.36 (both 95% values).

The normalizing effect is especially noticeable in the physiographic variables - AREA, STMFRQ, S1085 and TAYSLO. The effect is less in the climatic group of variables, less still for LAKE and URBAN (with SOIL reacting poorly) but rises again for those annual flood variables considered here. The histograms shown in Figure 1 illustrate the effect and bear out Clark's (1973) and Gardiner's (1974) plea for transformation of variables before multivariate analysis.

TABLE 1 Definitions of catchment characteristics

AREA	Catchment area in square kilometres
STMFRQ	The number of stream junctions, as shown on the 1:25 000 map, divided by the catchment area
S1085	The stream channel slope measured between two points, 10% and 85% of the stream length from gauge, expressed in m/km
TAYSLO	The Taylor-Schwarz slope of the channel. The stream is divided into n elementary reaches each with slope S_i. The Taylor-Schwarz slope S is calculated from $$\frac{1}{\sqrt{S}} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{\sqrt{S}}$$ and is expressed in m/km. (TAYSLO is the slope of the uniformly sloping channel which, assuming a Chézy law, has the same length and travel time as the original.)
SAAR	Standard period annual average rainfall in mm. The standard period used is 1916-1950 (1931-1960 in Ireland)
M52D	The catchment average two-day rainfall of five-year return period
SMDBAR	A weighted mean soil moisture deficit
RSMD	A net one-day rainfall of five-year return period. The rainfall is net in the sense that SMDBAR is subtracted from it
URBAN	The urban fraction of the catchment
LAKE	The fraction of the catchment draining through a lake or reservoir
SOIL	A soil index with values in the range 0.15 - 0.5. Soils are classified into five classes. If S_i is the fraction of the catchment covered by soil class i the index is given by $$\frac{0.15S_1+0.3S_2+0.4S_3+0.45S_4+0.5S_5}{S_1+S_2+S_3+S_4+S_5}$$
DD	Drainage Density in km of channel (1:25000 maps) per km^2 of catchment area
SHAPE	The K parameter of the lemniscate loop
SFD1	First moment of stream frequency diagram
MSL	Main stream length in km.
DVF	Dry valley between stream head and valley divide as proportion of MSL.
AMAF	Arithmetic mean of the annual maximum floods
CV	Coefficient of variation of the annual maximum floods
BESMAF	An estimate of the mean annual flood, derived wherever possible by extending the record by correlation with nearby records. If extension not possible, BESMAF = AMAF
CALMAF	Mean annual calendar day flood
T_p	Unit hydrograph time-to-peak

TABLE 2 Parameters of catchment and annual flood characteristics

		MEAN	STD DEVN	MIN	MAX	SKEWNESS	KURTOSIS
AREA	km^2	515.9	922.2	0.0482	9868.0	5.253	40.402
STMFRQ	jun/km^2	1.19	0.92	0.01	7.54	2.043	10.045
S1085		7.21	11.97	0.19	117.78	4.750	32.563
TAYSLO	m/km	6.14	10.17	0.26	93.99	4.915	33.060
SAAR	mm/yr	1113.9	443.5	551.0	3454.0	1.426	5.727
RSMD	mm	39.4	14.8	15.6	117.5	1.223	5.767
M52D	mm	65.1	18.8	42.5	175.0	1.850	8.301
SMDBAR	mm	6.8	4.0	1.0	17.0	0.629	2.211
SOIL		0.39	0.88	0.15	0.50	0.807	3.116
LAKE	see Table 1	0.98	0.22	0.00	1.00	2.865	10.630
URBAN		0.33	0.10	0.00	0.81	4.871	30.358
BESMAF	cumecs	105.8	148.3	0.11	997.5	2.603	11.417
CV	cumecs	35.7	19.3	2.21	169.9	2.386	12.911

TABLE 3 Parameters of logarithmically-transformed variables

	MEAN	STD DEVN	MIN	MAX	SKEWNESS	KURTOSIS
LOG10(AREA)	2.33	0.64	-1.32	3.99	-1.01	7.05
LOG10(STMFRQ)	-0.58	0.38	-2.00	0.88	-1.28	6.86
LOG10(S1085)	0.58	0.47	-0.71	2.07	0.31	3.26
LOG10(TAYSLO)	0.54	0.43	-0.59	1.97	0.53	3.48
LOG10(SAAR)	3.02	0.16	2.74	3.54	0.47	2.67
LOG10(RSMD)	1.57	0.16	1.19	2.07	0.101	2.69
LOG10(M52D)	1.80	0.11	1.63	2.24	0.898	3.747
LOG10(SMDBAR)	0.75	0.28	0.00	1.23	-0.350	2.615
LOG10(SOIL)	-0.43	0.12	-0.82	-0.30	-1.514	5.271
LOG10(LAKE)	0.34	0.70	0.00	0.30	2.544	8.766
LOG10 URBAN	0.01	0.03	0.00	0.28	4.305	23.963
LOG10 BESMAF	1.63	0.66	-0.96	2.99	-0.586	3.878
LOG10(CV)	1.50	0.21	0.34	2.23	-0.308	5.184

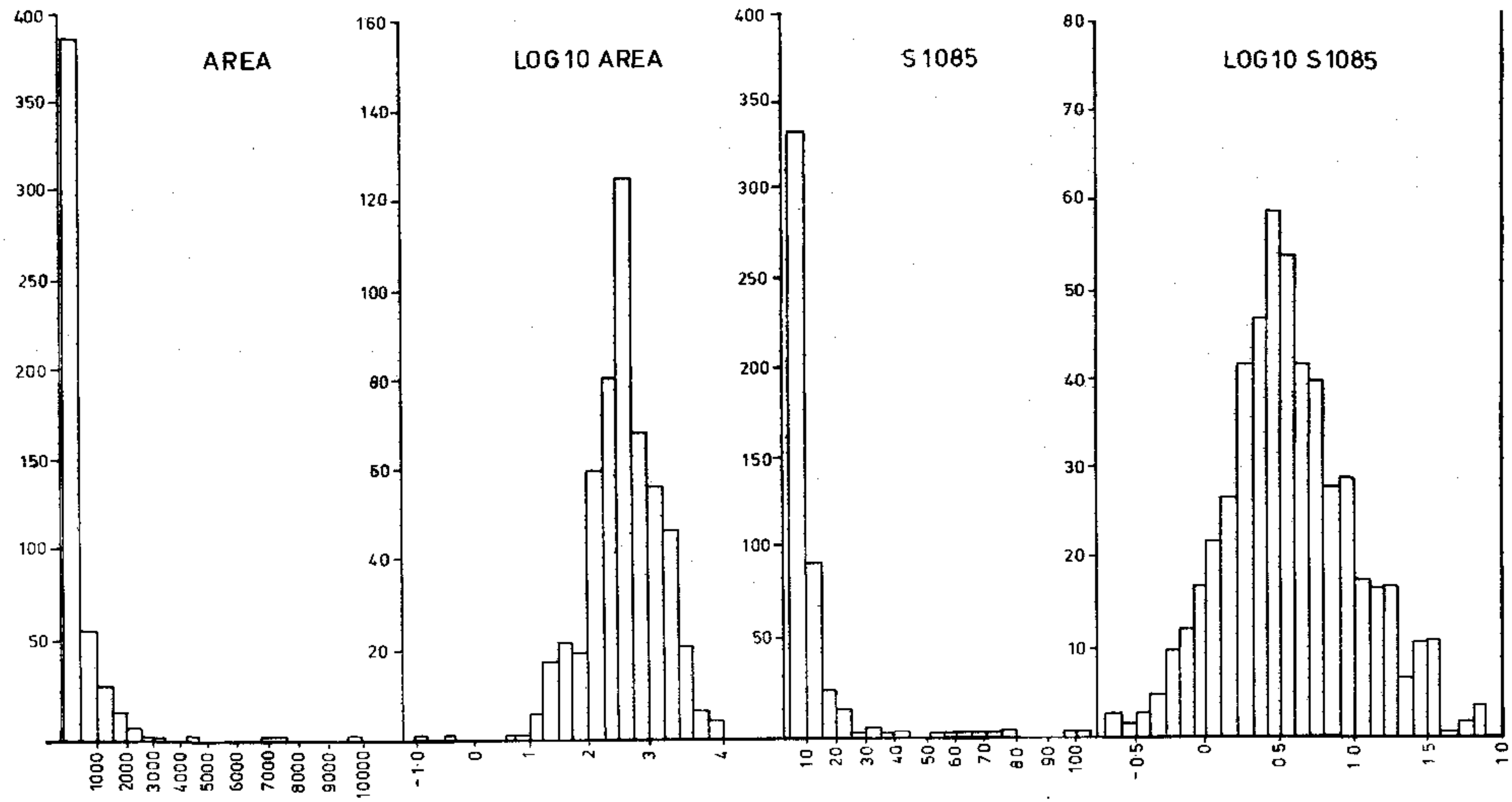

FIGURE 1 Histograms of untransformed and transformed data for two catchment variables, AREA and S1085

II THE INTERRELATIONSHIP OF VARIABLES DERIVED FOR PREDICTING MEAN ANNUAL FLOOD

Table 4 is reproduced from Table 4.8 of the Flood Studies Report. The use of 533 observations for each variable clearly gives the correlation coefficients in Table IV great strength; $r \gtrsim .090$ is significant at 0.05 and $r \gtrsim 0.117$ at 0.01.

Figure 2 shows the location of the sample catchments

Correlations between variables

As mentioned in Part I of this Report, the scaling effects of catchment area produce a wide range of correlations with other variables which inevitably complicate its use for predictor. The effect appears to be that area inevitably becomes the best prediction of discharge parameters through its integrating effect*. Of the variables shown in the table the most obvious strong relationships with area are those negative correlations with the two measures of channel slope (which are themselves closely related). Obviously the nesting of catchments leads to smaller catchment areas representing steeper watersheds. Steeper watersheds are also more densely drained, as can be judged from the relationship between the two slope measures and stream frequency. These inter-relationships are partly causal as evinced by the good correlations of rainfall, catchment wetness (RSMD) and slope variables but rather than attempting to unravel the causal links, as did Melton (1958), the present study has attempted to lump the related variables by principal components analysis (see Section IV).

Stream frequency is related to the interplay of relief and climatic influences; the good correlation with SOIL is perhaps an indication of Melton's view of feedback on the morphometric system.

The correlation of + 0.592 between SOIL and stream frequency is the strongest exhibited by the former variable (this means Winter Rain Acceptance Rate is lowest where stream frequency is highest - a logical outcome).

---

* To quote H. W. Anderson (1957) in full:

"area can well be called the devil's own variable. Almost every watershed characteristic is correlated with area so every characteristic that is left out as a separate variable is in part hidden in "area". Big watersheds are not like little watersheds and the differences may be disguised in the term area. Therefore it is dangerous to ascribe physical significances to the regression coefficients of the area variable".

TABLE 4 Correlation matrix of log transformed catchment characteristics for 533 catchments in Great Britain and Ireland

	BESMAF	CV	AREA	STMFRQ	TAYSLO	S1085	SOIL	URBAN	LAKE	SAAR	RSMD	M52D	SMDBAR
BESMAF	1.000	0.172	0.750	0.424	0.280	0.180	0.415	0.085	0.085	0.397	0.394	0.376	0.397
CV	0.172	1.000	0.069	0.240	0.001	0.065	0.130	0.068	0.095	0.341	0.300	0.243	0.377
AREA	0.750	0.069	1.000	0.076	0.717	0.675	0.022	0.088	0.106	0.122	0.122	0.135	0.007
STMFRQ	0.424	0.240	0.076	1.000	0.345	0.410	0.592	0.179	0.139	0.603	0.567	0.545	0.500
TAYSLO	0.280	0.001	0.717	0.345	1.000	0.888	0.242	0.042	0.171	0.390	0.429	0.444	0.192
S1085	0.180	0.065	0.675	0.410	0.888	1.000	0.309	0.014	0.104	0.514	0.548	0.560	0.289
SOIL	0.415	0.130	0.022	0.592	0.242	0.309	1.000	0.094	0.150	0.478	0.450	0.456	0.368
URBAN	0.085	0.068	0.088	0.179	0.042	0.014	0.094	1.000	0.087	0.213	0.210	0.197	0.204
LAKE	0.085	0.095	0.106	0.139	0.171	0.104	0.150	0.087	1.000	0.252	0.218	0.177	0.273
SAAR	0.397	0.341	0.122	0.603	0.390	0.514	0.478	0.213	0.252	1.000	0.928	0.915	0.809
RSMD	0.394	0.300	0.122	0.567	0.429	0.548	0.450	0.210	0.218	0.928	1.000	0.952	0.785
M52D	0.376	0.243	0.135	0.545	0.444	0.560	0.456	0.197	0.177	0.915	0.952	1.000	0.685
SMDBAR	0.397	0.377	0.007	0.500	0.192	0.289	0.368	0.204	0.273	0.809	0.785	0.685	1.000

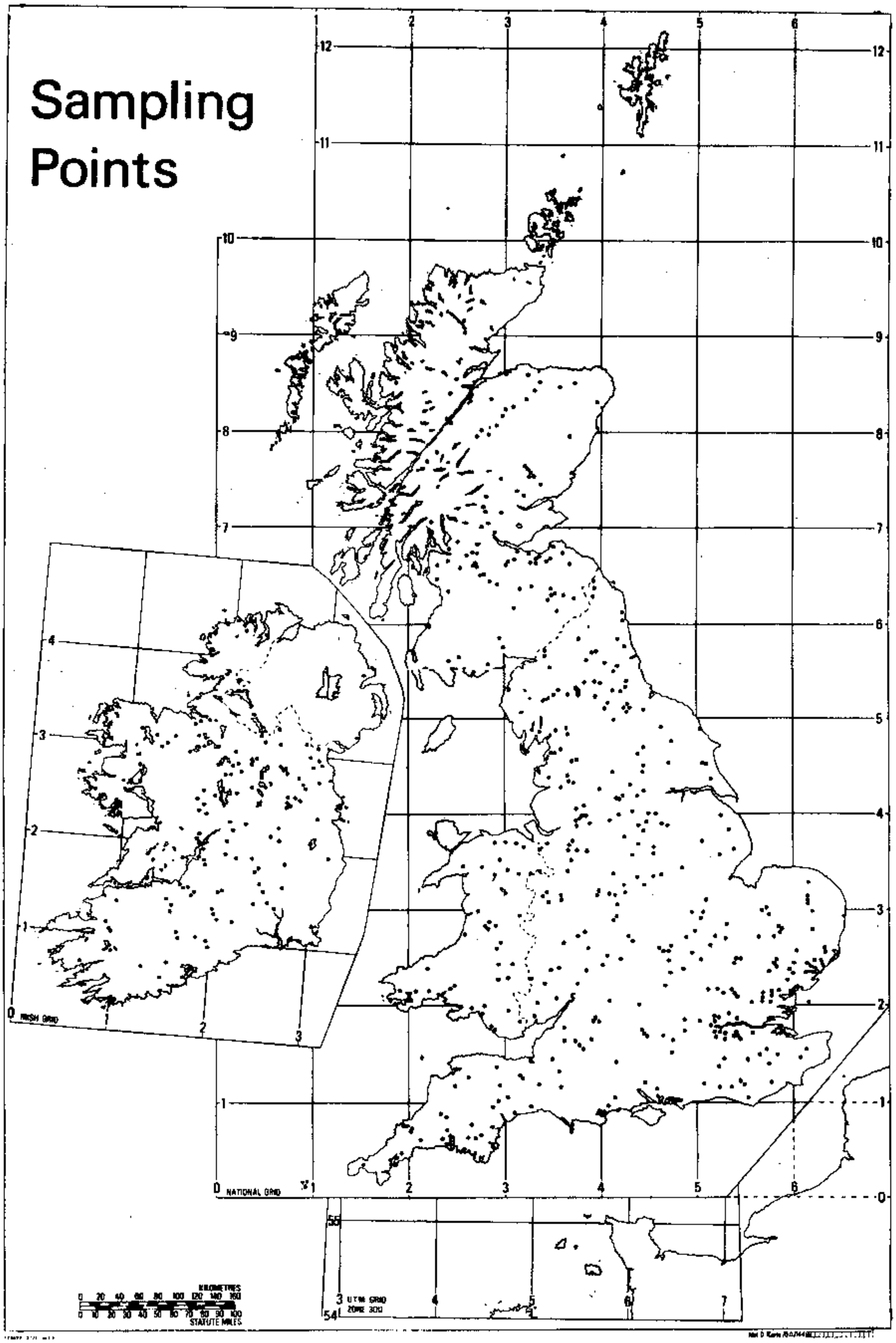

FIGURE 2 Location of gauging stations of sample catchments

The association of our major urban concentrations with lowland sites shows up in low but negative correlations of the URBAN variable with "wet" variables such as SAAR, RSMD and M52D. The firm positive inter-relationship of all the climatic variables is shown in the remaining columns of the matrix.

Not included in the matrix for the 533 stations but used in the analysis of 410 catchments in Great Britain alone (see Section IV), MSL (main-stream-length) has a very strong relationship with AREA (see Table 5)

TABLE 5 Correlations of MSL & DVF on 410 British catchments (these variables were omitted from the sample of 533)

	DVF	MSL
AREA	- 0.622	0.942
STMFRQ	- 0.287	- 0.056
S1085	0.299	- 0.620
TAYSLO	0.365	- 0.696
DVF	1.000	- 0.708
MSL	- 0.708	1.000
SAAR	- 0.109	- 0.079
M52D	- 0.122	- 0.050
RSMD	- 0.128	- 0.063
SMDBAR	0.164	- 0.019
SOIL	- 0.331	0.139
URBAN	0.111	- 0.105
LAKE	- 0.191	0.162

Ignoring Anderson's warning for a moment, the regression relationship obtained for MSL on AREA corresponds closely to those given in the literature by Hack (1957) and Gray (1961), and the summary given by Eagleson (1970).

$$MSL = 1.42\ AREA^{0.566} (R = 0.95) \qquad \text{..... (1)}$$

$$(\text{Hack: } MSL = 1.4\ AREA^{0.6}$$

$$\text{Gray: } MSL = 1.40\ AREA^{0.568})$$

Smart and Surkan (1967) point out that the exponent deviates from 0.5 due to increasing stream sinuosity with increasing catchment area and the separation of this phenomenon from one of changing catchment shape (i.e. longer catchments) as area increases is attempted here.

Figure 3 shows a plot of the relationship and deviations from it. Points below the regression line have a shorter mainstream than predicted by area, those above a longer one. Most of the large deviations are below the line and comprise a group of catchments in East Anglia and south-east England with Chalk in their headwater areas. Those larger rivers above the line have well-known meanders. To see if sinuosity is a real effect the possibility of a shape relationship with area is examined in Figure 4 when the k parameter of the lemniscate loop (see Section III) is related to AREA. The lack of relationship is clear.

Since Sokolov (1962) suggests the use of residuals from standard relationships such as MSL = f (AREA) as separate variables for flood analyses, a map has been prepared in Figure 5 to show the sign of residuals from equation (1) above for Great Britain. Comparing with residuals from BESMAF equation ((8) in this Report), it is clear that there is a resemblance in the midlands of Scotland, north-east England, and East Anglia, where the catchments with high sinuosities (positive deviation from equation (1)) have low annual floods (negative deviations from equation (8)). However, the same is not true on the western side of Britain and does not work in reverse in areas of shorter mainstreams. In areas such as Devon, East Cornwall and Lancashire perhaps rainfall influences are at work. (Fig. I 4.23, Volume V of the Flood Studies Report shows residuals from a nation-wide BESMAF equation. Later some of the regional differences shown were covered by regional analyses). Other high correlations followed up by regression analysis are as follows:

$$S1085 = 40.71\ MSL^{-0.684} (r = -0.63) \qquad \text{..... (2)}$$

showing the natural reduction of gradient with increasing distance from the watershed. Residuals from this equation, when mapped, show all of lowland Britain negative and all the highlands positive. Thus deviations here are of little value and result from the regional association of slope and the freedom from such of mainstream length.

$$S1085 = 0.0001\ SAAR^{1.540} (r = 0.59) \qquad \text{..... (3)}$$

$$STMFRQ = 0.00002\ SAAR^{1.548} (r = 0.61) \qquad \text{..... (4)}$$

Equations (3) and (4) were used to guide mapping of the individual variables S1085 and STMFRQ in areas where data were sparse (see

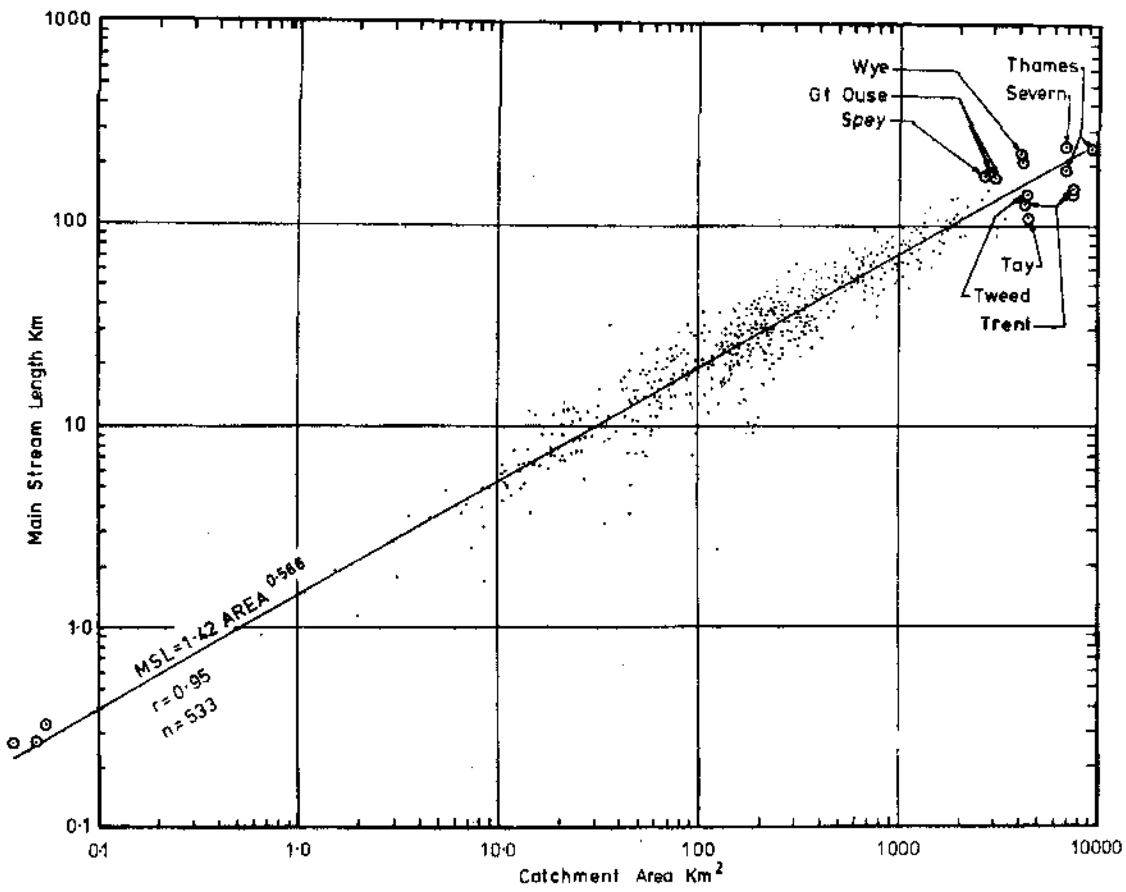

FIGURE 3 AREA's relationship with MSL

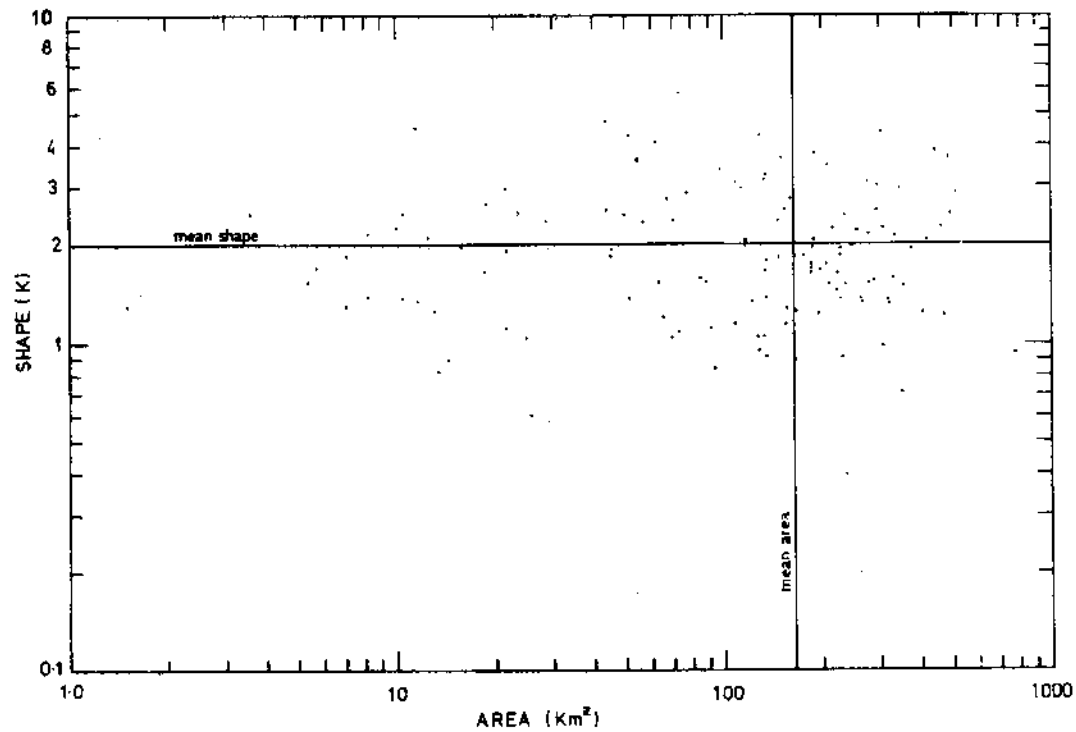

FIGURE 4 The relationship between SHAPE and AREA

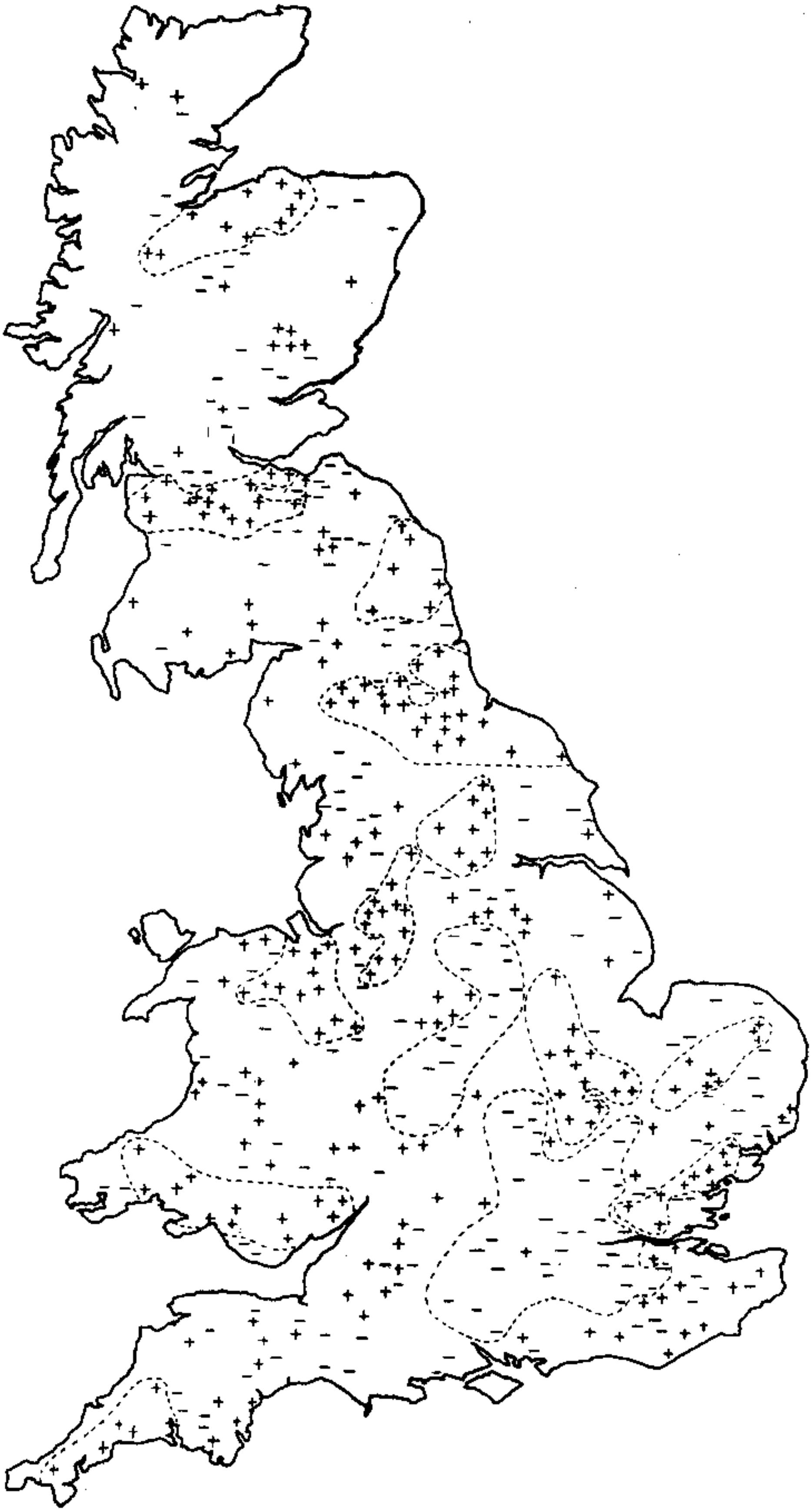

FIGURE 5 Residuals from Equation (1) plotted by sign, positive or negative. Significant regional groupings are ringed

Section IV). Residuals from equation (4) were mapped (Figure 6) to separate the climatic effect from stream frequency, leaving in Carlston's (1963) terms "terrain transmissibility". Very marked regional groupings result. The limestones of southern Britain give a broad belt of negative residuals from Dorset to the Wash, flanked by the Severn/Avon/Ouse lowlands, Essex and Sussex/Kent which have a higher stream frequency than rainfall predicts. The lower stream frequencies of S.W. England may be due to granite outcrops (see Gardiner, 1971). The Pennines, east of the watershed are positive, as is North Wales, possibly explained by having headwater areas with dense drainage, whilst SAAR is reduced sharply east of the divide. This argument ought to apply to east Scotland but doesn't, perhaps due to drift deposit differences.

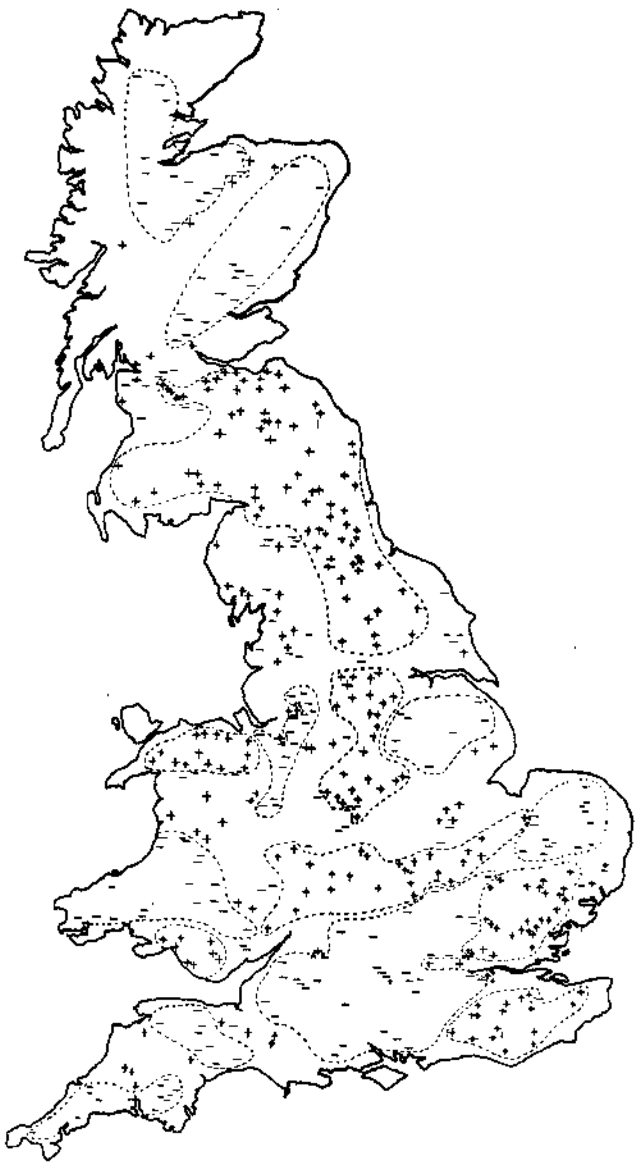

FIGURE 6 Residuals from Equation (4) plotted by sign, positive or negative

III VARIABLES DERIVED FOR UNIT HYDROGRAPH PREDICTION, ESPECIALLY THE STREAM FREQUENCY DIAGRAM

Certain extra catchment variables were derived for about 150 catchments under 500 km^2, used for catchment response studies (see Part I, pages 31-43). They included drainage density, the K parameter of the lemniscate loop (Chorley, Malm and Pogorzelski, 1957) and parameters of the distribution of channels at unit distances above the point of interest (stream frequency diagram).

Drainage density

The relationship between drainage density and stream frequency was used to index the former by the latter on the basis of Melton's work and a pilot study.

$$DD = 1.24\ STMFRQ^{0.500} \quad (r = 0.85) \qquad \ldots \qquad (5)$$

$$(n{=}130\ r \geqslant 0.230 \text{ sig. at } 0.01)$$

No other variables introduced to a multiple regression improved on this estimate. (TAYSLO, RSMD, SAAR and URB only raised the correlation coefficient to 0.86). The relationship is shown in Figure 7. The residuals from the relationship were investigated to discover explanations, especially since no other independent variables seem capable of reducing them. Figure 8 maps signs of the deviations with regional groupings.

The relationship between DD and STMFRQ has often been phrased in the reverse form to equation (5). Hence equation (5b) is given below.

$$STMFRQ = 0.651\ DD^2 \qquad \ldots \qquad (5b)$$

This compares almost exactly with Melton's (1957) result

$$STMFRQ = 0.6\ DD^2 \qquad \ldots \qquad (5c)$$

Melton says, "there is some reason to believe that degree of departure from maturity is reflected in values of F/D^2" (i.e. $STMFRQ/DD^2$).

It represents, he says, the completeness with which the channel network fills the basin outline for a given number of channel segments (junctions in our case). This theme is taken up by Abrahams (1972) and Wilcock (1975). The latter presents British results (26 catchments in the Bowland area, 19 in S.E. Scotland) which show that as the hypsometric integral approaches equilibrium, Melton's relation holds good (cf. the lower exponent which Wilcock derives for his third and fourth order basins). Thus on Figure 8, positive deviations (negative by equation 5(b)) mean immature basins in which the network does not fill the basin well. In practice the effect of geology is more pronounced.

Figure 9 shows that there is some systematic failure of equation (5) since high DD catchments have DD over-predicted and vice versa.

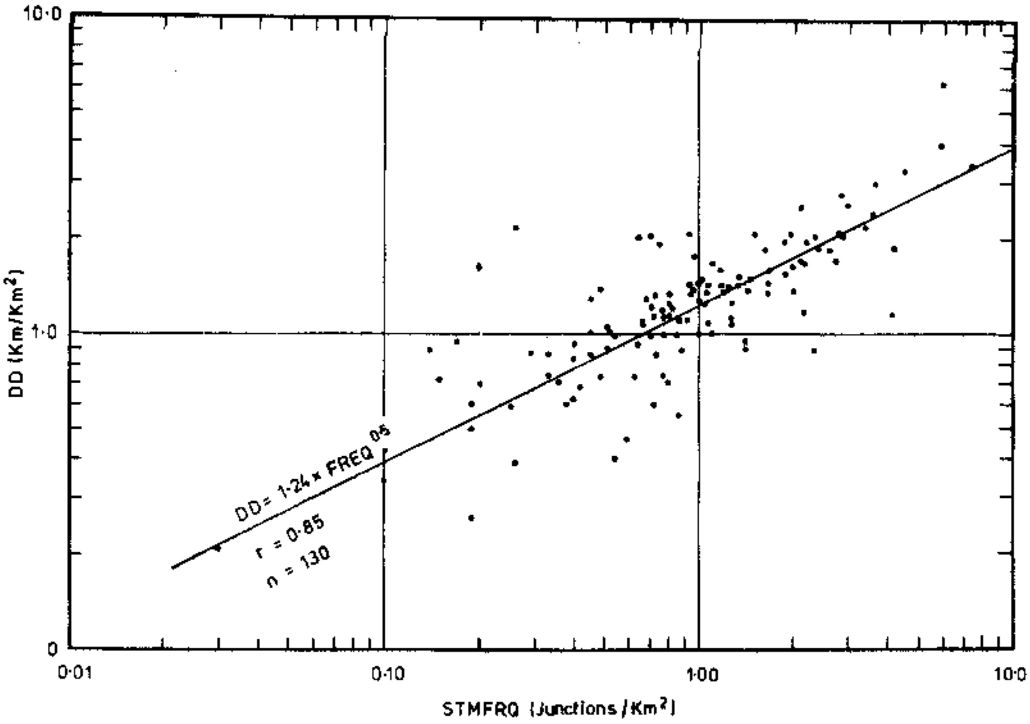

FIGURE 7 The relationship between stream frequency and drainage density

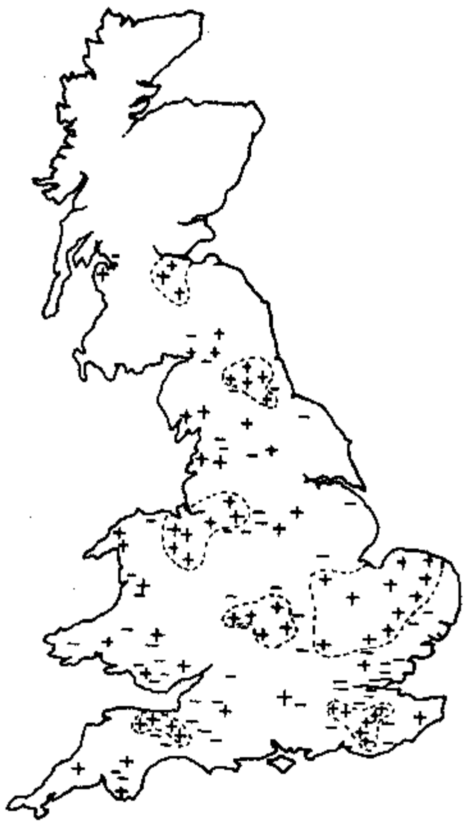

FIGURE 8 Residuals from Equation (5) by sign, positive or negative

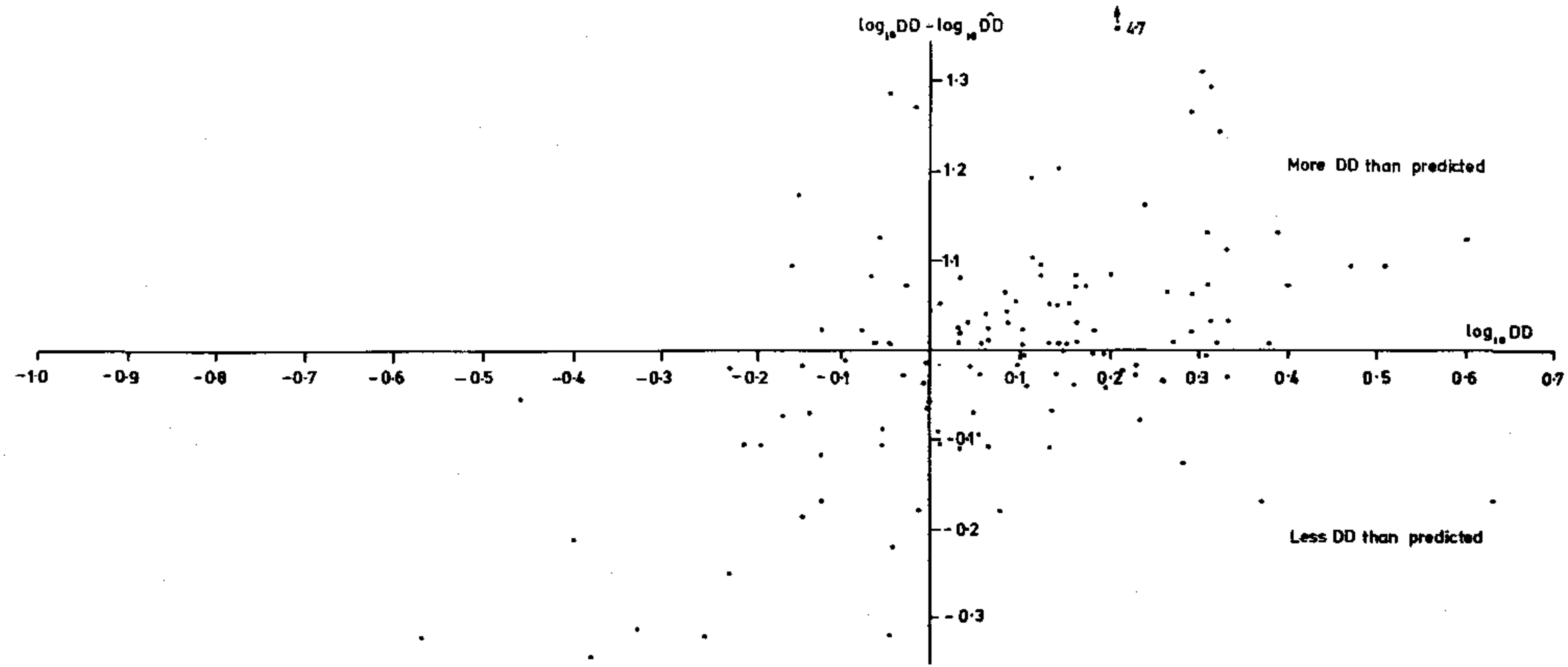

FIGURE 9 Residuals from Equation (5) plotted against drainage density

However, the pattern is again complicated by the inclusion of a wide range of geologies.

Table 3 shows that S1085 predicts STMFRQ (r = 0.60) better than DD (r = 0.46) which may also be a geological effect.

It should also be considered that map errors may contribute a great deal to scatter in DD and STMFRQ values.

The 'K' parameter

The value of this parameter increases as the model 'leaf' shape elongates. However, the relationship with MSL is clearly complex (see Table 6). Although MSL is the best single predictor, virtually no other variable is well correlated with SHAPE (apart from SFD1 which has the dimensions of stream length too). The following four-variable equation is required to index shape moderately well:

$$\text{SHAPE} = 0.885\ \text{AREA}^{-0.50}\ \text{MSL}^{0.91}\ \text{STMFRQ}^{-0.16}\ \text{TAYSLO}^{0.16} (r = 0.56) \quad \ldots\ldots (6)$$

It appears (Figure 4) that catchment shape is not strongly a function of area but of mainstream length. However, area does improve the prediction based on MSL alone.

$$\text{SHAPE} = 1.50\ \text{AREA}^{-0.43}\ \text{MSL}^{0.72}\ (r = 0.50) \quad \ldots\ldots (7)$$

The negative exponent on area, of course means there is some elongation of catchments as area decreases, but is also a complication created by the good correlation of AREA and MSL.

The stream frequency diagram

The diagram's construction is described in Part I. The chosen method of indexing the shape of the diagram was to calculate 1st, 2nd, 3rd and 4th moments, (later all but the first were dropped). However, to summarize the data, Figure 10 shows computer-produced histograms for the two most extremely skewed cases. The range of location of maximum network width, from 25% to 74% of network length agrees well with random model (Kirkby, 1974). By obtaining the distribution of networks at distance ordinates (both dimensionless by a %--total transformation) the whole picture of British channel network shapes is shown in Figure 11b; it can be seen that the average case shows a negative skewness although its general form has an unrepresentatively low kurtosis.

Turning to the relationship between SFD parameters and other catchment characteristics, SFD1 (centroid of the diagram) shows an obvious relationship with MSL since it retains the dimension of length. For this reason it also has a close relationship with AREA, S1085, and TAYSLO (see Table 6). It can be used instead of MSL to predict time-to-peak but takes longer to derive. In fact SFD2 enters regression

TABLE 6 Correlation matrix of unit hydrograph catchment characteristics

	AREA	MSL	STMFRQ	DDE	S1085	TAYSLO	SHAPE	SFD1E	RSMD	SAAR	URBAN	SOIL
AREA	1.000	0.929	- 0.275	- 0.207	- 0.632	- 0.713	- 0.013	- 0.843	- 0.190	- 0.139	- 0.219	0.046
MSL	0.929	1.000	- 0.125	- 0.029	- 0.574	- 0.688	0.167	0.912	- 0.135	- 0.113	- 0.281	0.091
STMFRQ	- 0.275	- 1.125	1.000	0.848	0.602	0.500	- 0.000	- 0.019	0.615	0.584	- 0.230	0.140
DDE	- 0.207	- 0.029	0.848	1.000	0.462	0.355	0.038	0.036	0.511	0.451	- 0.292	0.135
S1085	- 0.632	- 0.574	0.602	0.462	1.000	0.932	0.026	- 0.514	0.743	0.649	- 0.049	0.079
TAYSLO	- 0.713	- 0.688	0.500	0.355	0.932	1.000	0.005	- 0.607	0.597	0.518	0.065	0.037
SHAPE	- 0.013	0.167	- 0.000	0.038	0.026	0.005	1.000	0.162	0.050	- 0.017	- 0.022	- 0.071
SFD1	0.843	0.912	- 0.019	0.036	- 0.514	- 0.607	0.162	1.000	- 0.104	- 0.078	- 0.339	0.125
RSMD	- 0.190	- 0.135	0.615	0.511	0.743	0.597	0.050	- 0.104	1.000	0.872	- 0.228	0.095
SAAR	- 0.139	- 0.113	0.584	0.451	0.649	0.518	- 0.017	- 0.078	0.872	1.000	- 0.236	0.093
URBAN	- 0.219	- 0.281	- 0.230	- 0.292	- 0.049	0.065	- 0.022	- 0.339	- 0.228	- 0.236	1.000	- 0.415
SOIL	0.046	0.091	0.140	0.135	0.079	0.037	- 0.071	0.125	0.095	0.093	- 0.415	1.000

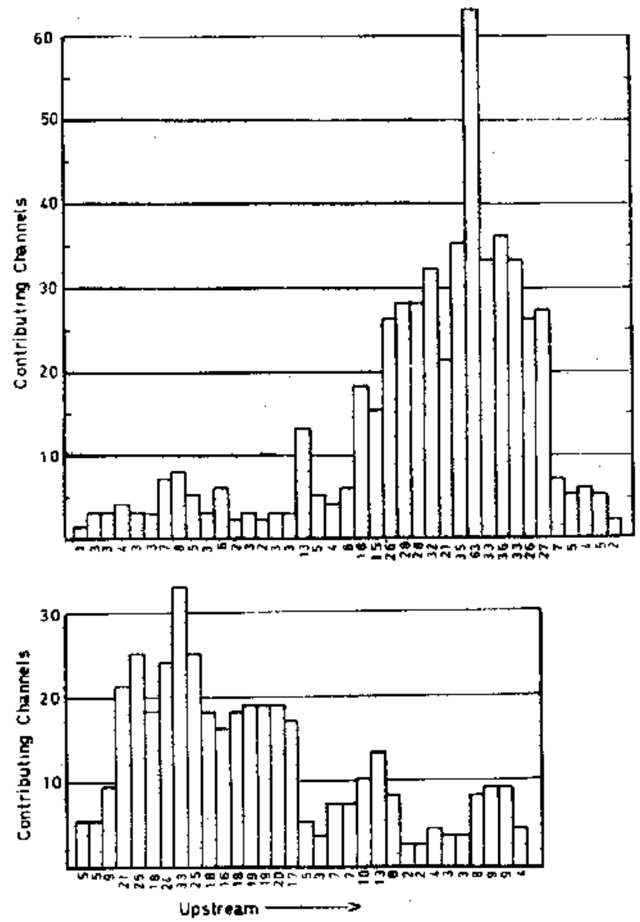

FIGURE 10 Histograms of stream frequency at 1 km intervals from the outlet (left-hand edge) of the two most skewed networks, Tone and Roden

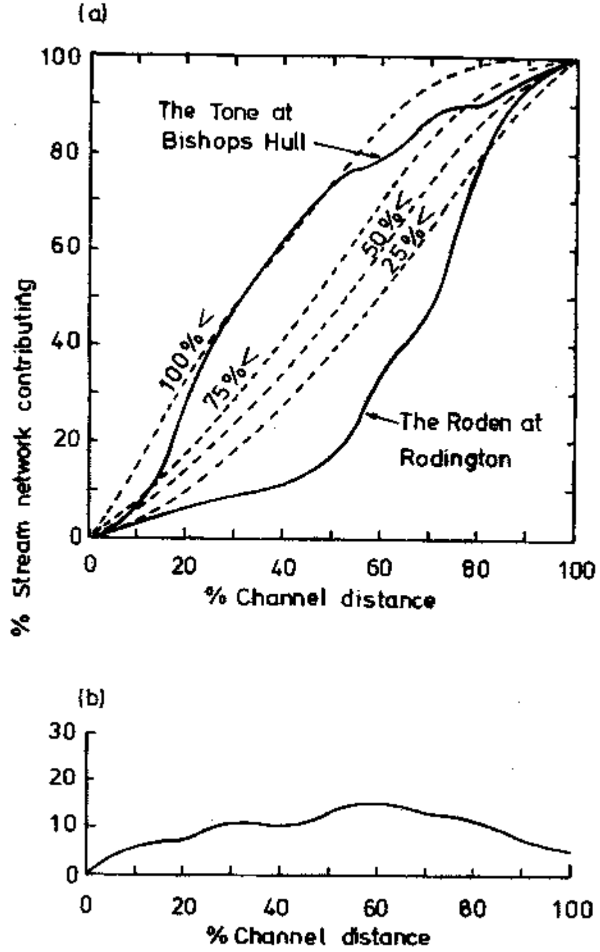

FIGURE 11 (a) Quartile ranges and extremes of cumulative stream frequency diagrams
(b) The median stream frequency diagram re-plotted

first, behind S1085 and URB, if all SFD parameters are used. None attains a significant t statistic.

Under the conditions investigated, the SFD does not obey the model of Rogers (1972) or Calver, Kirkby and Weyman (1972) in predicting outflow hydrographs even though theoretically it should. A possible explanation is that the use of distribution statistics of mapped channel flows to index actual flood flows is rather dangerous as indicated by an analysis of Blyth and Rodda's data (1973, plus personal communication) on the expanding network of the Ray at Grendon Underwood. Figure 12 shows how the distribution as well as the magnitude of the channel network changes in floods.

A further possible explanation is the dominance, in the sample catchments, of slope, climate and gross network properties.

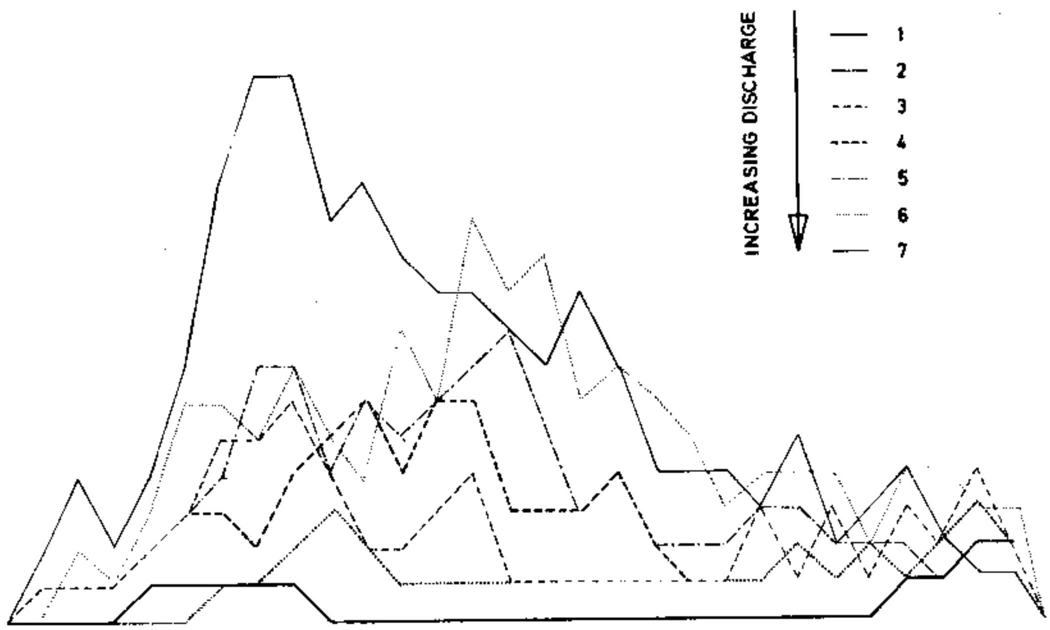

FIGURE 12 Stream frequency diagram of Ray at Grendon Underwood with changing flow conditions (field mapping)

IV REMAPPING VARIABLES AND PRINCIPAL COMPONENTS ANALYSIS

Remapping of derived data accomplishes two aims: it acts as a check on computation and the final maps are useful to obtain an approximate value of the catchment characteristic for an ungauged catchment. Of course, the procedure only applies to data derived from topographic maps: there would be little sense in re-mapping annual average rainfall.

Figure 13 consequently shows the maps of S1085 and STMFRQ for the British Isles. Both are given isopleth treatment although linear symbols should be used for main channel slope. However, there is a good correlation between all tributary slopes and that of the main channel (see Part I) so an areal interpolation is justified; the same view is taken of stream frequency.

Both maps show the obvious subdivision of Britain into highland and lowland zones with the Exe/Tees line as a rough Boundary. Thus an upland catchment would be expected to show a stream frequency of more than 1.0 and a channel slope of over 2m/km. As mentioned in Section II extrapolation in poorly-sampled areas has been achieved by means of regression on SAAR (a variable itself extrapolated by relief factors).

Several areas are displayed very clearly, for example the Chalk of Southern and Eastern Britain has very low STMFRQ values whilst the Avon/Thames/Ouse lowlands show up on the S1085 map. The less pervious soils of Essex and Kent show up as isolated high values for stream frequency. Stream frequency is a more conservative variable than S1085, only reaching peaks in the most pronounced of the uplands; S1085 values over ten can range over 100 as shown in Table 2.

Because the maps in Figure 13 resemble the relief, climate and soil maps of the British Isles, a multivariate analysis was conducted to investigate the possibility of dividing Britain (for which most data exists) by means of a multifactor variable. Lewis (1969) reports on a similar method of physiographic regionalization in Indiana. Principal components analysis was conducted on the matrix of 410 British catchments (preceded by $\log_{10}$ transformation of the variables). The results are shown in Table 7. Table 8 shows a more restricted analysis performed on all 533 catchments. In the former case, graduated values of each principal component were listed against catchment number and in Figure 14 a map was produced of Component I values. These describe the relief/climate/channel network set of variables for which reason the Principal Component has been called WET/DRY. Again the Exe/Tees line can be drawn, although the dividing line hugs the Welsh Marches before bulging east around the Pennines. North-east Scotland shows up as a comparatively dry area. Lewis, Mather and Doornkamp (1970) and Eyles (1971) all use more rigorous techniques for their final regional groupings, e.g. cluster analysis and multiple discriminant analysis, using more than one factor.

Component I has high negative loadings on 'wet' variables such as STMFRQ, S1085, TAYSLO, SAAR, M52D, and RSMD, with a high positive loading on SMDBAR. Component II is clearly one of scale; in other words catchments show most variance through WET/DRY but next through SIZE, with high positive

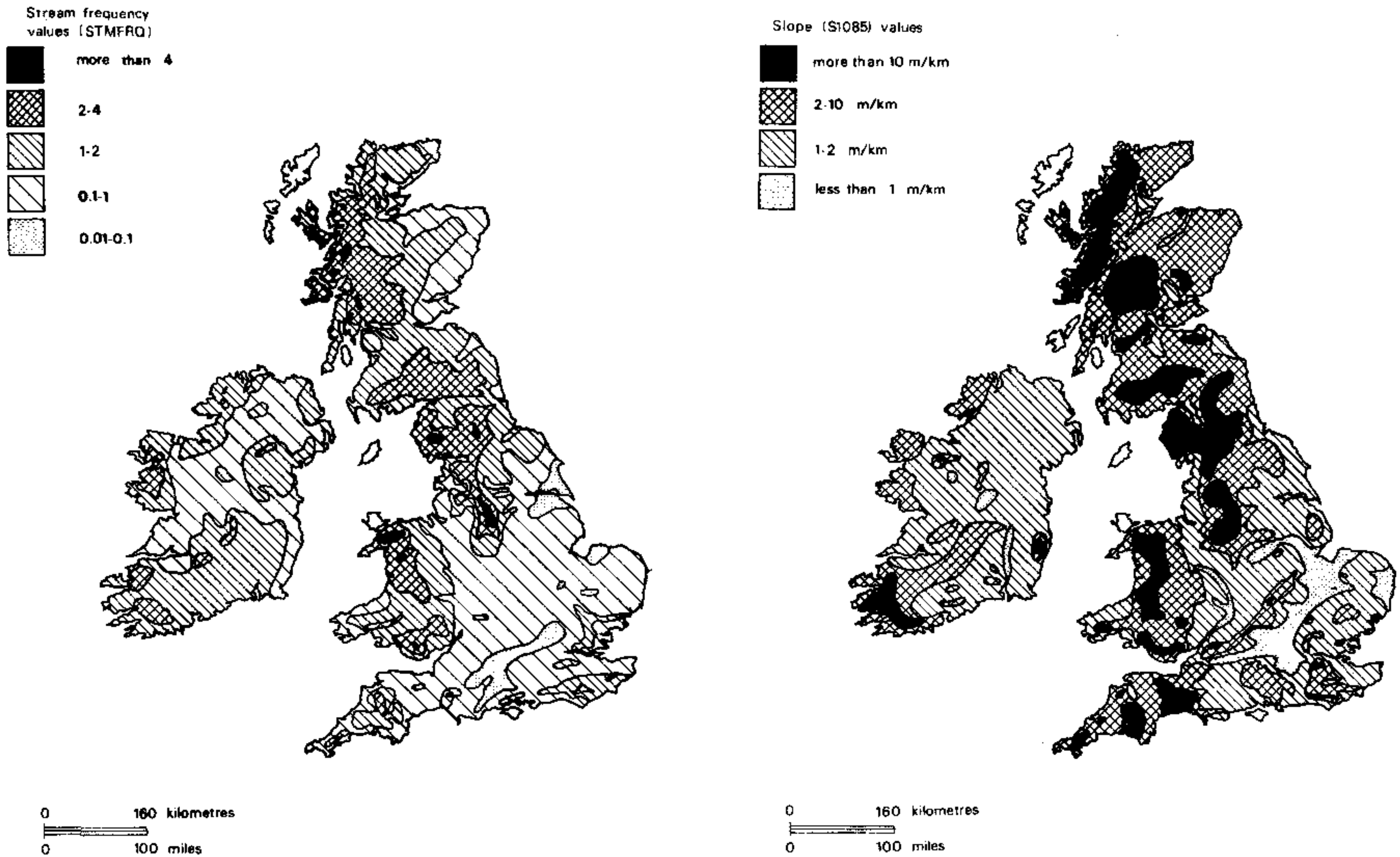

FIGURE 13 Maps of (a) STMFRQ and (b) S1085 for the British Isles

TABLE 7 <u>Principal Component Analysis of 410 British Catchments</u>

COMPONENT NUMBER	EIGENVALUE OR COMPONENT VARIANCE	ACCUMULATED VALUE AS % OF TOTAL VARIANCE
1	5.4337951	41.80
2	3.3767292	67.77
3	1.0173327	75.60
4	0.9476190	82.89
5	0.8116606	89.13
6	0.4596266	92.67
7	0.3309892	95.21
8	0.2215858	96.92
9	0.2099793	98.53
10	0.0721251	99.09
11	0.0583190	99.54
12	0.0444186	99.88
13	0.0158199	100.00

EIGENVECTORS	COMPONENT NO 1	COMPONENT NO 2	COMPONENT NO 3	COMPONENT NO 4	COMPONENT NO 5	COMPONENT NO 6	COMPONENT NO 7	COMPONENT NO 8	COMPONENT NO 9	COMPONENT NO 10	COMPONENT NO 11	COMPONENT NO 12	COMPONENT NO 13
VARIANCE	5.433795	3.376729	1.017333	0.947619	0.811661	0.459627	0.330989	0.221586	0.209979	0.072125	0.058319	0.044419	0.015820
AREA	0.163076	0.464233	0.109513	0.019577	0.187851	0.123670	0.147514	0.477507	0.001424	0.077365	0.366582	0.551987	0.010114
STMFRQ	- 0.307057	- 0.122947	- 0.405580	- 0.177426	- 0.087136	- 0.431114	- 0.597227	0.344128	0.116063	0.077889	- 0.033612	0.073399	- 0.003669
S1085	- 0.351637	- 0.242630	- 0.084854	- 0.059641	0.054083	- 0.107458	0.417998	- 0.292061	- 0.062787	- 0.546035	- 0.369747	0.292099	- 0.089996
TAYSLO	- 0.316433	- 0.303244	- 0.125316	- 0.049370	0.071443	- 0.145016	0.430819	0.291854	- 0.139419	0.446698	0.412893	- 0.321074	0.017558
DVF	- 0.009597	- 0.441816	0.275836	0.146695	- 0.119485	0.522370	- 0.362831	0.519438	0.088974	0.026504	- 0.065615	- 0.074144	0.018769
MSL	0.141970	0.486979	0.013012	0.018310	0.172799	0.014219	0.129065	0.386519	0.027709	- 0.117012	- 0.329126	- 0.650324	0.019260
SAAR	- 0.390905	0.134725	0.193062	0.045039	0.097116	0.061181	- 0.168208	- 0.116249	0.056380	- 0.577084	0.576740	- 0.242778	0.083572
M52D	- 0.382833	0.143216	0.224252	0.067302	0.085719	0.080806	0.054387	- 0.127325	0.506174	0.244197	- 0.101623	0.037907	- 0.646107
RSMD	- 0.393617	0.135103	0.202418	0.062965	0.129714	0.051897	0.068390	- 0.094466	0.274525	0.232195	- 0.259439	0.093479	0.737177
SMDBAR	0.351729	- 0.166476	- 0.161604	- 0.056676	- 0.242647	- 0.146431	0.216106	0.093416	0.771562	- 0.169155	0.183941	- 0.056203	0.150272
SOIL	- 0.192175	0.202315	- 0.594830	- 0.217023	- 0.245915	0.661469	0.110062	- 0.100922	0.012855	0.004423	0.018930	- 0.007825	0.017212
URBAN	0.092887	- 0.129029	0.192131	- 0.885421	0.368247	0.091168	- 0.067744	- 0.029578	0.062220	0.003823	- 0.010915	- 0.021179	0.001864
LAKE	- 0.107342	0.209730	0.423041	- 0.309004	- 0.785106	- 0.113736	0.116594	0.073732	- 0.142703	0.030354	- 0.009426	- 0.017579	0.001759

TABLE 8 Principal Component Analysis of logarithmically transformed catchment characteristics for 533 catchments in Great Britain and Ireland

Eigenvectors	Component No.1	Component No.2	Component No.3	Component No.4	Component No.5	Component No.6	Component No.7
Variance	2.61677	1.54151	0.92411	0.86008	0.51698	0.37390	0.16665
AREA	0.25068	0.62026	0.04361	0.37060	0.38707	0.07856	0.50715
STMFRQ	- 0.49125	0.20439	0.03616	0.27014	0.04750	0.79880	0.04769
S1085	- 0.47426	- 0.44002	- 0.04460	- 0.10076	0.12842	0.09129	0.73789
SOIL	- 0.43204	0.23762	0.20694	0.39108	- 0.55373	0.50320	- 0.03596
URBAN	0.13298	- 0.34338	0.86395	0.25853	0.21117	- 0.05196	0.06209
LAKE	- 0.11049	0.43306	0.45294	- 0.74294	- 0.14023	- 0.05225	0.14403
RSMD	- 0.50433	0.13402	- 0.02029	- 0.08826	0.67865	0.29796	0.41248

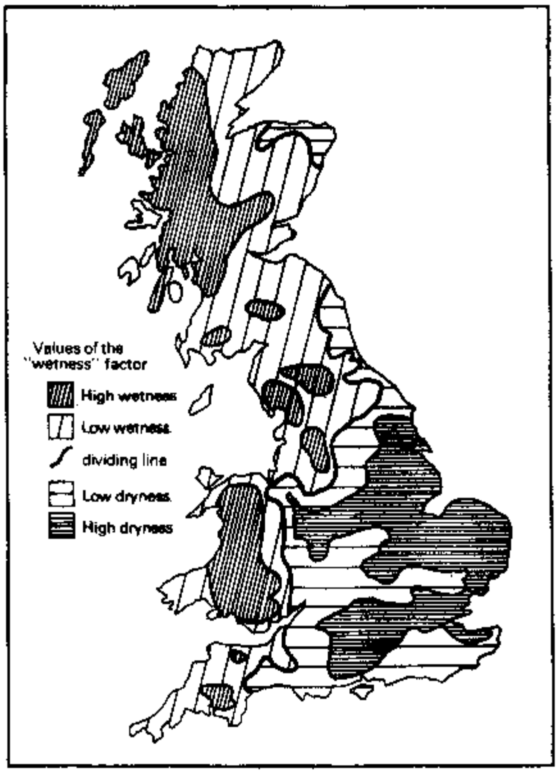

FIGURE 14 Mapping of Principal Component I - WET/DRY

loadings on catchment area and channel length variables. This component has not been mapped since catchment size has very little regional cohesion and cannot be interpolated.

Components I and II are shown in Figure 15 to be the only significant ones and account for 68% of catchment variance. Further explanation this method is treated by Craddock and Flood (1969) and Farmer (1969). Table 8 shows that loadings are similar for the larger sample. A brief examination of lower order components in Table 7 shows less obvious relationships with original variables although 3 appears to be connected with STMFRQ and SOIL, 4 with URBAN, 5 with LAKE with repetitions thereafter. The instability of these components is indicated by changes between Tables 7 and 8.

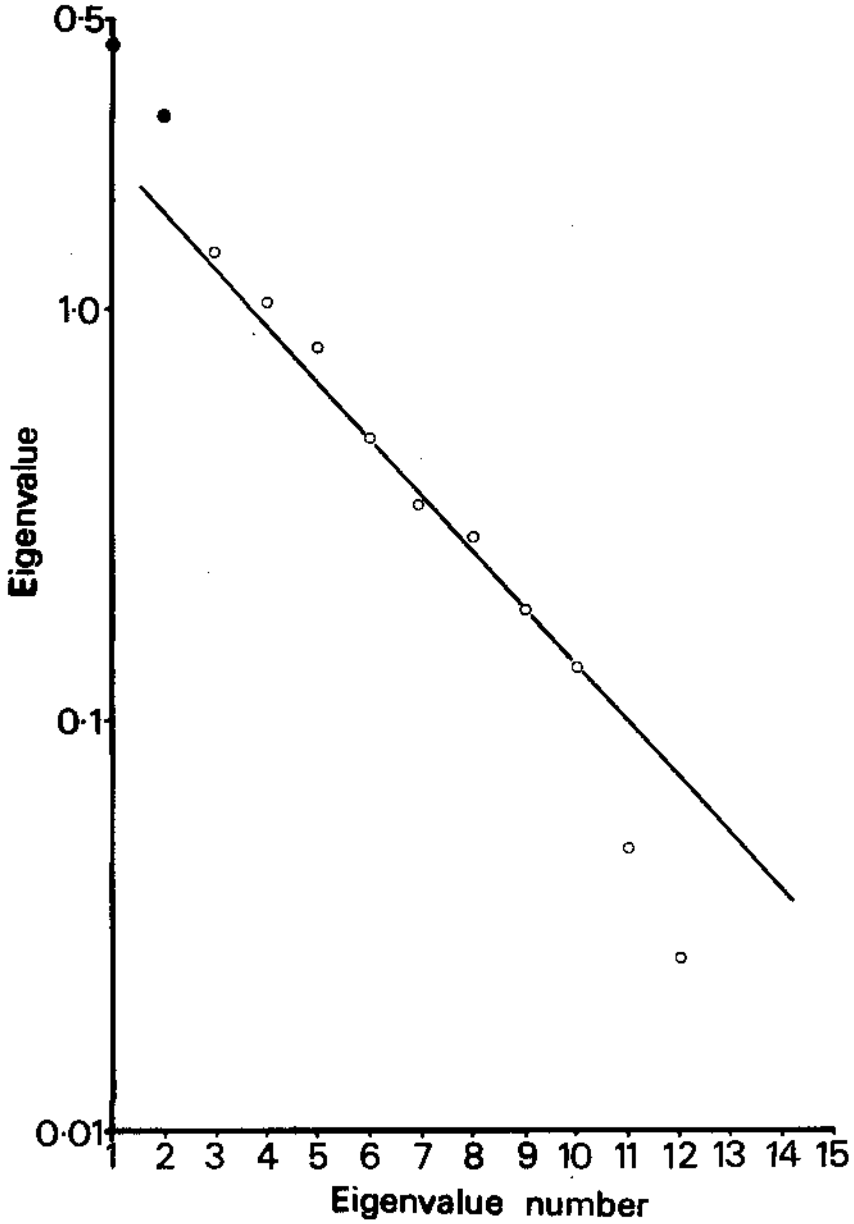

FIGURE 15 Separation of significant Principal Components

Table 9 shows that size, relief and channel network components dominate the outcome of multivariate analyses of catchment characteristics. Variations in their order and dominance are clearly related to sample size and type: most single region approaches relegate size factors down the list (e.g. 1,3) although Gardiner's intensive study of S.W. England lists size as the first component. Given a wide range of catchment sizes, AREA's *position* in the outcome depends on its logarithmic transformation (see Part I of this Report). The regional coverage of the sample is clearly important and the British sample relegates the size factor below those of the climate/relief/network complex.

A plot of catchments according to their scaling on components I and II is shown in Figure 16. This can effectively be used for picking large/wet, large/dry, small/wet and small/dry catchments for further analysis. The plot is not symmetrical, showing that the majority of small catchments are fairly dry (mostly urban) although a few are very wet (mostly research upland catchments). King (1966) uses such a plot for regionalization but this is avoided here because of the lack of clear groupings; further analysis is necessary, as suggested by Mather and Doornkamp (op. cit.). Figure 14 is left as the most effective regionalization available from these data, bearing in mind especially that the gauged catchments do not represent a totally objective sample.

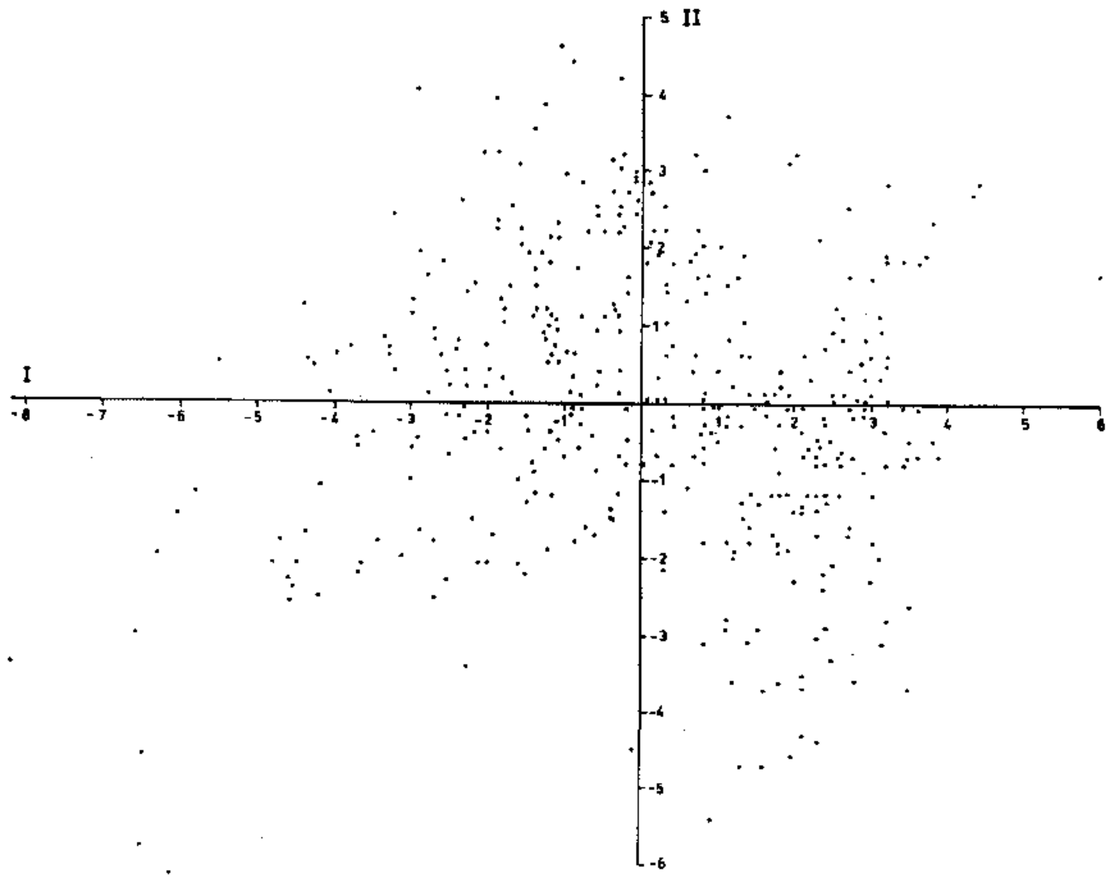

FIGURE 16 Catchments scaled against the Principal Components WET/DRY (I) and SIZE (II)

TABLE 9 Principal Components analysis of catchment characteristics - comparative listing of approach and results

	Authors	Region	No. of catchments	Variables	Size range of catchments	Log transformation	First 4 components and % explanation of total variance where stated	
1	Overton 1969	Mid-U.S.A.	37	9	< 111 km^2	?	1 slope 2 size 3 drainage 4 shape	(32) (23) (21) (8)
2	Wong 1963	New England	90	12	26-5180 km^2	✓	1 size 2 slope 3 wetness 4 random covariation	(47) (39) (7) (7)
3	Lewis 1969	Wisconsin	60	13	4th orders	✓	1 relief 2 size 3 network 4 shape	(36) (19) (12) (8)
4	Mather & Doornkamp 1970	Uganda	130	18	3rd orders	✓	1 size 2 network 3 stream length ratio 4 relief	(48) (21) (9) (8)
5	Abrahams 1972b	S. Australia	282	5	3rd orders	✓	1 planimetric (area/network) 2 relief 3 geological control	
6	Gardiner 1974	N.W. Devon	2000	23	1st, 2nd, 3rd orders	✓	1 size 2 relief 3 network 4 network	(39) (24) (11) (9)
7	Blake, Cook & Greenall 1970	N. Island N.Z.	14	39	2.5-238.0 km^2	?	1 soil 2 slope/veg. 3 shape/geology 4 network	
8	Rice 1970	San Dimas, Calif.	14	12	0.13-11.38 km^2	?	1 soil 2 shape 3 shape/geology 4 network	
9	This Report 1976	U.K.	410/533	13	0.05-9868 km^2	✓	1 wetness 2 size 3 soil/network 4 urban	(42) (26) (8) (7)

Conclusions

The derivation of catchment characteristics for a large number of sites in the British Isles was not a geomorphological exercise and the considerable effort involved must be judged finally on the degree to which the variables were successful in predicting flood parameters.

Although the final prediction equations are statistically (not mechanically) valid, some very logical combinations of variables appear. For instance:

$$\text{BESMAF} = 0.0201 \ \text{AREA}^{0.94} \ \text{RSMD}^{1.03} \ \text{SOIL}^{1.23} \ \text{STMFRQ}^{0.27} \ \text{S1085}^{0.16} \ (1 + \text{LAKE})^{-0.85} \quad \ldots (8)$$

shows the logical scaling of the input by AREA, its conversion to runoff by SOIL in conjunction with RSMD, the conversion of this runoff to a hydrograph by the channel network (STMFRQ and S1085) and the influence of storage on a large scale by LAKES. Slope and lakes have less effect on the longer duration "mean annual calendar day flood" (CALMAF):

$$\text{CALMAF} = 0.0056 \ \text{AREA}^{0.97} \ \text{RSMD}^{1.09} \ \text{SOIL}^{0.49} \ \text{STMFRQ}^{0.46} \quad \ldots (9)$$

but slope returns to explain the ratio of even longer duration floods, e.g. that of 10 days, to the calendar day flood:

$$\text{AR10} = 0.54 \ \text{S1085}^{-0.13} \quad \ldots (10)$$

The prediction of unit hydrograph time-to-peak logically includes channel length and slope (though these were not used as a combined index of velocity), plus a climatic and a land-use factor (RSMD and URBAN)

$$\text{Tp} = 46.6 \ \text{MSL}^{0.14} \ \text{S1085}^{-0.38} \ (1 + \text{URBAN})^{-1.99} \ \text{RSMD}^{-0.4} \quad \ldots (11)$$

The derivation of the SOIL index again proves of value in the prediction of standard percentage runoff (SPR).

$$\text{SPR} = 95.5 \ (\text{SOIL}) + 0.12 \ (\text{URBAN}) \quad \ldots (12)$$

Although this equation only explains 45% of the variance it represents an advance on the conclusion of Woodruff and Hewlett (1970) that quick-flow runoff volumes cannot be predicted, only mapped.

The appearance of URBAN in the prediction of unit hydrograph variables is the logical outcome of including certain London catchments in the sample. The effect of urban development on hydrograph shapes has been studied (e.g. by Hollis, 1974) and the appearance of the URBAN variable is satisfactory only if it is remembered that it is a very coarse variable and purports to represent a uniform effect whereas, of course, urban land-use and storm drain networks differ widely. At least URBAN appears with exponents of the correct sign; at one stage of the investigation it appeared that mean annual flood would be *inversely* related to URBAN through the latter's negative correlation with the flood-enhancing properties of slope, stream frequency and so on. Finally, however, in the prediction of mean annual floods for the Thames, Lee & Essex areas where such negative correlations are weak, URBAN proves a strongly positive variable:

$$\text{BESMAF} = 0.302\ \text{AREA}^{0.70}\ \text{STMFRQ}^{0.52}\ (1+\text{URBAN})^{2.5} \quad \ldots (13)$$

Further indication of a suitable outcome is the agreement between the BESMAF equation (8) and those produced by Benson (1962) for a humid temperature zone of the United States, Nash and Shaw (1965) and Cole (1965) for Britain (see Section 4.3.11 of the Flood Studies Report).

No detailed comparison of coefficients is attempted for statistical reasons quoted previously. The time-to-peak relationship includes the well-correlated size and slope indices (cf. Nash (1960) whose area and land-slope were not well correlated) and includes the urban variable which has been proved so important in American studies.

On the less optimistic side it must be pointed out that, both in the case of mean annual flood and unit hydrograph predictions, considerable residual errors exist. For instance the information content of even the detailed regional equations for mean annual flood is only equivalent to 1.1 years of data. The superior performance of Benson's (1962) equations is noticeable. If a measurement of catchment lag exists where the unit hydrograph is required, the error in predicting Tp can be halved from that inherent in predicting from catchment characteristics.

There is evidence both in the use of a subdivision into regions for BESMAF prediction and in the distribution of residuals that a further real factor is at work. Data errors can be discounted since flow data of various grades were combined in different ways during trials without significantly affecting the outcome. Various stratifications by catchment area were also tried. Of the less available catchment characteristics used for trials with unit hydrograph predictions, none was successful. SFD1 was a poor surrogate for MSL, shape appeared only in Tp regressions for catchments under 50 km^2 and DD similarly for catchments between 200 km^2 and 500 km^2. It therefore seems unlikely that, if derived for all catchments, an improved BESMAF prediction would result.

It seems difficult to spot a catchment characteristic which will produce further refinements if derived. However, certain research topics are worthy of attention.

(a) The indexing of the partial contributing area by means of more detailed soils, land use or natural vegetation mapping is the next logical step following the inclusion of stream network variables. This would incorporate a concept of flood runoff widely proved in humid, especially upland, catchments.

(b) Whilst good correlations between channel slope and catchment ('overland') slope were proved in pilot studies of the

Bristol region and for Ireland, the relationship is not so strong on a national scale. It is now felt that a corollary of partial area indices would be the use of maps of ground slope categories when available. The form of the hypsometric integral might also be useful. The markedly weaker relationship of overland slope with size variables makes it more statistically valid in situations such as equation (11). For this reason it is preferred to channel slope by Wong (1971). The problem of intercorrelations of AREA and S1085 are shown by the technique of ridge regression (see Flood Studies Report, Vol. I, page 362). Refinements may still be possible to channel slope measurements. British catchments are all relatively small and steep, with "polycyclic" channel profiles. Although TAYSLO, which samples the whole profile, is closely related to S1085 which picks only two points, there remains much information in the whole profile and the channel slope is much easier to work on with automated methods than the ground slope. An attempt has been made to plot the relief - stream frequency diagram (King, 1972) and so derive overland slope. This works quite convincingly in Zambia but no similar orientation of the points could be discovered in British data. Much more spread occurred in relief than in stream frequency. Consequently, only data for E. Anglia, Essex/London and the South gave a spread within King's range of values. This is clearly not a short cut to overland slope values although trial comparisons have not been made.

(c) No time was available to include channel capacity measures in the Flood Study. Evidence is growing that incorporation of field survey data (which often exist near gauging stations) could well improve predictions. The author is still convinced of the theoretical validity of the stream frequency diagram and hopes that the diagram will achieve more success if combined with the nature of the channels sampled, which must be very variable in capacity, roughness and slope.

It seems unlikely that progress in morphometry from maps, even using automated techniques, will be rapid enough to provide national scale data to refine flood prediction. More likely is the arrival of remote sensing and the automatic analysis of monochrome, colour, UV and I-R photography which will provide the next new opportunity for significant progress.

Hollyday (1975) claims a reduction of standard errors by over ten percent in 15 out of 40 prediction equations for streamflow characteristics when using ERTS data, automatically retrieved, compared with map catchment characteristics.

Finally, the catchment characteristics assembled for the Flood Study have been shown to exhibit the regularity of interrelationships between physiography, land-use, soils, climate and hydrology that the systems concept of drainage basins has led us to expect and which led to their initial choice (see Part 1 of this Report).

Acknowledgements

The work described here formed part of a team exercise and the author wishes to acknowledge all his colleagues, specifically, Dr J. B. Miller encouraged the study of inter-relationships between variables held on the massive computer master-list, Mrs A. Sekulin advised on histogram graphics by line-printer, Mr. K. Blyth provided field data on the Grendon Underwood channel network. Miss B. J. Glover and Miss J. Ellis worked hard on the derivation of indices. These are in addition to those acknowledged in Part One.

REFERENCES

ABRAHAMS A. D. 1972. Environmental constraints on the substitution of space for time in the study of natural catchments. *Bull. Geol. Soc. Amer.*, 83, 1523-1530.

ABRAHAMS A. D. 1972(b). Factor analysis of drainage basin properties: evidence for stream abstraction accompanying degradation of relief. *Water Resour. Res.* 8, 624-633.

ANDERSON H. W. 1975. Relating sediment yield to watershed variables. *Trans. Amer. geophys. Un.* 38, 921-924.

BENSON M. A. 1962. Factors influencing the occurrence of floods in a humid region of diverse terrain. *U.S.G.S. Water Supply Paper*, 1580-B.

BLYTH K. and RODDA J. C. 1973. A stream length study. *Water Resour. Res.* 9, 1454-1461.

CALVER A, KIRKBY M. J. and WEYMAN D. R. 1972. Modelling hillslope and channel flows. Spatial analysis in Geomorphology. Ed. R. J. Chorley, Methuen.

CLARK D. 1973. Normality transformation and the principal components solution: an empirical note. *Area*, 5(2), 110-113.

COLE G. 1965. An application of the regional analysis of flood flows. River Flood Hydrology: *Proc. Symp. Inst. Civ. Engrs.* London, 39-57.

CRADDOCK, J. M. and FLOOD, C. R. 1969. Eigenvectors for representing the 500 mb geopotential surface over the northern hemisphere. *Quart. J. Roy. Met. Soc.*, *95*, 576-593.

EAGLESON P. S. 1970. Dynamic Hydrology. McGraw-Hill.

EYLES R. J. 1971. A classification of West Malaysian drainage basins. *Annals of Assoc. Amer. Geogrs*, 61(3), 460-467.

FARMER, S. A. 1969. An investigation into the results of principal components analysis of data derived from random numbers. *The Statistician*, *20*(4), 63-72

GARDINER V. 1971. A drainage density map of Dartmoor. *Trans. Devonshire Assn*, 103, 167-180.

GARDINER V. 1974. Drainage basin morphometry. *British Geomorph. Research Gp.*, *Tech. Bull.* 14.

GRAY D. M. 1961. Interrelationships of watershed characteristics. *J. Geophys. Res.*, 66, 1215-1223.

HACK J. T. 1957. Studies of longitudinal stream profiles in Virginia and Maryland. *U.S.G.S. Prof. Paper* 294-B.

HOLLIS 1974. The effect of urbanisation of floods in the Canon's Brook, Harlow, Essex. In: 'Fluvial Processes in Instrumented Watersheds, Ed. K. J. Gregory and D. E. Walling. Inst. Br. Geogrs., Spec. Pubn. 6, 123-139.

HOLLYDAY E. F. 1975. Drainage basin characteristics from ERTS data. Progress Rept. U.S.G.S. Reston, Virginia.

KING C. A. M. 1966. A morphometric example of factor analysis. In: Morphometric Analysis of Maps., *Br. Geomorph. Res. Gp., Occ. Paper 4*, 41-48.

KING R. B. 1972. Relief-stream frequency (R-F) diagram: method of displaying physiographic regions. *J. Geol.*, 80, 740-743.

KIRKBY M J. 1974. Generation structure and topology of channel networks. Working Paper 88, Dept. of Geography, University of Leeds.

LEWIS L. A. 1969. Analysis of surficial landform properties - the regionalization of Indiana into units of morphometric similarity. *Proc. Indians Acad. Science*, 78, 317-328.

MATHER P. M. and DOORNKAMP J. C. 1970. Multivariate analysis in geography with particular reference to drainage basin morphometry. *Inst. Br. Geogrs. Trans.*, 51, 163-187.

MELTON M. A. 1957. An analysis of the relations among elements of climate, surface properties and geomorphology. Office of Naval Research, Geography Branch, Project NR 389-042, *Tech. Rep.* 11.

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

NASH J. E. 1960. A unit hydrograph study with particular reference to British catchments. *Proc. Instn. Civ. Engrs.*, 17, 249-282.

NASH J. E. and SHAW B. L. 1965. Flood frequency as a function of catchment characteristics. River Flood Hydrology: *Proc. Symp. Inst. Civ. Engrs.*, London 115-135.

N.E.R.C. 1975. Flood Studies Report. London.

OVERTON D. E. 1968. Eigenvector analysis of geomorphic interrelations of small agricultural watersheds. *Trans. Amer. Soc. ag. Engrs.* 12(4), 506-508.

ROGERS W. F. 1972. New concept in hydrograph analysis. *Water Resour. Res.* 8(4), 973-981.

SOKOLOV A. A. 1962. Interrelationships between geomorphological characteristics of a drainage basin and the stream. *Soviet Hydrology: Selected Papers*, 1, 16-22.

SMART J. S. and SURKAN A. J. 1967. The relation between mainstream length and area in drainage basins. *Water Resour. Res.* 3(4), 963-974.

WILCOCK D. N. 1975. Relations between planimetric and hypsometric variables in third and fourth order drainage basins. *Geol. Soc. Amer. Bull.* 86(1), 47-50.

WONG S. T. 1963. A multivariate statistical model for predicting mean annual flood in New England. *Ann. Ass. Amer. Geogr.*, 53, 298-311

WONG S. T. 1971. Effect of stream size and average land slope on the recurrence of floods in New Zealand. *Geographical Analysis*, 3, 77-83.

WOODRUFF J. F. and J. D. HEWLETT 1970. Predicting and mapping the average hydrologic response for the Eastern United States. *Water Resour. Res.* 6(5), 1312-1326.