GREAT-ER Working Note No. 5

VELOCITY ESTIMATION
AT
UNGAUGED SITES

Final Report

This report is an official document prepared under contract between ECETOC and the Natural Environment Research Council. It should not be quoted by third parties without permission of both ECETOC and the Institute of Hydrology.

Institute of Hydrology
Crowmarsh Gifford
Wallingford
Oxfordshire
OX10 8BB

Tel: 01491 838800
Fax: 01491 692424
Telex: 849365 Hydrol G

February 1997

VELOCITY ESTIMATION
AT
UNGAUGED SITES

C. Round & A. Young

Contents

Executive summary

A velocity estimation model is required for GREAT-ER which is regionally applicable throughout Europe and uses readily available data. Therefore, the primary objective of this report is to identify the best possible method for estimating flow velocities at ungauged sites.

The data used in the study was derived from two sources; detailed field studies and velocity-area gauging stations. The field studies were based upon sites with catchment areas of less than 200 km^2, thus archive data was sourced for velocity area gauging stations on river stretches with catchment areas greater than 200 km^2. For each of the 111 sites information of slope, catchment area, BFI, the mean, Q10 and Q95 flows and their corresponding channel widths, depths, hydraulic radii, and velocities were collated.

By comparison with gauged catchments held on the UK National river flow archive and the European water archive, it was identified that the catchments used in this study were from areas with generally higher rainfall.

A multivariate regression analysis found that the average velocity can be best estimated from a power function relating velocity to both discharge and discharge standardised by mean flow. Using this equation the velocity can be estimated, with 68% confidence, to within ±1.76 of the estimate. The model implies that velocity increases with discharge.

Due to its universal application for velocity estimation, an analysis was undertaken to assess the feasibility of Manning for regional velocity estimation. Two analyses were undertaken; the first based upon the back calculation of n and, the second based upon the estimating velocity using estimates of both hydraulic radius and Manning's n derived from regression models relating these to readily available catchment characteristics.

From the first analysis it was found that 70% of n values were back calculated to within the range expected for these channels, as given in the literature. The 68% confidence interval of the estimates of n implied that n, and thus velocity, could be estimated to within ±2.6 of the estimate.

The second analysis showed that hydraulic radius and n can be estimated from flow and channel data, and that by substituting these estimates and measured values of slope into Manning's equation estimates of velocity could be made. Using this approach would enable the velocity to be estimated, with 68% confidence, to within ±2.25 of the estimate.

It is not recommended that hydraulic radius is estimated from the regression equation, due to the gauging stations used in this study being at locations with natural and uniform properties. Many channels have anthropogenic influences which are not generally mapped and therefore are not reflected in this study.

The best estimate of velocity, based upon the 68% confidence intervals, can therefore be derived using the regression model.

An analysis of the relative importance of uncertainties in C_o and velocity in controlling the response of a first order biodegradation model indicated that, for a readily degradable substance such as LAS, it is fundamental to accurately estimate the pollutant mass loads to rivers in order to estimate aquatic concentrations in all but the longest rivers. The resulting dilution and the estimation of the degradation may be a secondary influence under these conditions. In longer rivers the uncertainty introduced by an uncertainty in velocity is likely to be masked by the much greater uncertainties in the decay constant, k.

1 Introduction

The objectives of this report are to:

1. summarise the existing velocity estimation techniques (presented in full in the IH GREAT-ER Progress Report: June-September 1996);

2. detail the analysis which has been undertaken in order to identify the best possible method for estimating flow velocities at ungauged sites;

3. conclude which approach should be adopted for use within GREAT-ER.

4. demonstrate, in the case of LAS, the relative importance of uncertainties in initial concentrations, C_o, and the product of the decay constant, k, and time of travel derived from velocity, t, in controlling the uncertainty in the response of the first order degradation model.

A summary of a literature review which details the techniques in existence for estimating flow velocities and/or channel geometry at ungauged sites is presented in Chapter 2. These techniques are based upon either power functions derived from linear regressions of log transformed regional variables, such as those derived by Leopold and Maddock (1953), or empirical formulae, such as Manning's (1891) and Simons and Albertsons' (1960) formulae, or mathematical formulae, such as Chézy (1775).

Chapter 3 details the construction of the data set which was used in the analysis. The data set was derived from two sources; detailed field measurements and archive data.

The analysis is described in Chapter 4, which first identifies the variables which are important in the estimation of velocity by undertaking multivariate regression analyses of log transformed data, and then discusses the derivation of the best possible model for use within GREAT-ER. A comparison of the performance of this regression derived equation with the performance of Manning's equation is made by comparing the confidence interval for velocity derived from the regression equation to the confidence interval associated with the velocity derived using Manning's. The comparison was made with Manning's because it is the most widely used uniform flow formula for open channel flow computations (Chow, 1959).

Chapter 5 discusses the relative importance of uncertainties in C_o and velocity, in order to determine their relative impacts on the response of the first order degradation equation.

Conclusions and recommendations are discussed in Chapter 6.

2 Background to velocity estimation

From the literature review it is evident that there are four principal methods used for estimation of velocity. These are described in the progress report (June-September 1996) and are summarised below.

2.1 REGRESSION BASED ESTIMATION

In its simplest term, channel flow (Q) at a point can be approximated as the product of the average velocity (V) and the cross-sectional area (A).

$$Q = AV$$

Leopold and Maddock (1953) showed that the relationships between channel measures; width, depth, and velocity, and discharge at a channel cross section are described best by the power functions:

$$W = aQ^b \qquad D = cQ^f \qquad V = kQ^m$$

where:

W	= top width (m)
D	= mean depth (m)
V	= mean velocity (m)
a, b, c, f, k, m	= numerical constants and exponents

These power functions are derived by fitting a linear regression between the log of the dependent and the log of the independent variable. Taking the anti-log of the resultant regression yields a power function, as above.

Following the approach of Leopold and Maddock (1953), numerous authors have derived power functions for specific geographical regions or catchments (Betson, 1979; Charlton, 1978; Gregory & Walling, 1974; Nixon, 1959; Park, 1976). The estimates of the constants and parameters are specific to those geographic regions or catchments and are not directly applicable elsewhere due to contrasts in fluvial and climatological environments.

If this approach were to be adopted for GREAT-ER it would be necessary to ensure that the relationships, described by the power functions, were applicable on a European scale as opposed to regional.

2.2 SIMONS AND ALBERTSON

Simons and Albertson (1960) theorised that velocity could be estimated using:

$$V = 9.355 \cdot \sqrt[3]{HR^2 \cdot S}$$

where:

V	is velocity (ms^{-1})
S	is slope
HR	is hydraulic radius (m)

2.3 SIMCAT

SIMCAT is a hybrid statistical/deterministic water quality model which calculates the quality of river water throughout a catchment. Warn (1995) highlights that, in theory, the equations of River Chemistry need to know the time taken for a body of water to travel between different points within the catchment and that velocity is a function of river flow. SIMCAT uses an empirical model to estimate velocity from flow, which has the form:

$$V = \alpha \cdot \left(\frac{Q_d}{\bar{Q}} \right)^{\beta}$$

where:

$\bar{Q}$	= average flow (m^3s^{-1})
α	= velocity associated with $\bar{Q}$ (ms^{-1})
Q_d	= river flow on day d (m^3s^{-1})
β	= constant which typically has a value of 0.5

2.4 CHÉZY

Chézy mathematically derived a formula and verified it by experiments made on an earthen canal and on the Seine River, and is given by:

$$V = C \sqrt{HR \ s}$$

where:

V	is velocity (ms^{-1})
HR	is hydraulic radius (m)
s	is slope
C	is Chézy's factor of flow resistance and is equal to:

$$C = \sqrt{\frac{(\rho_w \cdot g)}{k}}$$

2.5 MANNING

The Manning equation (1891) relates discharge to the channel slope, roughness and channel dimensions and is given by:

$$V = \frac{1.49 \ . \ HR^{2/3} \ . \ S^{1/2}}{n}$$

where:

V	= average cross-section velocity (ms^{-1})
HR	= hydraulic radius (m)
n	= Manning's roughness coefficient
S	= channel slope

Manning's formula was derived from 7 different formulas and verified by 170 observations, and due to its simplicity and satisfactory results it has become the most widely used uniform flow formula for open channel flow computations (Chow, 1959). The greatest difficulty lies in the determination of the roughness coefficient n. Guidance is given for the determination of n, based upon four general approaches:

1. understanding the factors affecting n;
2. consulting a table of n values for various channel types;
3. becoming acquainted with the appearance of typical channels with a known n;
4. determining n by an analytical procedure based on the theoretical velocity distribution in the channel cross section and on the velocity or roughness measurements.

However without a good knowledge of basic channel hydraulics, even with guidance, the selected value of n is likely to be a reflection of subjective judgement.

3 Data set construction

Two sources of data were available for use within this study. These were:

1. detailed field study data;

2. archive data.

The first data source consisted of 14 river sites with detailed channel and current metering measurements. 13 of the 14 sites have a catchment area of less than 200 km^2, thus, as GREAT-ER is concerned with larger catchments, further catchments with larger areas were required for use within the analysis.

The second source of available data were the archives, kept at the Institute of Hydrology, which hold information for every gauging station in the UK. The gauging stations selected for this study are on a relatively natural section of river and gauge catchment areas of more than 200 km^2.

The two sets of data were therefore collated to give a data set for a range of catchment areas from 3.5 km^2 to 6850 km^2.

For each site the data required included channel slope (S), BFI, catchment area (A), 3 discharges (MF, Q10, Q95), and for each of these discharges an associated velocity (V), width (W), hydraulic radius (HR) and depth (D). The MF, Q10 and Q95 flows were chosen to represent a wide range of discharges which would enable the derivation of a model from which velocity could be estimated over a range of flows.

The two sources of data are considered in detail in Sections 3.1 and 3.2.

3.1 FIELD STUDY DATA

As a part of the Physical Habitat Simulation (PHABSIM) study at the Institute of Hydrology detailed field studies have been undertaken at numerous sites, of which 14 were suitable for use within this analysis. At each of these sites current meter and channel geometry measurements have been taken for a series of transects along a river stretch for high, medium and low flow conditions. The river stretches used vary between 0 m and 5000 m in length, with an average length of 534 m. From this information the discharge, average velocity, average depth and width are available for each transect on each river stretch. At some of the sites the three flows were not significantly different and were therefore of limited value. In order to use this data in this study it was necessary to derive values of MF, Q10 and Q95 and associated values of V, HR, W and D. These discharges were derived by relating flows to downstream gauging stations, using:

$$Q_{site} = Q_{gs} \times \left(\frac{A_{site}}{A_{gs}} \right)$$

where:

Q_{site} is the discharge (MF, Q10 or Q95) at the PHABSIM site (m^{3}s^{-1})
Q_{gs} is the discharge (MF, Q10 or Q95) at the nearest gauging station (m^{3}s^{-1})
A_{site} is the catchment area at the PHABSIM site (km^2)
A_{gs} is the catchment area at the gauging station (km^2)

By fitting a regression to the log transformed flows and velocities measured at each PHABSIM site (and similarly for D, HR and W) power functions were derived for each site from which V, D, HR and W were calculated at the MF, Q10 and Q95 flows. Whilst this data set of 14 sites (Appendix 1) is the most detailed available, it is biased towards smaller catchments; 13 of the 14 sites having a catchment area of less than 200 km^2, with only the Thames site exceeding this at 3487km ^{2}. Smaller catchments, however, are more likely to be truly representative, if physical scale is used as a definition, as there are many more shorter rivers than longer ones. GREAT-ER, however, is concerned with a wide range of catchment sizes, thus further catchments with larger areas were required for use within the analysis.

3.2 ARCHIVE DATA

At the Institute of Hydrology information regarding flows, channel geometry and stage-discharge relationships is held for every gauging station in the UK. Gauging stations vary in their form and where there is a gauging structure, such as a weir or flume, the velocity is altered from that which would occur naturally. As the analysis requires velocities of natural river stretches alone, only VA[1], US[2] and EM[3] gauging stations were used in this study. In order to obtain a wider range of catchment areas only those gauging stations with catchment areas greater than 200 km^2 were used, yielding 97 stations in total.

The archives hold information on the mean, Q10 and Q95 flows, derived from long term flow records, for which there are corresponding stages (ST), from the stage-discharge relationship, and W, from scale drawings of the channel cross-section. The assumption of a rectangular channel was made and, based upon the scale drawings and measured stage, estimates of width and depth for the rectangular channel at each flow were made. From this information the V was calculated for each gauging station, using:

$$V = \frac{Q}{W \times ST}$$

Slope was calculated from 1:50 000 Ordnance Survey maps by finding the nearest upstream and downstream contours crossing the river and dividing the drop in height by the distance between the two contours. Using the Institute of Hydrology's Digital Terrain Model (DTM) S was also calculated over a 2 km and 10 km distance, centred on the gauging station, in order to see if any considerable differences existed. The slopes were found to be very similar for distances of 2 km and 10 km around the gauging stations and also for those calculated from the OS maps. As the slopes calculated directly from the OS maps are more likely to pick up any rapid changes in the gradient around the gauging station these slopes were used in the analysis.

3.3 DATA SET FOR ANALYSIS

The data set used in the analysis consists of 111 sites, Figure 3.1, for which S, BFI, A, and flows; MF,

[1] Velocity-area (VA) gauging station includes natural section, open channel, river section and rated section. The stage-discharge relation is obtained by measuring the stream velocity and cross-sectional area at points throughout the flow range at a site characterised by its ability to maintain the relationship.

[2] Ultrasonic gauging station (US) which computes flow data on-site by measuring times for acoustic pulse to traverse a river section along an oblique path in both directions. The mean river velocity is related to the difference in the two timings and the flow is assessed using the river's cross-sectional area.

[3] Electromagnetic gauging station (EM) requires the measurement of the electromotive force (emf) induced in flowing wave as it cuts a vertical magnetic field generated by a large coil beneath the river bed or constructed above it. The emf is directly proportional to the average velocity in the cross-section.

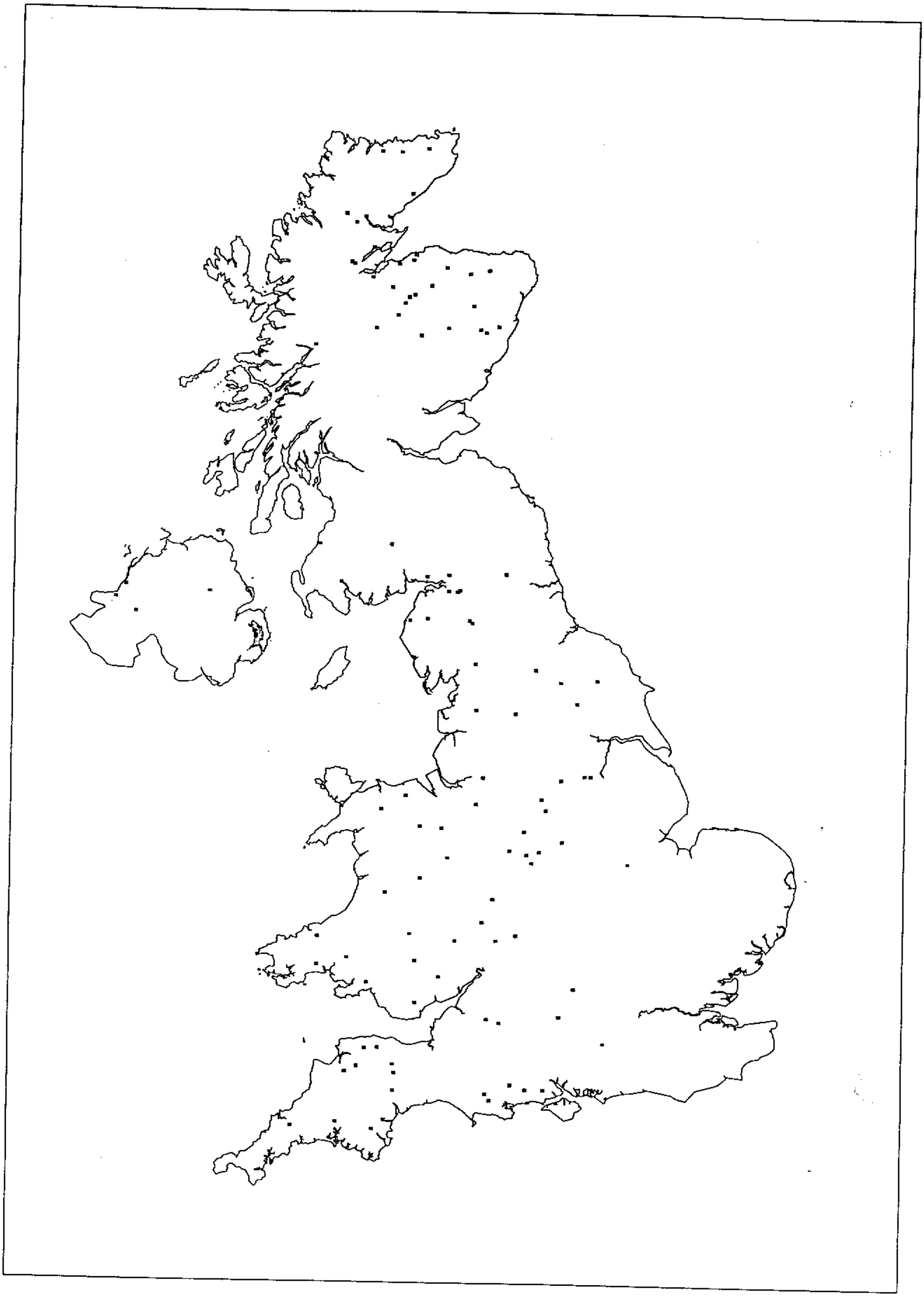


Figure 3.1 *Locations of sites used in the analysis*

Q10 and Q95, with corresponding values of W, D, HR, and V are available. The data set is summarised below in Table 3.1 and plotted as histograms in Appendix 2.

Table 3.1 *Summary of flow and channel geometry statistics*

	MF					Q10					Q95				
	Q	D	V	W	HR	Q	D	V	W	HR	Q	D	V	W	HR
max	87.1	3.8	2.51	184	2.07	213	4.2	3.22	96	3.16	19.2	3.4	1.88	352	1.92
min	0.11	0.12	0.04	1.62	0.10	0.28	0.17	0.16	1.86	0.15	.005	.098	.005	1.35	.086
med	13.7	0.71	0.51	33.5	0.68	31.6	1.10	0.82	34.6	1.06	1.62	0.33	0.21	31.5	0.32
CV	0.96	0.55	0.61	0.73	0.44	0.96	0.53	0.51	0.58	0.45	1.21	0.89	1.08	1.20	0.66

where:

Q is discharge (m^3s^{-1})
D is depth (m)
V is velocity (ms^{-1})
W is width (m)
HR is hydraulic radius (m)

From the coefficients of variations (CV) for each set of statistics it can be seen that the variation in flow is much greater than that in velocity and, contrary to Davisian theory, there is a general trend for flow velocity to increase with area and not vice versa.

3.4 REPRESENTATIVENESS OF THE UK DATA SET

An analysis was undertaken in order to determine the representativeness of the data set used in this study, when compared to Europe. European gauging station information was available from the FRIEND European Water Archive, from which data was available regarding MF and Q95 flows for 3401 and 2878 stations respectively. The countries included in the analysis were the UK, Ireland, Denmark, Germany, Austria, Italy, France, Spain, Netherlands, and Belgium.

MF/A and Q95/A were calculated to yield the unit runoff in mm yr^{-1} of the European catchments, the UK data set catchments, and all of the UK catchments. The median MF/A and Q95/A for each data set are given in Table 3.2. The frequency distributions are given in Appendix 3.

Table 3.2 *Median values for all data sets*

	Europe	UK study data set	UK
Q95/A	67.21	115.38	77.08
MF/A	420.10	713.79	483.45

From Table 3.2 it is evident that, from the comparison between the UK study data set and the whole of the UK, the study data have generally higher unit runoff and are thus biased towards the wetter catchments in the UK, as can be seen in Figure 3.1. This bias is also apparent when comparing the study data set to Europe.

4 Analysis

4.1 MULTIVARIATE REGRESSION ANALYSIS

In order to ascertain the variables which are important in determining V a multivariate regression analysis was undertaken, for which two types of regression were utilised (Freund & Littell, 1986). The first incorporates all of the independent variables in to the regression, which in doing so can force the R^2 value to 1 simply by adding superfluous variables to the model but with no real improvement to the fit. The second type, stepwise regression, begins by finding the variable that produces the optimum one variable subset (i.e. the variable with the largest R^2). In the second step the procedure finds the variable which, when added to the first variable, results in the largest increase in R^2 (i.e. the largest decrease in the residual sum of squares). After each addition of a variable the procedure examines the resulting equation to determine whether a backward elimination should be implemented. This procedure is continued until no further additions or deletions are indicated according to a defined significance level. In this analysis a significance level of 0.15 was used. The C(P) statistic (Mallows, 1995) was used to assess the merit of a selected subset. C(P) is a measure of the total squared error for a subset model containing p independent variables. The total squared error is a measure of the error variance plus the bias introduced by failing to include important variables into the model.

$$C(P) = (SSE(P)/MSE) - (N - 2P) + 1$$

where:

MSE	is the error mean square for the full model
SSE(P)	is the error sum of squares for the subset model containing p independent variables
N	is the total sample size

When $C(P)>(p+1)$ there is evidence of bias due to an incompletely specified model, and similarly if $C(P)<(p+1)$ the full model is overspecified and probably contains too many variables. Also, for any given number of selected variables, larger C(P) values indicate equations with larger error mean squares.

In this analysis it was necessary to identify those variables which are important for estimating V. This was achieved by applying a linear multivariate regression analysis to log transformed data. With reference to Appendix 4 it is evident that the log-log plots between V and the independent variables imply a linear relationship, which is in keeping with the reviewed literature. The conditions required for linear regression are that the variable must be normally distributed and the variable is not hetetoscedastic; in other words the variances should be the same for the whole population. Both of these conditions are implied by the plots in Appendix 4.

4.1.1 Parameter identification

Initially seven possible variables for entry in the multivariate regressions were considered. These were:

- Q discharge (m^3s^{-1})
- A catchment size (km^2)
- BFI base flow index, which is the ratio of the volume of water under the hydrograph derived from base flow to the total volume of water under the hydrograph, which is indicative of the geology and therefore the 'response' of the catchment (dimensionless)
- S the slope (dimensionless)
- HR hydraulic radius (m)

- Q/A runoff depth over the catchment and is indicative of the 'wetness' of the catchment (ms^{-1})
- Q/MF which reduces the influence of A and rainfall on Q, by standardizing by the MF at each site, and is therefore indicative, for a given Q, of the 'flashiness' of the catchment. It also enables differentiation between high and low flows at a given site (dimensionlesss)

It was considered probable that regression models for predicting V for a given flow, such as MF, Q10 or Q95, may be different to each other and also different to a model which predicts V for all flows. Thus, individual regressions were fitted to V and the independent variables for the four flow scenarios; MF, Q10 and Q95 and all flows, in order to identify whether the variables significant in determining V for each of the four flow scenarios were the same.

All of the analyses showed that BFI and S were insignificant in the estimation of V, and that HR was not significant in every model. The model fits, as indicated by the R^2 values, for each of the above scenarios suggested that using a single model for the estimation of V for any Q is a feasible approach. This is advantageous because it enables estimates of V to be made for any catchment and any flow percentile within a given catchment.

The regression which was fitted to the log of V and the log transformed variables, for all of the flows, yielded the parameter estimates given in Table 4.1.

Table 4.1 *Parameter estimates from the normal and stepwise regressions*

Variable	Parameter estimate	
	Normal	Stepwise
Q	580.86	573.26
Q/MF	0.1911	0.1820
A	-580.53	-572.94
Q/A	-580.42	-572.83
HR	-0.2839	-0.2760
S	-0.0312	
BFI	-0.0436	
Intercept	-0.6051	-0.5347

Thus, V could be estimated from:

$$V = -0.5347 \; . \; Q^{573.26} \; . \left(\frac{Q}{MF}\right)^{0.1820} . \; A^{-572.94} \; . \left(\frac{Q}{A}\right)^{-572.83} . \; HR^{-0.2760}$$

The R^2 and f.s.e.[4] values of 0.6059 and 1.77 respectively, derived from the normal regression, suggest that the model is appropriate for the estimation of V, because 60% of the variability in V can be

[4] The standard error (s.e.) is referred to as an f.s.e. because the model is linear (in terms of log(V), log(Q), log(A), log(Q/A), log(Q/MF), log(R) and log(S)), so the error is + or -, but when the antilog of the equation is taken the error becomes a multiplicative or divisive error. The advantage of using the f.s.e. is that the error scales with the estimated value of V.

explained by the variability in the independent variables and the error is within an acceptable level. However, the large negative parameter values for both Q and Q/A imply that these are unimportant in the model, because they will always tend towards 1. The analysis showed that S and BFI did not occur in the stepwise model, indicating that they did not meet the 0.15 significance level for entry into the models. The C(P) value of 5.581 is less than $p+1$ which indicates that the model is overspecified.

As the BFI was not included in the stepwise model it was omitted from further analysis. For this same reason S should also have been omitted. However, the literature review revealed several authors who considered S to be an important variable in the estimation of V, such as Manning whose equation is the most widely used uniform flow formula for open channel flow computations (Chow, 1959), thus S was retained within the subsequent analysis.

From this analysis it was concluded that the best model could be derived by incorporating a combination of the variables, Q, A, Q/A, Q/MF, HR and S, although Q and Q/A were believed to be unimportant in the estimation of V.

4.1.2 Model development

Before considering which variables would provide the best model the data was analysed to determine whether the independent variables were correlated with each other and also with the dependent variable V. The correlation coefficients and their probabilities are presented in Table 4.2.

Table 4.2 Correlation coefficients of the model variables

	V	Q	HR	Q/MF	Q/A	S	A
V	1 0.0	0.70223 0.0001	0.5147 0.0001	0.60591 0.0001	0.64010 0.0001	-0.15522 0.0047	0.30729 0.0001
Q		1 0.0	0.76325 0.0001	0.70373 0.0001	0.75598 0.0001	-0.34142 0.0001	0.62350 0.0001
HR			1 0.0	0.62745 0.0001	0.60560 0.0001	-0.30265 0.0001	0.44167 0.0001
Q/MF				1 0.0	0.86691 0.0001	-0.05417 0.3266	0.03881 0.4803
Q/A					1 0.0	0.03401 0.5281	-0.04042 0.4622
S						1 0.0	-0.56051 0.0001
A							1 0.0

where the correlation coefficient is:

	<|0.5|
	|0.5| to |0.6|
	|0.6| to |0.7|
	|0.7| to |0.8|
	>|0.8|

In a given row and column the upper number is the estimated correlation coefficient between the row variable and column variable, and the lower number is the significance probability for testing that the corresponding population correlation is zero. For example, the estimated correlation between V and S is -0.15522 which is significantly different from zero at the $p = 0.0047$ level. This correlation therefore indicates that there is a negative relationship between V and S. The correlations between the dependent variable, V, and the six independent variables are all significantly different from zero ($p<0.05$), suggesting that all of these factors could be useful by itself in estimating velocity. S however has the weakest correlation with V and, as indicated by the significance probability of 0.0047, is the least significant.

The highest of the correlations may indicate that it is not necessary to include both of the correlated variables, for example Q/MF and Q/A, in the same model. This can be explained by taking two variables, A and B, with a correlation coefficient of 1, indicating that all of the variability in A can be explained by all of the variability in B and vice versa. If A was a function of B, then A would be rejected and vice versa. If A and B were both a function of a third variable, C, then either A or B should be rejected.

As the correlations between V and the independent variables are all significant a model was proposed which included all six variables:

V = fn{Q, A, Q/MF, Q/A, HR, S}

By including highly correlated variables within one model it was possible to identify from the parameter estimates which of the variables should be included in the model. The model performance is summarised in Table 4.3.

Table 4.3 *Regression summary*

Model	Regression	R^2	Variables included	m.s.e.	f.s.e.	C(P)
V = fn{Q, A, Q/MF,	Normal	0.6049	All	0.249	1.77	
Q/A, HR, S}	Stepwise	0.6040	Q, A, Q/A, HR, Q/MF			5.725

In terms of R^2 this model is not affected by the removal of BFI, and the C(P) of 5.725 value indicates that the model is less overspecified, although its continued overspecification is perhaps a consequence of the inclusion of A and Q/A, which again have very small parameter estimates (Table 4.4).

Table 4.4 *Parameter estimates*

Variable	Parameter estimate	
	Normal	Stepwise
Q	581.53	572.14
Q/MF	0.1796	0.1821
A	-581.20	-71.80
Q/A	-581.08	-571.69
HR	-0.2844	-0.2760
S	-0.0308	

The regression yielded an f.s.e. of 1.77, therefore it can be assumed that with 68% confidence:

$$0.56 \times V_{est} < V_{est} < V_{est} \times 1.77$$

giving a factor of 3.2 difference between the upper and lower confidence limits.

Based upon the above findings Q/A and A were omitted from the subsequent regression analyses. Three models were analysed, based upon removing, one by one, the least important factor in estimating V. These models are:

V = fn{Q, Q/MF, HR}

V = fn{Q, Q/MF}

V = fn{Q}

The regression analysis results for each model are summarised in Table 4.5.

Table 4.5 *Regression summary*

Model	Regression	R^2	Variables included	m.s.e.	f.s.e.	C(P)
V = fn{Q, Q/MF, HR}	Normal	0.5228	All	0.2717	1.869	
	Stepwise	0.5228	All			4
V = fn{Q, Q/MF}	Normal	0.5179	All	0.273	1.875	
	Stepwise	0.5179	All			3
V = fn{Q}	Normal	0.4931	All	0.2792	1.902	
	Stepwise	0.4931	All			2

It can be seen that none of the models are over or under specified as C(P) = p+1 and thus, based upon the highest R^2 value and the lowest f.s.e., the first model is the best. The f.s.e. of 1.869 indicates that from the model it can be assumed that with 68% confidence:

$$0.535 \times V_{est} < V_{est} < V_{est} \times 1.869$$

which gives a factor of 3.49 difference between the upper and lower confidence limits.

However, if this model were to be used, an estimate of HR would have to be made. HR is calculated using:

$$HR = \frac{CSA}{WP}$$

where:

CSA is the channel cross-sectional area, which is the product of W and D
WP is the wetted perimeter, which for a rectangular channel is the product of 2D and W

The estimate of HR would thus require an estimate of W and D, for which an existing method is not available, and its dependency upon local channel geometry would render one model for estimating HR inapplicable; not many channels are a perfect rectangle. Thus, as HR is unknown and not readily available, propagation of the error in any method used to estimate HR would cause the velocity model to be less reliable than the second model.

It was therefore concluded that V = fn{Q, Q/MF} provides the best model for the estimation of V. This regression yielded an f.s.e. of 1.875, thus from the model it can be assumed that with 68% confidence:

$$0.533 \times V_{est} < V_{est} < V_{est} \times 1.875$$

which gives a factor of 3.52 difference between the lower and higher confidence limits, which is similar to the model with HR, suggesting that HR is unimportant in the estimation of V in the data set used.

The equation derived from the regression is:

$$V = 10^{-0.599}\, Q^{0.286} \left(\frac{Q}{MF} \right)^{0.165}$$

which is of the same form as that used in the SIMCAT model:

$$V = \alpha \left(\frac{Q}{\bar{Q}} \right)^{\beta}$$

The average velocity, α, is replaced by $10^{-0.599}\, Q^{0.286}$ which improves the model by giving an α value which scales with discharge. Thus, the mean V increases with Q, and at a given site V increases with $Q/\bar{Q}$.

4.2 ANALYSIS OF THE RESPONSE OF MANNING'S EQUATION

Due to its universal application for velocity estimation, an analysis was undertaken of the Manning's equation in order to assess, by estimating 68% confidence intervals, whether the use of Manning's equation for the regional estimation of velocity is feasible, and if the approach is more accurate than that derived using the multivariate regression analysis.

Manning's empirical formula gives velocity as a function of S, HR and n:

$$V = \frac{1.49\, S^{1/2}\, HR^{2/3}}{n}$$

4.2.1 Analysis of Manning based upon back calculation of n

In the above equation the only unknown is n. Assuming that the measurements of V, S and HR are accurate it may be implied that n can be calculated with no associated error. The roughness coefficients were therefore back calculated and examined in order to determine whether they fell between the maximum and minimum values of n possible for the rivers in the data set. From Chow (1959) the maximum and minimum n values were found for excavated, dredged or natural streams, which are assumed to reflect the channels within the data set. The highest values of n are for flood plains of natural channels, although these were disregarded because none of the Q10 flows exceeded bankful level. Thus, the maximum and minimum n values are 0.15 and 0.016 respectively.

The n values were plotted in log-log space against the measured velocity, and the maximum and minimum n values were plotted on the same graph in order to determine whether the calculated values of n fell within the correct range (Fig. 4.1). From the plot it can be seen that 70% of the n values calculated from the Manning's equation fall within the range.

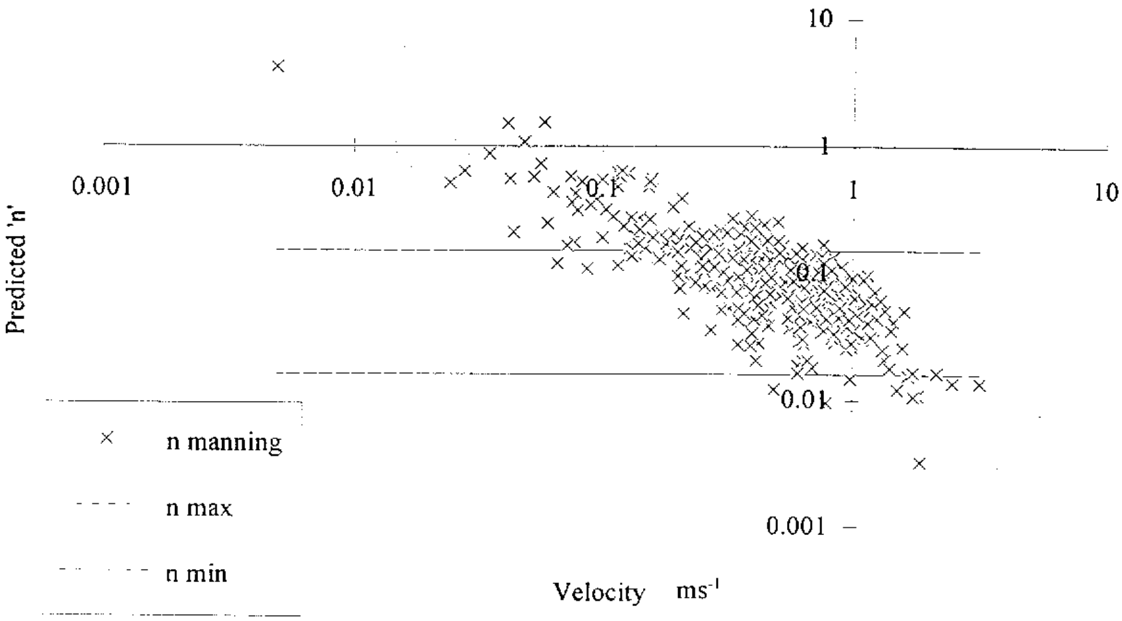


Figure 4.1 *Manning's n in relation to maximum and minimum n values*

The n values derived from the Manning's equation were ranked and the upper and lower 68 percentiles were found to be 0.2027 and 0.0390 respectively. It can therefore be implied that with 68% confidence:

$$0.0390 < n_{est} < 0.2027$$

which represents a factor of 5.2. Uncertainty in V would also be a factor of 5.2 for V if these upper and lower values of n were substituted into Manning's equation, which is explained in the following example.

Example

$S = 0.001$ $n_{lower} = 0.0390$
$HR = 2.2$ $n_{upper} = 0.2027$

Substituting S, R and n_{lower} into Manning's equation gives the upper 68% confidence limit for V of 2.23 ms^{-1}; and, substituting S, R and n_{upper} into Manning's equation gives the lower 68% confidence limit for V of 0.428 ms^{-1}.

This gives a factor difference of 5.2 between the upper and lower confidence limits for V.

4.2.2 Analysis of Manning based upon estimates of HR and n

A further investigation was undertaken in order to determine whether n and HR can satisfactorily be estimated from catchment characteristic and/or channel geometry data, and, if so, to determine the 68% confidence interval for an estimate of V.

A regression analysis was performed between the log of n and the log transformed independent variables, where:

n = fn{Q, HR, S, BFI, A, Q/MF, Q/A}

The R^2 and f.s.e. values of 0.646 and 1.77 respectively suggest that an estimate of n can be made from catchment and channel characteristics. However, the stepwise regression gave S, HR, Q, Q/MF and BFI as significant in the estimation of n and the C(P) value of 5.7 indicated that the model was overspecified.

As BFI was the least significant of the variables used in the previous stepwise regression it was omitted along with A and Q/A, thus a second model was proposed, where:

n = fn{Q, HR, S, Q/MF}

This gave the regression equation:

$$n = 10^{0.9116} \cdot Q^{-0.377} \cdot HR^{0.963} \cdot S^{0.499} \cdot \left(\frac{Q}{MF}\right)^{-0.262}$$

$$R^2 = 0.6291 \qquad \text{f.s.e.} = 1.607$$

The R^2 value dropped slightly, but the f.s.e. improved considerably to 1.607, giving a factor of 2.58, as opposed to 3.13, difference between the 68% upper and lower limits of the n estimate. Ideally, the model would exclude HR because it is not a readily available variable, thus its estimation would increase the model error. However, removing HR decreased the R^2 to 0.4898 and the factor difference between the 68% upper and lower limits increased to 4.18.

Similarly, the same analysis was undertaken for HR. Initially, a regression based on:

HR = fn{Q, A, S, Q/MF, Q/A, BFI}

was undertaken to identify which variables are important in the estimation of HR. This yielded R^2 and f.s.e. values of 0.6436 and 1.53 respectively, although the C(P) value indicated that the model was overspecified and that Q/A was not significant. By removing the least significant variable in the model, one by one, the best R^2 and f.s.e. values were derived for the model:

HR = fn{Q, Q/MF, S}

giving the regression equation:

$$HR = 10^{-0.4529} \cdot Q^{0.270} \cdot \left(\frac{Q}{MF}\right)^{0.133}$$

$$R^2 = 0.6416 \qquad \text{f.s.e.} = 1.53$$

To test the performance of Manning's equation, V was estimated using values of n and HR calculated from the above regression models, in conjunction with measured values of S. In order to include the error of the estimate of HR into the estimate of n, the estimated values of HR were substituted into the regression equation, as opposed to the measured values of HR.

The resultant estimates of V were ranked and the upper and lower 68% confidence limits were found to be 0.8394 and 0.1867, representing a factor of 4.50 difference.

5 Model uncertainty

5.1 PRECISION OF FLOW AND VELOCITY INFORMATION

As discussed in Chapter 3 a data set for 111 catchments was compiled for analysis for which data on the following are held for each catchment:

- Slope
- catchment area
- Q95, Q10 and mean flows
- V95, V10 and mean velocities
- Channel geometry at each flow

These data were derived from station details for 97 velocity-area gauging stations within the UK and 14 case study sites for the application of models for estimating physical habitat availability. The variability in flow and velocity for the data set were summarised in Table 3.1.

The distinction should be made between particle velocity and kinematic wave velocity. At mean flow and Q95 flow the measured velocity is going to be a close approximation to particle velocity. From the coefficients of variations (CV) for each set of statistics it can be seen that the variation in flow is much greater than that in velocity and, contrary to Davisian theory, there is a general trend for flow velocity to increase with area and not vice versa.

At the last UK sub group meeting concerns were raised regarding the relative impact of uncertainties in flow estimation (and hence dilution) and velocity estimation on the response of first order degradation equations. The response of the first order degradation equation:

$$C = C_0 \ e^{-kt}$$

has been investigated with the coefficient k set at 0.4 d^{-1}. This value is the middle value from the range quoted in the recent document circulated by ECETOC. The uncertainty in the initial concentration downstream of a discharge will be a function of uncertainties in the estimation of river flows/pollutant concentrations and/or effluent flows/concentrations. Assuming the pollutant load of a river upstream of a discharge point to be negligible, then, at t=0, the initial in stream concentration downstream of the discharge is given by the product of the dilution ratio and the pollutant concentration in the discharge.

The uncertainty in C_o is inversely proportional to any uncertainties in the measurement of river flow. The impact of uncertainties in river flow estimation on downstream concentrations were investigated for the three catchments in Table 3.1 at the mean and Q95 flow conditions. These uncertainties were based on the standard uncertainties associated with estimates of MF and Q95 from the hydrological model, furthermore it was assumed that the downstream flow accretion profile was zero and velocity remained constant. Figure 5.1 shows a example plot for the median catchment, the River Idle at Mattersley Bridge, at the Q95 flow. This demonstrates that although the relative uncertainty in downstream concentrations remains unchanged the absolute uncertainty decreases with distance, as would be expected.

Considering the upper confidence limit, it is at a distance of some 30 km downstream that this concentration decays to one, the mean initial concentration. If the river is flowing at a velocity, V, the time of travel for 30 km represents the uncertainty in time of travel that would be equivalent to the uncertainty in C_o, represented by the upper confidence limit for C_o. This semi-quantitative inspection of the relative sensitivity of the first order decay equation to uncertainties in C_o and t (and consequently V) are explored in the following paragraphs.

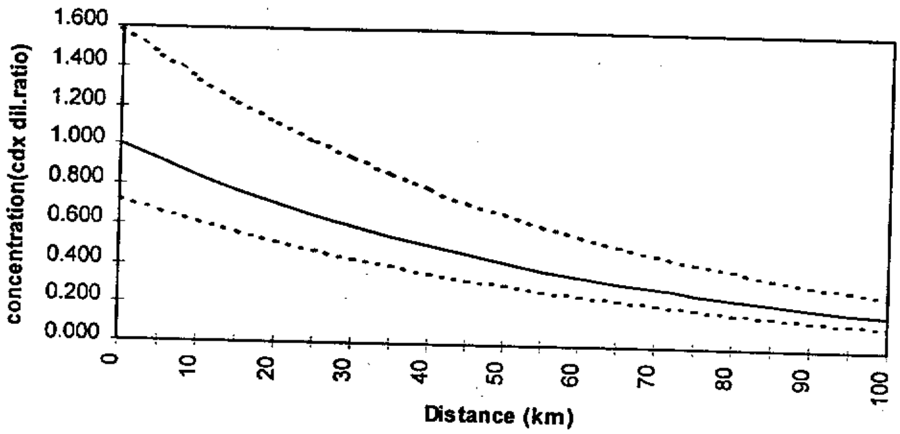


Figure 5.1 *Dilution uncertainty for the River Idle at Q95*

A discharged pollutant, travelling at constant velocity, V, will travel a distance, X, downstream in time, t. The uncertainty in the predicted downstream concentrations at X will be a function of the uncertainty in estimating the initial concentration, C_o. If the uncertainty in C_o is ΔC_o then the question may be posed; how long would it take for in stream concentrations to decay from $(C_o + \Delta C_o)$ to C_o? This time, Δt, represents the additional degradation time required if the predicted pollutant concentration at X is to be correct. If the uncertainty were then transferred from C_o to t the resultant effective time of travel at distance X downstream can be calculated by re-arranging the first order degradation equation which gives:

$$t = \frac{\ln C - \ln C_o}{-k}$$

The effective time of travel is then given by:

$$t = \frac{x}{v}$$

The equivalent time of travel is then given by:

$$\Delta t = \pm \left(\frac{\ln C_o - \ln (C_o + \Delta C_o)}{-k} \right)$$

and hence:

$$t \pm \Delta t = \frac{x}{v} \pm \left(\frac{\ln C_o - \ln (C_o + \Delta C_o)}{-k} \right)$$

for $t > \Delta t$, and thus an effective velocity over distance X can be calculated from:

$$V \pm \Delta V = X \cdot \left(\frac{X}{V} \pm \frac{\ln C_o - \ln (C_o + \Delta C_o)}{-k} \right)^{-1}$$

The logarithmic terms have not been transformed so as to retain the clarity of their origins. Figures 5.2 a, b and c demonstrate the relationships obtained by taking the upper confidence interval for an uncertainty in initial concentration and plotting the proportional uncertainty in velocity, $(V - \Delta V)/V$, as a function of the proportional uncertainty in initial concentration, $(C_o + \Delta C_o)/C_o$, for a range of downstream distances and mean velocities of 2, 0.2 and 0.02 ms^{-1}. Conditions of constant downstream velocity and zero flow accretion profiles were assumed for this analysis.

From the plots in Figure 5.2 it can be seen that, generally, the response of the first order degradation equation is more sensitive to uncertainty in the estimation of initial concentrations than the equivalent uncertainty in velocity. The relative importance is controlled by the following factors:

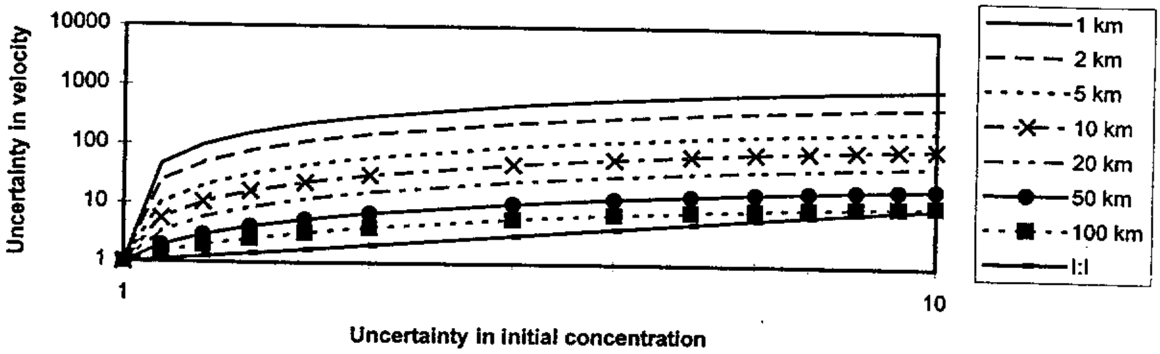
1. magnitude of mean V - at higher velocities the response of the equation is much more sensitive to uncertainties in C_o than those in velocity;

2. magnitude of the uncertainty in C_o - the response of the equation is much more sensitive to uncertainties in C_o than those in velocity when the uncertainty in C is small (however, the overall magnitude of the uncertainty in predicted concentrations is obviously directly proportional to the magnitude of the uncertainty in initial concentration);

3. distance downstream - the uncertainty in C_o is more important than the equivalent uncertainty in velocity at shorter distances. At longer distances, when the time of travel is considerably longer, the relative importance of uncertainties in velocity, and hence time of travel, increase.

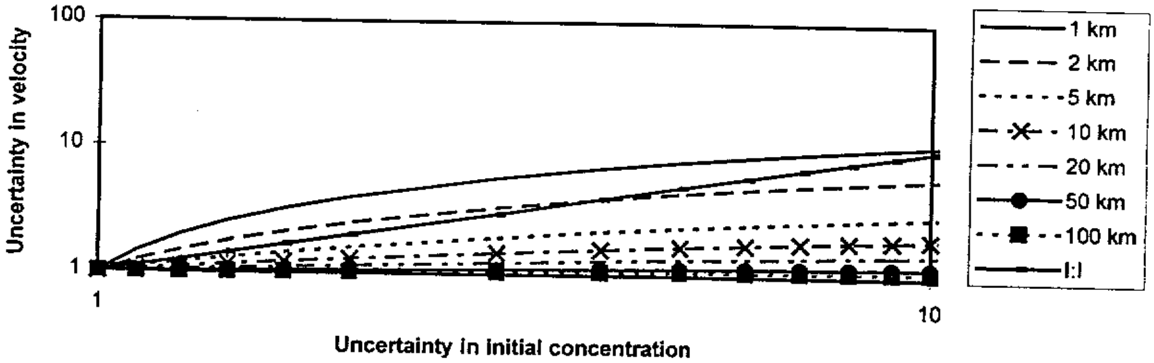
In reality flow accretion profiles are not zero and velocity profiles are not constant. Within the context of GREAT-ER is has been suggested that rivers will be split into n 2 km reaches in series and with each reach being modelled as a completely stirred tank reactor. It is more appropriate to consider the relative importance of uncertainties in C_o to those in V at this scale. Figure 5.3 presents the uncertainty relationship for different velocities at a distance of 2 km. From this plot it is evident that the uncertainty in C_o is dominant for velocities greater than 0.05 ms^{-1} which corresponds to 4.3 km d^{-1}. This velocity is equivalent to the lowest velocity in the UK data set.

V = 2 ms^{-1}; k = 0.4 d^{-1}



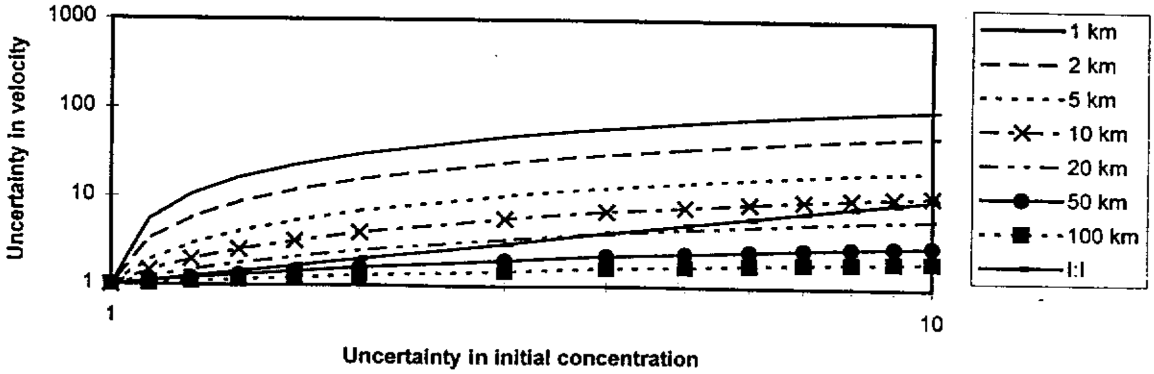
(a)

V = 0.2 ms^{-1}; k = 0.4 d^{-1}



(b)

V = 0.02 ms^{-1}; k = 0.4 d^{-1}



(c)

Figure 5.2 *Relationships between uncertainties in V and C_o as a function of distance and velocity*

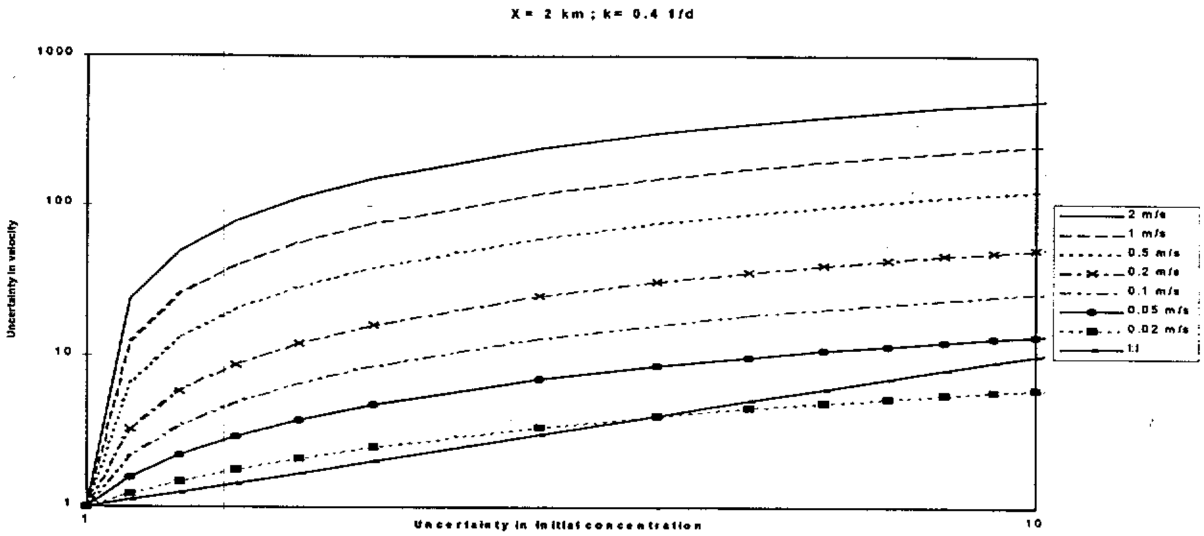


Figure 5.3 *Relationships between uncertainties in V and C_o as a function of velocity*

5.2 IMPLICATIONS FOR GREAT-ER

This analysis has concentrated on assessing the relative importance of uncertainties in C_o and V as a function of V and distance. The analysis is also equally as valid for considering the relative importance of uncertainties in C_o and the product kt. Considering the influence of velocity on k; as velocity increases turbulence, and hence re-aeration due to mixing, the exposure rate of the benthos to the pollutant will also increase. These factors are likely to increase bio-degradation and volatilisation k factors. The time of travel t will decrease with increasing V and thus, the product kt is subject to a negative feedback loop, where an increase in t will lead to a decrease in k, and vice versa. It is suggested that the product kt is likely to be more invariant than might initially be thought, although the major uncertainties in predicting k should always be considered.

The analysis indicates that the key to estimating aquatic concentrations of the target chemical may lie in the correct estimation of pollutant mass loads to rivers and the consequent dilution of those mass loads and the correct estimation of the degradation, whilst obviously important, may be a secondary influence. With the natural flow estimation procedures in the UK the 68% confidence interval for mean flow are approximately 20%MF, whilst those for Q95 (expressed as a percentage of mean flow) are 7.4%MF. However, in artificially influenced catchments the uncertainties may be in excess of one order of magnitude. It is therefore proposed that the correct estimation of dilution is the dominant factor in the determination of environmental concentrations and that the correct estimation of river flows is one of the key factors in determining dilution ratios and thus the IH work package is correctly positioned within GREAT-ER. Obviously the other key components of ensuring discharged volumes and associated pollutant masses are correct are also as equally important in determining C_o and should not be overlooked in the project.

6 Conclusions

The analysis aimed to provide a model, from which velocity can be estimated on a European scale using readily available data and without a requirement for detailed knowledge of catchment fluvial geomorphology. Readily available variables in this study were Q, A, S, BFI, W, D, and HR.

Multivariate regression analysis on log transformed was undertaken in order to identify which of the variables, from Q, Q/MF, A, Q/A, HR, S and BFI, were important in the estimation of V. This was performed for four flow scenarios; MF, Q10, Q95 and all flows. All of the analyses showed that BFI and S were insignificant in the estimation of V and that HR was not significant in every model. It was concluded that using a single model for estimating velocity at any flow was feasible and that the model fit actually improved. The parameter estimates showed that Q and Q/A were unimportant in the model when compared to the other variables.

Although S was found to be insignificant in the estimation of V it was left in the subsequent model due to its importance in the Manning's equation. Based upon correlations, between V and the independent variables, all were found to be significantly different from zero, and thus a model was proposed which included Q, Q/MF, A, Q/A, HR and S.

Using the six variables Q, Q/A and S were found to be insignificant in the estimation of V and as a consequence the model was overspecified. Removing these three variables resulted in a model based upon Q, Q/MF and HR, which gave an f.s.e. of 1.869. However, this model was rejected on the basis that the error associated with making an estimate of HR would be greater than the increase in model error if HR was excluded. Removing HR from the model gave an f.s.e. of 1.875 and a completely specified model. The factor difference between the 68% upper and lower confidence limits of the velocity estimate increased slightly from 3.49 to 3.52.

From the multivariate regression analysis of all of the data the best model is therefore V as a function of Q and $Q/\bar{Q}$, which takes the form:

$$V = 10^{-0.599} \cdot Q^{0.286} \cdot \left(\frac{Q}{MF}\right)^{0.165}$$

The variables within this model are readily available and therefore makes its application easy for the user. Also, it can be used for any flow in any physiographic region or catchment and therefore meets the aim of deriving a model with regional applicability on a European scale. The model is of the same form as that used within the SIMCAT water quality model, except the value for the average time of travel at MF, α, is replaced by $10^{-0.599}\,Q^{0.286}$, which enables α to scale with discharge.

This model implies that mean velocity increases with discharge. Whilst the mean velocity may increase downstream due to increasing catchment area and discharge, the velocity distributions at upstream and downstream river sites will show that there is more variability in the velocity at the upstream site at a given discharge than there is at a downstream site. This is due to differing primary controls over velocity. The channel morphology, which changes rapidly in headwater, controls velocity in these stretches, whereas the more stable nature of the downstream channel morphology results in velocity being controlled by hydrological processes.

In contrast to the above regression analysis, in which HR and S were relatively unimportant in the estimation of V, Manning empirically related V to HR, S and n. An analysis was undertaken to determine whether Manning is a feasible approach for estimating V based upon two approaches.

Firstly, the back calculation of n using measured values of V, R and S enabled a comparison to be made with expected values of n for the channel types in the study, thus indicating whether Manning's is valid. It was found that 70% of the values of n fell within the expected range and the upper and lower 68 percentile values of n, derived by ranking these estimates of n, indicated a factor of 5.2 difference which would also be reflected in the estimates of V.

Secondly, the calculation of V using measured values of S and values of HR and n derived from regression models was undertaken to determine whether Manning's equation provides better estimates than the multivariate regression equation. The analysis indicated that HR can be estimated from Q and Q/MF, with a factor difference of 2.34 between the upper and lower confidence limits, and that n can be estimated from Q, Q/MF, HR and S, with a factor difference of 2.58 between the upper and lower confidence limits. Substituting the regression equations for HR and n into Manning's equation gave the upper and lower 68% confidence limits of V which represented a factor of 4.5 difference.

Although it has been shown that HR can be estimated from readily available data, it is not recommended that this regression equation is used for such a purpose. This is due to the gauging stations used in this study being of the VA type, the locations of which are chosen for their natural and uniform properties. The channels of many rivers have been modified over time for flood alleviation or for navigation purposes, where water levels are controlled by weirs. These anthropogenic influences are not generally mapped and therefore are not reflected in the catchment characteristics used in this study.

In conclusion it is recommended that, based on the best f.s.e. and C(P) values and the limitations of estimating HR, the regression equation, relating V to Q and Q/MF, should be used for estimating stream velocity.

The analysis of the relative importance of uncertainties in C_o and V, as a function of V and distance indicates that it is fundamental to correctly estimate pollutant mass loads to rivers in order to estimate aquatic concentrations of the target chemical. The resulting dilution of the mass loads and the estimation of the degradation may be a secondary influence in all but the longest European rivers. The distribution of catchment stream length in Europe is a continuous distribution bounded by zero at one extreme and the total stream length of the Danube at the other. The correct estimation of dilution is the dominant factor in determining environmental concentrations and the accurate estimation of river flows is a key factor in determining dilution ratios.

The analysis indicates that the minimisation of uncertainty should concentrate on the correct characterisation of river flow, discharge volumes and discharge mass loadings. The impacts of uncertainties in k and V at a reach scale are secondary to equivalent uncertainties in these input data. The uncertainty in k may in many instances be much larger than those in C_o. Under these circumstances and in long rivers the uncertainty in k may become dominant.

References

Betson, R. 1979. A geomorphic model for use in streamflow routing. *Water Resources Research*, 15(1), 95-101.

Charlton, F., Brown, P., Benson, R. 1978. *The hydraulic geometry of some gravel rivers in Britain.* Report INT-180, Wallingford: Hydraulics Research Station.

Chézy, A. de 1775. Memoire sur la vitesse de l'eau conduite dans une régole. MS, reprinted in *Annals des Ponts et Chaussées*, 60, 1921.

Chow, V.1953. *Open channel hydraulics*. McGraw-Hill.

Freund, R.J., Littell, R.C. 1986. *SAS System for Regression.* SAS Institute Inc., US.

Gregory, K., Walling, D. 1974. *Drainage basin form and processes: a geomorphological approach.* John Wiley, New York.

Leopold, L., Maddock, T. 1953. The hydraulic geometry of stream channels and some physiographic implication. *US Geological Survey Professional Paper 252*, Washington, D.C.

Mallows, C. L. 1995. More comments on C_p. *Technometrics*, 37(4), 362-372.

Manning, R. 1891. On the flow of water in open channels and pipes. *Transactions, Institution of Civil Engineers of Ireland*, 20, 161-207.

Nixon, M. 1959. A study of the bankful discharges of rivers in England and Wales. *Proceedings of the Institution of Civil Engineers*, 12, 157-175.

Park, C. 1976. The relationship of slope and stream channel form in the River Dart, Devon. *Journal of Hydrology*, 29, 139-147.

Post D.A. & Jakeman A.J. 1996. Relationships between catchment attributes and hydrological response characteristics in small Australian Mountain Ash catchments. *Hydrol. Proc.* 10, 877-892

Simons, D.T., Albertson, N.L. 1960. Uniform water conveyance channel in alluvial materials. *ASCE*, 86 (HY5).

Singh V.P. 1995. *Watershed Modelling*. Chapter 1, Computer Models of Watershed Hydrology. Editor V.J. Singh, Water Resources Publications, Colorado, USA.

Warn, T. 1995. *SIMCAT V4.13 - A guide for users*. National Rivers Authority.

Appendix 1

PHABSIM sites

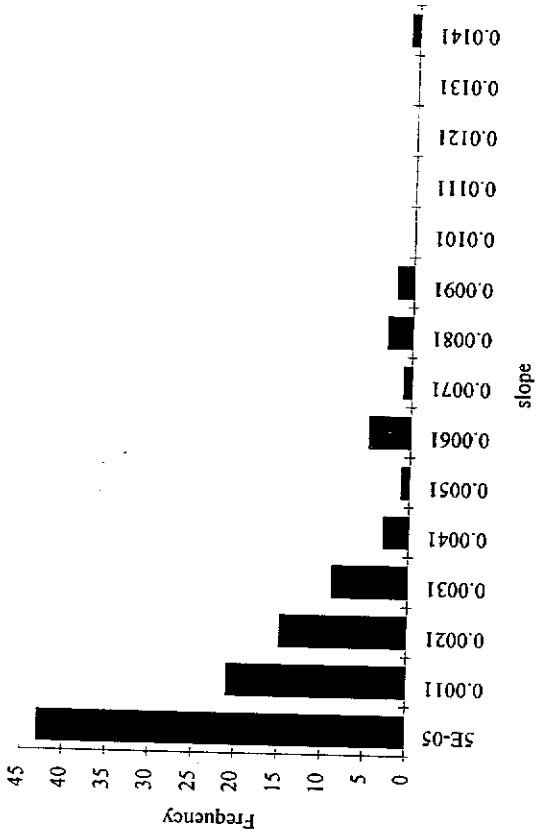
River	PHABSIM Grid Ref.	Gauging station Grid Ref.	Gauging station No.	$Area_{gs}$	$Area_{site}$	$\bar{Q}_{gs}$	$\bar{Q}_{site}$	Slope m/km	BFI
Exe (SW) Warren Farm	SS791 407	Pixton SS935 260	45009 D/S	147.5 (147.6)	3.5	4.562 (4.478)	0.160	8.3	0.5
Wye (Welsh) Pant Mawr	SN847 823	Pant Mawr SN843 825	55010 U/S	27.75 (27.2)	27.75	1.685 (1.649)	1.665	45.3	0.31
Hodder (NW) Hodder Bank	SD655 487	Stocks Res. SD719 546	71002 U/S	37 (37)	180.25	1.34	6.601	14.95	0.31
Lymington (S) Balmerlawn	SU302 033	Brokenhurst Park SU318 019	42003 D/S	103 (98.9)	77.75	1.173 (0.983)	0.910	4.59	0.36
Lambourn (Th) Hunts Green	SU 435 701	Welford SU470 682	39019 D/S	237.25 (234.1)	183.25	1.875 (1.683)	1.487	2.59	0.98
Gwash (Ang) Belmesthorpe	TF 041105	Belmesthorpe TF 038 097	31006 D/S	154.5 (150)	150.25	0.889 (0.849)	0.872	3.04	0.75
Barle (SW) Perry Weir	SS 931 256	Brushford SS 927 258	45011 U/S	128.25 (128)	131.75	4.984 (4.413)	5.065	6.8	(< 0.5)
Bray (SW) Leehamford	SS 678 399	Leehamford Bridge SS 677 399	50821 (No data)	18.5 (17.6)	18.5	0.726 (0.67)	0.726	8.39	(< 0.5)
Piddle (WES) Briantspuddle	SY 823 933	Baggs Mill SY 913 876	44002 D/S	182.75 (183.1)	115.5	2.833 (2.33)	1.93	3.3	0.89
Devils Brook (WES)	SY 779 990	Dewlish T 778 000	44702 (No data)	21.25	21.25	0.133 (.104)	0.133	5.1	(0.89)
Allen U/S (WES)	SU 007 080	Loverley Mill SU 006 085	43010 U/S	101.25 (94)	101.25	1.404 (0.967)	1.404	2.3	0.89
Allen D/S (WES)	SU 003 075	Loverley Mill SU 006 085	43010 U/S	101.25 (94)	101.25	1.404 (0.967)	1.404	3.45	0.89

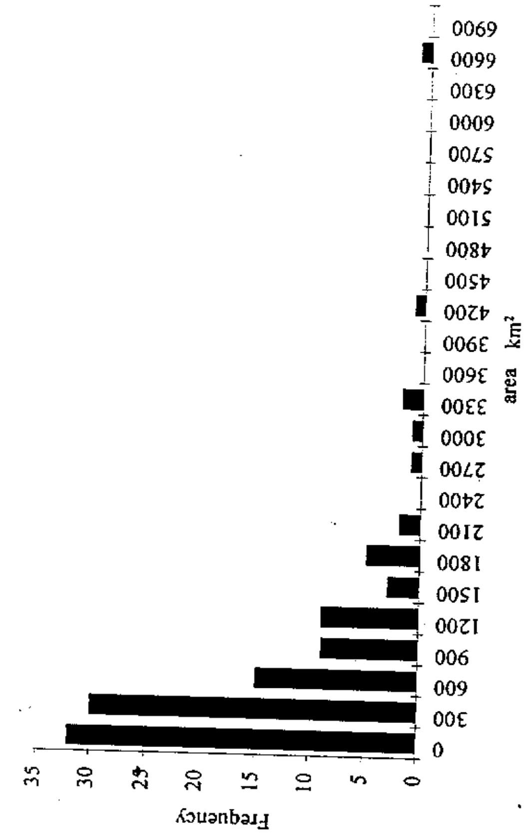
River	PHABSIM Grid Ref.	Gauging station Grid Ref.	Gauging station No.	$Area_{gs}$	$Area_{site}$	$\bar{Q}_{gs}$	$\bar{Q}_{site}$	Slope m/km	BFI
Thames (TH)	Clifton Hampton SU 565 955	Days Weir SU 568 935	39002 D/S	3489.25 (3444.7)	3486.5	26.746 (28.03)	26.735	0.37	0.65
Wey (TH)	Farnham SU 838 463		39078 U/S	194 (191.1)	196	2.402 (0.734)	2.420	10.0	0.72

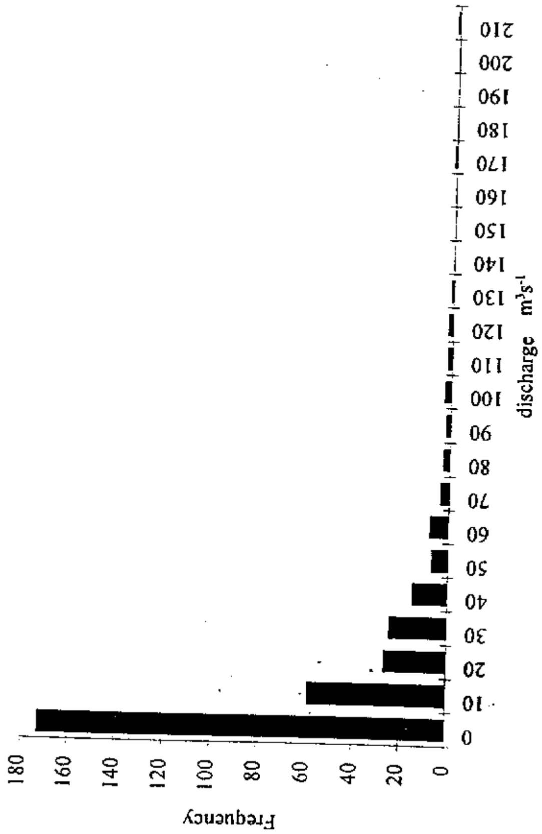
Figures in ( ) are the actual measured areas and mean flow, and estimated BFI

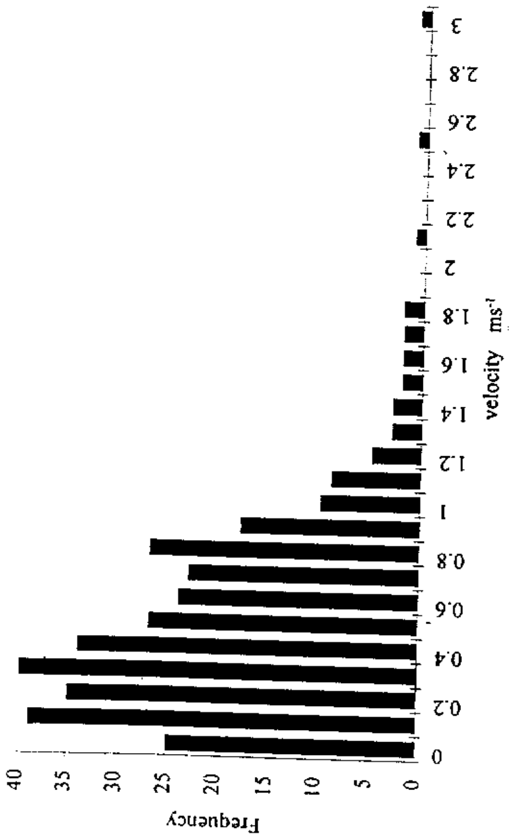
Appendix 2

Frequency distributions of Q, V, S and A


Frequency
slope

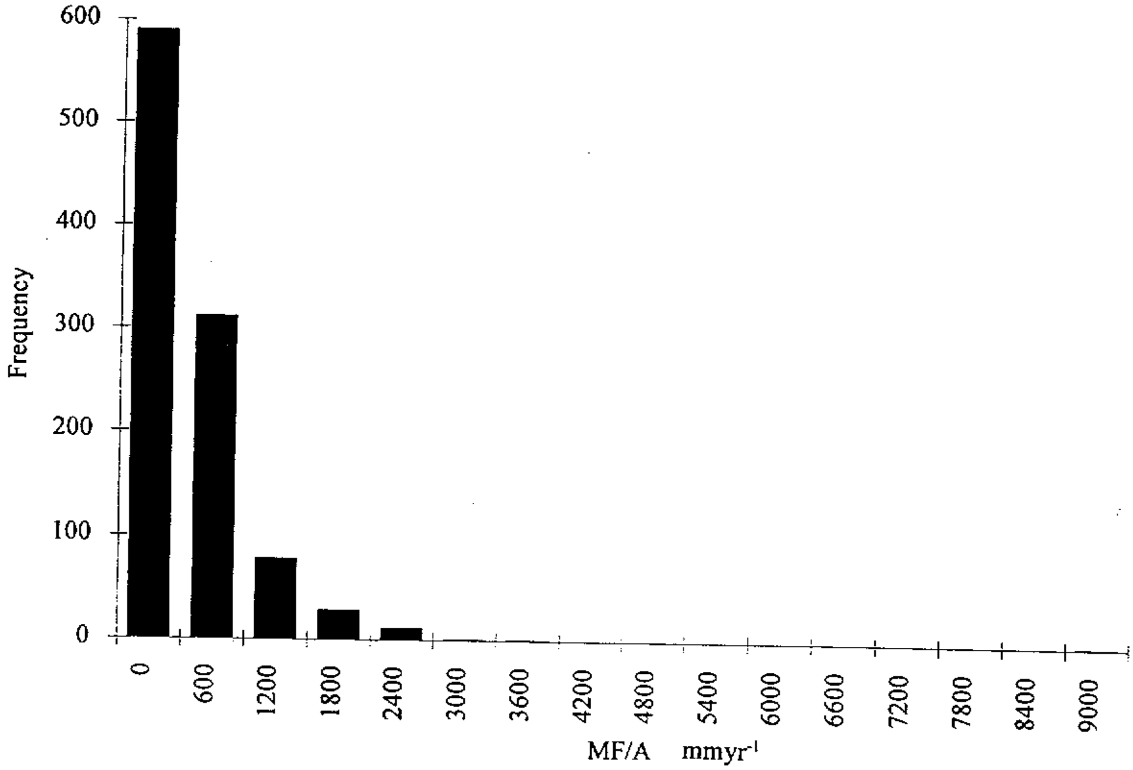

Frequency
area km^2


Frequency
discharge m^{3}s^{-1}


Frequency
velocity ms^{-1}

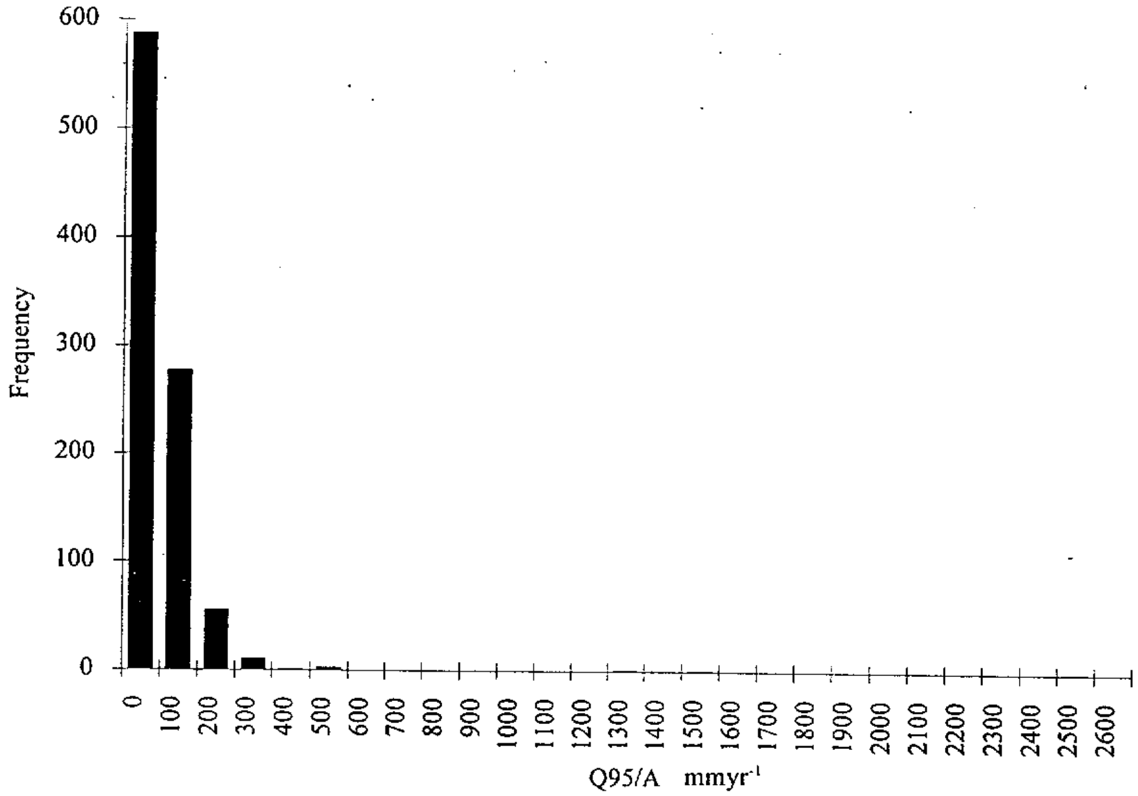
Appendix 3

Frrequency distributions of MF/A and Q95/A for:
(a) UK gauging stations
(b) European gauging stations
(c) Study data set

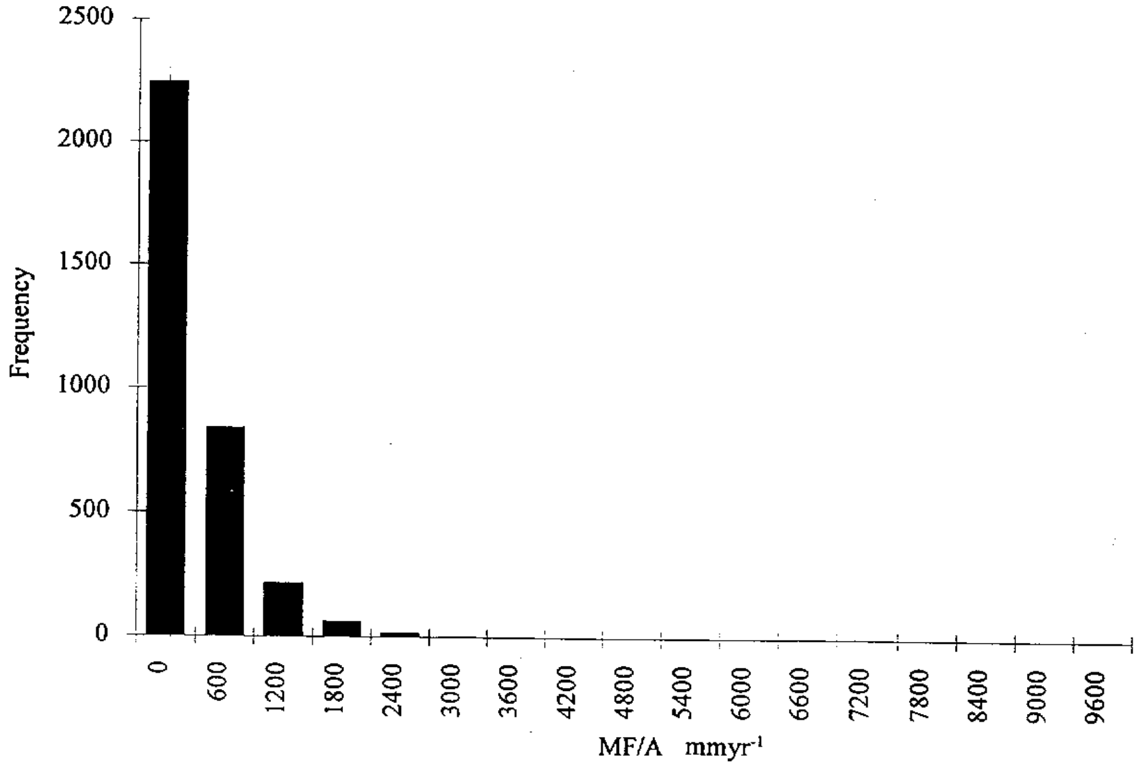


600
500
400
300
200
100
0
Frequency
0
600
1200
1800
2400
3000
3600
4200
4800
5400
6000
6600
7200
7800
8400
9000
MF/A mmyr-1

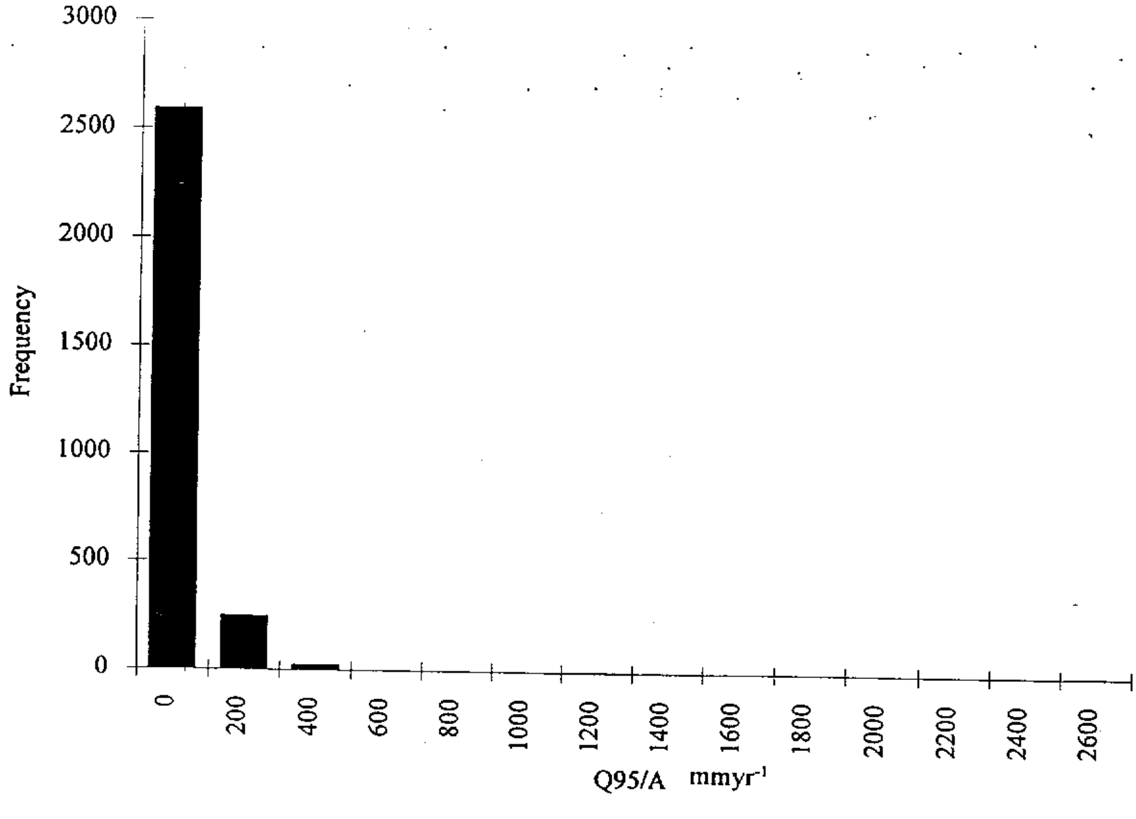
(a)

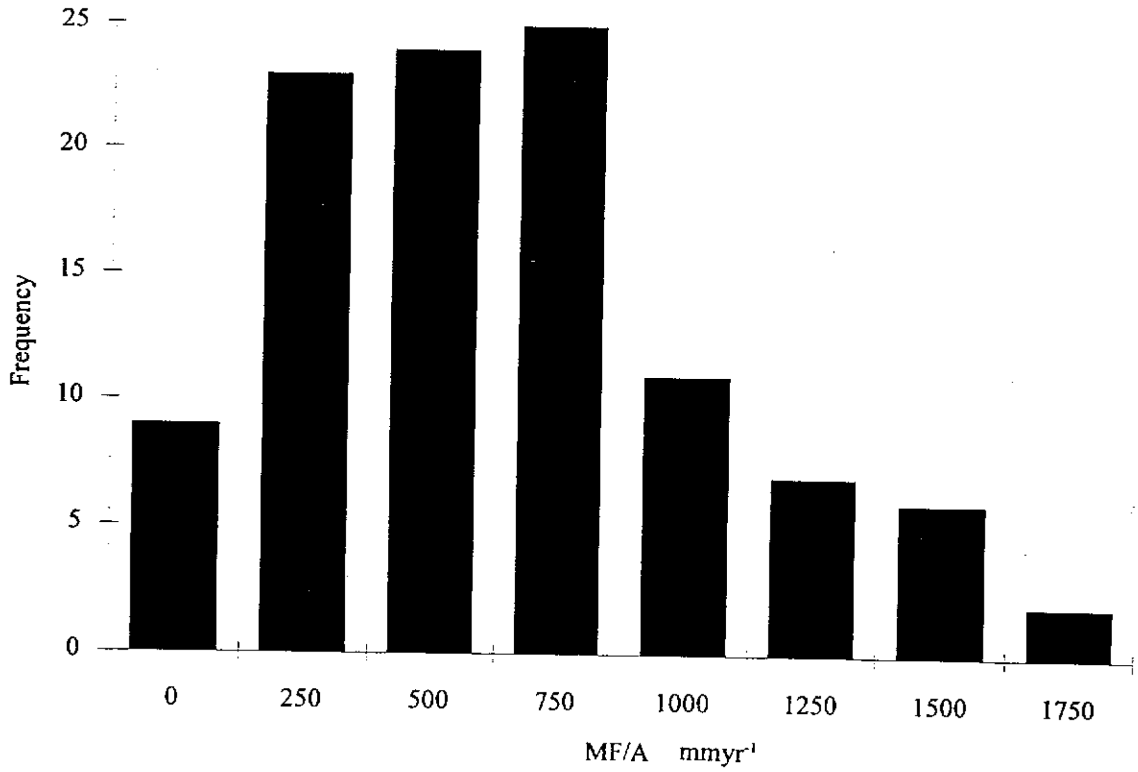


600
500
400
300
200
100
0
Frequency
0
100
200
300
400
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
2000
2100
2200
2300
2400
2500
2600
Q95/A mmyr-1


2500
2000
1500
1000
500
0
Frequency
0
600
1200
1800
2400
3000
3600
4200
4800
5400
6000
6600
7200
7800
8400
9000
9600
MF/A mmyr-1

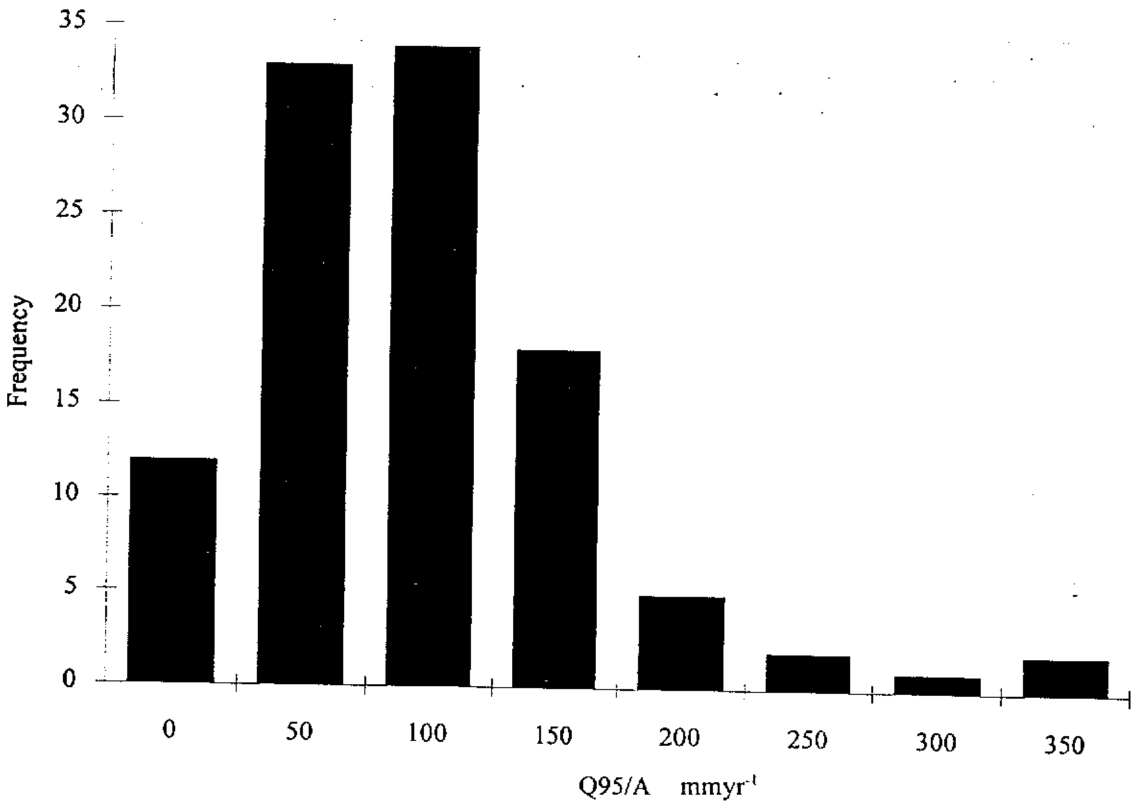
(b)


3000
2500
2000
1500
1000
500
0
Frequency
0
200
400
600
800
1000
1200
1400
1600
1800
2000
2200
2400
2600
Q95/A mmyr-1



25
20
15
10
5
0
Frequency
0
250
500
750
1000
1250
1500
1750
MF/A mmyr-1

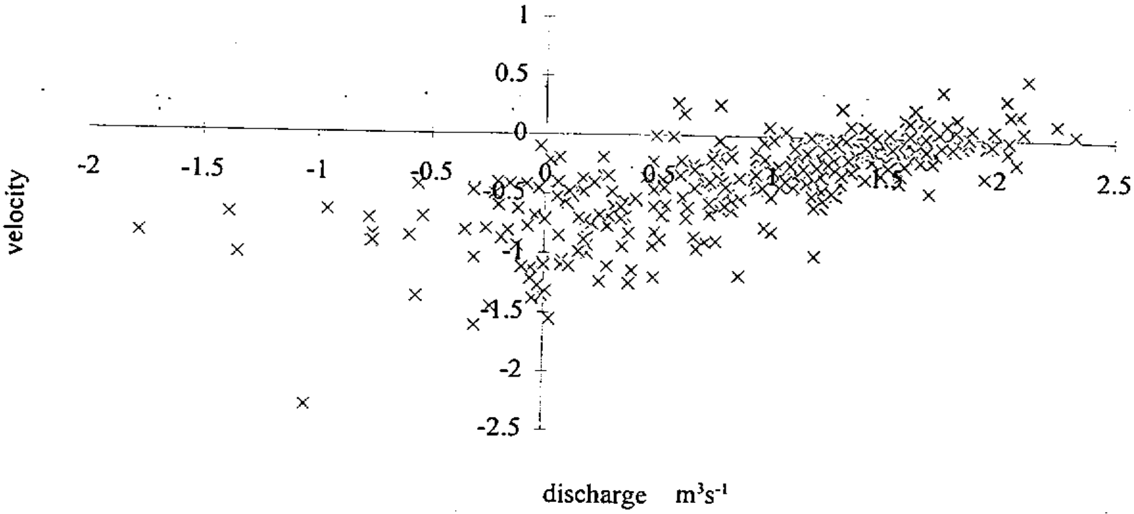
(c)

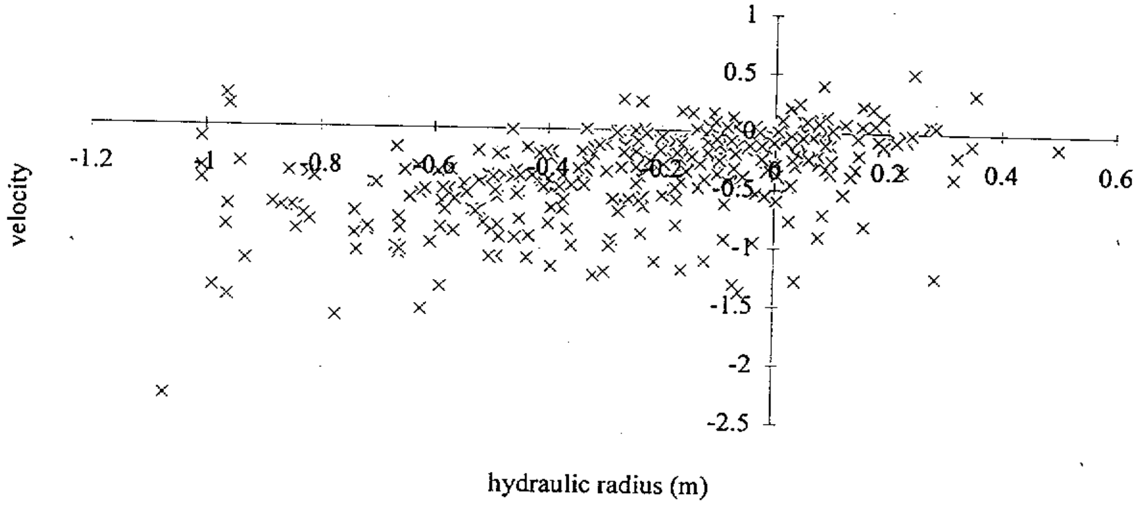


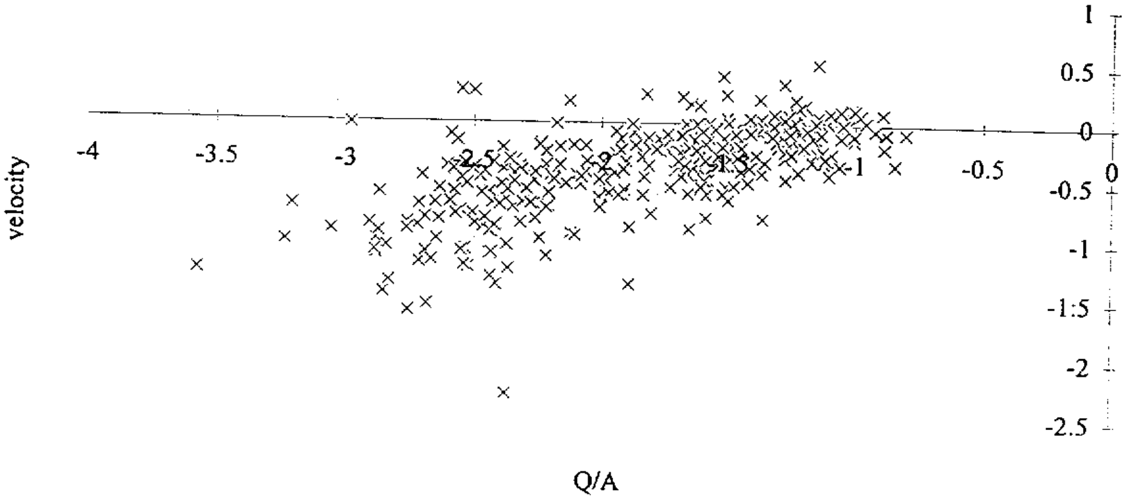
35
30
25
20
15
10
5
0
Frequency
0
50
100
150
200
250
300
350
Q95/A mmyr-1

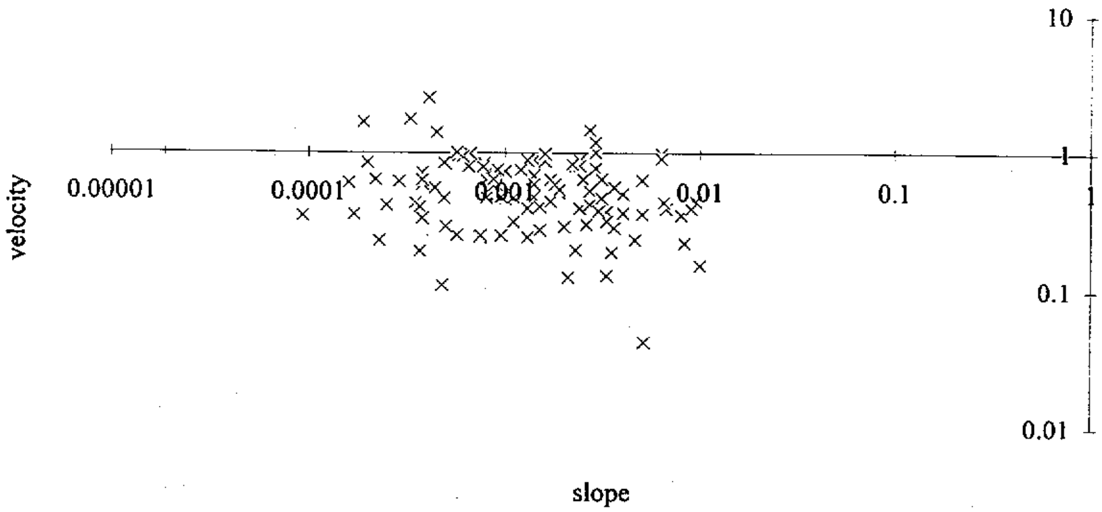
Appendix 4

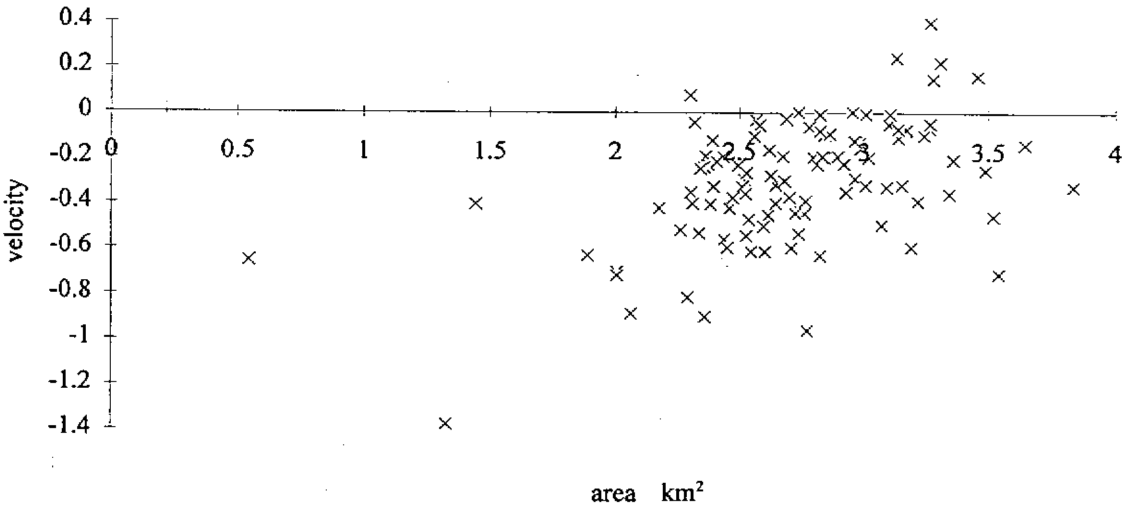
Example log-log plots implying linearity


velocity
discharge m³s⁻¹


velocity
hydraulic radius (m)


velocity
Q/A


10
1
0.1
0.01
0.00001
0.0001
0.001
0.01
0.1
velocity
slope


0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
0
0.5
1
1.5
2
2.5
3
3.5
4
velocity
area km²