HOW MUCH WATER COMES INTO THE OKAVANGO DELTA SYSTEM? - A STUDY OF THE DISCHARGE GAUGING DATA FOR THE OKAVANGO RIVER AT MOHEMBO

F.T.K. SEFE
DEPARTMENT OF ENVIRONMENTAL SCIENCE

National Institute of Development Research and Documentation (NIR)
UNIVERSITY OF BOTSWANA

August 1998

Published by the National Institute of Development Research and
Documentation (NIR)
University of Botswana
P/Bag UB00708
Gaborone
Botswana

NIR was established in 1975 and is an integral part of the University of Botswana. Its main aims are to promote, coordinate and conduct research on issues of socio-economic, environmental and cultural development affecting Botswana; to develop national research capacity within Botswana; and to document, publish and disseminate results of such research.

First published in 1998

ISBN 99912-2-125-5

Sefe, F.T.K. *How much water comes into the Okavango Delta System?* - A study of The Discharge Gauging Data forThe Okavango River at Mohembo, 1998, 26p.

/surface water/ /water resources/ /resources potential/
/mohembo/ /okavango delta/ /okavango river/ /botswana/

Printed and bound by NIR Printing Unit.

PREFACE

This report emanated from a fieldtrip I undertook with final year hydrology students (1993/94 academic year) to Maun. The students visited the Maun office of the Department of Water Affairs to study discharge measurement procedures. We were briefed about the 'instability' of the rating curve of the Okavango River at Mohembo. Discussions with officials of the Department of Water Affairs in Gaborone confirmed this view. In the months which followed the fieldtrip, I carried out a study into this apparent instability of the rating curve. This report is the culmination of that study.

ACKNOWLEDGEMENT

I am grateful to the Department of Water Affairs for providing the data for this study. I wish to express my gratitude to colleagues who read through this report or with whom I discussed methodologies. In particular I am indebted to Dr E. Lungu, Dr S. Ringrose and Dr A. Gieske.

The assistance of the NIR in accepting to publish this report is gratefully acknowledged.

A shorter version of this report appeared in the Hydrological Sciences Journal, Vol. 41, Number 1, February 1996; pp97-116.

TABLE OF CONTENTS

List of Tables

List of Figures

INTRODUCTION

The stage-discharge (H-Q) relationship is a fundamental technique employed in discharge data gathering. Typically the relationship is established from periodic measurements of stream discharge and corresponding water surface elevation or stage. While recent developments (e.g. Gawne and Simonovic, 1994) may take the tedium out of rating curve determination, the fact still remains that a good and representative rating curve requires quality data. The H-Q relationship at a particular river cross-section, even under conditions of meticulous observation, is not unique as rivers are often influenced by factors not always understood nor easy to quantify. This natural variability in the H-Q relationship is often aggravated by considerable errors of human origin.

The Okavango River system has, over the years, attracted considerable interest as a major source for water and other ecological resources. Recently a major water scheme focused on this system had to be abandoned by the Botswana government as a result of widespread concern about the impact of the project on the unique ecosystem of the area (SMEC, 1987; IUCN, 1992). However, considering the projected increases in water demand in northeast Botswana, it is only a matter of time before the project is resurrected again. Undoubtedly, the management of any resource abstraction scheme in a unique environment such as this, depends on a thorough understanding of the working of the entire system. In this regard, accurate discharge data on various time scales are required for resource assessment, environmental monitoring and flood forecasting. Given the primary importance of river gauging activities to discharge data gathering, this study focused on stage-discharge relationships for the Okavango at Mohembo, which is the point where the flow from the Cubango system enters the Okavango Delta system.

The study reported here emanated from concern within the Department of Water Affairs, the government department responsible for the collection of hydrological data in Botswana, about the existence of "inexplicable" outliers in the stage-discharge plot. Outliers occur commonly in discharge data. However, in this particular case, huge discharge values were recorded which corresponded to stages to which much lower discharges were previously associated - a phenomenon which is referred to here as instability, with no implication for the condition of the control at the gauging site. Moreover, the records did not show any corresponding change in velocity or cross-section. Considering the fact that gauging at this remote site is largely carried out by inadequately trained staff, it seems likely that human error could be one of the causes of the occurrence of the outliers. This however, cannot be assumed. The aim of this study, therefore, is to investigate the reason for this instability and how a useful rating curve can be extracted from the available data. It describes and characterises the nature of the H-Q relationship, and then investigates the use of regression methods to verify the existence of errors in the data and to derive a useful rating curve. Finally, it investigates the possibility of fitting the rating curve using stochastic analyses.

THE STUDY AREA

The Okavango River is the extension of the Cubango River into Botswana. On entering Botswana the river becomes confined between two parallel faults at the end of which it spreads out to form the Okavango Delta. Mohembo is at the northern end of the panhandle - as the parallel faults are called (Fig. 1). The panhandle is a broad well-defined channel with a clearly defined floodplain. The catchment area up to Mohembo is about 150,000 km^2 (IUCN, 1992) with most of this area contributing very little or no flow. The location of the gauging site at the northern end of the pan-handle provides natural control for river gauging, with negligible bedslope and a well defined channel. The gradient of the river between Mohembo and Shakawe (Fig.1) is only about 1:7000.

THE DATA

River gaugings at Mohembo started on a more or less regular basis in 1974. There are irregular data dating as far back as 1930's (SMEC, 1987). Data of variable length are available on the following parameters: stage, discharge (gauged or estimated from rating curve), cross-sectional area, surface width, mean velocity, mean depth, and maximum depth. Discharge is gauged with an AOTT current meter from a boat, with velocity in the vertical measured by a mixture of the six-tenths and two-point methods. The sixth-tenths method was usually used when depth in the vertical was below 1 m. The number of verticals varied, although rarely, according to the width of the cross-section at the time of gauging. Usually the distance between verticals was typically 5 m, but this reduced to 2-3 m for the end verticals. The discharges were calculated using the mid-section method (Morsley and McKerchar, 1993). Although the number of verticals varied, the reliability of the gaugings were not assessed.

For this study, the primary gauging data sheets (the HY1 forms) were studied. In a few cases the discharges were recomputed. Overall, the following characteristics were observed about the data:

i)There are many years during which surface width was recorded as the same, irrespective of season. This is due to the fact that the river is incised between the two parallel faults, giving a stable gauging section.

ii) Most often the HY1 forms were not available and the records available (HY2 forms) showed only stage and discharge. Enquiries revealed that most of the discharge data on the HY2 forms may have been estimated from a previously defined H-Q relationship, and the rest may have come from HY1 forms which may now have been lost.

iii) The same magnitude of stage yielded different discharges at various times. There is no evidence to suggest that this could be attributed to variable channel storage, particularly as the plotted rating curve did not show a definite loop. Nor is there evidence of backwater effects. While channel aggradation and degradation cannot be ruled out, it seems likely that errors in discharge measurement might be a contributory factor. In one of the files kept on the gauging station, it was discovered that an attempt was made at one time to verify the different discharges at similar stages by undertaking a joint gauging exercise with the South African Department of Water Affairs using their equipment as well as that of the Botswana Department of Water Affairs. Over the two days of gauging, it was found that the Botswana equipment gave about 20% more discharge.

iv) Where the records indicated a change in stage during a gauging exercise, the discharge plotted or recorded is not the steady-state discharge.

v) The same surface width yielded different cross-sectional areas and discharges at various times. This could be due to a number of reasons, namely: a) the depth soundings during gauging were not accurate, b) the verticals were not located the same distance apart for the same cross-section width, which happened in a few of the data sheets examined, and, c) bed movement.

It was not possible to use all the available data for this study. The portion of data used for this study is listed in Table 1.

Table 1: Length of data used for various relationships

Relationship	No of data points
H vs Q	1183[1]
Mean depth vs Q	70
Mean vel vs Q	354
Surface Width vs Q	331

1. Note that not all these are actual gaugings.

METHODOLOGY

Hydraulic geometry

One way in which to verify the quality of the available data is to examine the nature of the relationship between mean depth (MND), mean velocity (MNV) and surface width (SW) on one hand, and discharge (Q) on the other. Along with slope and friction factors, those parameters determine the nature of the cross-section and, therefore, H-Q relationship. The relationships are expressed as follows (Leopold and Maddock, 1953):

$$MND = cQ^f \qquad (1)$$
$$MNV = kQ^m \qquad (2)$$
$$SW = aQ^b \qquad (3)$$

with the condition that $f+b+m=1$, and $a.c.k.=1$. These relationships are investigated by simple linear regression in which those conditions are maintained. The particular way in which the regression analyses are handled in this study is described later.

The stage-discharge relationship
The relationship between stage (H) and discharge (Q) is normally represented simply by a plot of H against Q. The plot can be either in natural or log space, and an equation is fitted to it as below:

$$Q = aH^b \tag{4}$$

or

$$\log Q = \log a + b \log H \tag{5}$$

These were the initial approaches adopted in the assessment of the stage-discharge data.

Linear regression
It is assumed that the data available contained errors of unknown magnitude that cannot be ascribed to the inherent variability in natural systems. It is also assumed that there are a 'reasonable' amount of 'correct' observations in the data set. Under these assumptions, the linear regression technique was employed in investigating the relationships among the various parameters as outlined in the above equations.

The basic method of achieving linearity was logarithmic transformation, although in the case of the H-Q relationship, the power and Box and Cox transformations were used. Where it is necessary to choose between a number of regression equations, the decision was based on three statistics - standard error of estimate (SEE), the sample coefficient of determination (R^2) (Gawne and Simonovic, 1994), and the extent to which the distribution of residuals approached normality as indicated by the normal probability plot and the mean and standard deviation of the standard residuals. For equations (1,2,and 3) these criteria were superseded by the conditions that f+b+m = 1 and a.c.k. = 1.

Outliers were identified by the Mahalanobi's distance (D_i) routine (Norušis, 1990), which is a measure of the distance of an individual observation from average values of the independent variable in a regression analysis given as

$$D_i = \left(\frac{X_i - \overline{X}}{S_x}\right)^2 \tag{6}$$

where X_i is the *i*th observation, and MN_x and S_x the mean and standard deviation of the sample respectively. This method was supplemented by a physical examination of the plots of standardized residuals. It was assumed that identified outliers were observations which contained substantial error. They were therefore removed after each regression run until the results were found acceptable in terms of the three statistics outlined above.

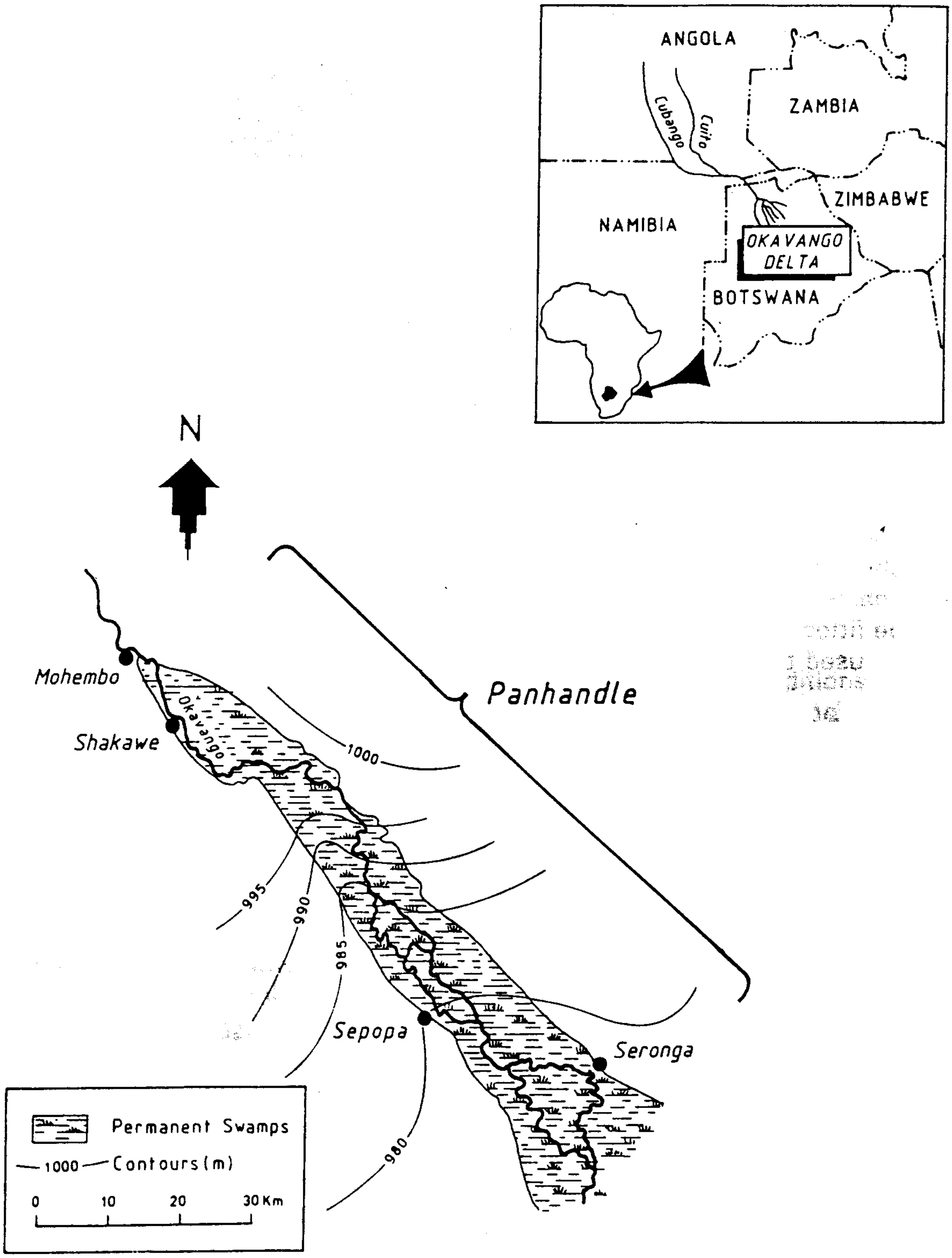

Fig. 1: Study area

It is recognised that the method of treatment of outliers adopted here may in fact lead to discarding of important information which may have physical impact on the various relationships studied. However, it was judged that in this situation removal of outliers would lead to more valid relationships, especially since many of these outliers may have resulted from human or procedural errors.

Stochastic modelling of gauging data

Although flow gaugings are not carried out daily throughout the year, the observed gaugings can be treated as a time series as it is possible to treat the days without gauging as days with missing values. These missing values can then be estimated by one of the routines incorporated in the SPSS statistical package (SPSS Inc.,1990). This was done in this case and the resultant time series subjected to an ARIMA (p,d,q) analysis.

Discussion of ARIMA models abound in the literature (eg. Box and Jenkins, 1976; Kendall and Ord, 1990; Wei, 1993). The general ARIMA (p,d,q) model may be expressed as (Wei, 1993):

$$\phi_p(B)(1-B)^d Z_t = \theta_0 + \theta_q(B)a_t \qquad (7)$$

where the stationarity AR operator $\phi_p(B) = (1-\phi_1 B - \ldots - \phi_p B^p)$ and the invertible MA operator $\theta_q(B) = (1-\theta_1 B - \ldots - \theta_q B^q)$ share no common factors; Z_t is a realisation of a stationary process and a_t is a zero mean white noise process; B is the backward shift operator. The modelling was done using the SPSS Trends package. This analysis was done with the purpose of investigating whether an ARIMA model can be fitted to observed discharge data, for, it was possible, the technique could also be used to estimate the inevitable missing data.

Effect of changing stage on the rating curve

The discharge plotted against stage to obtain the rating curve should be the steady-state discharge. Where stage changes during the period of gauging, the measured discharge is often adjusted to obtain a steady-state discharge. The methodology for carrying out such an adjustment is covered in standard texts in hydrology (e.g. Wilson, 1990). The method is based on the Manning formula:

$$Q=AV=\frac{AR^{2/3}S^{1/2}}{n} \qquad (8)$$

where Q = steady state discharge ($m^3 s^{-1}$), A = cross-sectional area, V = average water velocity through the given cross-section, R = hydraulic radius, and S = water surface slope. It is obvious from equation (8), that the steady state discharge, Q is:

$$Q \propto \sqrt{S} \qquad (9)$$

During a rising (+) or a falling (-) stage the actual discharge measured is given by

$$Q_a \propto \sqrt{(S \pm \Delta S)} \qquad (10)$$

where ΔS represents the change in stage during gauging which can be expressed in terms of flood celerity, U (m s^{-1}), as:

$$\Delta S=\frac{dh}{Udt}=\frac{dh/dt}{U} \qquad (11)$$

where dh/dt is the change in stage over time interval t.
From equations (8, 9 and 10), the steady-state discharge, Q can be expressed in terms of the measured discharge, Q_a as:

$$\frac{Q_a}{Q}=\sqrt{(1+\frac{dh/dt}{US})} \qquad (12)$$

The flood celerity U is often evaluated in terms of average water velocity as

$$U = 1.3\ Q_a/A \qquad (13).$$

From equation (12) the steady-state discharge can be found as

$$Q=\frac{Q_a}{\sqrt{(1+\frac{A.dh/dt}{1.3Q_aS}}} \qquad (14)$$

The data files examined contained records of gaugings carried out under conditions of rising or falling stage, but the computed flows were not adjusted. Accordingly, equation (14) was used to compute steady-state discharges in order to assess the impact of this omission on the rating curve.

RESULTS AND DISCUSSIONS

Hydraulic geometry

The results of the regression of MNV, MND and SW on Q are summarised in Table 2 and Figs. 2, 3, 4, 5, and 6. Table 2 shows the regression equations obtained for the various hydraulic parameters at each iteration of the regression analysis and subsequent removal of outliers. Fig 2 shows a plot of MNV vs Q of the original data. The first observation to make is the single data-point representing a log discharge in excess of 2.7 (more than 500 m^3s^{-1}) and log MNV very close to minus 1.5 (i.e. about 0.03 m/s). It is inconceivable that this data-point was actually observed. It would require a cross-section area of 16,000 m^2 to produce this flow. Clearly this data-point must be an error. A second area of interest on Fig. 2 is the small cluster of data-points between the larger cluster above and the single outlier below. Some of these points occur in bands, but others appear to be scattered randomly through the data. The time between bands ranges from a few months to two years. However, there is no evidence in the files to indicate that these points represent different hydrological characteristics. It is inconceivable that hydraulic characteristics of the cross-section could change at such a magnitude and so rapidly to produce the sudden variations in flow indicated by these points. It seems highly likely that some observation or calculation error might have been made.

Logarithmic transformation produced a better linear plot. Regression analyses identified those points referred to earlier as outliers. In line with the assumption that such outliers would be erroneous observations, they were removed on each occasion. The final relationship between mean velocity and discharge is shown in Fig. 3.

Fig. 4 shows the relationship between mean depth (MND) and discharge (Q). Here it can be seen that a discharge of about 300 $m^3 s^{-1}$ is associated with two vastly different depths of flow. It is difficult to imagine the flow conditions that would have produced such a relationship. Once again, there is strong indication of error in the data. It is intriguing that the errors occur only at a discharge of about 300 $m^3 s^{-1}$. The relationship between MND and Q was investigated in a similar manner as for mean velocity. The final relationship is shown in Fig. 5.

The relationship between SW and Q (Fig. 6), excluding the single outlier, provides further confirmation of a confined section. However, the scatter of points suggests two different surface width-discharge relationships, about 10 m apart. There is no evidence in the files of the gauging station to indicate that this separation has any physical significance. From Figs. 2 and 4, it can be seen, for example, that a flow of 100 $m^3 s^{-1}$ corresponds to a velocity and depth of approximately 0.4 $m s^{-1}$ and 3 m respectively. From Fig. 6, with a width of 85 m, the discharge (computed as vwd) is 102 $m^3 s^{-1}$. At a higher flow of 700 $m^3 s^{-1}$, but with the same surface width of 85 m, the corresponding velocity of 1.4 $m s^{-1}$ and depth of 6 m, the computed discharge is 714 $m^3 s^{-1}$. Thus it seems that with a low river gradient of about 1:7000 in this area, the depth of flow, rather than velocity, is the dominant control on flow magnitude.

An inspection of the gauging site indicated that at higher flows some discharge passes through the swamps on the left bank of the river. It is possible that this flow may be missed on a gauging mission, but that would not affect the relationships shown in Figs. 2, 4 and 6 in any appreciable manner.

The relationship between SW and Q was analyzed in similar manner as for MNV and MND. Logarithmic transformation was carried out as before. Subsequent regression and removal of outliers did not improve the correlation and the exercise was abandoned after the third run.

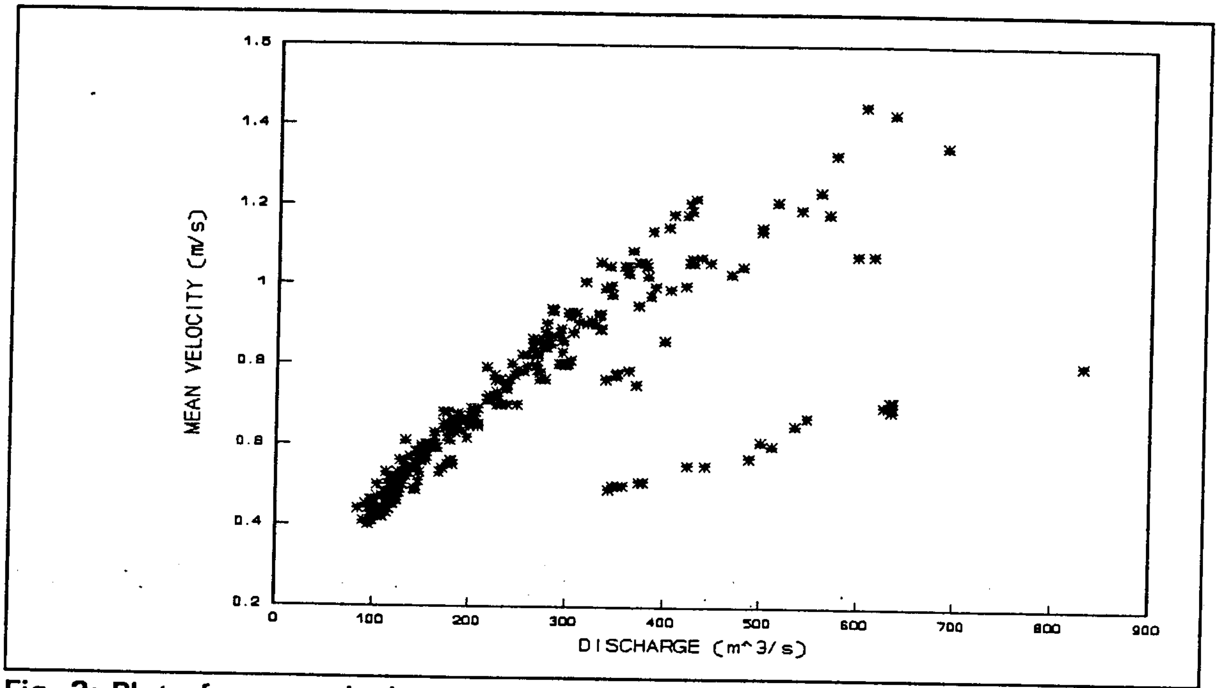

Fig. 2: Plot of mean velocity vs discharge

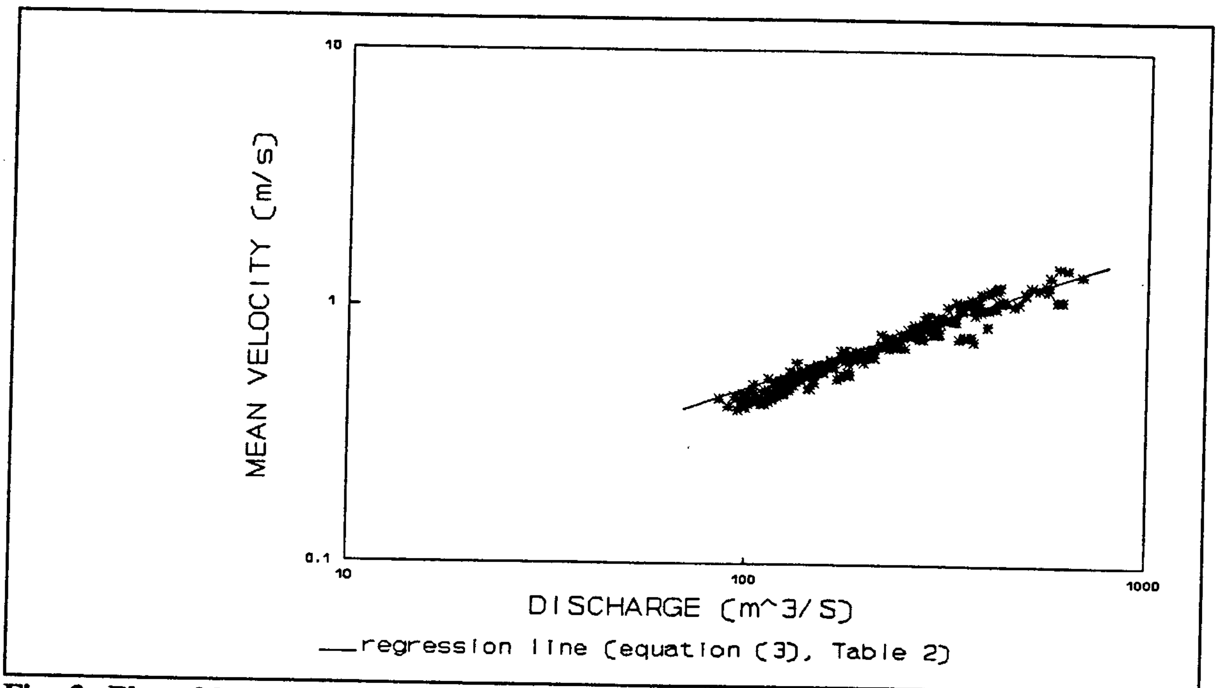

Fig. 3: Plot of log mean velocity vs log discharge (excluding outliers)

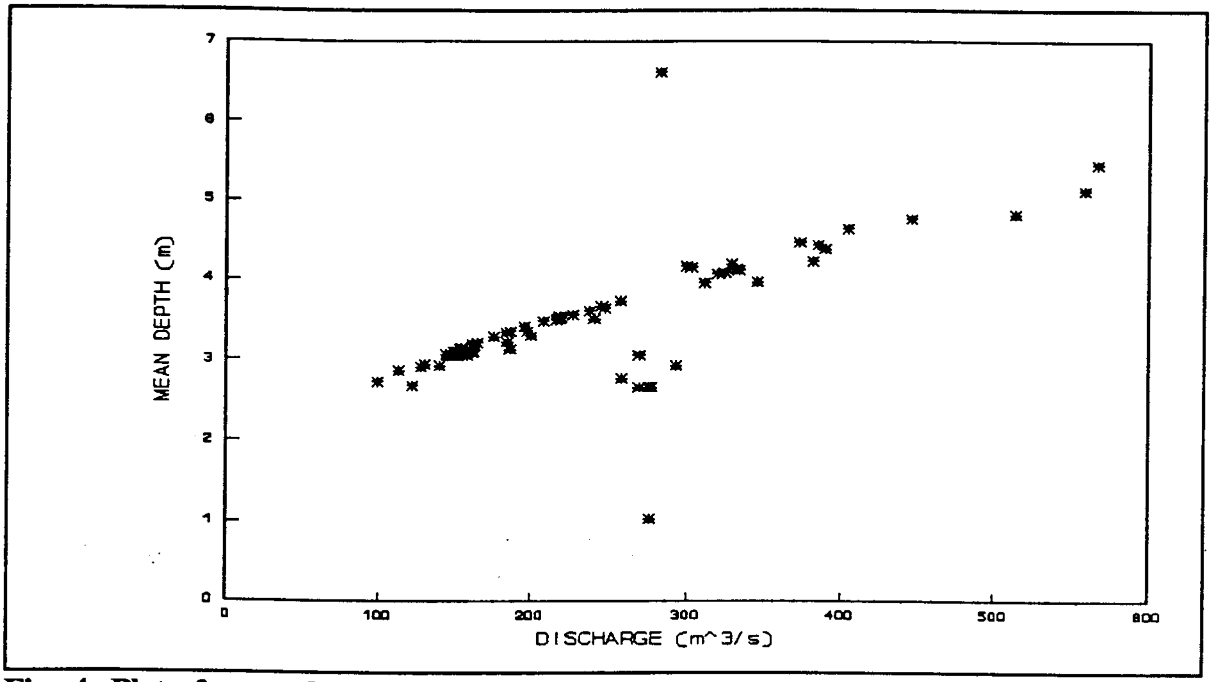

Fig. 4: Plot of mean depth vs discharge

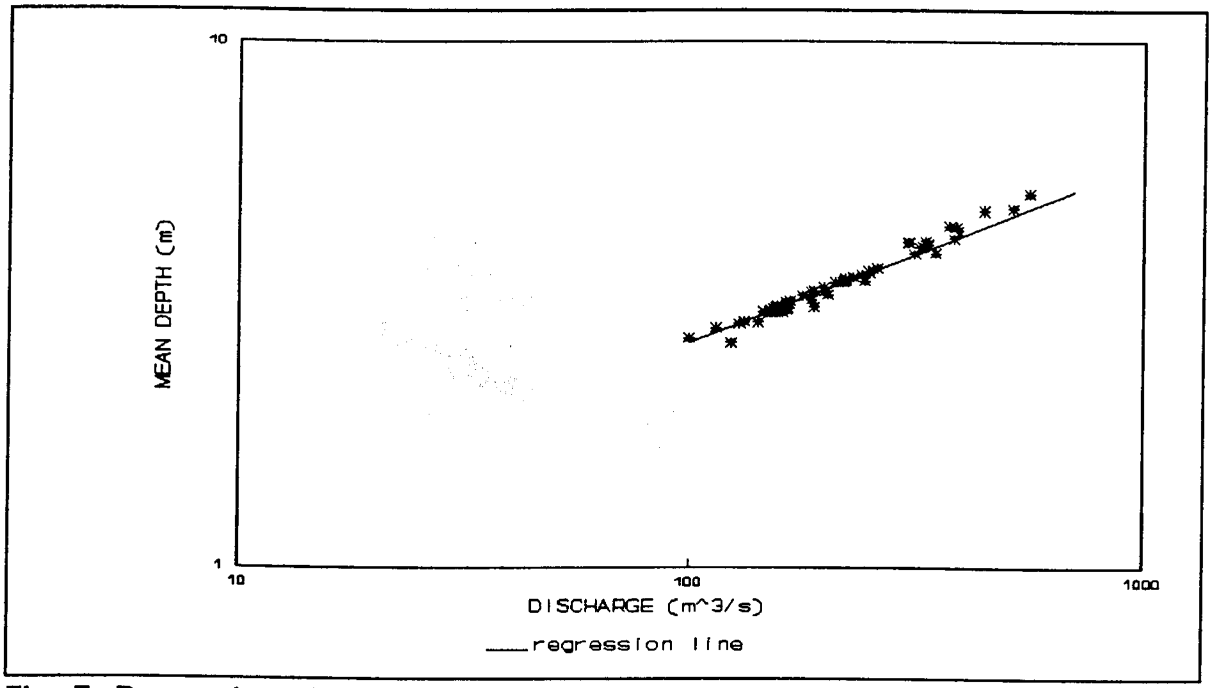

Fig. 5: Regression of mean depth on discharge (excluding outliers)

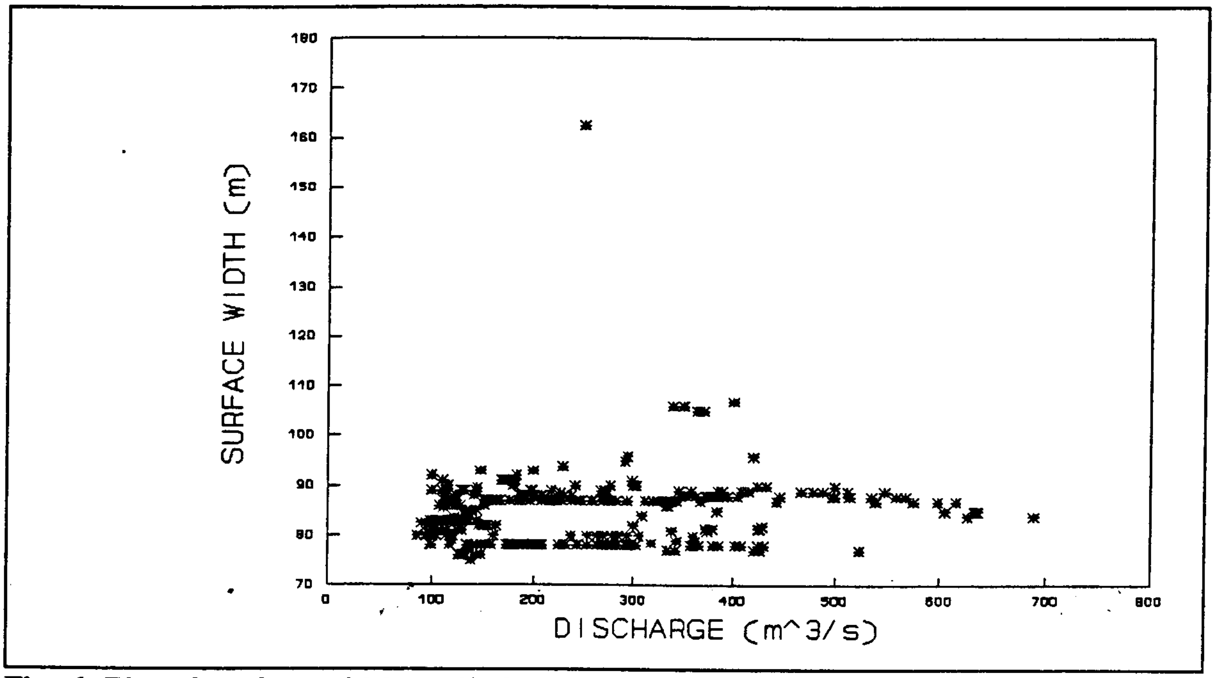

Fig. 6: Plot of surface width vs discharge

The relationships between MNV, MND and SW on one hand, and Q on the other hand which approximately satisfy the conditions (f + b + m = 1; a*c*k = 1) are:

$$MNV = 0.04\ Q^{0.54} \tag{15}$$

(R = 0.91; standard errors: regression equation = 0.05, constant = 1.07, exponent = 0.01; df = 302)

$$MND = 0.56\ Q^{0.34} \tag{16}$$

(R = 0.57; standard errors: regression equation = 0.08, constant = 1.38, exponent = 0.06; df = 68)

$$SW = 76.0\ Q^{0.02} \tag{17}$$

(R = 0.19; standard errors: regression equation = 0.02, constant = 1.03, exponent = 0.01; df = 329)

Equation (17) is, in fact, not valid as the exponent is effectively zero. This is further confirmation of errors in the surface width measurements.

Table 2: Results of regression of hydraulic parameters on discharge.

Relationship	Regression equation	Correlation coefficient of logs
MNV vs Q	(1) MNV $= 0.07\ Q^{0.43}$ (2) MNV $= 0.05\ Q^{0.49}$ (3) MNV $= 0.04\ Q^{0.54}$ (4) MNV $= 0.02\ Q^{0.62}$	R = 0.67[1] R = 0.85 R = 0.91 R = 0.98
MND vs Q	(5) MND $= 0.56\ Q^{0.34}$ (6) MND $= 0.46\ Q^{0.38}$	R = 0.57 R = 0.99
SW vs Q	(7) SW $= 73.32\ Q^{0.03}$ (8) SW $= 76.68\ Q^{0.02}$ (9) SW $= 75.57\ Q^{0.02}$	R = 0.19 R = 0.16 R = 0.19

1. Note that R applies to the association of the logarithms, not the actual values.

The nature of the data errors

It is obvious from the foregoing that the discharge gauging data contain numerous erroneous values. In order to investigate the nature of these errors, a sample of the data which had measurements of mean velocity and cross-section area were extracted. Discharge can be estimated as the product of cross-section area and mean velocity, i.e. Q = A x V. Fig. 7 shows the plot of discharge against stage for this data subset. With the exception of the two points at about 400 $m^3\ s^{-1}$ corresponding to stage of about 1.5 and 3 m respectively, the figure resembles the normal scatter that one would expect to find on stage-discharge plot. Comparison of Fig. 7 with Fig. 8 gives a completely different picture. The plot of estimated discharge (i.e. Q = A x V) against observed stage is shown in Fig. 8. Fig. 9 shows a comparison between observed discharge and discharge estimated as above. If there are no substantial errors in the measurements, Figs. 7 and 8 should look alike, and all the points on Fig. 9 should lie on the 45° line. However, as can be seen from the figures, this is not the case. From Fig. 8, it can be seen that there is a cluster of points lying to the left of the main cluster, and on Fig. 9 these points appear to the right of the 45° line. These cluster of points are claerly erroneous values which could only be due measurement errors. Thus it seems very likely that the errors in the data are largely of human origin.

The H - Q relationship

The question that arises from the foregoing is whether it is possible to obtain a useable rating curve from the data. The H - Q plot characteristically shows curvature when plotted in natural units, but often the plot can be fitted with a linear regression equation after some form of data transformation; thereby yielding a linear rating curve which greatly facilitates extrapolation. Fig. 10 shows a plot H against Q. A number of observations can be made about the stage-discharge relationship depicted. The figure suggests that there may be some curvation in the H-Q relation. Secondly, in view of the foregoing discussion, several of the points plotted cannot be explained by normal or natural variation inherent in the discharge process . Particularly, the situation where a stage of less that 2 m yielded discharge in excess of 800 m^3s^{-1} must be one of such errors that cannot be attributed to the inherent variability of the rating curve. Note also that a similar magnitude of discharge corresponds to a stage of about 3 m. Thirdly, it was found that logarithmic transformation did not remove the curvature in the H -Q relationship (Fig.11). This is quite interesting in view of the fact that some of the discharges plotted might have been obtained by linear extrapolation of an existing rating curve. However, it was not possible to determine from the available records what proportion of the data would have been obtained in this way.

Transformations applied to the H - Q relationship

With the failure of the simple log transformation to yield a linear H - Q relationship, other models were resorted to. In particular, the power and Box - Cox transformations (Box and Cox, 1964; Gawne and Simonovic, 1994), and stochastic modelling were investigated.

Power transformation

In the power transformation method, the transformation is applied to stage, H. Thus, designating the transformed stage as H^λ,

$$QM = b_0 + b_1H^\lambda \tag{18}$$

where QM is the discharge estimated from the model; b_0 and b_1 are parameters of the regression and λ is the exponent with which H was transformed.

As λ was evaluated by trial and error, there was the need to choose between the models produced by the various values of λ tried. This choice was aided by considering the closeness to normality of the residuals from each trial value of λ as indicated by the skewness and kurtosis coefficients (Gawne and Simonovic, 1994). The value of λ adopted is 1.5. Thus H was transformed to $H^{1.5}$, and as can be seen from Fig.12 the relationship is approximately linear.

The transformed stage and discharge were then subjected to regression analyses and the outliers removed as previously described. The result is the final power-transformed model shown in Fig. 13. The regression equation is:

$$QM = 102.44 + 98.61\ H^\lambda \tag{19}$$

(R = 0.97; standard errors: on R = 33.59, on constant = 1.60, slope of regression line = 0.70; degrees of freedom = 1161).

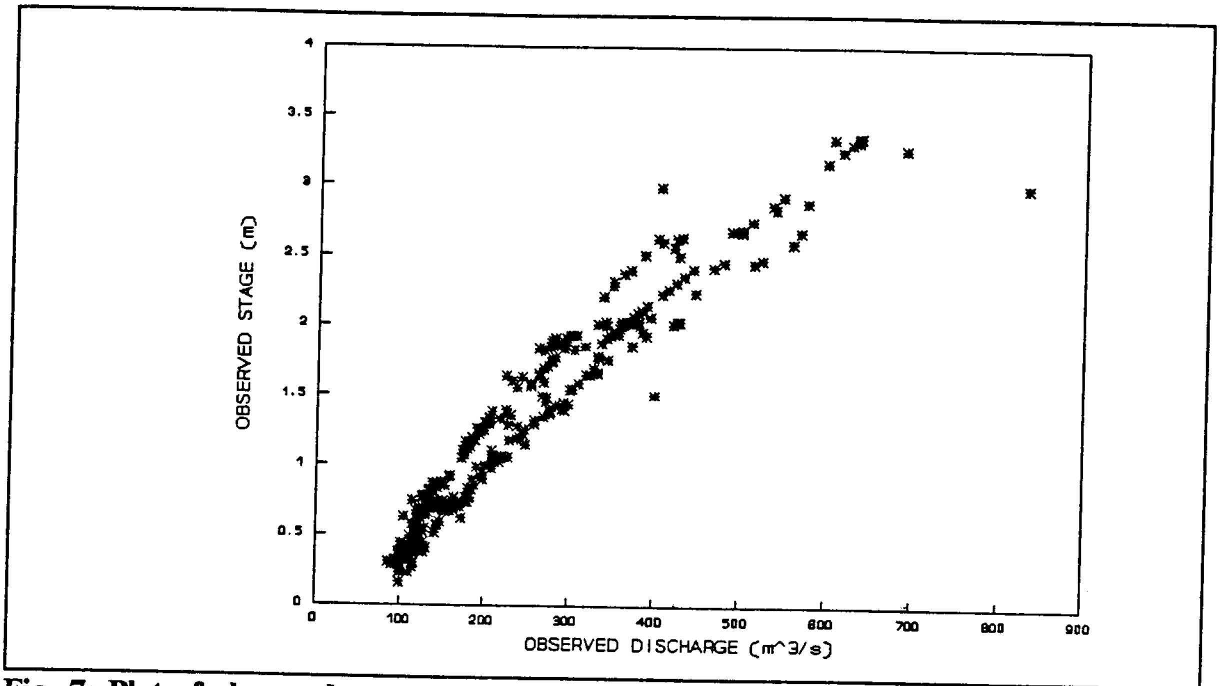

Fig. 7: Plot of observed stage vs observed discharge - subset of data

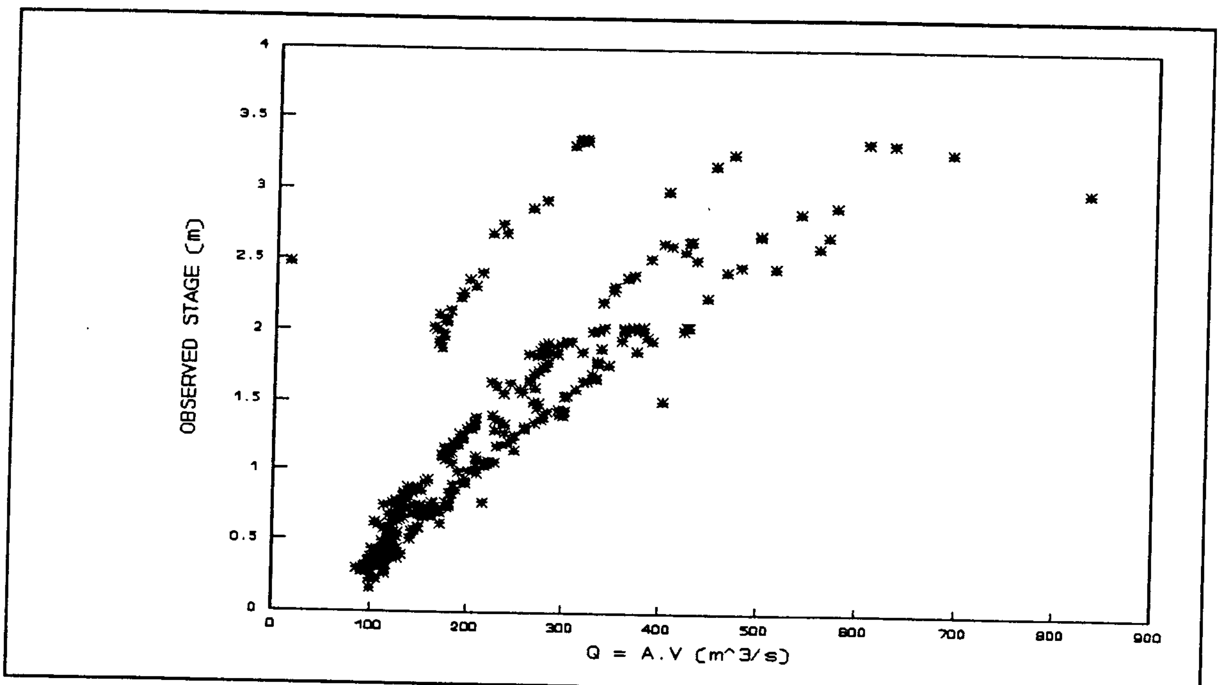

Fig. 8: Plot of observed stage vs estimated discharge (Q = A.V)

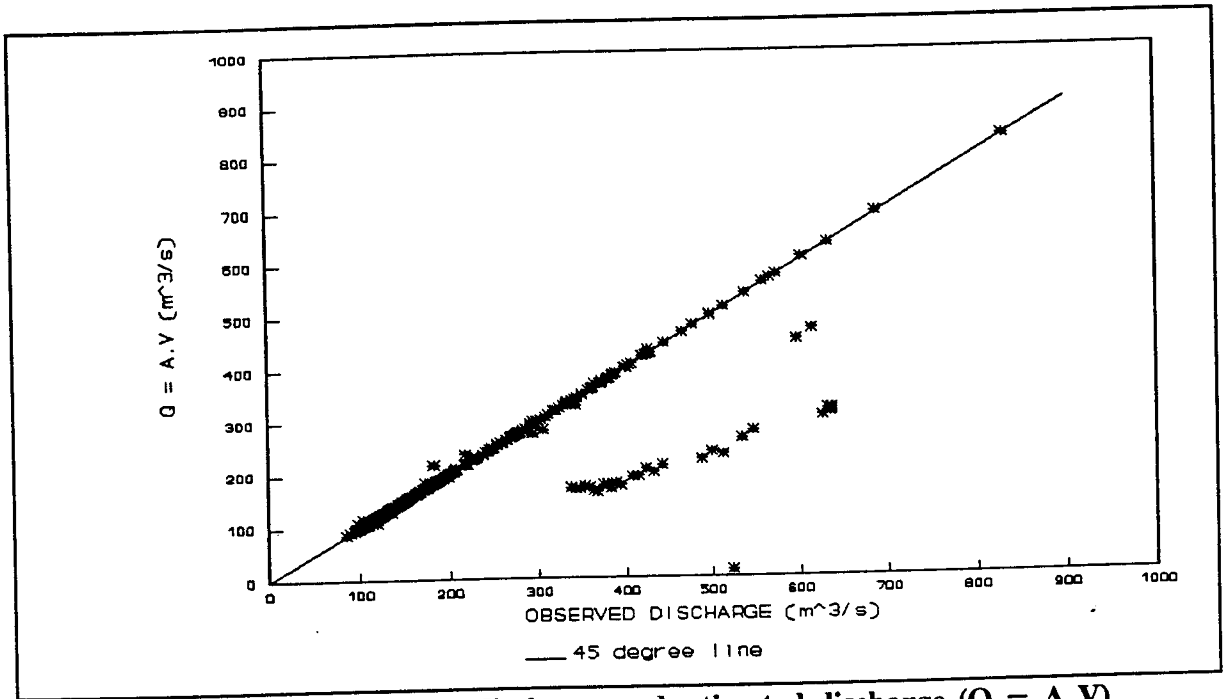

Fig. 9: Comparison of observed discharge and estimated discharge (Q = A.V)

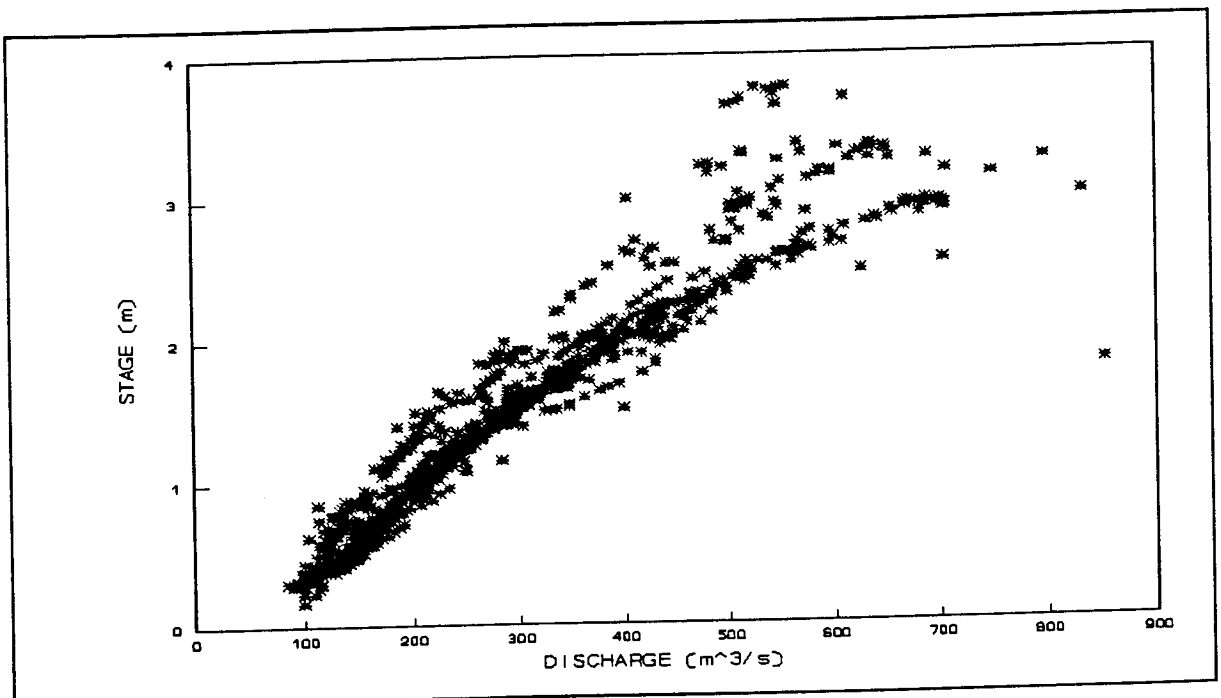

Fig. 10: Plot of stage vs discharge

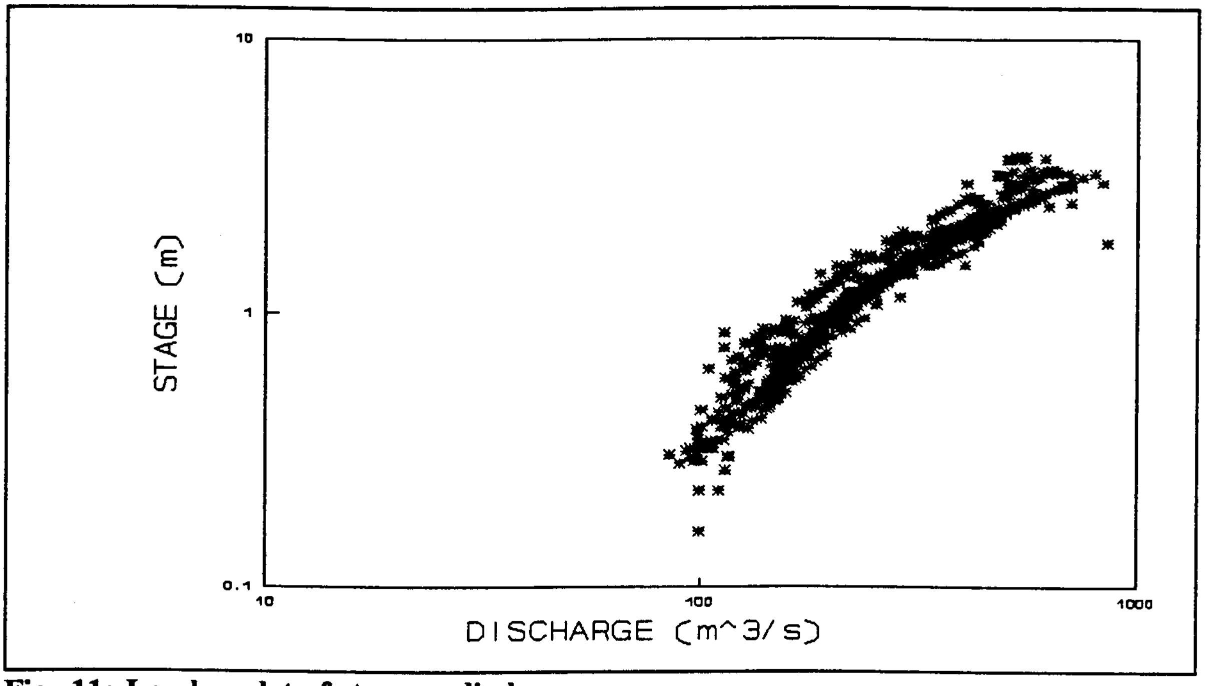

Fig. 11: Log-log plot of stage vs discharge

The Box and Cox transformation.

The next transformation tried was the Box and Cox transformation (Box and Cox, 1964; Gawne and Simonovic, 1994). This transformation is generalised as

$$QT = [(Q + K)^{\lambda} - 1]/\lambda \qquad (20)$$

where K is any constant and λ is the transformation parameter. As in the case of the power transformation, the value of λ adopted was found by trial and error to be 0.5 while K = 0 (Fig. 14).

Repeated regression analyses and removal of outliers as in the previous transformation case led to the final H - Q relationship shown in Fig. 15 with R = 0.98.(Standard errors: on R = 1.77, on constant = 0.11, slope of regression line = 0.07; degrees of freedom = 1161.) The regression model is

$$Q_T = 15.79 + 10.61H \qquad (21)$$

Stochastic modelling of discharge

Owing to technical limitations it was not possible to use the entire discharge data in the stochastic analysis. A sample consisting of 635 (Box-Cox transformed) observations was selected from the available data in such a way that gaps in the data sequence were kept to a minimum. The missing values, totalling 51 cases, were estimated by the linear interpolation routine incorporated in the SPSS statistical package (SPSS Inc., 1990). It is likely that linear interpolation would introduce some bias to the resulting ARIMA model. In order to account for the association of stage with the discharge measurements, the sampled series of (Box-Cox) transformed discharges were divided by the corresponding stage, yielding a transformed discharge-stage ratio series.

Preliminary non-seasonal differencing led to an ARIMA (3,2,1) model being applied to the transformed discharge-stage ratio series. The final parameters are listed in Table 3. It can be seen that the model fits the data adequately. The Durbin-Watson statistic of 1.9819 for the residual series fulfils the requirement of absence of autocorrelation. This is also confirmed by the autocorrelation and partial autocorrelation of the residual series (Fig. 16). The Box-Ljung statistic (SPSS Inc., 1990) is nowhere significant. The efficiency of the fit of the model to the observed discharges (as transformed) can also be seen in Fig. 17 where a new (transformed) discharge data series derived from the ARIMA (3,2,1) model (note that these are not ratios of discharge to stage) is plotted against the transformed observed discharges. It can be seen that most of the points lying on the 45° line.

The ARIMA modelled discharges (as transformed by the Box-Cox procedure) are plotted against stage in Fig. 18. It can be seen from the figure that the ARIMA model preserved the range of variability in the data as shown by the preservation of the 'outliers' - those points that lie farthest away from the 45° line (seen in Fig. 17). The problem of a lower stage producing an inexplicably large discharge is also noticeable in Fig. 18.

The rating equation associated with this plot is (with $R = 0.98$; standard errors: on $R = 1.35$, on constant $= 0.13$, on slope of regression line $= 0.09$; df $= 478$)

$$QTM = 15.64 + 11.29H \tag{22}$$

where QTM is the Box-Cox transformed ARIMA-derived discharge. Although the residuals from the regression indicated the existence of outliers, it was not considered worthwhile to remove them and rerun the analysis as before because of the high value of R. Two conclusions can be drawn from the above results. Firstly, they suggest that discharge in the Okavango at Mohembo on intervening days when no measurements were undertaken can be estimated from an ARIMA (3,2,1) model. Secondly, the retention of the outliers in the ARIMA modelled discharge series implies that if the outliers were genuinely observed values, say, flash floods, for example, ARIMA model would reproduce them reasonably well. This is a finding that may have significant implications in a region in which flash floods are prevalent, yet difficult to gauge. However, if erroneous values are suspected to be present in a suite of time series data, stochastic analysis cannot be employed as a technique to reduce the influence of the erroneous observations.

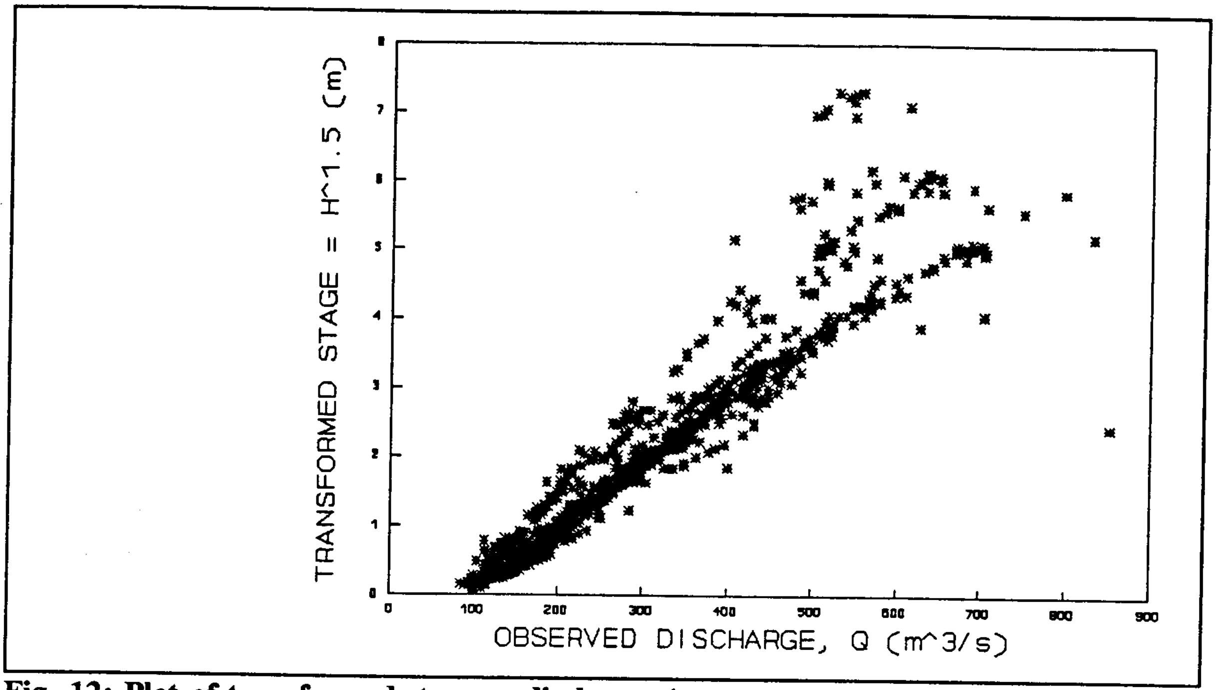

Fig. 12: Plot of transformed stage vs discharge (power transformation)

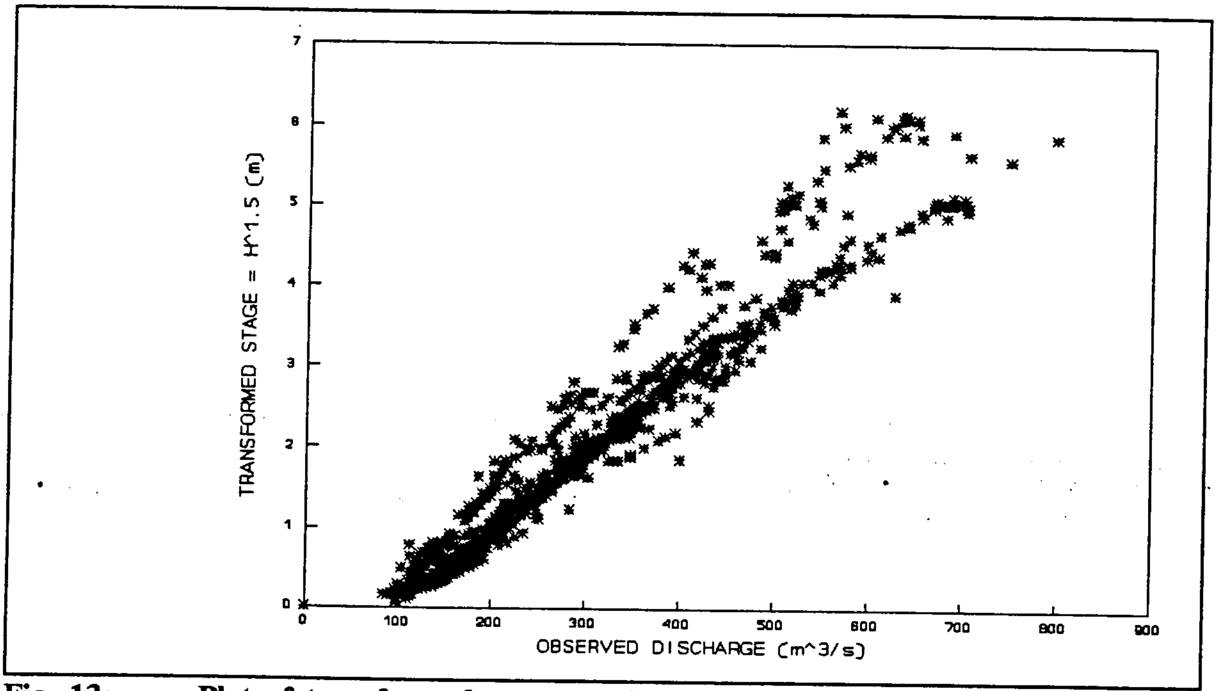

Fig. 13: Plot of transformed stage vs discharge: excluding outliers identified on second regression run. (Power transformation)

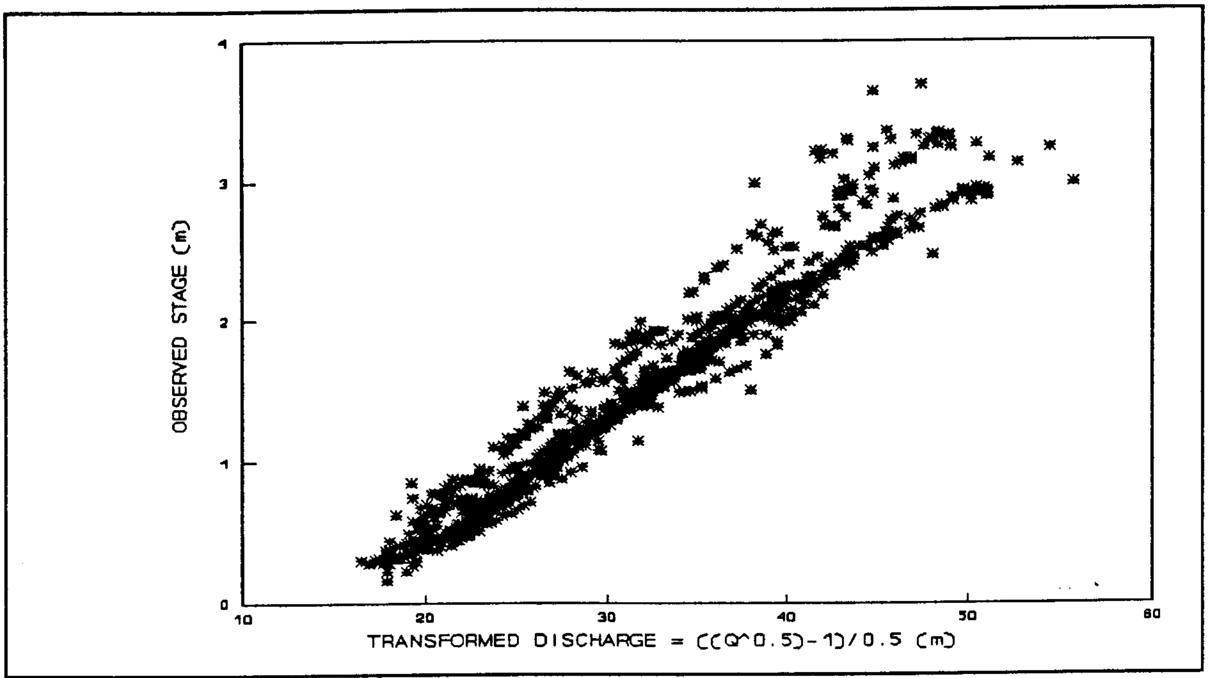

Fig. 14: Plot of observed stage vs transformed discharge (Box-Cox transformation)

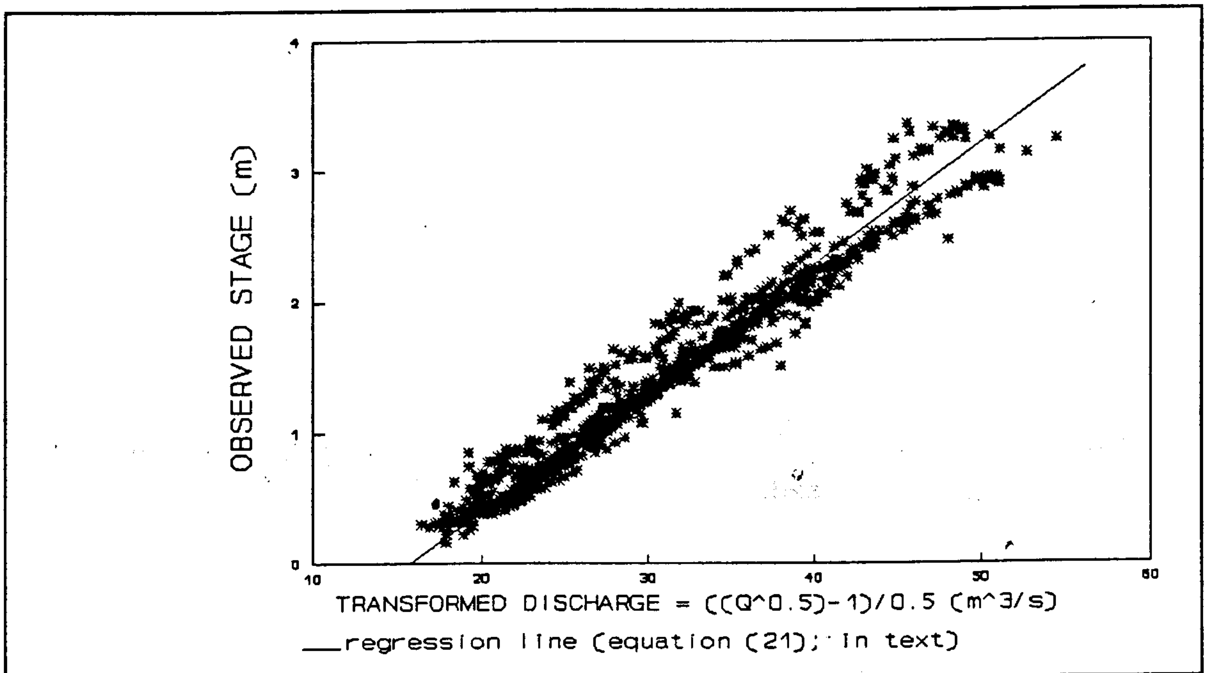

Fig. 15: Plot of observed stage vs transformed discharge: excluding outliers identitified on second regression run. (Box-Cox transformation)

Table 3: Final parameters of the ARIMA (3,2,1) model

Variables in the model:

	B	SEB	T-RATIO	APPROX. PROB.
AR(1)	-0.4020	0.0473	-8.5038	0.0000
AR(2)	-0.2156	0.0507	-4.2525	0.0000
AR(3)	-0.0997	0.0452	-2.2070	0.0277
MA(1)	0.8560	0.0275	31.1405	0.0000
CONSTANT	-0.0009	0.0029	-0.3053	0.7603

Covariance matrix:

	AR(1)	AR(2)	AR(3)	MA(1)
AR(1)	0.0022	0.0013	0.0008	0.0007
AR(2)	0.0013	0.0026	0.0012	0.0008
AR(3)	0.0008	0.0012	0.0020	0.0006
MA(1)	0.0007	0.0008	0.0006	0.0008

Correlation matrix:

	AR(1)	AR(2)	AR(3)	MA(1)
AR(1)	1.0000	0.5548	0.3822	0.5307
AR(2)	0.5548	1.0000	0.5286	0.5473
AR(3)	0.3822	0.5286	1.0000	0.4674
MA(1)	0.5307	0.5473	0.4674	1.0000

Fit error statistics:

Number of cases: 630; Degrees of freedom: 625

Mean error: -0.0150; Mean absolute error: 0.3912

RMS: 0.8613; Durbin Watson statistic: 1.9819

Influence of changing stage on discharge measurements

An averaged value of water surface slope was estimated by using occasionally available water level observations at Shakawe (see Fig. 1), about 10 km downstream of Mohembo. Average water surface slope was estimated as 0.01%. Equation (7) was then used to compute the steady-state discharges for the few occasions on which changes in stage were recorded. All the instances dealt with represented rising stage and steady-state discharge should have been 0.12 to 4.31% below the measured discharge (Table 4). Although only a crude estimate of water surface slope was used, it is likely that, overall, the non-adjustment of the measured discharges will have only minimal direct impact on the rating curve. Nevertheless, because the existing rating curve is used at times to convert observed stage to discharge, this impact may become compounded with time leading to overestimation of discharges.

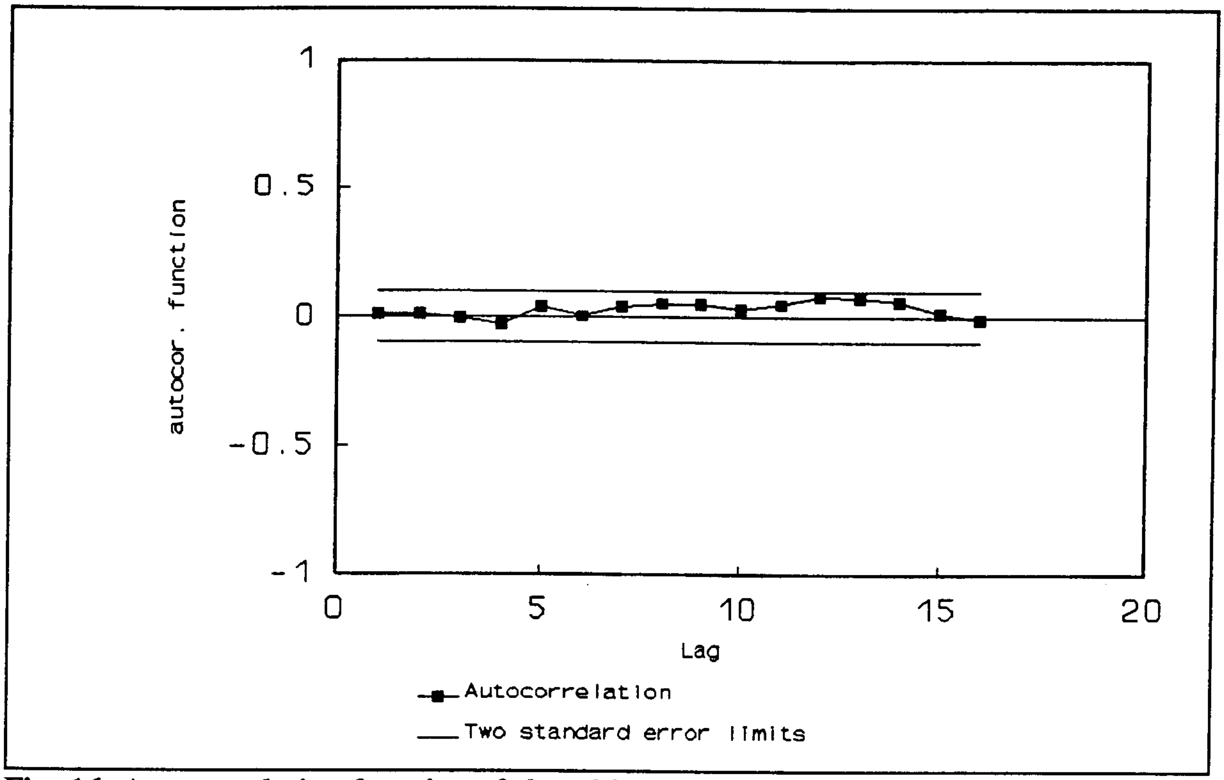

Fig. 16: Autocorrelation function of the white noise component

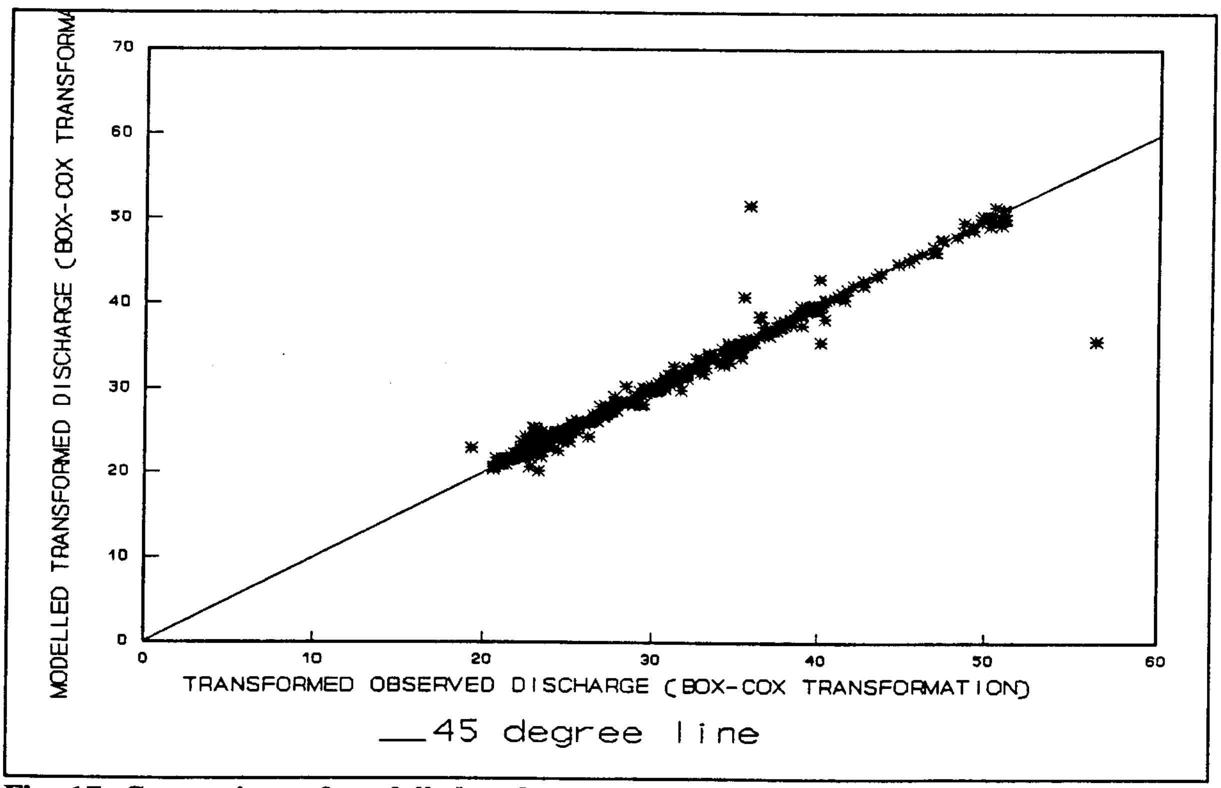

Fig. 17: Comparison of modelled and observed discharges

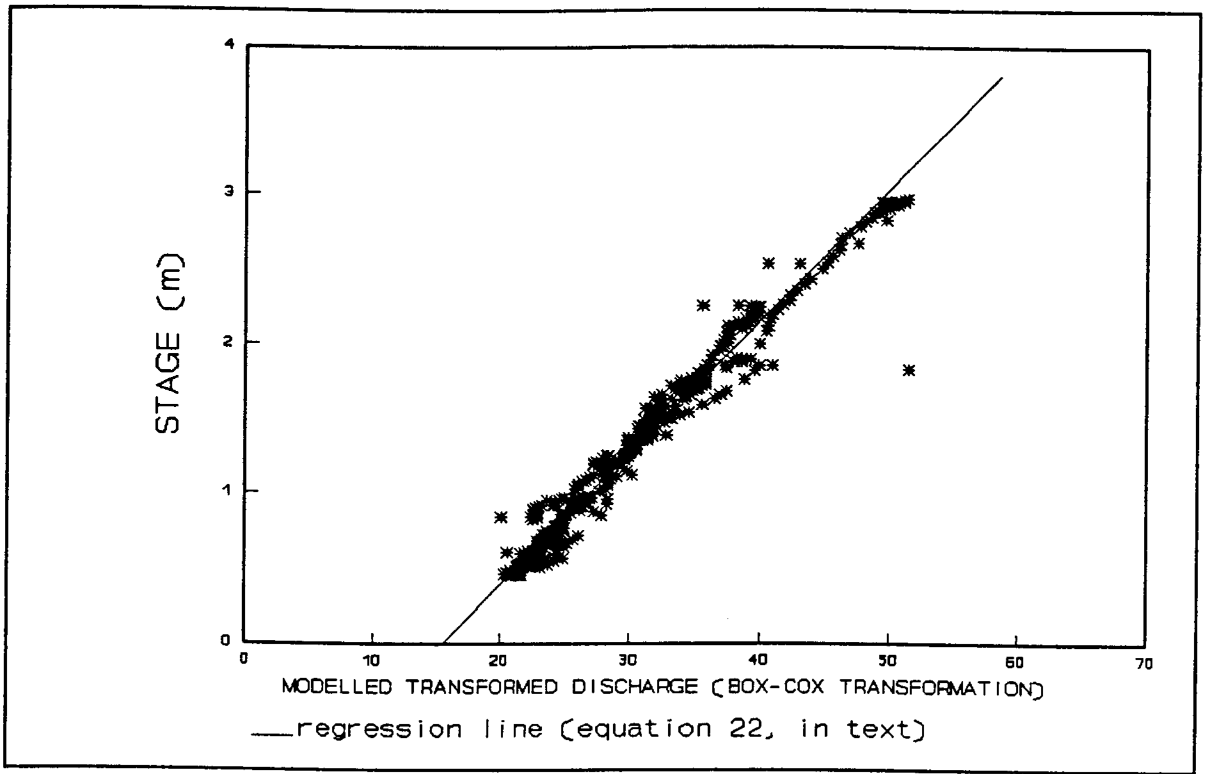

Fig. 18: **Stage discharge relationship - transformed ARIMA modelled discharge (Box-Cox transformation)**

Choice of model

Firstly, it is important to observe that the analyses undertaken here did not remove all the erroneous values of stage and discharge. The treatment of outliers undertaken here only removed those glaring "unreasonable" values that exerted considerable influence on the regression. The remaining data points may be taken as containing "allowable error" (Gawne and Simonovic, 1994). Secondly, it should also be noted that the transformations investigated only provide alternatives to conventional means of fitting an algebraic relation to rating curves. Their value lies in the greater flexibility they provide and for choosing a more appropriate relation, especially in situations as this one where there are doubts about the quality of the data.

It is necessary to look at equations (14) and (15) again. Equation (14) is the rating relation with the basic Box-Cox transformation of discharge while equation (15) gives the relation between the ARIMA model-derived discharge and stage. Each of these equations would suit a different purpose. For example, if the assumption that outliers represent observations with considerable error, and so it is desirable to remove them, then equation (14) would yield an H - Q relationship with considerably less scatter. On the other hand, if the outliers were unusual discharge events, say, flash floods, then it would be desirable to retain then in the H - Q relationship. In that case, the ARIMA-derived relationship would be more suitable.

Table 4: Measured and computed steady-state discharges

DATE	MEASURED DISCHARGE, m^3s^{-1}	COMPUTED STEADY-STATE DISCHARGE	%CHANGE
21/3/85	358.473	357.551	-0.257
22/3/85	367.862	366.885	-0.266
25/3/85	386.355	385.298	-0.274
2/5/85	519.085	518.397	-0.132
3/5/85	521.446	520.835	-0.117
7/5/85	556.386	555.399	-0.177
29/1/86	160.607	159.663	-0.588
3/2/86	177.904	177.103	-0.450
5/2/86	182.563	182.106	-0.250
8/4/86	681.256	677.854	-0.499
11/4/86	707.209	676.760	-4.306
22/11/89	117.586	117.133	-0.385
23/11/89	116.717	116.221	-0.425
6/2/90	218.525	218.125	-0.183
13/290	251.073	249.696	-0.549
15/2/90	263.046	262.356	-0.262
17/2/90	285.991	285.300	-0.242
19/2/90	299.329	298.368	-0.321
20/2/90	304.206	303.324	-0.290
22/2/90	319.187	316.456	-0.856
25/2/90	360.01	359.015	-0.276

Reasons for instability of the rating curve

Rating curve instability, as defined in this study, can result from human and physical causes. From the evidence so far presented, especially the review of the data and the results of the analysis shown in Figs. 2 - 6, a number of observations regarding the reasons for the instability of the rating curve can be made. The gauging site is a stable confined section in which depth of flow is the dominant hydraulic variable controlling discharge magnitude. Although bed movement cannot be completely ruled out, it is difficult to it the large magnitude outliers observed the data series. It seems, therefore, that physical causes can largely be excluded. The implication is that the instability of the rating curve can be attributed to human errors. Some of the evidence in support of this position has already been presented in the review of data. Additional evidence that can be cited include the assumption of linearity of the logarithmic rating curve, the failure to adjust the measured discharge for falling or rising stages, and the use of faulty equipment at times. One can also cite the loss of field sheets (HY1 forms) from which doubtful values can be verified.

CONCLUSION

This study has described the nature of available river gauging data for the Okavango at Mohembo. Many data points were found to be out of the range of possible stage-discharge relationship. By repeated regression and removal of outliers, a relationship was established between hydraulic parameters and discharge that approximately fulfilled the conditions that $f + m + b = 1$ and $a * c * k = 1$. Of three hydraulic parameters investigated, the relationship between surface width and discharge did not conform with an exponential relationship. The relationship between the respective hydraulic parameters and discharge conforms to a confined gauging section.

The relationship between stage and discharge was found to be non-linear and to require a transformation to become linear. The Box and Cox transformation in which $\lambda = 0.5$ was found to be suitable. The use of linear regression and removal of outliers led to the establishment of a rating curve represented by equation (14). It was also found that the series of gauged and stage-estimated daily discharges can be fitted with an ARIMA (3,2,1) model. The quality of the fit obtained means that missing data for this site can be estimated with an ARIMA (3,2,1) model.

Overall, the conclusion can be reached that the cumulative effects of the errors outlined will likely lead to over-estimation of discharges for the Okavango at Mohembo. Reasons for this conclusion can be found in the fact that linear extrapolations (in log-space, hopefully) were used to estimate discharges for some observed stages, when, in fact, as shown by the results of this study, there is curvature in the H - Q plot even in log-space; the values of discharge plotted on the rating curve were not necessarily the steady-state discharges as no adjustment for rising or falling stages were carried out; and, the discrepancy in discharge measurements reported during the gauging exercise undertaken jointly with the South African Department of Water Affairs.

The findings of this study have implications for the water resources potential and other ecological resources. The discharge data of the Okavango at Mohembo constitutes an input into hydrological models of the delta. The ability to represent the hydrological processes in the delta adequately depends to a large extent on the accuracy of the input data. Thus the overestimation of the discharge at Mohembo may lead to an erroneous assessment of such processes in the delta as evapotranspiration, groundwater recharge, and sedimentation.

RECOMMENDATIONS

A number of recommendations can be made from this study as detailed below:
1) Training in hydrometry
The extensive review of the HY1 data sheets in this study led to the conclusion that staff who undertake the gauging exercises require some refresher courses. Areas of particular interest are measurement of cross-section parameters and the adjustment of the computed discharge for falling or rising stage. The latter would require the measurement of water surface slope. This is not done at the moment. This course should also cover techniques of rating curve extension.

2) In view of the findings in this study, it will be necessary to carry out a similar review of other gauging stations. This exercise should be carried out urgently especially as consultants employed by the Department of Water Affairs usually end up modifying the original data to a large extent.

3) As a matter of routine, measured discharges should be continually plotted on the rating curve. That way, in consistencies can be discovered early and new gaugings carried out to rectify the errors.

REFERENCES

Box, G.E.P. and Cox, D.R. (1964): An analysis of transformations. Royal Statistical Society, Series B, 26, p211-252.

Box, G.E.P. and Jenkins, G.M. (1976): Time Series Analysis, Forecasting and Control. Holden Day. (McGraw-Hill Book Co. New York and Maidenhead, England) Revised Edition.

Gawne, K.D. and Simonovic, S.P. (1994): A computer-based system for modelling the stage-discharge relationships in steady state conditions. Hydrological Sciences Journal, Vol.39, No. 5, p487-506.

IUCN (1992): The IUCN Review of the Southern Okavango Integrated Water Development Project.

Kendall, Sir M. and Ord, J.K. (1990): Time Series. Edward Arnold, UK.

Leopold, L. B. and Maddock, T. (1953): The hydraulic geometry of stream channels and some physiographic implication. US Geol. Survey Prof. Paper, 252.

Morsley, P.M. and McKerchar, A.I. (1993): Streamflow. In Maidment, D.R. (ed.) Handbook of Hydrology, McGraw-Hill Inc., New York.

Norušis, M.J. (1990): SPSS/PC+ Statistics 4.0 for the IBM PC/XT/AT and PS/2. SPSS Inc., Chicago.

SMEC (1987): Southern Okavango Integrated Water Development Phase 1. Department of Water Affairs, Government of Botswana.

SPSS Inc. (1990): SPSS/PC+ Trends for the IBM PC/XT/AT and PS/2. SPSS Inc., Chicago.

Wei, W.S. (1993): Time Series Analysis. Addison-Wesley Publishing Company, Inc. Redwood City, California.

Wilson, E.M. (1990): Engineering Hydrology. MacMillan, UK.