Nordic Hydrology, 28 (3), 1997, 189-200

Comparison of Sediment Rating Curves Developed on Load and on Concentration

Margareta B. Jansson
Institute of Earth Sciences, Uppsala Univ., Sweden

A sediment rating curve developed as a linear regression on logged values which is back-transformed must be corrected for the bias introduced by the log transformation. This article shows that the variances are identical for linear regressions based on values of logged load and logged concentration from the same data set. This means that the bias correction factors $10^{1.1513\sigma^2}$ for the back-transformed regressions are equivalent. Therefore a back-transformed log regression based on loads corrected for bias gives identical sediment discharges to a back-transformed log regression on concentrations corrected for bias. Regression equations from gauging stations in two neighbouring basins in Costa Rica confirm this conclusion.

Mean loads for individual discharge classes were plotted on diagrams in log scales to find the points where the sediment rating curve changes direction. When sediment rating curves were developed on logged mean concentrations, water discharge weighted mean concentrations had to be determined in order to produce equations comparable to those on logged mean loads. Consequently, discharge weighted mean concentrations must be used in a plot to determine the change in direction of a sediment rating curve and to check the goodness of fit of a regression developed by any model employing concentration as the dependent variable.

Introduction

A sediment rating curve is a relationship established between sediment concentration, C, and water discharge, Q, so that $C=f(Q)$, or between load, L, and discharge so that $L=f(Q)$. This relationship is in most cases defined as a power equation $L=aQ^b$ where a and b are constants. As sediment concentration or load has a lognormal distribution it has been common practice to logtransform the data to obtain a normal distribution and to develop a linear regression equation on the logarithms using the least-squares method.

$$\log L = b \log Q + \log a \tag{1}$$

where L is load, a, b are constants

The definition of load gives

$$\begin{aligned} \log(\lambda\, C Q) &= \log\lambda + \log C \log Q = b \log Q + \log a \\ \log C &= (b-1)\log Q + \log\left(\frac{a}{\lambda}\right) \end{aligned} \tag{2}$$

C is concentration, λ is conversion factor for load units, a, b are constants in load equation. Eqs. (1) and (2) are then back-transformed to

$$L = aQ^b \quad \text{and} \quad C = \frac{a}{\lambda} Q^{b-1} \tag{3}$$

The two variables yield identical estimates of sediment load at this stage, before the bias correction. Nevertheless, the use of load as a dependent variable has been criticised because Q is included both in the dependent variable and in the independent variable of the regression equation, and therefore gives an increased correlation coefficient (McBean and Al-Nassri 1988). However, the conclusion by McBean and Al-Nassri that the correlation between concentration and discharge is correct and that between load and discharge incorrect has been contradicted (Annandale 1990; Demissie and Fitzpatrick 1990; Gilroy *et al.* 1990; Milhous 1990; Nordin 1990). The possibility that correlations and hence regressions are meaningfully determined between variables containing a common variable was acknowledged (Gilroy *et al.* 1990). What was considered wrong was the incorrect interpretation of a higher correlation for regressions on different response variables, such as log-load and log-concentration, as indicative of a better prediction. But the coefficient of determination, r^2, of a regression on logged loads can be compared with other regressions on logged loads and r^2 of a regression on logged concentrations can be compared with other regressions on logged concentrations to see which one has the highest explained variation. However, regression equations corrected for bias were not discussed in McBean and Al-Nassri (1988). The present article will show that power functions developed as regressions on log-loads and log-concentrations give equal sediment transport also when correction for bias is performed.

It can also be pointed out that r^2 values (regression sum of squares/total sum of squares) are seldom calculated for the regressions back-transformed to power equations and corrected for bias. There is no general rule for what is too low an r^2 for a useful regression equation (Helsel and Hirsch 1992, p. 231). However, load and concentration equations giving equal sediment discharges along their regression lines, are equally good for predicting sediment transport irrespective of r^2.

By definition, a regression should go through the means of the dependent variable. A sediment rating curve developed as a linear regression on logged values which is back-transformed must be corrected for the bias introduced by the log transformation, that is to say caused by the difference in means between back-transformed means of logarithms and means of nonlogged values (Jansson 1996). In this study, the usual practice of log transformation of all individual values to produce a linear regression is used to verify mathematical expressions.

This article will not discuss what types of regressions might be better than the chosen regression model in certain cases. However, a second object of this paper is to demonstrate how mean loads or mean concentrations for individual discharge classes should be used when sediment rating curves are constructed, both for determining where the sediment rating curve changes direction and for checking how well the corrected sediment rating curve fits the data. It will be concluded from the calculations what type of concentration means should be used.

A regression model based on logged mean loads for individual discharge classes can also be applied to develop sediment rating curves and often gives the best regression (Jansson 1985, 1996). This article compares power equations developed using regressions on logged mean loads and on logged mean concentrations by using data from one of the gauging stations employed in the study.

Sediment rating curves for five gauging stations in two neighbouring drainage basins in Costa Rica are used for the verification of compatibility between rating curves developed on log-load and on log-concentration and for the demonstration of the utilization of means. The gauging stations Dos Montañas and Pacuare are located in the Pacuare river basin and Oriente, Angostura and Guayabo in the Reventazón river basin (Fig. 1). Both rivers drain to the Atlantic. The study areas have tropical climates without a dry season. At higher elevations the climates become subtropical. Mean annual rainfall in the basins varies between 1,500 and 7,500 mm. Most rainfall is convective but there is also frontal rain and orographic rain during the season when trade winds are strong. These differences in rainfall conditions and the high intensity of local convective rainfall contribute to great variations in sediment concentration for given values of water discharge at a station. An additional factor contributing to high variation in sediment concentration is the relatively high frequency of landslides and slumps within the basins. In addition, two stations are located below a regulated reservoir and a power station, with Angostura approx. 30 km and Guayabo approx. 40 km below the dam.

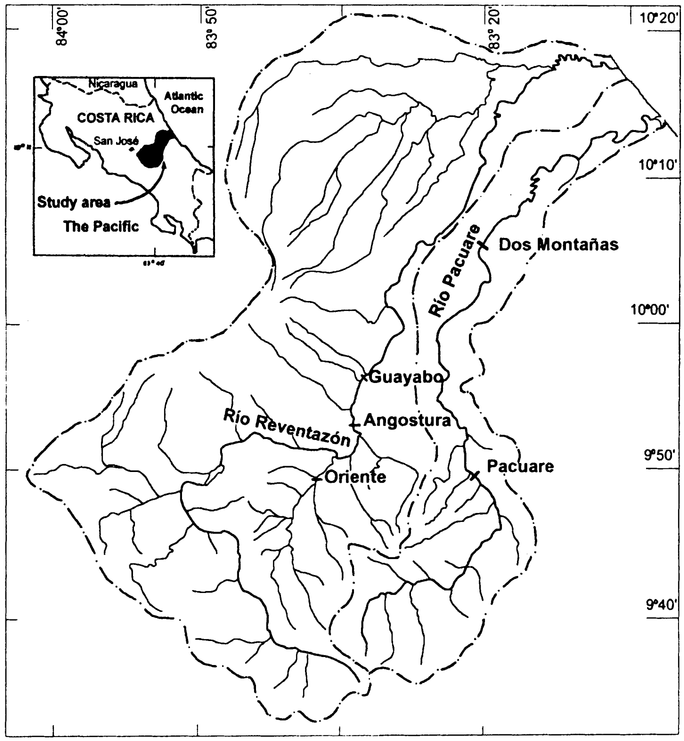

Fig. 1. The Reventazón and the Pacuare drainage basins in Costa Rica. The gauging stations are indicated.

Methods

A mathematical expression is presented which implies equality of sediment rating curves developed as linear regressions on log-transformed load and concentration, and which are back-transformed and corrected for bias. The findings are verified with sediment rating curves from field measurement data at gauging stations in two neighbouring drainage basins in Costa Rica.

The procedures used to develop regression equations in this article for verification of the mathematical implications are as follows:

Water samples were taken for suspended sediment concentration analysis at five gauging stations in the Pacuare and Reventazón basins (Fig. 1). Depth-integrated samples were taken at three profiles in the river. Mean sediment concentration was calculated from the three determined concentrations. If one of the concentrations was an outlier it was excluded from the mean calculation. The water level was recorded at the sampling occasion and a discharge rating curve was used to calculate the water discharge. The discharge rating curves were changed during the year. These data have been used to establish regressions on suspended sediment load in t/day, and on suspended sediment concentration in mg/l. For each measuring station the water discharges have been divided into discharge classes. The definition of the width of the water discharge classes depends on the data base of the station in question. If the data base is small the discharge classes must be wide. At low discharges there may be up to 50-70 data in a discharge class. The data diminish at higher discharge classes and when the data diminish to less than 7-10 data in a class the class width has to be increased. This means that the class width may be changed twice if the data are sparse at high discharge ranges.

Mean sediment load and mean water discharge are calculated for each discharge class. These mean values are not used to develop the regression but mean loads are plotted against mean water discharges in a diagram to see where the row of means and thus the sediment rating curve has a slope change. Regression equations are calculated for the segment between these slope changes. This demonstration how to utilize means when sediment rating curves are to be constructed is a second aim of this article. However, the regression model used in this article is that which is used most commonly for sediment rating curves, namely log-transformation of all values of L or C and Q, a linear regression of the log-transform, and back-transformation to a power equation (*e.g.* Nilsson 1971; Loughran 1976; Walling 1977; Ferguson 1986; Crawford 1991; Helsel and Hirsch 1992).

The back-transformed equations are then corrected for bias (Jansson 1985; Ferguson 1986; Jansson 1991; Helsel and Hirsch 1992). The correction factor used in this study is $10^{1.1513\sigma^2}$ where σ^2=variance of the log regression.

The power function based on logged concentrations in mg/l is multiplied by 0.0864 Q to get a power function for load in t/day. This load equation derived from that for C is compared with the equation calculated directly from the load data.

All decimal places provided by the computer program are indicated in the tables to demonstrate that the values are identical and that differences are not hidden by rounding. By definition a regression curve should go through the arithmetic mean loads at every water discharge. For practical purposes a check of the goodness of fit of the curve can be made, superimposing it on a plot of the arithmetic mean loads for individual discharge classes. To check whether the regression lines fit the mean loads, the curves of the corrected regressions developed on all values are plotted on a diagram which also shows the arithmetic mean loads in discharge classes. Again, this will illuminate the second aim of this article, the one of the utilization of means when sediment rating curves are developed.

The compatibility of equations developed on logged means in discharge classes is also tested but only with data from the Oriente gauging station. Equations developed as regressions on logged mean loads, logged mean concentrations, and on logged mean water discharge weighted concentrations in discharge classes are compared.

Results

The variance of a logged load regression is

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}(\log L_i - \log L'_i)^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left((\log C_i + \log Q_i) - (\log C'_i + \log Q'_i)\right)^2 =$$

$$\frac{1}{N-1}\sum_{i=1}^{N}(\log C_i - \log C'_i)^2 \qquad (4)$$

where L' is the load value of the regression equation. $L'=C'\,Q'$. $Q' = Q$ for all values.

Consequently, the expression of the variance of the linear regression on logged loads equals the variance of the regression on logged concentrations. The identical variances of the two variables imply identical correction factors as the correction factor used is $10^{1.1513\sigma^2}$. The equivalence of variances and correction factors are verified using data from the gauging stations.

The sediment rating curve of the station Dos Montañas is analysed first. The data base is indicated in Fig. 2. The data base has been sorted according to water discharge and is divided into water discharge classes. Mean water discharge and mean sediment load have been calculated for each discharge class, as shown in Fig. 2.

As the rating curve should go through the means, Fig. 2 shows where the row of means changes direction. Based on changed direction of means, the data base of Dos Montañas is divided into three groups of data for the construction of three regression lines. Regressions are developed on all separate data in each group, not on the means (Table 1).

The original concentration data and the same division into groups of data are used for the three regressions on logged concentration, C, for Dos Montañas (Table 1 continued).

As can be seen from Table 1 the variances of the regressions on logged loads are identical with the variances of the regressions on logged concentrations (columns 3 and 8) and the correction factors (columns 5 and 10) are equal. Consequently, the regression equations developed on log-loads corrected for bias are identical with the corresponding load equations developed as regressions on log-concentrations (columns 6 and 12). It is also interesting to compare r^2 of the log regressions. Sometimes the difference in r^2 between load and concentration regressions is great, sometimes it isn't.

The sediment rating curves of the stations in Table 1, except for Dos Montañas (Fig. 3), are indicated in Fig. 4 in order to check how well the curves fit the means. As can be seen from Fig. 4 the rating curve at Oriente goes too high compared with the means at the highest discharges. The curve of Angostura does not fit the data very well either. However, the choice of regression model is not the subject of this study and is discussed elsewhere (Jansson 1996).

Table 1 – Comparison of sediment rating curves at five stations in the Pacuare and Reventazón river basins. Regressions developed on loads.

Station 1	Back-transformed regr. equation on load 2	σ^2 of log regr. 3	r^2 of log regr. 4	Correction factor 5	Corrected regr. equation 6
Dos Montañas	L_1=0.062302 $Q^{2.001551}$	0.054226	0.66	1.154596	L_1=0.071933 $Q^{2.001551}$
	L_2=0.000388 $Q^{3.366771}$	0.090546	0.88	1.271292	L_2=0.000493 $Q^{3.366771}$
	L_3=6.585015 $Q^{1.778356}$	0.002323	0.95	1.006178	L_3=6.625698 $Q^{1.778356}$
Pacuare	L_1= 0.107896 $Q^{1.823471}$	0.067046	0.58	1.194512	L_1= 0.128883 $Q^{1.823471}$
	L_2= 0.0000029 $Q^{4.820294}$	0.213848	0.74	1.762803	L_2= 0.0000051 $Q^{4.820294}$
Oriente	L_1= 0.193232 $Q^{1.611973}$	0.078557	0.48	1.231522	L_1= 0.237969 $Q^{1.611973}$
	L_2= 0.0000035 $Q^{4.725815}$	0.249995	0.52	1.94008	L_2= 0.0000068 $Q^{4.725815}$
	L_3= 0.018138 $Q^{2.806526}$	0.17387	0.49	1.585536	L_3= 0.028758 $Q^{2.806526}$
Angostura	L_1= 0.00011 $Q^{3.444674}$	0.262671	0.73	2.00638	L_1= 0.000221 $Q^{3.444674}$
	L_2= 0.320742 $Q^{2.072513}$	0.070564	0.60	1.205705	L_2= 0.38672 $Q^{2.072513}$
Guayabo	L_1= 0.409055 $Q^{1.363165}$	0.082544	0.33	1.244609	L_1= 0.509114 $Q^{1.363165}$
	L_2= 0.000000057 $Q^{4.983622}$	0.192951	0.70	1.667801	L_2= 0.000000095 $Q^{4.983622}$
	L_3= 0.162233 $Q^{2.278721}$	0.056074	0.47	1.160266	L_3= 0.188233 $Q^{2.278721}$

Table 1 – Regressions developed on concentration.

Back-transformed regr. equation on concentration 7	σ^2 of log regr. 8	r^2 of log regr. 9	Correction factor 10	Corrected regr. equation 11	Corrected regr. equation on conc. times 0.0864 Q 12
C_1= 0.721086 $Q^{1.001551}$	0.054226	0.33	1.154596	C_1= 0.832562 $Q^{1.001551}$	L_{C_1}= 0.071933 $Q^{2.001551}$
C_2= 0.004486 $Q^{2.366771}$	0.090546	0.78	1.271292	C_2= 0.005704 $Q^{2.366771}$	L_{C_2}= 0.000493 $Q^{3.366771}$
C_3= 76.21545 $Q^{0.778356}$	0.002323	0.79	1.006178	C_3= 76.68632 $Q^{0.77838}$	L_{C_3}= 6.625698 $Q^{1.778356}$
C_1= 1.248795 $Q^{0.823471}$	0.067046	0.22	1.194512	C_1= 1.4917 $Q^{0.823471}$	L_{C_1}= 0.128883 $Q^{1.823471}$
C_2= 0.000034 $Q^{3.820294}$	0.213848	0.64	1.762803	C_2= 0.000059 $Q^{3.820294}$	L_{C_2}= 0.0000051 $Q^{4.820294}$
C_1= 2.236481 $Q^{0.611973}$	0.078557	0.12	1.231522	C_1= 2.754276 $Q^{0.611973}$	L_{C_1}= 0.237969 $Q^{1.611973}$
C_2= 0.00004 $Q^{3.725815}$	0.249995	0.41	1.94008	C_2= 0.000078 $Q^{3.725815}$	L_{C_2}= 0.0000068 $Q^{4.725815}$
C_3= 0.20993 $Q^{1.806526}$	0.17387	0.28	1.585536	C_3= 0.332851 $Q^{1.806526}$	L_{C_3}= 0.028758 $Q^{2.806526}$
C_1= 0.001276 $Q^{2.444674}$	0.262671	0.57	2.00638	C_1= 0.00256 $Q^{2.444674}$	L_{C_1}= 0.000221 $Q^{3.444674}$
C_2= 3.712288 $Q^{1.072513}$	0.070564	0.29	1.205705	C_2= 4.475922 $Q^{1.072513}$	L_{C_2}= 0.38672 $Q^{2.072513}$
C_1= 4.734437 $Q^{0.363165}$	0.082544	0.03	1.244609	C_1= 5.892524 $Q^{0.363165}$	L_{C_1}= 0.509114 $Q^{1.363165}$
C_2= 0.00000066 $Q^{3.983622}$	0.192951	0.60	1.667801	C_2= 0.0000011 $Q^{3.983622}$	L_{C_2}= 0.000000095 $Q^{4.983622}$
C_3= 1.877698 $Q^{1.278721}$	0.056074	0.22	1.160266	C_3= 2.178628 $Q^{1.278721}$	L_{C_3}= 0.188233 $Q^{2.278721}$

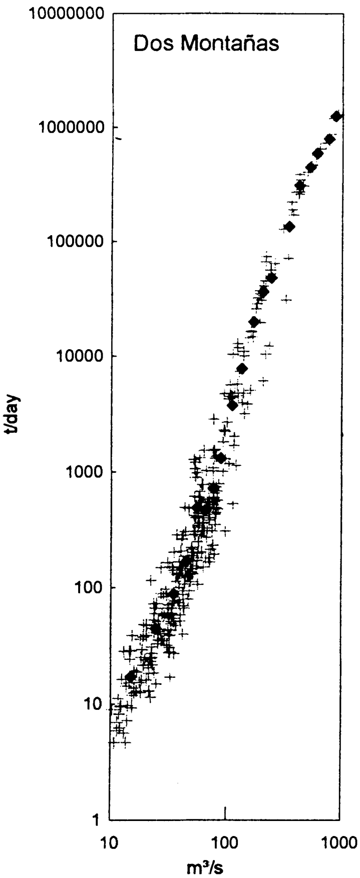

Fig. 2. Dos Montañas. Measurements made during 1970-1993. Samples are also taken during the passage of some highwater events in 1993-94. Mean loads for discharge classes.

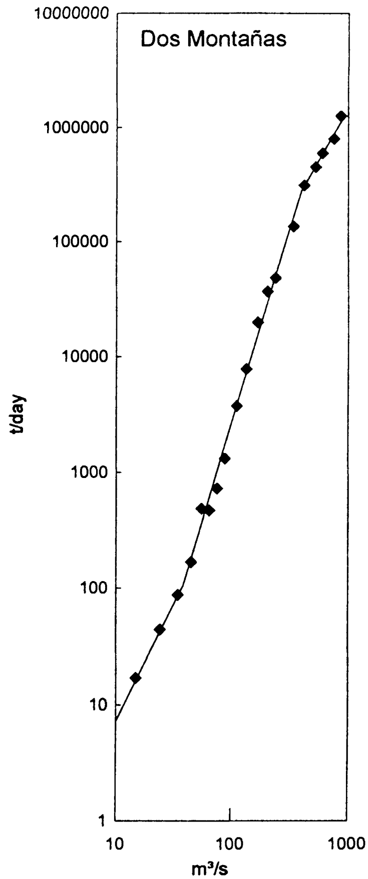

Fig 3. Dos Montañas. Mean load in water discharge classes. Back-transformed log regressions corrected for bias by the correction factor $10^{1.1513\sigma^2}$

The compatibility of equations developed on logged means is tested on data from the Oriente gauging station. Power equations developed as regressions on logged mean loads in discharge classes, $\log L_M$, logged mean concentrations, $\log C_M$, and logged mean discharge weighted concentrations, $\log C_{MD}$, are indicated in Table 2. The squared correlation coefficients are high both for log-mean load and log-mean concentration regressions. As can be seen, regressions on mean discharge weighted concentrations give comparable equations as regressions on mean loads. Consequently, water weighted means must be used when a sediment rating curve is developed on mean concentrations. Moreover, water weighted mean concentrations must be used in a plot for determining the change in direction of a rating curve on concentration and for checking the goodness of fit of regressions developed by any model

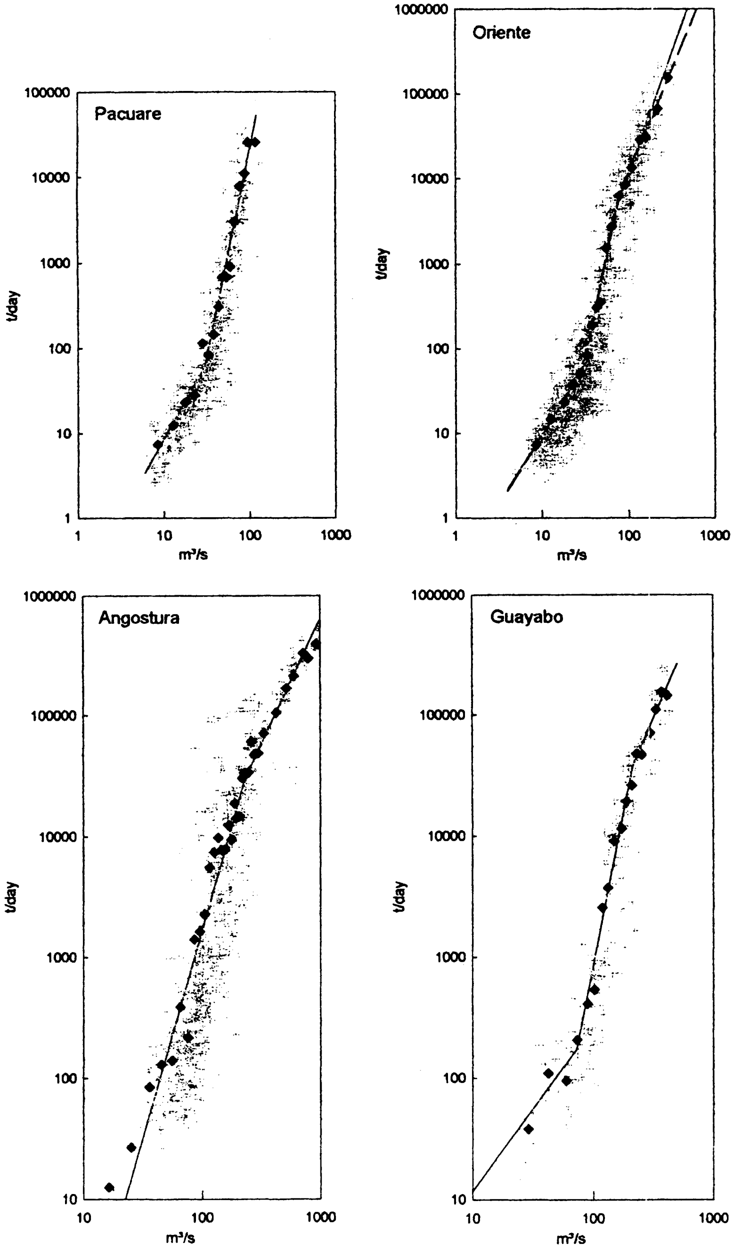

Fig. 4. Sediment rating curves developed on all separate values (crosses). Here these sediment rating curves are compared with the arithmetic mean values of load within water discharge classes. A second sediment rating curve developed on mean loads for individual discharge classes is marked on the diagram of Oriente (hatched line).

Table 2 – Comparison of sediment rating curves at Oriente developed on logged mean loads L_M, logged mean concentrations C_M, and on logged mean water discharge weighted concentrations C_{MD}.

Back-transformed regr. equation	σ^2 of log regr.	r^2 of log regr.	Correction factor	Corrected regr. equation	Corrected regr. equation on conc. times 0.0864 Q
Mean loads					
$L_M = 0.207281\ Q_M^{1.665349}$	0.000431	0.997	1.001143	$L_M = 0.207519\ Q_M^{1.665349}$	
$L_M = 0.0000041\ Q_M^{4.864309}$	0.012109	0.98	1.03262	$L_M = 0.0000042\ Q_M^{4.864309}$	
$L_M = 0.153305\ Q_M^{2.433066}$	0.002923	0.99	1.007778	$L_M = 0.154497\ Q_M^{2.433066}$	
Mean concentrations					
$C_M = 2.3576\ Q_M^{0.668294}$	0.000428	0.98	1.001137	$C_M = 2.36028\ Q_M^{0.668294}$	$L_{C_M} = 0.203928\ Q_M^{1.668294}$
$C_M = 0.000048\ Q_M^{3.856057}$	0.011835	0.97	1.03187	$C_M = 0.00005\ Q_M^{3.856057}$	$L_{C_M} = 0.0000043\ Q_M^{4.856057}$
$C_M = 1.693442\ Q_M^{1.440704}$	0.002672	0.97	1.007109	$C_M = 1.70548\ Q_M^{1.440704}$	$L_{C_M} = 0.147353\ Q_M^{2.440704}$
Mean disch. weighted conc.					
$C_{MD} = 2.399093\ Q_M^{0.665349}$	0.000431	0.98	1.001143	$C_{MD} = 2.401837\ Q_M^{0.665349}$	$L_{C_{MD}} = 0.207519\ Q_M^{1.665349}$
$C_{MD} = 0.000047\ Q_M^{3.864309}$	0.012109	0.97	1.03262	$C_{MD} = 0.000049\ Q_M^{3.864309}$	$L_{C_{MD}} = 0.0000042\ Q_M^{4.864309}$
$C_{MD} = 1.774359\ Q_M^{1.433066}$	0.002923	0.97	1.007778	$C_{MD} = 1.788161\ Q_M^{1.433066}$	$L_{C_{MD}} = 0.154497\ Q_M^{2.433066}$

employing concentrations as the dependent variable. A sediment rating curve based on loads can easily be converted to a rating curve on concentration if concentration diagrams are preferred for other purposes.

The sediment rating curve based on mean loads (Table 2) is indicated as a hatched line in the diagram Oriente of Fig. 4. The regressions on means are not corrected for bias (*cf* Jansson 1996).

Conclusions

Using the formula for the variance, Eq. (4), it was demonstrated that the variances of regressions on logged loads and logged concentrations are equivalent. This means that the bias correction factors $10^{1.1513\sigma^2}$ for the back-transformed regressions are equivalent. Regressions developed as linear regressions on logged loads and on logged concentrations with data from five gauging stations confirm that the variances are identical and that the correction factors of the back-transformed regressions are equivalent. That also means that regressions developed on logged loads and on logged concentrations, both of which are back-transformed and corrected for bias, give equal values of sediment load. The equations based on load in t/day are identical to the equations based on concentration (mg/l) times 0.0864 Q.

In order to determine where the sediment rating curve has a slope change, mean load and mean water discharge for individual discharge classes were calculated and plotted on a diagram in log scales. Based on changed directions of the plotted means, the data base was divided into groups of data for which regression equations were calculated.

When regression equations were developed on logged mean values for individual discharge classes, water discharge weighted mean concentrations had to be used in order to obtain equations comparable to the regressions developed on logged mean loads. Consequently, water weighted mean concentrations for discharge classes must also be used in a plot for the determination of changes in direction of a sediment rating curve on concentration.

For the purpose of calculation of sediment transport it is unimportant whether a relationship based on log-loads or on log-concentrations is used. A regression line should go through the arithmetic means of the variable and in practice the curve can be checked against the plot of means to see if the curve fits the means and in what range it does not fit the means. However, the goodness of fit of a sediment rating curve developed on concentrations must be checked against values of discharge-weighted mean concentrations in the discharge classes. When performing this check, mean loads are less time consuming to calculate than mean water weighted concentrations. However, for purposes other than sediment transport calculation, plots of concentration are often preferable to plots of load.

References

Annandale G. W. (1990) Uncertainty in suspended sediment transport curves. Discussion, *Journal of Hydraulic Engineering, Vol. 116*(1), pp. 140-141.

Crawford, C. G. (1991) Estimation of suspended-sediment rating curves and mean suspended-sediment loads, *J. Hydrol., Vol. 129*, pp. 331-348.

Demissie, M., and Fitzpatrick, W. P. (1990) Uncertainty in suspended sediment transport curves, Discussion, *Journal of Hydraulic Engineering, Vol. 116*(1), pp. 141-143.

Ferguson, R. I. (1986) River loads underestimated by rating curves, *Water Resour. Res., Vol. 22*(1), pp. 74-76.

Gilroy, E. J., Kirkby, W. H., Cohn, T. A., and Glysson, G. D. (1990) Uncertainty in suspended sediment transport curves. Discussion, *Journal of Hydraulic Engineering, Vol. 116*(1), pp. 143-145.

Helsel, D. R., and Hirsch, R. M. (1992) *Statistical Methods in Water Recources.* Developments in Water Science Series, Elsevier, Amsterdam.

Jansson, M. B. (1985) A comparison of detransformed logarithmic regressions and power function regressions, *Geogr. Ann., Vol. 67A*(1-2), pp. 61-70.

Jansson, M. B. (1996) Estimating a sediment rating curve of the Reventazón river at Palomo using logged mean loads within discharge classes, *J. Hydrol., Vol. 183*, pp. 227-241.

Loughran, R. J. (1976) The calculation of suspended-sediment transport from concentration v. discharge curves: Chandler river, NSW, *Catena, Vol. 3*, pp. 45-61.

McBean, E. A., and Al-Nassri, S. (1988) Uncertainty in suspended sediment transport curves, *Journal of Hydraulic Engineering, Vol. 114*(1), pp. 63-74.

Milhous, R. T. (1990) Uncertainty in suspended sediment transport curves. Discussion, *Journal of Hydraulic Engineering, Vol. 116*(5), pp. 730-732.

Nilsson, B. (1971) Sediment transport i svenska vattendrag. Ett IHD-projekt. Del 1: Metodik (in Swedish), *UNGI Rapport 4*, Uppsala.

Nordin, C. F. (1990) Uncertainty in suspended sediment transport curves. Discussion, *Journal of Hydraulic Engineering, Vol. 116 (1)*, pp. 145-148.

Walling, D. E. (1977) Limitations of the rating curve technique for estimating suspended sediment loads, with particular reference to British rivers, *IAHS Publ., Vol. 122*, pp. 34-48.

Received: 17 February, 1997
Revised: 7 July, 1997
Accepted: 15 July, 1997

Address:
Institute of Earth Sciences,
Uppsala University,
Norbyvägen 18 B,
S-752 36 Uppsala,
Sweden.