Sediment and the Environment (Proceedings of the Baltimore Symposium, May 1989)
IAHS Publ. no. 184, 1989.

Developing accurate and reliable stream sediment yields

Krishan P. Singh & Ali Durgunoğlu
Illinois State Water Survey, 2204 Griffith Drive, Champaign, Illinois 61820-7495, U.S.A.

ABSTRACT Quantitative estimates of stream sediment yields have usually been made by using empirical relations. The most commonly used relation, or rating curve, is the power function, $S = a\,Q^b$, where S = sediment load, and Q = discharge. The parameters a and b are conventionally estimated from least-squares regression analysis on logarithms of S and Q. This leads to considerable underestimation of annual sediment yields. Estimating a and b by using nonlinear regression can reduce such bias. Other factors which lead to incorrect estimation of sediment yields are seasonality, nonlinear rating curves, use of average daily flows, and high ratio of peak to daily average flows in small drainage basins. Some remedies are identified for improving stream sediment yield estimates by reducing biases and errors caused by various factors.

INTRODUCTION

The understanding of hydrological processes and prediction of hydrological variables are important for better management of our water resources. Because of the complexity of hydrological processes, predictions of most hydrological variables are obtained by statistical analyses of short-term and discrete data by using empirical relations. For example, the sediment yield from a drainage basin is related to temporal and spatial distribution of precipitation, soil properties, land use, and geomorphic properties of the drainage basin, but it is extremely difficult to combine all the relevant variables into one theoretical relation. As a result, the sediment yield from a drainage basin is usually estimated by developing empirical sediment rating curves from instantaneous records of stream discharges and sediment concentrations, and then using them with the flow-duration curves.

The use of empirical relations for estimation of sediment yields provides information which can be used in the design of economically feasible reservoirs and efficient sediment control structures. However, sediment yield estimates obtained by using the conventional methods are substantially less than observed (Osterkamp, 1976; Cheetham & Wilke, 1976). This would lead to faulty economic evaluation and building of reservoirs which may not be economically or environmentally justified. Some of the factors which cause underestimation of the sediment yields, as well as the range of the underestimation will be illustrated. Techniques for improving the estimates will also be provided.

RATING CURVE

If instantaneous sediment concentration or sediment load S, and stream discharge Q, are available at sample intervals of Δt, then the total sediment load, L, for a period T can be calculated by

$$L = \sum_{i=1}^{T/\Delta t} S_i \, \Delta t \tag{1}$$

However, for accurate estimates of L, Δt may need to vary from hours to days for various basin sizes, whereas only average daily streamflow data are usually available. On the other hand, most sediment load data are collected periodically, usually on a weekly or monthly basis. As a result the total load L is estimated by replacing the unknown sediment loads, S_i, by estimates $\hat{S}_i$, which are based on the rating curve between observed values of S and Q. In this study, the total sediment load L is estimated by the flow duration curve method. This involves the summation of sediment loads as shown in Equation 1 over a number of discharge bands, from high to low, weighted by the frequency of their occurrence. The sediment load estimates used in Equation 1 are obtained from a rating curve.

The rating curve is almost always expressed as a power function of the form $S = a\,Q^b$ where S is the sediment load and Q is the discharge. The parameters a and b are usually estimated by the ordinary least squares regression of the log-transformed variables S and Q. In this case the sum of squares of the residuals

$$SSR = \sum_{i=1}^{N} (\log S_i - \log \hat{S}_i)^2 \tag{2}$$

of the log-transformed sediment loads is minimised. The sum of the log-transformed residuals $\Sigma(\log S_i - \log \hat{S}_i)$ is zero, but since the least squares regression was done in the logarithmic space, the sum of the untransformed residuals $\Sigma(S_i - \hat{S}_i)$ is always positive, unless the regression is a perfect fit, leading to considerable underestimation of the annual sediment yields. This underestimation is proportional to the magnitude of the mean square error

$$MSE = \frac{1}{N-2} \sum_{i=1}^{N} (\log S_i - \log \hat{S}_i)^2 \tag{3}$$

of the log-transformed regression. Several forms of bias correction factors are suggested for correcting the coefficient a (Ferguson, 1986; Koch & Smillie, 1986; Jansson, 1985), but the exponent b is not modified. All these bias correction factors vary with the mean square error of the log-transformed regression. As an example, Ferguson (1986), gives the corrected value of a as

$$a' = a\, e^{2.65\, MSE} \tag{4}$$

However, a nonlinear least squares regression procedure for estimating the parameters α and β in the relation $S = \alpha\, Q^{\beta}$ will obviate the bias present in log-transformed regression.

Moreover, the sum of squares of the untransformed residuals $SSR^* = \Sigma(S_i - \hat{S}_i)^2$ is found to be consistently lower with nonlinear regression than with log-transformed linear regression. As an example, streamflow and sediment load data for Salamonie River near Warren, Indiana (Crawford & Mansue, 1988) (Fig. 1), and Elkhorn Creek near Frankfort, Indiana (Flint, 1983) are analyzed by using log-transformed and nonlinear regression, and Ferguson's method. Results for Salamonie River which drains 1100 km^2, and Elkhorn Creek which drains 1225 km^2 are shown in Table 1. The results indicate that the SSR^* of the untransformed sediment loads obtained with nonlinear regression estimates is much lower than the SSR^* obtained with log-transformed linear regression estimates. With the flow duration curve method, the annual sediment yields obtained with log-transformed regression for Salamonie River, and Elkhorn Creek are only 53 and 57 percent of the values computed by using nonlinear regression estimates, respectively.

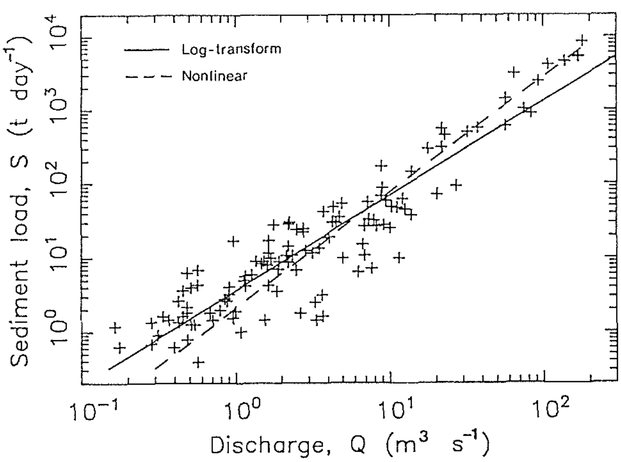


Fig. 1 Streamflow versus sediment load data for Salamonie River near Warren, Indiana.

Table 1 Comparison of results obtained by log-transformed regression, Ferguson's method, and nonlinear regression

Stream	Method	Parameters		L (t $year^{-1}$)	SSR^*
Salamonie R.	Log-transform	a=3.60	b=1.28	42 460	7.27 10^7
N=118	Ferguson's	a′=5.27	b=1.28	62 200	3.90 10^7
MSE=0.144	Nonlinear	α=2.01	β=1.57	80 880	1.13 10^7
Elkhorn Cr.	Log-transform	a=0.38	b=1.61	34 500	3.38 10^7
N=35	Ferguson's	a′=0.52	b=1.61	46 800	1.53 10^7
MSE=0.115	Nonlinear	α=0.17	β=1.90	60 780	5.60 10^5

SEASONALITY

There are other factors that contribute to incorrect estimation of sediment yields. For example, sediment load and stream discharge relation for various seasons tend to differ substantially due to seasonal patterns of precipitation, land cover, and land use. In the Midwest region of the U.S.A. for example, sediment loads for the months April-June are higher than for the period July-March. The runoff in the latter season, which is generally from low-intensity rainfall and snowmelt, or from rainfall over basin soils deficient in soil moisture, causes very little or no erosion and low sediment concentrations. During the April-June period, on the other hand, the ground is usually disturbed due to agriculture, and the runoff is caused by precipitation in the form of intense storms over rather saturated soils, leading to higher sediment concentrations. Seasonality of sediment transport is illustrated in Figure 2 for East Nishnabotna River at Red Oak, Iowa (Schuetz & Wilbur, 1977), which drains an area of 2315 km^2.

If seasonal behaviour is thought to exist, then determinations for the different seasons needs to be done on a region-by-region basis. Seasonality in the data should be investigated, especially if the mean square error of the regression is high. The most practical solution to this problem would be to separate the data into seasons, fit different sediment rating curves for each season, and estimate yields by using seasonal flow-duration curves.

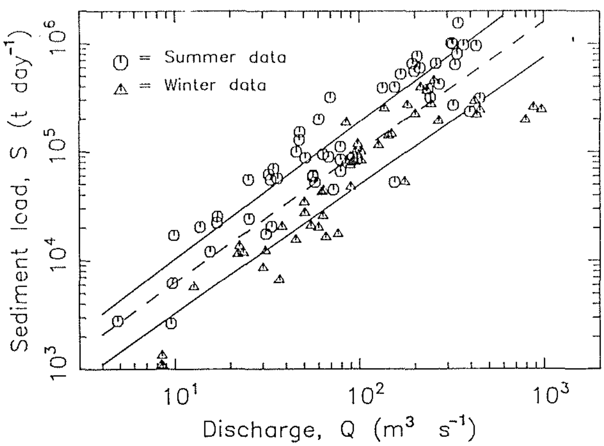


Fig. 2 Seasonal streamflow versus sediment load curves for East Nishnabotna River at Red Oak, Iowa.

SHARP BREAKS IN RATING CURVES

In some cases high mean square errors do not necessarily imply seasonality. For example the plot of Q versus S data for Rapid Creek near Iowa City, Iowa (Schuetz & Wilbur, 1977), is shown in Figure 3. The scatter diagram clearly indicates an obvious break in the rating curve. This nonlinear behaviour is difficult to estimate by using

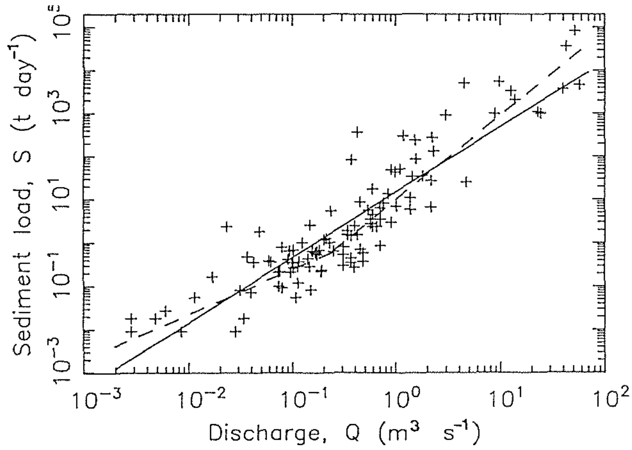


Fig. 3 Nonlinear behaviour indicated by the sediment load versus streamflow data for Rapid Creek near Iowa City, Iowa.

conventional regression methods because the point of inflection is not known a priori.

Underestimation of sediment loads in this situation is a function of the amount of deflection in the slope. If the changing slope is ignored and a straight line regression is estimated (shown by solid line), the sediment loads are underestimated at the lower and higher ends of the flow spectrum. Underestimation at the higher end is especially important because this region contributes a very substantial portion of the total sediment load in a given period. This condition can be remedied by fitting two regression lines to the data, intersecting at the inflection point. This, however, requires an optimisation procedure for determining the optimum inflection point. Another alternative is to fit a curvilinear rating curve in the form of $S = \alpha\, Q^{\beta+\gamma Q}$. The parameters α, β, and γ can be estimated by using a nonlinear regression method.

AVERAGE DAILY FLOW

Quite often, sediment loads are estimated from sediment rating curves by using average daily flows, rather than instantaneous flows. This almost always results in underestimation of the sediment load. The level of underestimation is directly proportional to the variation of discharge during the day, and to the magnitude of the exponent b in the sediment rating equation. The variation in discharge during a day is usually related to the size of the drainage basin. With large basins (2000 to 2500 km^2), the ratio of average daily flow Q_{av} to peak flow Q_p (r_p) can vary between 0.9 to 0.6, but for small basins (40 to 50 km^2), r_p can be as low as 0.1. Thus, for small drainage basins the runoff hydrograph rises and falls within a few hours. The sediment load for streams with small drainage basins can be drastically underestimated if average daily flows are used since the sediment load S is not a linear function of Q.

The effect of using average daily flows for the estimation of sediment loads can be analyzed by using three imaginary basins, and appropriate values for b. It is assumed that r_p values vary from 0.83 to 0.6 for the first basin (large), from 0.63 to 0.33 for the second basin (medium), and from 0.18 to 0.10 for the third basin (small). For the first basin, the flow peaks and returns to its initial level within 24 hours. For medium and small basins these periods are 12 and 4 hours, respectively. The total daily sediment loads are calculated for both the true flow hydrograph and the average daily flow by

$$L = \int_{t=0}^{24} a\, Q^b\, dt \qquad (5.a) \quad ; \qquad L' = \int_{t=0}^{24} a\, Q_{av}^b\, dt \qquad (5.b)$$

respectively, and the percentage underestimation is given by [100 (L-L')/L].

The results, shown in Figure 4, indicate that the underestimation is small for large basins (8%), whereas it may be as high as 68% for small basins.

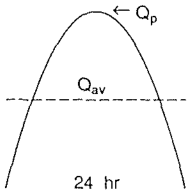
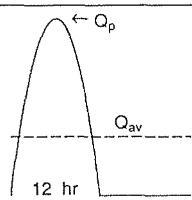
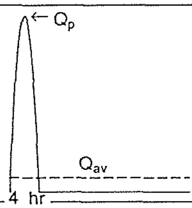
Basin Size	Hydrograph Shape	Underestimation = [100 (L-L')/L]		
		b = 1.3	b = 1.5	b = 1.7
Large		(0.3-3.0)	(0.5-5.0)	(1.0-8.0)
Medium		(1.2-12)	(2.5-22)	(4.0-31)
Small		(15-35)	(27-54)	(40-68)

Fig. 4 Underestimation of daily sediment loads for different scenarios.

CONCLUSIONS

The bias in the estimation of sediment loads by rating curve due to using log-transformed estimates can be significantly reduced by using nonlinear regression. Further improvements can be achieved by identifying seasonalities and breaks in slopes of the rating curves. Finally, the underestimation caused by using average daily flows with the rating curve can be eliminated by using sub-daily flow data, if available.

The sources of underestimation in sediment yields identified in this paper are not exhaustive. The problems associated with the rising and falling stages of the flow hydrograph (hysteretic effect), armouring effects due to depletion of erodible material after successive storms, distribution and movement of rainfall over a drainage basin, and the combined effects of all these factors, need to be dealt with. The main problem that concerns both the log-transformed and the nonlinear regression methods is that a very limited number of sediment observations are made during high flows. Considering that 80-90% of the sediment load is carried during the highest 10-15% of the flows, it is essential that more sediment samples be collected at high flows.

ACKNOWLEDGEMENTS The authors wish to thank Ganapathi S. Ramamurthy and Nandakishore Rajagopalan for their help in analysing data.

REFERENCES

Crawford, C.G. & Mansue, L.J. (1988) Suspended-sediment characteristics of Indiana streams, 1952-84. USGS Open-File Report 87-527, Indianapolis, Indiana, USA.

Cheetham, R.N. Jr. & Wilke, R.F. (1976) Sedimentation in Birch Lake, Iowa County, Wisconsin. In: Proceedings of the Third Federal Inter-Agency Sedimentation Conference. Denver, Colorado, USA.

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

Flint, R.F. (1983) Fluvial sedimentation in Kentucky. USGS Water Resources Investigations Report 83-4152, Louisville, Kentucky, USA.

Jansson, M. (1985) A comparison of detransformed logarithmic regressions and power function regressions. Geografiska Annaler 67 A(1-2), 61-70.

Koch, R.W. & Smillie, G.M. (1986) Bias in hydrologic prediction using log-transformed regression models. Wat. Resour. Res. 22(5), 717-723.

Osterkamp, W.R. (1976) Variation and causative factors of sediment yields in the Arkansas River Basin, Kansas. In: Proceedings of the Third Federal Inter-Agency Sedimentation Conference. Denver, Colorado, USA.

Schuetz, J.R., & Wilbur, J.M. (1977) Fluvial sediment data for Iowa: suspended-sediment concentrations, loads and sizes; bed-material sizes; and reservoir siltation. Iowa Geological Survey Technical Information Series 6, Des Moines, Iowa, USA.