Discharge indices for water quality loads

Richard M. Vogel

Department of Civil and Environmental Engineering, Tufts University, Medford, Massachusetts, USA

Jery R. Stedinger

School of Civil and Environmental Engineering, Cornell University, Ithaca, New York, USA

Richard P. Hooper[1]

U.S. Geological Survey, Northborough, Massachusetts, USA

Received 29 November 2002; revised 21 May 2003; accepted 22 July 2003; published 1 October 2003.

[1] Effective discharge has been used to describe the streamflow level that is responsible for transporting the most sediment over the long term. Careful inspection reveals that this concept may not have been well defined, and different interpretations have led to conflicting representations. Because total load is ultimately the quantity of interest, we define a new index, the half-load discharge, which is that discharge above and below which half the total long-term load is transported. The value of the half-load discharge is derived for a reasonable model of flows and constituent concentration. The effective discharge has generally been thought to be a relatively common or frequent flood. The half-load discharge is generally a much greater and less frequent flow than commonly used estimators of the effective discharge. Relations provided here for the frequency and magnitude of the half-load discharge provide evidence that it is relatively rare floods that transport most of the sediment over the long term. These ideas apply to other constituents as well. *INDEX TERMS:* 1815 Hydrology: Erosion and sedimentation; 1824 Hydrology: Geomorphology (1625); 1860 Hydrology: Runoff and streamflow; 1871 Hydrology: Surface water quality; *KEYWORDS:* effective discharge, transport, sediment, constituents, rating curve, half-load

Citation: Vogel, R. M., J. R. Stedinger, and R. P. Hooper, Discharge indices for water quality loads, *Water Resour. Res.*, *39*(10), 1273, doi:10.1029/2002WR001872, 2003.

1. Introduction

[2] Since the pioneering article by *Wolman and Miller* [1960], hundreds of studies have investigated the relations between river sediment loads and the magnitude and frequency of discharge that give rise to those sediment loads for rivers around the world (see *Sichingabula* [1999] for a list of citations). These studies have employed the concept of "effective" discharges Q_e, defined as the range of river discharges that transport most of the sediment in the long term. *Wolman and Miller* [1960] argued that the amount of sediment transported by river flows of a given magnitude depends on the form of the relationship between river discharge and sediment as well as on the form of the frequency distribution of the river discharges.

[3] The notion of effective discharge as a useful and compelling concept is evidenced by over 300 refereed journal articles citing the study by *Wolman and Miller* [1960]. *Andrews and Nankervis* [1995] argue that the conceptual model introduced by *Wolman and Miller* [1960] describing the influence of flow magnitude and frequency on the relative sediment-transporting effectiveness of various discharges in natural channels has become one of the fundamental paradigms of geomorphology.

[4] The site-to-site variability in the probability distribution of river discharge and the relationship between discharge and sediment concentrations have led investigators to a wide variety of conclusions regarding the frequency and magnitude of the effective discharge. For example, estimation of Q_e for a wide range of U.S. rivers led *Benson and Thomas* [1966] to conclude that this discharge is exceeded 12% of the time, or about 44 days per year. *Wolman and Miller* [1960], *Andrews* [1980], and *Leopold* [1994] concluded that the effective discharge is closer to the "bankfull" discharge that occurs roughly once or twice per year. More recently, *Sichingabula* [1999] found that Q_e is exceeded between 0.02 and 19.6% of the time or between 7 and 72 days per year. However, *Kirchner et al.* [2001] found that it is the very rare catastrophic erosion events that dominate the long-term sediment yield and not the incremental yet frequent erosion events.

[5] The lack of agreement on the frequency of Q_e led *Nash* [1994] to introduce an analytical approach to evaluating the overall behavior and the recurrence interval associated with the effective discharge. The mathematical definition of Q_e introduced by *Nash* [1994] is shown here to be quite different from the estimators introduced by *Wolman and Miller* [1960]. Furthermore, we show that the value of Q_e estimated using the mathematical expressions introduced by *Nash* [1994] generally exceeds the value of

[1]Now at Consortium of Universities for the Advancement of Hydrologic Science, Inc., Washington, DC, USA.

0043-1397/03/2002WR001872$09.00

Q_e based on the approach suggested by *Wolman and Miller* [1960].

[6] The primary objectives of this paper are to (1) provide a rigorous mathematical definition and justification of several discharge indices including Q_e, (2) extend the definition of these discharge indices to all constituent loads, and (3) evaluate the theoretical behavior of these indices. Relations between the frequency and magnitude of the discharge indices are derived, and an application is provided to the Susquehanna River at Harrisburg, Pennsylvania.

2. Transport Effectiveness of River Discharge

[7] *Wolman and Miller* [1960] introduced the concept of work done by sediment as the product of the amount of sediment carried by a given flow and the frequency of that flow. We term that work done more generally as the transport effectiveness because it represents the mass load of any contaminant carried by a particular value of streamflow. If we let L be load and Q be discharge, then the expected load can be defined as

$$\mu_L = \int_0^{\infty} E[L|Q = q] f_Q(q) dq, \tag{1}$$

where $f_Q(q)$ is the probability density function (pdf) of streamflow Q and $E[L|Q = q]$ is the conditional mean load given $Q = q$, where q denotes a particular value of discharge. *Parker and Troutman* [1989] employed a similar conditional analysis in their investigation of the probability distribution of annual sediment loads. The mean load $E[L]$ is the integral of an expression, which we term the transport effectiveness $e(q)$, defined as

$$e(q) = E[L|Q = q] f_Q(q). \tag{2}$$

The mathematical definition of transport effectiveness in equation (2) is analogous to the definition of transport effectiveness introduced by *Wolman and Miller* [1960], which was

$$e_{WM}(q) = (\text{sediment carried by flow})(\text{frequency of flow}),$$

which can be written mathematically as

$$e_{WM}(q) = L(q) f_Q(q). \tag{3}$$

When one's interest is in the long-term contaminant load, the concept of transport effectiveness introduced by *Wolman and Miller* [1960] appears useful because it represents the net load in units of mass.

[8] The definition in equation (3) is quite different from the definition of transport effectiveness introduced by *Nash* [1994], which was

$$e_N(x) = L(x) f_X(x), \tag{4}$$

where $f_X(x)$ represents the pdf of the logarithms of streamflow $X = \ln(Q)$. A physical basis for use of Nash's definition of transport effectiveness is provided in section 3.

3. Effective Discharge for Lognormal Streamflows and Power Law Rating Curves

[9] Let effective discharge Q_e be that streamflow which maximizes the long-term transport effectiveness $e(q)$. The expressions for transport effectiveness derived in section 2 ($e(q)$, $e_{WM}(q)$, and $e_N(X)$) can be differentiated with respect to discharge to determine their maxima. Analyses of this type have been performed in hundreds of sediment transport studies. By making a few simplifying assumptions about the load-discharge relationship and the probabilistic behavior of discharges, we can make general statements about the magnitude and frequency of observing the effective discharge that cannot be made by empirical studies. We follow *Nash* [1994] and assume that daily river discharge Q follows a two-parameter lognormal (LN2) distribution so that

$$f_Q(q) = \frac{1}{q\sigma_x\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln(q) - \mu_x}{\sigma_x}\right)^2\right], \tag{5}$$

where $X = \ln(Q)$ and μ_x and σ_x are the mean and standard deviation, respectively, of X. The lognormal assumption plays a central role in the following analyses; section 3.1 provides further background.

3.1. Properties of Daily Streamflow in the United States

[10] Computation of daily sediment, nutrient, and other constituent loads requires information regarding daily streamflow. Extensive statistical analyses have been performed on series of annual maximum, minimum, and average streamflows (*Vogel and Wilson* [1996] provide a review), yet few studies have summarized the behavior of daily streamflow across wide regions. Using over 25 million observations of daily streamflow at 1571 gauged rivers in the United States, *Limbrunner et al.* [2000] used L moment diagrams to show that daily streamflows are well approximated by both a two- and three-parameter lognormal distribution. *Vogel and Fennessey* [1993] found that ordinary product moment estimates of the coefficient of variation and skewness of daily streamflow are remarkably downward biased and should not be employed even with very long time series of daily streamflow. Instead, estimates of the coefficient of variation of daily streamflow C_Q were obtained using L moment estimators for a three-parameter lognormal distribution. It was necessary to fit a three-parameter lognormal distribution in this case because many sites exhibit zero daily streamflows and the logarithm of zero is undefined. L moment estimators of C_Q, like ordinary moment estimators, are bound to exhibit bias but are not bounded above like ordinary moment estimators of C_Q. Figure 1 summarizes the relation between sample estimates of the coefficient of variation of daily streamflow C_Q and drainage area for the entire United States, the northeastern water resource region 1, and California region 18. Regions 1 and 18 correspond to the water resource regions developed by the U.S. Water Resource Council. Across the United States the median value of C_Q is 10 with an interquartile range from 3 to 33. For humid regions in the northeast, C_Q

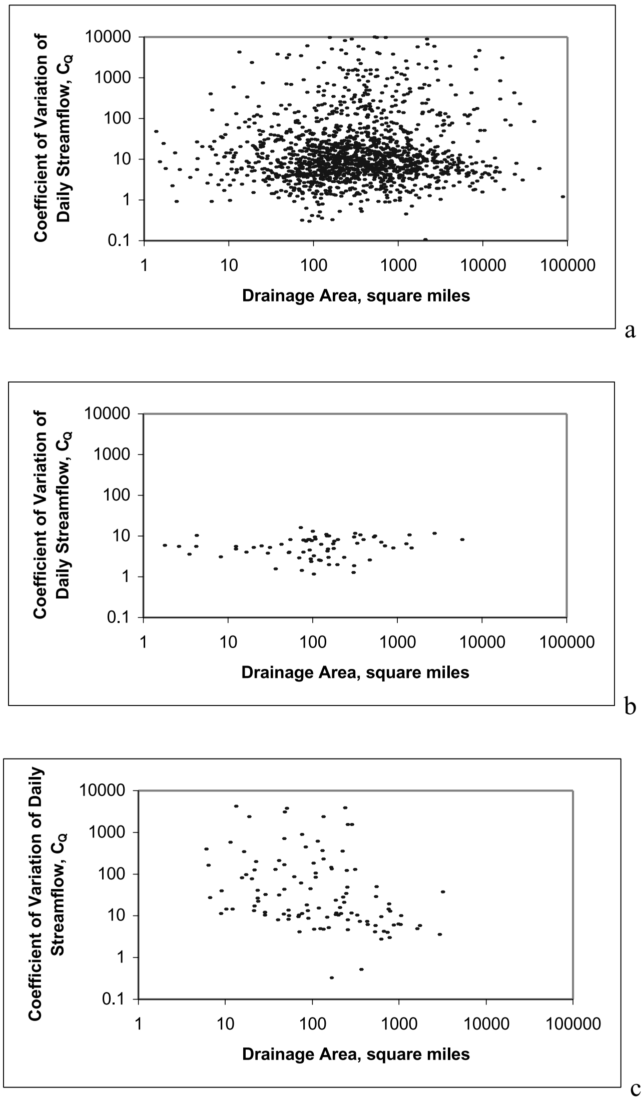


Figure 1. Coefficient of variation of daily streamflow C_Q for (a) all regions of the United States; (b) region 1, northeast; and (c) region 18, California.

does not seem to be scale-dependent, whereas in arid regions such as in region 18, C_Q tends to increase as drainage area decreases. The extremely large values of C_Q correspond to intermittent rivers that include zero flows.

3.2. Theoretical Expressions for Effective Discharge

[11] Using well-known transformations for the lognormal distribution, the mean and variance of the logarithms are related to the mean and variance of the flows via

$$\mu_x = \ln\left[\frac{\mu_Q}{\sqrt{1+C_Q^2}}\right] \tag{6a}$$

$$\sigma_x^2 = \ln\left(1 + C_Q^2\right), \tag{6b}$$

where μ_Q and C_Q are the mean and coefficient of variation, respectively, of the daily streamflows in real space. If a sample estimate of C_Q is desired, either L moments or equation (6b) should be used to obtain $\hat{C}_Q = \sqrt{\exp(s_x^2) - 1}$, where s_x^2 is a sample estimate of variance of $x = \ln(q)$.

[12] As in studies by *Cohn et al.* [1992], *Nash* [1994], *Syvitski et al.* [2000], and many others, we assume contaminant loads can be described using the power law model

$$L = e^a Q^b e^{\varepsilon} = \exp(a + bX + \varepsilon), \tag{7}$$

where $X = \ln(Q)$, L has units of mass per day, a and b are model parameters, and the ε are assumed to be normally distributed model errors with zero mean and constant variance σ_ε^2. The use of a regression model that relates streamflow to load is termed a "rating curve" method. This rating curve is probably the most widely used approach for estimating both sediment and nutrient concentrations from continuous streamflow records, and it has proven satisfactory for a wide range of rivers and constituents (see *Cohn et al.* [1992] and *Cohn* [1995] for review).

[13] Since load L is the product of flow Q and concentration C, the power law rating curve in equation (7) implies that $C = \exp[a + (b - 1)X + \varepsilon]$. Dissolved constituents tend to dilute hence for such constituents $b - 1 < 0$ or $b < 1$ [*O'Connor*, 1976]. Suspended constituents, including nutrients and sediment, tend to exhibit values of $b > 1$. For example, *Nash* [1994] reported values of b corresponding to suspended sediment load relations at 55 rivers across the United States that ranged from 1.2 to 3.0 with a median value of $b = 1.76$. Similarly, *Syvitski et al.* [2000] reported values of b corresponding to suspended sediment load relations for 48 rivers in North America that are approximately normally distributed with a mean of 2.15 and standard deviation of 0.42. *Rudolph* [2002] reports values of b of 0.78, 1.15, 1.25, and 1.32 on the Grand River, Honey Creek, Maumee River, and Sandusky River, respectively, for phosphorus load-discharge relationships. If the bivariate power law model given in equation (7) is not adequate, a multivariate power law model can usually be fit that accounts for seasonal variations and time trends [*Cohn et al.*, 1992] as well as hysteresis effects [*House and Warwick*, 1998].

[14] Using the two assumptions (1) LN2 probability distribution of streamflows and (2) the power law load-discharge model, the transport effectiveness of *Wolman and Miller* [1960] can be compared analytically with that of *Nash* [1994]. Figure 2 illustrates the transport effectiveness for the case where the mean daily flow $\mu_Q = 1$, the coefficient of variation of the daily flows $C_Q = 5$, $a = 0$, and $b = 1.5$ in equation (7). The transport effectiveness e curves computed using the two different approaches can lead to very different conclusions regarding both e and the discharge corresponding to the peak value of the transport effectiveness, that is, the effective discharge Q_e. Taking the derivative of e_N and e_{WM} with respect to Q and setting it equal to zero leads to expressions for the effective discharge Q_e based on *Nash*'s [1994] approach,

$$Q_e(\text{Nash}) = \exp\left(\mu_x + b\sigma_x^2\right), \tag{8a}$$

and based on *Wolman and Miller*'s [1960] approach,

$$Q_e(\text{Wolman \& Miller}) = \exp\left(\mu_x + (b-1)\sigma_x^2\right). \tag{8b}$$

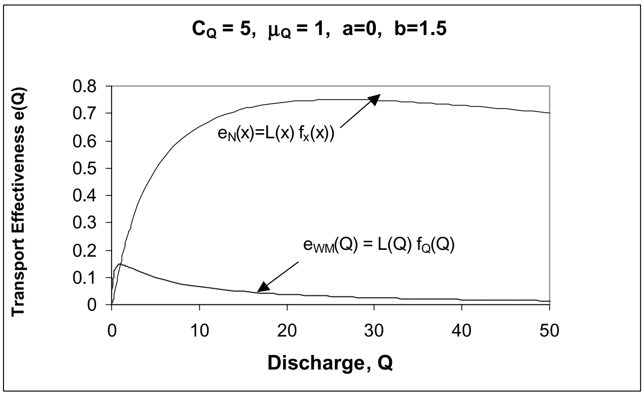


Figure 2. Relation between transport effectiveness $e(q)$ and discharge Q for a particular case.

[15] Note that Q_e(Nash) $\geq$ Q_e(Wolman & Miller) and often is considerably so. For the case illustrated in Figure 2, Q_e(Nash) = 26, whereas Q_e(Wolman & Miller) = 1, with daily exceedance probabilities of 0.0033 and 0.18, respectively. Each of these definitions of effective discharge has different interpretations because each discharge corresponds to the maximum of a different definition of transport effectiveness. Over all possible discharges Q and associated discharge intervals $Q \pm \Delta$ (where Δ is some small change in discharge), the Wolman and Miller effective discharge estimator transports the maximum load. Similarly, over all discharges associated with the ranges defined by $[Q(1 - \Delta'),\ Q(1 + \Delta')]$ (where Δ' is a fraction of discharge), the Nash effective discharge transports the maximum load. Neither of these interpretations is as easy to understand as the interpretation that follows.

4. Half-Load and *f* Load Discharge for Lognormal Streamflows and Power Law Rating Curves

[16] The half-load discharge $Q_{1/2}$ is defined as that value of discharge above and below which half the long-term sediment load is transported. A very simple estimator of this quantity, in practice, would be obtained by ordering all the daily flows by magnitude and adding up the corresponding sediment loads in order. When the sum of the sediment carried reaches half the total, we have reached the half-load discharge. Unfortunately, such a simple estimator can work only if flows on the order of magnitude of the half-load discharge and larger are well represented in the historical record so that the sediment load-frequency relation is well represented by the empirical experience. As will be demonstrated, this is not commonly the case because the half-load discharge is a very infrequent event.

[17] Under the distributional and power law assumptions described above, we want to estimate $Q_{1/2}$ from

$$\int_0^{Q_{1/2}} E[L|Q=q] f_Q(q) dq = \frac{\mu_L}{2}, \tag{9}$$

where the mean load μ_L is defined in equation (1). The conditional expectation in equations (1) and (9) is $E[L|Q = q] = \exp(a + b \ln(q))\exp(\sigma_\varepsilon^2/2) = e^a q^b e^{\sigma_\varepsilon^2/2}$, where σ_ε^2 is the variance of the residuals in equation (7) and the factor $\exp(\sigma_\varepsilon^2/2)$ represents the bias introduced from the retransformation of the power law model [*Ferguson*, 1986]. Substitution of this and the pdf in equation (5) into equations (1) and (9) leads to the general expression for cumulative load

$$\int_0^Q e^a q^b \frac{1}{q\sigma_x\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln(q)-\mu_x}{\sigma_x}\right)^2\right] dq = \mu_L \Phi\left[\frac{\ln(Q)-\mu_x-b\sigma_x^2}{\sigma_x}\right], \tag{10a}$$

where

$$\mu_L = \exp\left(a + b\mu_x + b^2\sigma_x^2/2 + \sigma_\varepsilon^2/2\right),$$

where $x = \ln(q)$ and Φ is the cumulative distribution function for a standard normal variable. Combining equations (9) and (10a) yields the needed equation for $Q_{1/2}$:

$$\Phi\left[\frac{\ln(Q_{1/2})-\mu_x-b\sigma_x^2}{\sigma_x}\right] = \frac{1}{2}. \tag{10b}$$

This leads to the following closed-form equation for the half-load discharge for use with lognormally distributed daily discharges:

$$Q_{1/2} = \exp\left(\mu_x + b\sigma_x^2\right), \tag{10c}$$

which is identical to the expression for effective discharge introduced by *Nash* [1994]. Previously, the interpretation of Nash's expression for effective discharge was that it corresponded to the value of the logarithm of discharge that maximizes the long-term contaminant load. We have shown in section 3 that for a two-parameter lognormal model, Nash's definition of effective discharge is also that discharge above and below which half the long-term load is transported. We believe the concept of half-load discharge is more interpretable than the definition given by *Nash* [1994]; hence we drop further reference to Nash's index so that in the remainder of this paper Q_e always refers to *Wolman and Miller*'s [1960] index.

[18] The concept of half-load discharge is easily generalized to the *f* load discharge by replacing ½ in equation (10b) with the generalized fraction *f*. The *f* load discharge is that discharge above which a fraction *f* of the long-term load is transported. It is given by

$$Q_f = \exp\left[\mu_x + \sigma_x\left(b\sigma_x + z_f\right)\right], \tag{11}$$

where z_f is the percentile of a standard normal variable with exceedance probability *f*. Solving equation (11) for the fraction of load *f* carried by discharges in excess of Q_f leads to the expression

$$f = 1 - \Phi\left[\frac{\ln(Q_f)-\mu_x-b\sigma_x^2}{\sigma_x}\right]. \tag{12}$$

5. Relationships Among Discharge Indices

[19] The lognormal transformations given in equation (6) can be combined with the discharge indices to yield

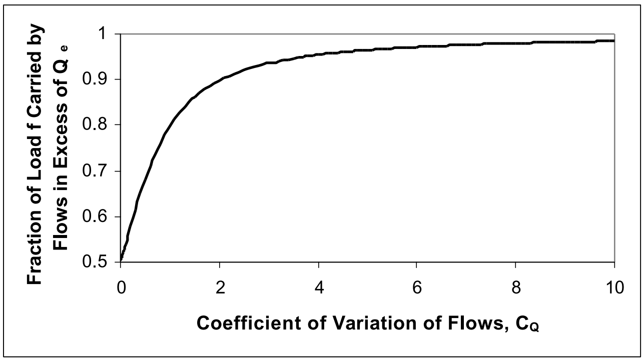


Figure 3. The fraction of load f carried by streamflows in excess of the effective discharge Q_e as a function of the coefficient of variation of daily flows C_Q.

$$Q_{1/2} = \mu_Q\left[1 + C_Q^2\right]^{b-0.5} \tag{13a}$$

$$Q_e = \mu_Q\left[1 + C_Q^2\right]^{b-1.5}. \tag{13b}$$

In general, the ratio of these two discharge indices is

$$\frac{Q_{1/2}}{Q_e} = 1 + C_Q^2 \tag{14}$$

so that the difference between these two indices depends on streamflow variability alone and not the parameters of the load-discharge relation. The fraction of the load carried by discharges greater than Q_e may be found by setting Q_f equal to the effective discharge Q_e, which leads to

$$f = 1 - \Phi\left[-\sqrt{\ln\left(1 + C_Q^2\right)}\right], \tag{15}$$

where $\Phi(\ )$ denotes the cumulative distribution function of a standard normal variable. The expression in equation (15) is illustrated in Figure 3. Note that f in equation (15) does not depend on b. Interestingly, for sites with C_Q in excess of 4, over 95% of the long-term load is carried by discharges in excess of the effective discharge. For watersheds with $C_Q > 4$, which are quite common, the effective discharge is not a very useful index for describing the ability of the river to transport loads over the long term. This is because the effective discharge is simply the maximum of effectiveness rather than its integral, which is the more meaningful quantity describing the cumulative total mass of sediment transported for flows up to some level.

6. Average Return Period of Discharge Indices

[20] Because daily streamflow is assumed to be lognormal, the average return period T of a particular streamflow q, in years, is given by

$$T(q) = \frac{1/365}{1 - \Phi\left[\frac{\ln(q) - \mu_x}{\sigma_x}\right]}, \tag{16}$$

where $\Phi(\ )$ represents the cumulative distribution function of a standard normal variable. Because daily streamflows exhibit significant persistence and seasonality, one should be careful when interpreting the average return period $T(q)$. It can be correctly understood as the average time between exceedances of the streamflow level q, or one can consider that the fraction of flows that will exceed the half-load discharge is $1/T(q)$ in the long run.

[21] Combining equation (16) with the various discharge indices defined in sections 4 and 5 leads to average return periods (in years) of the effective discharge, the half-load discharge, and the f load discharge:

$$T(Q_e) = \frac{1/365}{1 - \Phi\left[(b-1)\sqrt{\ln\left(1 + C_Q^2\right)}\right]}, \tag{17a}$$

$$T\left(Q_{1/2}\right) = \frac{1/365}{1 - \Phi\left[b\sqrt{\ln\left(1 + C_Q^2\right)}\right]}, \tag{17b}$$

$$T\left(Q_f\right) = \frac{1/365}{1 - \Phi\left[z_f + b\sqrt{\ln\left(1 + C_Q^2\right)}\right]}, \tag{17c}$$

respectively. Figure 4 illustrates the average return period (in years) corresponding to $Q_{1/2}$ for values of C_Q in the range 0–25 for a range of values of b. In general, the average return period is >1 year for $b > 1$ and <1 year for $b < 1$. For sediment and nutrient transport problems in which b is normally greater than unity, the half-load discharge can easily exhibit return periods of several decades or even centuries for most commonly observed values of C_Q. Recall that typical values of the exponent b for suspended sediment loads range from 1.2 to 3 [*Nash*, 1994; *Syvitski et al.*, 2000]. Thus, for a common case in which $b = 2$ and $C_Q = 10$, Figure 4 gives a return period of about 300 years for the half-load discharge. For larger

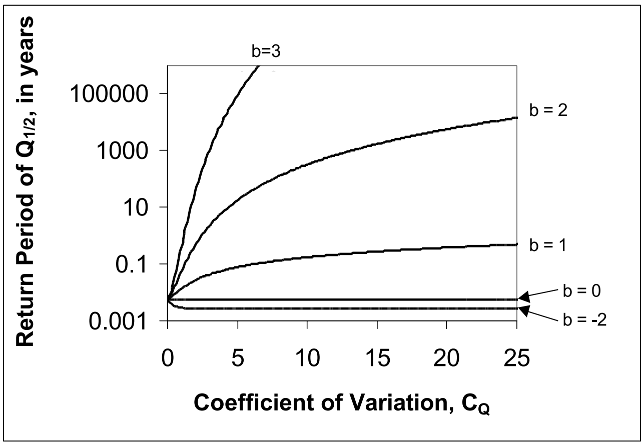
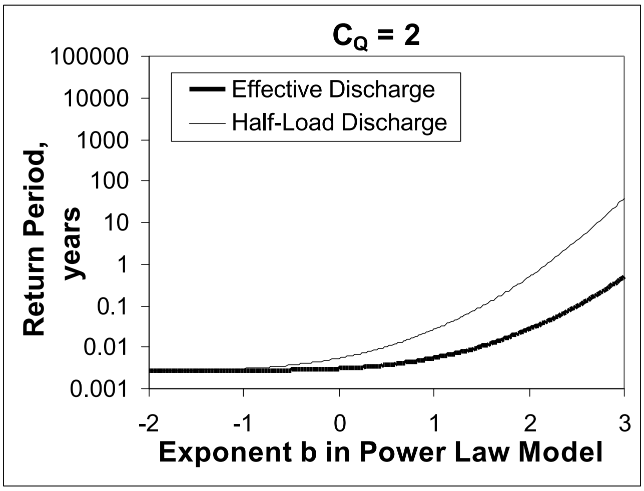
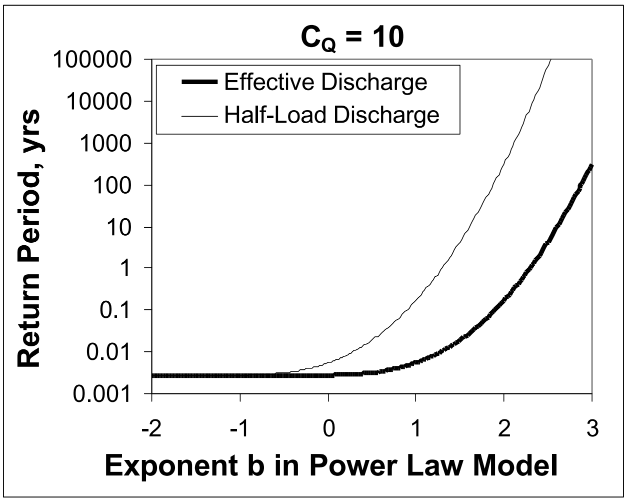


Figure 4. The average time between exceedances of the half-load discharge $Q_{1/2}$, in years, as a function of the coefficient of variation of streamflow C_Q and the exponent b in the power law model.



(a)



(b)

Figure 5. The average time, in years, between exceedances of the half-load discharge $Q_{1/2}$ and the effective discharge Q_e as a function of the exponent b in the power law model for two different values of coefficient of variation of streamflow: (a) $C_Q = 2$ and (b) $C_Q = 10$.

values of C_Q and/or b the half-load discharge can easily have a return period in excess of 1000 years. In such cases the need for a theoretical model such as the one introduced here becomes essential because the empirical approach provides only information regarding a small fraction of the long-term load.

[22] Figure 5 compares the average return period of Q_e and $Q_{1/2}$. Naturally, because Q_e is always less than $Q_{1/2}$, $T(Q_e)$ is always less than $T(Q_{1/2})$. For commonly observed values of C_Q in the range of 2–10, Figure 5 shows that the average return period associated with the effective discharge is generally <1 year, except for situations when C_Q and b are both large. For values of $b > 3$ and values of $C_Q > 10$, the return period of the effective discharge is a rare flood. Figure 5 demonstrates that the dramatic variations associated with the return period of the effective discharge found by previous investigators is to be expected. Such variations are largely due to variations in the values of C_Q and b.

[23] Figure 6 illustrates the average return period corresponding to the f load discharge Q_f as a function of the exponent b for the case when $C_Q = 5$. We observe that for the case when $C_Q = 5$ and $b > 2$, over 80% of the load is carried by flows that have average return periods of 1 year or greater.

7. Water Quality Monitoring Network Design

[24] Each curve in Figure 6 may also be thought of as a water quality monitoring design curve for a particular contaminant because each curve illustrates how much of the long-term load is carried by daily flows of various recurrence intervals. This allows one to determine the type of flows that must be monitored in order to monitor a certain fraction of the nutrient or contaminant load. However, one must keep in mind that our analyses focus on the long-term constituent load. Further research is needed to develop discharge indices that characterize the more common constituent loads that occur in most seasons or years, which might be more relevant to water quality management.

[25] In the context of the design of a sampling strategy it is interesting to note that *Gilroy et al.* [1990, Appendix p. 2077] found that to minimize the root-mean-square error of the estimated total annual load, the average of the logarithms of the sampled discharges should equal $\log(Q_{1/2})$. Thus *Gilroy et al.* [1990] show that the half-load discharge is indeed a critical value around which one should sample to get an accurate estimate of annual loads. Their result is a mathematical one related to the minimization of sampling error; the definition of and interpretation given here to the half-load discharge illustrate why this quantity is indeed critical: It is critical because it is at the center of the distribution of constituent loads.

8. Application to the Susquehanna River

[26] In this section, we illustrate the application of these ideas to the Susquehanna River at Harrisburg, Pennsylvania,

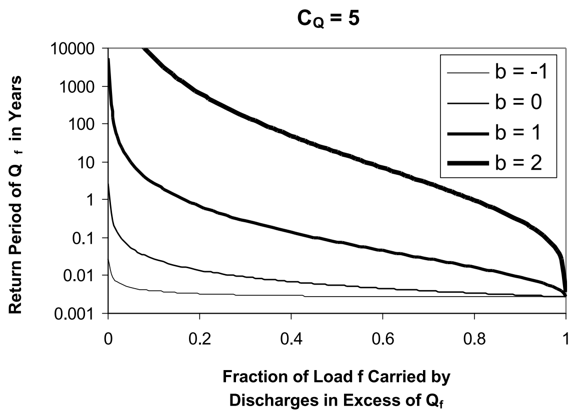


Figure 6. The average time, in years, between exceedances of the half-load discharge $Q_{1/2}$ as a function of the fraction of the long-term load f carried by discharges in excess of the f load discharge Q_f.

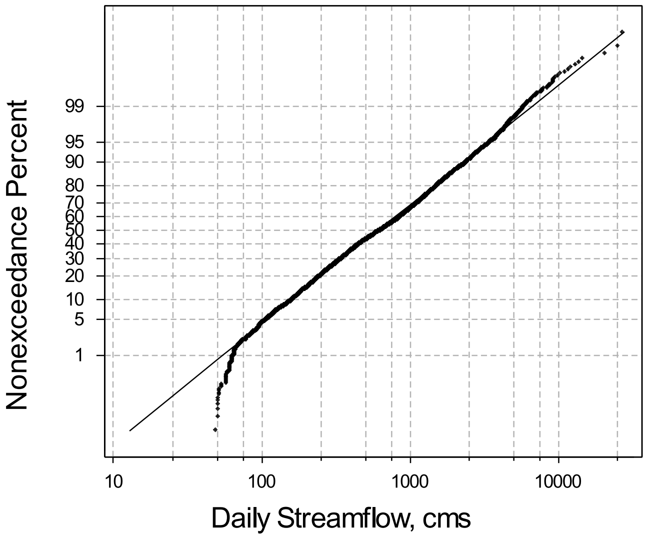


Figure 7. Lognormal probability plot of the average daily discharges on the Susquehanna River based on the record from 1962 to 1981.

a 55,425 km^2 basin. The river was gauged continuously over the period 1890–2001, and 5626 suspended sediment measurements are available for the period 1962–1981. Figure 7 uses a probability plot to illustrate that a two-parameter lognormal distribution provides an excellent fit to the 19-year (1962–1981) sequence of average daily streamflows for this site, with the exception of the very low flows, which have negligible impact on this analysis. Figure 8 illustrates the power law relation between daily suspended sediment load and daily streamflow along with the fitted model $L = 0.00613Q^{1.84}$, where L is load in mg/d and Q is streamflow in m^3/s. The correlation between ln(L) and ln(Q) is 0.943.

[27] Figure 9 illustrates the empirical and theoretical relations between the fraction of the cumulative load f and discharge Q. Here the empirical relation is computed by simply ordering the flows and the corresponding measured sediment loads and plotting the ordered flows versus the cumulative fraction of the total load transported over this

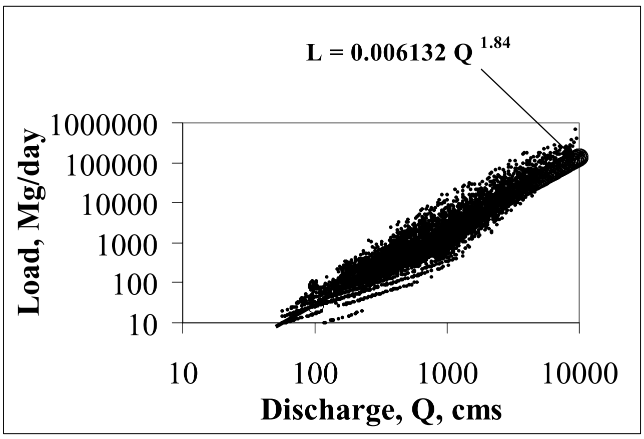


Figure 8. The relation between the suspended sediment load and river discharge for the Susquehanna River based on 5626 observations for the period 1962–1981.

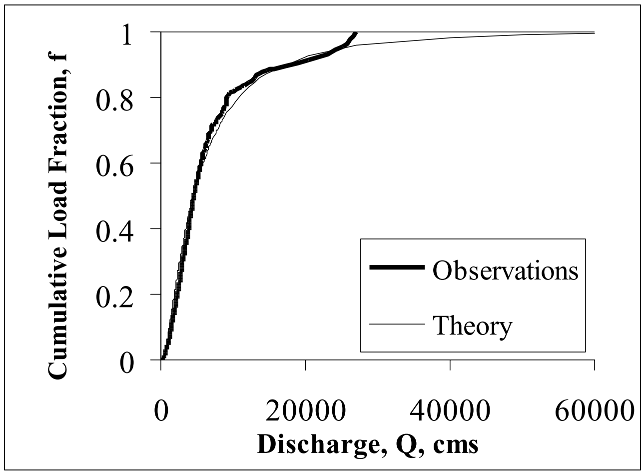


Figure 9. Theoretical and empirical relation between the cumulative long-term load fraction f and river discharge Q for the Susquehanna River.

19-year period. Note that Figure 9 is based only on the 5626 suspended sediment and associated streamflow measurements and not the entire 365 $\times$ 19 = 6935 days. The theoretical relation between f and Q is obtained from equation (12). The empirical relation is truncated at the largest flow, leading one to conclude that the entire long-term load is carried by discharges lower than the largest flow on record of 28,883 m^3/s, which occurred in 1972. The 1972 flood resulted in the largest flow on record since the flood stage record began in 1786.

[28] For this case, the *Wolman and Miller* [1960] effective discharge Q_e = 1510 m^3/s, using equation (8b). The empirical half-load discharge $Q_{1/2}$ = 4560 m^3/s, and the theoretical half-load discharge $Q_{1/2}$ = 4480 m^3/s based on equation (10a); a bias correction factor to reflect the sampling uncertainty in the three model parameters was not employed (for such estimators, see *Bradu and Mundlak* [1970]). In this case, given the relatively large number of sediment and daily flow measurements, such a correction should be small. In this case, the empirical and theoretical half-load discharges are quite similar, and it is only for discharges in excess of about 27,000 m^3/s that the empirical and theoretical relations between f and Q differ. To examine the likelihood of these discharges, we fit a three-parameter lognormal distribution to the series of annual maximum discharges at this site. The empirical and theoretical half-load discharges have average return periods of approximately 1 year. By comparison, the effective discharge Q_e = 1510 m^3/s was significantly lower than the lowest annual maximum streamflow on record (which was 3653 m^3/s); thus Q_e was exceeded every year. From Figure 7, 19% of the average daily streamflows exceed Q_e, whereas only 2.6% of the average daily streamflows exceed $Q_{1/2}$ over the long term. Viewing the series of annual maximum streamflows for this site for the entire period of record 1890–2001, the annual maximum streamflow was greater than the half-load discharge in just eight of those 117 years. Clearly, in most years, suspended sediment measurements could only provide information regarding less than half the long-term load, yet without the theory introduced above, this would not be so clear. For rivers in which C_Q or b, or both, are greater than for the Susquehanna River, the theoretical

model would provide even more information than for the case considered here.

[29] Estimation of return periods associated with half-load discharges will in many cases require extrapolation of both the assumed power law load-discharge model as well as the assumed lognormal probability distribution for daily streamflow. Extrapolation can be dangerous, and we know that sediment density cannot continue to increase with flow, for eventually a cubic meter of water and sediment would weigh more than the corresponding volume of concrete. Even in this example, while the half-load discharge was within the data set, it was based on a computation that used the entire daily flow-frequency relationship and assumed that the load-flow relationship was valid over that entire range. Extrapolation of either of these two models beyond the available data may not always be wise, and future research should evaluate the validity and robustness of such extrapolations. Clearly, results like those in Figures 5 and 6 that suggest that the half-load discharge has a return period of 100,000 years should be taken with a grain of salt. However, it is also clear, as shown in Figure 6 and for the Susquehanna River in Figure 9, that large infrequent flows are responsible for the movement of most sediment. That is the point of the analysis.

9. Conclusions

[30] Until publication of the study by *Wolman and Miller* [1960] it was commonly thought that relatively infrequent floods were responsible for carrying the bulk of the sediment load. After the introduction of the concept of the effective discharge by *Wolman and Miller* [1960] a consensus emerged that a wide range of relatively common or frequent floods is responsible for the work involved in shaping our landscape [e.g., *Wolman and Miller*, 1960; *Benson and Thomas*, 1966; *Ashmore and Day*, 1988; *Sichingabula*, 1999]. Hundreds of studies since *Wolman and Miller* [1960] have identified the dominant or effective discharge as that discharge interval that maximizes the transport effectiveness. Such analyses provide a measure of the effectiveness of individual flow rates in terms of their ability to transport sediment over the long term. The problem is that these analyses do not measure the effectiveness of the entire flow distribution. This is because the effective discharge is simply the maximum of effectiveness rather than its integral, which is the more meaningful quantity describing the cumulative total mass of sediment transported for flows up to some level. The analyses here show that when one accounts for the work done by the entire flow distribution, one finds that most of the sediment is transported by rather infrequent floods, particularly for basins with large values of C_Q and b (see Figures 4–6).

[31] Figure 3 demonstrates that for sites with streamflow coefficient of variation C_Q in excess of about 4 (which is most rivers in the United States, as shown in Figure 1), nearly all of the long-term nutrient or contaminant load is carried by flows in excess of the effective discharge index Q_e, introduced by *Wolman and Miller* [1960], regardless of the exponent b. Since nearly all load transport occurs for streamflows in excess of Q_e, this index is not a very useful descriptor of the ability of a river to transport loads over the long term. The half-load and f load discharge indices introduced here appear to be more meaningful than Q_e in such situations because they clearly document which discharges are responsible for carrying the bulk of the long-term load.

[32] An empirical estimator of the f load discharge Q_f is obtained by simply ranking the streamflows and loads and summing the loads until a fraction f is reached, at which point the corresponding discharge becomes the f load discharge. This empirical approach to estimation of Q_f and $Q_{1/2}$ seems attractive because it does not require lumping discharges into arbitrary intervals and does not require assumption of a model structure. However, we have shown that in most situations the half-load discharge is such a rare event that it cannot be reliably estimated without use of a flow-frequency model. In most cases, even long records of streamflow and load will be insufficient in length; our analyses have shown that it is unlikely we will have even observed the half-load discharge. To evaluate such situations, it becomes necessary to employ the type of theoretical analysis described here to obtain a more reliable estimate of the frequency and magnitude of the half-load discharge than can be obtained from the empirical approach on the basis of ranking the loads and streamflows.

[33] For typical sediment transport problems the data presented here demonstrate that the half-load discharge can easily exhibit return periods of several decades to centuries. This implies that it is relatively rare floods that are responsible for carrying most of the sediment over the long term, which was the consensus prior to the introduction of the effective discharge concept in 1960. This conclusion is consistent with measurements of sediment transport by *Kirchner et al.* [2001], who found that incremental erosion prevails most of the time but accounts for a small fraction of the total sediment yield, whereas by contrast, catastrophic erosion events are rare and brief but dominate the long-term sediment yield.

[34] **Acknowledgments.** The authors are indebted to Brent Troutman, Tim Cohn, Thomas E. Adams, and Robert Hirsch for their thoughtful review of an early version of this manuscript. The authors also are indebted to James Limbrunner for his assistance in the creation of Figure 1.

References

Andrews, E. D., Effective and bankfull discharges of streams in the Yampa River basin, Colorado and Wyoming, *J. Hydrol.*, *46*, 311–330, 1980.

Andrews, E. D., and J. M. Nankervis, Effective discharge and the design of channel maintenance flows for gravel-bed rivers, in *Natural and Anthropogenic Influences in Fluvial Geomorphology: The Wolman Volume*, *Geophys. Monogr. Ser.*, vol. 89, edited by J. E. Costa et al., pp. 151–164, AGU, Washington, D. C., 1995.

Ashmore, P. E., and T. J. Day, Effective discharge for suspended sediment transport in streams of the Saskatchewan River basin, *Water Resour. Res.*, *24*(6), 864–870, 1988.

Benson, M. A., and D. M. Thomas, A definition of dominant discharge, *Bull. Int. Assoc. Sci. Hydrol.*, *11*, 76–80, 1966.

Bradu, D., and Y. Mundlak, Estimation in lognormal linear models, *JASA J. Am. Stat. Assoc.*, *65*, 198–211, 1970.

Cohn, T. A., Recent advances in statistical methods for the estimation of sediment and nutrient transport in rivers, *U.S. Natl. Rep. Int. Union Geod. Geophys. 1991–1994*, *Rev. Geophys.*, *33*, 1117–1123, 1995.

Cohn, T. A., D. L. Caulder, E. J. Gilroy, L. D. Zynjuk, and R. M. Summers, The validity of a simple statistical model for estimating fluvial constituent loads: An empirical study involving nutrient loads entering Chesapeake Bay, *Water Resour. Res.*, *28*(9), 2353–2363, 1992.

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

Gilroy, E. J., R. M. Hirsch, and T. A. Cohn, Mean square error of regression-based constituent transport estimates, *Water Resour. Res.*, *26*(9), 2069–2077, 1990.

House, W. A., and M. S. Warwick, Hysteresis of the solute concentration/discharge relationship in rivers during storms, *Water Res.*, *32*(8), 2279–2290, 1998.

Kirchner, J. W., R. C. Finkel, C. S. Riebe, D. E. Granger, J. L. Clayton, J. G. King, and W. F. Megahan, Mountain erosion over 10 yr, 10 k.y., and 10 m.y. time scales, *Geology*, *29*(7), 591–594, 2001.

Leopold, L. B., *A View of the River*, Harvard Univ. Press, Cambridge, Mass., 1994.

Limbrunner, J. F., R. M. Vogel, and L. C. Brown, Estimation of harmonic mean of a lognormal variable, *J. Hydrol. Eng.*, *5*(1), 59–66, 2000.

Nash, D. B., Effective sediment-transporting discharge from magnitude-frequency analysis, *J. Geol.*, *102*, 79–95, 1994.

O'Connor, D. J., The concentration of dissolved solids and river flow, *Water Resour. Res.*, *12*(2), 279–294, 1976.

Parker, R. S., and B. M. Troutman, Frequency distribution for suspended sediment loads, *Water Resour. Res.*, *25*(7), 1567–1574, 1989.

Rudolph, B. E., Incorporating hysteresis into concentration-discharge models, M.S. thesis, Tufts Univ., Medford, Mass., 2002.

Sichingabula, H. M., Magnitude-frequency characteristics of effective discharge for suspended sediment transport, Fraser River, British Columbia, Canada, *Hydrol. Processes*, *13*, 1361–1380, 1999.

Syvitski, J. P., M. D. Morehead, D. B. Bahr, and T. Mulder, Estimating fluvial sediment transport: The rating parameters, *Water Resour. Res.*, *36*(9), 2747–2760, 2000.

Vogel, R. M., and N. M. Fennessey, *L* moment diagrams should replace product moment diagrams, *Water Resour. Res.*, *29*(6), 1745–1752, 1993.

Vogel, R. M., and I. Wilson, Probability distribution of annual maximum, mean and minimum streamflows in the United States, *J. Hydrol. Eng.*, *1*(2), 69–76, 1996.

Wolman, M. G., and J. P. Miller, Magnitude and frequency of forces in geomorphic processes, *J. Geol.*, *68*, 54–74, 1960.

R. P. Hooper, CUAHSI, 2000 Florida Avenue, NW, Washington, DC 20009, USA. (rhooper@cuahsi.org)

J. R. Stedinger, Cornell University, Department of Civil and Environmental Engineering, Hollister Hall, Ithaca, NY 14853, USA. (jrs5@cornell.edu)

R. M. Vogel, Tufts University, Department of Civil and Environmental Engineering, 113 Anderson Hall, Medford, MA 02155, USA. (rvogel@tufts.edu)