WATER RESOURCES RESEARCH, VOL. 25, NO. 5, PAGES 937–942, MAY 1989

Estimating Constituent Loads

TIMOTHY A. COHN, LEWIS L. DELONG, EDWARD J. GILROY,
ROBERT M. HIRSCH, AND DEBORAH K. WELLS

U.S. Geological Survey, Reston, Virginia

Several recent articles have called attention to the problem of retransformation bias, which can arise when log linear regression models are used to estimate sediment or other constituent loads. In some cases the bias can lead to underestimation of constituent loads by as much as 50%, and several procedures have been suggested for reducing or eliminating it. However, some of the procedures recommended for reducing the bias can actually increase it. This paper compares the bias and variance of three procedures that can be used with log linear regression models: the traditional rating curve estimator, a modified rating curve method, and a minimum variance unbiased estimator (MVUE). Analytical derivations of the bias and efficiency of all three estimators are presented. It is shown that for many conditions the traditional and the modified estimator can provide satisfactory estimates. However, other conditions exist where they have substantial bias and a large mean square error. These conditions commonly occur when sample sizes are small, or when loads are estimated during high-flow conditions. The MVUE, however, is unbiased and always performs nearly as well or better than the rating curve estimator or the modified estimator provided that the hypothesis of the log linear model is correct. Since an efficient unbiased estimator is available, there seems to be no reason to employ biased estimators.

INTRODUCTION

There has recently been considerable discussion related to methods for estimating the loads of constituents carried by streams [*Ferguson*, 1986, 1987; *Koch and Smillie*, 1986; *Richards and Holloway*, 1987; *Young et al.*, 1988]. In general, these papers have considered the case where one is interested in estimating loads for a single stream. In such cases, the use of biased estimators has been justified by virtue of their simplicity and, perhaps, by some desirable sampling properties [*Lee*, 1982]. However, biased estimators can introduce substantial errors in many situations. In particular, where multiple tributaries must be considered, as is often the case when estimating total loads to an estuary, the relative importance of bias tends to increase; the random errors in the estimates for each tributary will tend partially to offset one another, while the error due to bias will tend to accumulate.

Recent articles by *Ferguson* [1986, 1987] and *Koch and Smillie* [1986] have pointed out that the traditional rating curve method for estimating mean constituent loads can be highly biased and may lead to the severe underestimation of loads. The bias is introduced in the retransformation from "log space," where regression estimates are derived, to "real space," generally the realm of interest. *Ferguson* [1986] recommends a simple correction for reducing the bias. Unfortunately, the recommended correction, although satisfactory in many practical situations, does not eliminate bias and can lead to severe overestimation of loads.

This paper summarizes the sampling properties of the uncorrected estimator and the "bias-corrected" estimator and describes an exact, minimum variance unbiased (MVUE) procedure [*Finney*, 1941; *Bradu and Mundlak*, 1970; *Lee*, 1982].

Paper number 88WR03863.

A GENERAL MODEL FOR CONSTITUENT CONCENTRATIONS OR LOADS

Many methods for estimating concentration or load assume that a relationship exists between the logarithm of constituent concentration [ln (C)] and the logarithm of flow [ln (Q)] or other explanatory variables. The relation is usually expressed as a linear model

$$Y \equiv \ln (C) = \beta_0 + \beta_1 \ln (Q) + \varepsilon = \mu + \varepsilon \tag{1}$$

where ln ( ) represents the natural logarithm function, β_0 and β_1 are the coefficients of the model, μ is the conditional mean of Y given Q, and ε is a normal random variable. In practice, concentrations are measured relatively infrequently, and it may be assumed that each of the ε_i in a sample is independent of the others.

Equation (1) corresponds to the multiplicative model

$$\begin{aligned} C &= \exp (\mu)\eta \\ &= \exp (\beta_0)Q^{\beta_1}\eta \end{aligned} \tag{2}$$

where η is a lognormal random variable. It is important to note that C is a random variable. The conditional (upon Q) median of C is given by $\exp (\beta_0)Q^{\beta_1}$, whereas the conditional mean is given by $\exp (\beta_0)Q^{\beta_1} \exp (\sigma_\varepsilon^2/2)$, where σ_ε^2 is the variance of ε [*Aitchison and Brown*, 1981].

The load L corresponding to C and Q is given by

$$L = KCQ \tag{3}$$

where K is a unit conversion factor. Thus the expected value of L is

$$\begin{aligned} E[L] &= KQE[C] \\ &= K \exp (\beta_0)Q^{\beta_1}E[\eta]Q = K \exp (\beta_0)Q^{\beta_1+1}E[\eta] \end{aligned} \tag{4}$$

Because K and Q are considered fixed, the analysis presented here for estimating concentrations applies equally well to estimating loads. For example, the biases and standard deviations of the concentration estimators, when expressed in percentage terms, are identical to those of the model relating ln (L) to ln (Q). The results in this paper also apply to more complex regression models that might include one or more different functions of flow, or other variables related to seasonality, long-term trend, or hydrograph hysteresis as explanatory variables. However, a cautionary note is required: all of the procedures described below are based on the assumption that constituent concentrations obey a specified log linear model. This appears to be approximately true in many cases, but in practice one should always assess the validity of the model before applying it. The consequences of model misspecification can be much more serious than the effect of retransformation bias.

ESTIMATING THE $\hat{\beta}$ VECTOR

Assume there exists a sample of constituent concentration data $\mathbf{C} = \{C_1 \cdots C_N\}$, and the corresponding instantaneous flow data $\mathbf{Q} = \{Q_1 \cdots Q_N\}$, measured at times $\{t_1 \cdots t_N\}$. Also, assume there exists essentially continuous measurement of flow. To simplify the notation, let $\mathbf{Y}$ be a vector containing the logarithms of $\mathbf{C}$, and let $\mathbf{X}$ be an $N \times 2$ matrix containing a vector of ones in the first column and the logarithms of $\mathbf{Q}$ in the second column. When estimating $k = 2$ parameters, one can derive ordinary least squares estimates for σ_ε^2 and

$$\hat{\boldsymbol{\beta}} = \{\hat{\beta}_0, \hat{\beta}_1\}' = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \tag{5}$$

where the apostrophe indicates transposition [see *Draper and Smith*, 1981]. An unbiased estimate of the expected value of the logarithm of C, denoted $\hat{\mu}$, for arbitrary flow Q^*, is given by

$$\hat{\mu} = \hat{\beta}_0 + \hat{\beta}_1 \ln (Q^*) = \underline{\mathrm{X}}\hat{\beta} \tag{6}$$

where $\underline{\mathrm{X}} = \{1, X^*\} \equiv \{1, \ln (Q^*)\}$. It can be shown that $\hat{\mu}$ is a normal random variable with mean $\mu = \underline{\mathrm{X}}\beta$ and variance $V\sigma_\varepsilon^2$, where $V \equiv \underline{\mathrm{X}}(\mathbf{X}'\mathbf{X})^{-1}\underline{\mathrm{X}}'$. For the bivariate case considered here,

$$V = \left\{\frac{1}{N} + \ln^2 (Q^*) \bigg/ \sum_{i=1}^{N} (\ln (Q_i) - \overline{\ln Q})^2\right\}$$

$$V = \left\{\frac{1}{N} + \ln^2 (Q^*) \bigg/ \sum (\ln (Q_i) - \overline{\ln Q})^2\right\}$$

where

$$\overline{\ln Q} = \sum_{i=1}^{N} \ln (Q_i)/N$$

The problem of estimating the logarithms of concentrations is completely solved by (6). However, log concentrations are seldom of interest; it is usually desirable to provide results in real space. It is in the retransformation from log space to real space that the bias complications are introduced. We will now consider some estimators for $E[C] = \exp (\mu + \sigma_\varepsilon^2/2)$.

RATING CURVE RETRANSFORMATION METHOD

The simplest and probably most widely used retransformation method is to exponentiate $\hat{\mu}$:

$$\hat{C}_{RC} = \exp (\hat{\mu}) = \exp (\underline{\mathrm{X}}\hat{\beta}) = \exp (\hat{\beta}_0)Q^{\hat{\beta}_1} \tag{7}$$

This is frequently called the "rating curve" method. *Ferguson* [1986] points out, as did *Sichel* [1952], *Lane* [1975], *Landwehr* [1978], *DeLong* [1982], *Lee* [1982], and *DeLong and Wells* [1987], among others, that $\hat{C}_{RC}$ is a biased estimator for $E[C]$. Since $\hat{\mu}$ is a normal $N(\mu, V\sigma_\varepsilon^2)$ variate, $\exp (\hat{\mu})$ is lognormally distributed with parameters $(\mu, V\sigma_\varepsilon^2)$. Thus the mean of $\exp (\hat{\mu})$ is given by [*Aitchison and Brown*, 1981]

$$\begin{aligned} E[\hat{C}_{RC}] &= \exp [\mu] \exp [V\sigma_\varepsilon^2/2] \\ &= E[C] \exp [(V-1)\sigma_\varepsilon^2/2] \end{aligned} \tag{8}$$

In most cases, this implies that $\hat{C}_{RC}$ is biased downward. However, $E[\hat{C}_{RC}]$ can be greater than $E[C]$ if Q^* happens to be much higher or lower than the sampled $\mathbf{Q}$ on which the rating curve is based.

Similarly, the variance of $\hat{C}_{RC}$ is computed from the moments of a lognormal variate:

$$\begin{aligned} \mathrm{Var}\,[\hat{C}_{RC}] &= E[(\hat{C}_{RC})^2] - E^2[\hat{C}_{RC}] \\ &= \exp [2\mu + 2V\sigma_\varepsilon^2] - \exp [2\mu + V\sigma_\varepsilon^2] \\ &= \exp [2\mu + V\sigma_\varepsilon^2]\{\exp [V\sigma_\varepsilon^2] - 1\} \end{aligned} \tag{9}$$

QUASI MAXIMUM LIKELIHOOD ESTIMATOR (QMLE)

Ferguson [1986] recognizes the bias of the rating curve method and recommends multiplying $\hat{C}_{RC}$ by $\exp (s^2/2)$, where

$$s^2 = \sum_{i=1}^{N} \frac{(Y_i - \hat{Y})^2}{N-k}$$

to eliminate the bias:

$$\hat{C}_{QMLE} = \exp (\hat{\mu}) \exp (s^2/2) \tag{10}$$

The estimator behaves well if V is small and $\exp (s^2/2)$ is a satisfactory estimator for $\exp (\sigma_\varepsilon^2/2)$.

The expected value of $\hat{C}_{QMLE}$ can be obtained by noting that s^2 and $\hat{\mu}$ are independent, and thus

$$\begin{aligned} E[\hat{C}_{QMLE}] &= E[\exp (\hat{\mu})]E[\exp (s^2/2)] \\ &= E[C] \exp [(V-1)\sigma_\varepsilon^2/2]\left[1 - \frac{\sigma_\varepsilon^2}{m}\right]^{-m/2} \end{aligned} \tag{11}$$

where $m = N - k$ is the number of degrees of freedom in the error distribution. The expected value of the second factor is obtained from the moment generating function of a χ^2 random variable.

The variance of $\hat{C}_{QMLE}$ can be computed as above, after noting that

$$\begin{aligned} E[(\exp (s^2/2))^2] &= E[\exp (2s^2)] \\ &= \left[1 - \frac{2\sigma_\varepsilon^2}{m}\right]^{-m/2} \end{aligned} \tag{12}$$

```
      FUNCTION GM(DF,ARG)
C=================================================================
C
C     FUNCTION TO COMPUTE FINNEY'S gm(t)
C
C     AUTHOR.........TIM COHN
C     DATE...........OCTOBER 1, 1986
C
C     DF      R*4      NUMBER OF DEGREES OF FREEDOM OF RESIDUALS
C     ARG     R*4      ARGUMENT TO FINNEY'S FUNCTION
C
C=================================================================
C

      DATA TOL/1.E-7/

       IF(ABS(ARG) .GT. 50.0) THEN
          WRITE(*,*) 'MAGNITUDE OF ARG IS TOO LARGE (GM)'
          GM  =  0.0
          RETURN
       ENDIF
          GM        =  1.0
          IF(DF .LE. 0.0) RETURN
          BT        =  ARG*DF**2/(2.0*(DF+1.0))
          TERM      =  1.0
      DO 10 P=1,1000
          TERM      =  TERM * BT/((DF/2.0+P-1.0)*P)
          GM        =  GM+TERM
          IF(P .GT. 1.0 .AND. ABS(TERM) .LT. TOL)  RETURN
   10 CONTINUE
          WRITE(*,*) 'GM DID NOT CONVERGE'
          RETURN
      END
```

Fig. 1. The FORTRAN function above evaluates $g_m(\quad)$ for an argument in the range (−50, 50). If necessary, up to 1000 terms are evaluated; however, fewer than 10 terms are usually needed to satisfy the convergence criterion (1.0*E*-7).

$$\begin{aligned}\text{Var}\,[\hat{C}_{QMLE}] &= E[(\hat{C}_{QMLE})^2] - E^2[\hat{C}_{QMLE}] \\ &= \exp\,[2\mu + 2V\sigma_\varepsilon^2]\left[1 - \frac{2\sigma_\varepsilon^2}{m}\right]^{-m/2} \\ &\quad - \exp\,[2\mu + V\sigma_\varepsilon^2]\left[1 - \frac{\sigma_\varepsilon^2}{m}\right]^{-m} \\ &= \exp\,[2\mu + V\sigma_\varepsilon^2]\left\{\exp\,[V\sigma_\varepsilon^2]\left[1 - \frac{2\sigma_\varepsilon^2}{m}\right]^{-m/2}\right. \\ &\quad \left. - \left[1 - \frac{\sigma_\varepsilon^2}{m}\right]^{-m}\right\} \end{aligned} \tag{13}$$

The mean of $\hat{C}_{QMLE}$ is undefined for $\sigma_\varepsilon^2 > N - k$, and in general the pth moment is undefined for $\sigma_\varepsilon^2 > (N - k)/p$.

TABLE 1. Quantiles of Flow Assuming a Standard Lognormal Distribution

Quantile	Log Flow (X^*)	Flow (Q^*)
$Q_{(0.01)}$	−2.32	0.1
$Q_{(0.10)}$	−1.28	0.3
$Q_{(0.50)}$	0.00	1.0
$Q_{(0.90)}$	1.28	3.6
$Q_{(0.99)}$	2.32	10.2
$Q_{(0.999)}$	3.09	22.0
$Q_{(0.9999)}$	3.72	41.2

Although it is unlikely that the mean or variance of $\hat{C}_{QMLE}$ will be undefined in most hydrologic applications, it will be shown that the bias and variance of the estimator can be large under plausible circumstances.

Bradu-Mundlak Estimator

Bradu and Mundlak [1970] derive an estimator, based on work by *Finney* [1941], that has desirable properties although it is somewhat more complicated than the QMLE estimator. However, one of the authors [*DeLong*, 1982] has used the estimator in practice in a nonregression context. Define

$$\hat{C}_{MVUE} = \exp\,(\hat{\mu})g_m\left(\frac{m+1}{2m}\{(1 - V)s^2\}\right) \tag{14}$$

where

$$g_m(z) = \sum_{p=0}^{\infty} \frac{m^p(m + 2p)}{m(m + 2)\cdots(m + 2p)}\left(\frac{m}{m + 1}\right)^p\left(\frac{z^p}{p!}\right) \tag{15}$$

Bradu and Mundlak show that $\hat{C}_{MVUE}$ is an unbiased estimator for $E[C]$. As an unbiased function of sufficient statistics, it is optimal, in the sense of minimum variance, among all unbiased estimators, regardless of sample size (for discussion, see *Bradu and Mundlak* [1970]).

The variance of $\hat{C}_{MVUE}$ is [*Bradu and Mundlak*, 1970; *Likeš*, 1980]

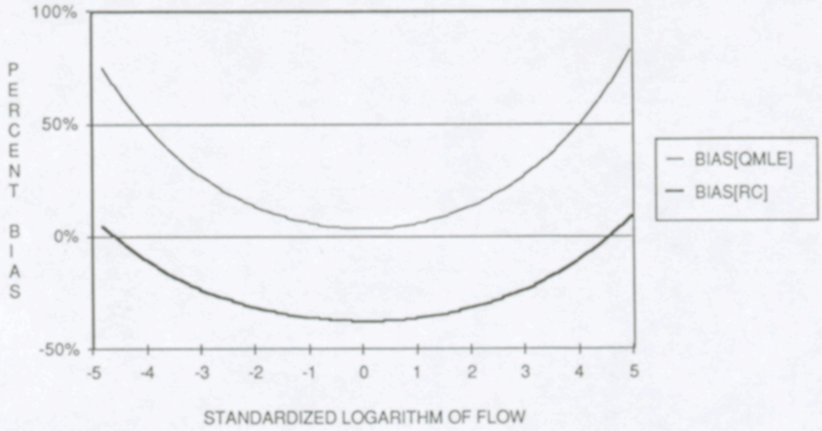

Fig. 2. Bias of $\hat{C}_{RC}$ and $\hat{C}_{QMLE}$ as a function of the standardized logarithm of the predictor variable Q^*, expressed as a percent of the true load. Here $\sigma_\varepsilon^2 = 1.0$ and $N = 22$.

$$\mathrm{Var}\,(\hat{C}_{MVUE}) = \exp\,(2\mu)\{\exp\,(2\sigma_\varepsilon^2 V) \cdot G_m((1-V)\sigma_\varepsilon^2/2) - \exp\,(\sigma_\varepsilon^2)\} \quad (17)$$

where

$$G_m(z) = \sum_{h=0}^{\infty} \frac{\Gamma(m/2)}{\Gamma(m/2+h)} \binom{m+2h-2}{h} z^h = \exp\,(2z) g_m\left[\frac{2(m+1)}{m^2}\, z^2\right] \quad (18)$$

While the equations may be somewhat intimidating, they have been tabulated [*Bradu and Mundlak*, 1970] and can be computed easily (see Figure 1 for FORTRAN listing).

Discussion

An example may help to illustrate the percent bias and root-mean-square error (rmse) of the estimators. Suppose we want to regress the logarithm of concentration {ln (C)} on the logarithm of flow {ln (Q)}. This corresponds to estimating $k = 2$ unknown parameters: β_0, the constant; and β_1, the coefficient relating ln (C) and ln (Q). Without loss of generality, we can assume that

$$\sum_{i=1}^{N} \ln\,(Q_i) = 0 \text{ and}$$

$$\sum_{i=1}^{N} \frac{\ln^2\,(Q_i)}{N} = 1$$

and that $\{\beta_0, \beta_1\}' = \{0, 1\}'$. The reader may choose to think of the flows as lognormal variates, whose quantiles are given in Table 1.

Given a particular model form, the percent bias and rmse of estimators for $E[C]$ are functions of three factors: the sample size (N) of the data set used to calibrate the model; the proximity of the desired explanatory variables (in this case X^*, the logarithm of flow) to the values used in model calibration; and the variance of the residuals (σ_ε^2).

Figure 2 displays the percent bias of the various estimators of concentration as a function of the value of Q^*, assuming there are $N = 22$ observations of constituent concentration for estimating the model parameters with log space standard error of residuals $\sigma_\varepsilon^2 = 1.0$ (CV = 131%). As anticipated, $\hat{C}_{RC}$ shows negative bias for most values of Q^*, while for very large or small values of Q^*, $\hat{C}_{RC}$ is positively biased. $\hat{C}_{QMLE}$ is nearly unbiased for moderate values of Q^*, but

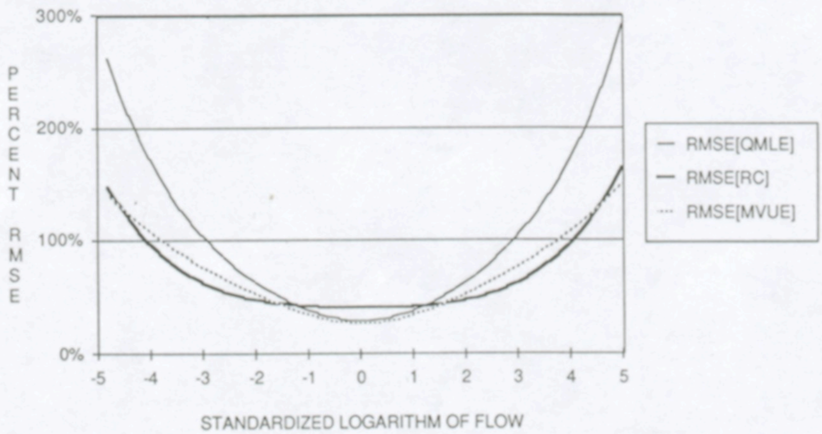

Fig. 3. The rmse of $\hat{C}_{RC}$, $\hat{C}_{QMLE}$, and $\hat{C}_{MVUE}$ as a function of the standardized logarithm of the predictor variable Q^*, expressed as a percent of the true load. Here $\sigma_\varepsilon^2 = 1.0$ and $N = 22$.

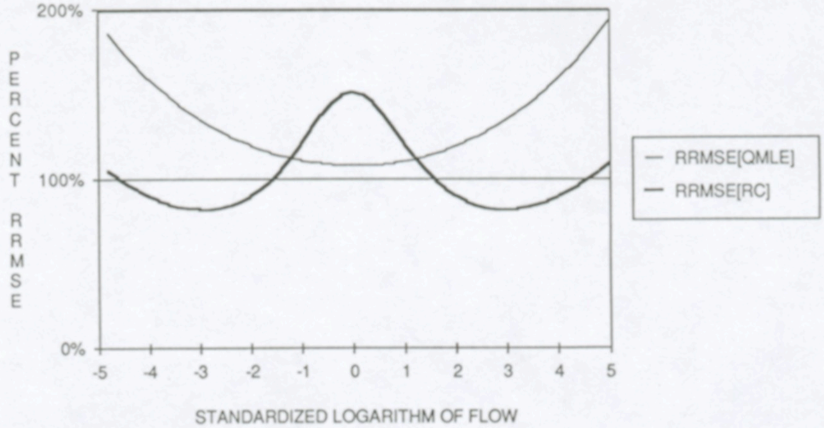

Fig. 4. Relative rmse of $\hat{C}_{RC}$ and $\hat{C}_{QMLE}$ as a function of the standardized logarithm of the predictor variable Q^*, expressed as a percent of the rmse of $\hat{C}_{MVUE}$. Here $\sigma_\varepsilon^2 = 1.0$ and $N = 22$.

shows increasingly positive bias for either large or small values of Q^*. The bias becomes very large for large values of Q^*. Such values of Q^* may correspond to those times when accurate estimates are most needed. The MVUE is not shown because it is unbiased in all cases.

Figure 3 displays the percent rmse of the estimators as a function of the value of the logarithm of Q^*. It is apparent that the rmse of $\hat{C}_{QMLE}$ always exceeds the rmse of $\hat{C}_{MVUE}$. $\hat{C}_{RC}$ has the lowest rmse for many values of Q^*. In fact, there exist estimators with even smaller rmse [*Rukhin*, 1986]. However, the rmse of $\hat{C}_{RC}$ is at most only slightly lower than that of $\hat{C}_{MVUE}$. $\hat{C}_{MVUE}$ has the lower rmse for either large or small values of Q^*.

Figure 4 displays the same information as Figure 3, except that in Figure 4 the rmses of $\hat{C}_{RC}$ and $\hat{C}_{QMLE}$ are expressed as a percentage of the rmse of $\hat{C}_{MVUE}$. This is called the relative root-mean-square error (rrmse). The case in Figure 5 corresponds to that of Figure 4 except that Q^* and σ_ε^2 are both set equal to unity, and the sample size N is varied from 5 to 100. As the sample size increases one sees that $\hat{C}_{QMLE}$ converges rapidly to $\hat{C}_{MVUE}$, although the rmse of $\hat{C}_{QMLE}$ always exceeds that of $\hat{C}_{MVUE}$. For large N, the $\hat{C}_{RC}$ is dominated by its substantial bias, and it performs much worse than the other estimators.

Figure 6 illustrates the effect of σ_ε^2 on the performance of the estimators, with N fixed at 22, and Q^* equal to unity. Here one sees that $\hat{C}_{QMLE}$ does well for low or moderate values of σ_ε^2, but has considerably higher rmse than $\hat{C}_{MVUE}$ for σ_ε^2 greater than 1.0. At $\sigma_\varepsilon^2 = 1.0$, $\hat{C}_{RC}$ has a rmse that is about 50% above the rmse of $\hat{C}_{MVUE}$ or $\hat{C}_{QMLE}$.

It should be noted that the discussion has centered on fully parametric estimators. These are efficient where the assumed model adequately describes the physical situation. However, there are no guarantees that the assumed model will be adequate. In many cases, particularly with hydrologic data, it may be desirable to employ a nonparametric estimator, such as the "smearing estimator" of *Duan* [1983]. Also, *Thomas* [1985] provides a probability-based sampling procedure that provides unbiased load estimates even if the model is misspecified.

CONCLUSIONS

Several recent articles about retransformation bias have called attention to a serious problem that may lead to underestimation of constituent loads by as much as 50%. However, several of the procedures suggested for reducing the bias can actually increase the bias in some cases. In this paper, the bias and variance have been derived for three

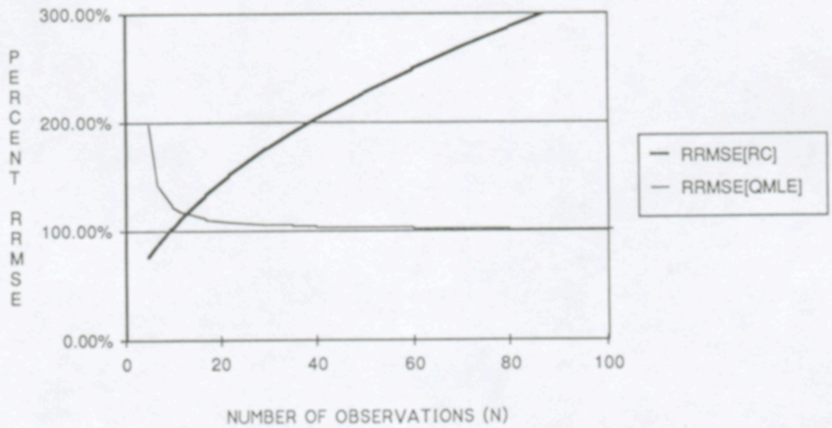

Fig. 5. Relative rmse of $\hat{C}_{RC}$ and $\hat{C}_{QMLE}$ as a function of the number of observations N, expressed as a percent of the rmse of $\hat{C}_{MVUE}$. Here $\sigma_\varepsilon^2 = 1.0$ and $Q^* = 1.0$.

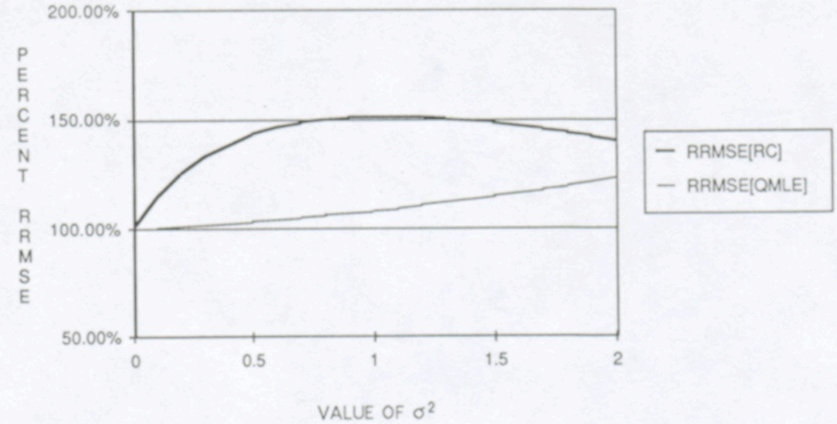

Fig. 6. Relative rmse of $\hat{C}_{RC}$ and $\hat{C}_{QMLE}$ as a function of the value of σ_ϵ^2, expressed as a percent of the rmse of $\hat{C}_{MVUE}$. Here $Q^* = 1.0$ and $N = 22$.

estimators that can be used with log linear regression models: $\hat{C}_{RC}$, $\hat{C}_{QMLE}$, and $\hat{C}_{MVUE}$. It has been shown that for many conditions $\hat{C}_{RC}$ and $\hat{C}_{QMLE}$ can provide satisfactory estimates. However, other conditions exist where they have substantial bias and a large mean square error. These conditions commonly occur when sample sizes are small, or when loads are estimated during high-flow conditions. If one can assume that the hypothesis of the log linear model is correct, $\hat{C}_{MVUE}$ always performs nearly as well, and in many cases much better, than $\hat{C}_{RC}$ or $\hat{C}_{QMLE}$. Since $\hat{C}_{MVUE}$ is unbiased and reasonably easy to use, there seems to be no reason to employ biased estimators.

Acknowledgments. For reviewing early drafts of the paper, and sharing their thoughts and suggestions, the authors would like to thank Dennis Helsel, Bill Kirby, Jurate Landwehr, Jery Stedinger, and Gary Tasker. The authors would also like to thank three anonymous reviewers, whose suggestions have substantially improved the manuscript. For their continuing support of the River Input Monitoring Program for Chesapeake Bay, the authors would like to thank Robert M. Summers and the Maryland Department of the Environment.

REFERENCES

Aitchison, J., and J. A. C. Brown, *The Lognormal Distribution With Special Reference to Its Uses in Economics*, Cambridge University Press, New York, 1981.

Bradu, D., and Y. Mundlak, Estimation in lognormal linear models, *J. Am. Stat. Assoc.*, *65*(329), 198–211, 1970.

DeLong, L. L., Water quality of streams and springs, Green River Basin, Wyoming, *U.S. Geol. Surv. Water Resour. Invest. Rep.*, *82-4008*, 36 pp., 1982.

DeLong, L. L., and D. K. Wells, Estimating average dissolved-solids yields from basins drained by ephemeral and intermittent streams—Green River Basin, Wyoming, *U.S. Geol. Surv. Water Resour. Invest. Rep.*, *87-4222*, 29 pp., 1987.

Draper, N., and H. Smith, *Applied Regression Analysis*, 2nd ed., 709 pp., John Wiley, New York, 1981.

Duan, N., Smearing estimate: A nonparametric retransformation method, *J. Am. Stat. Assoc.*, *78*(383), 605–610, 1983.

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

Ferguson, R. I., Accuracy and precision of methods for estimating river loads, *Earth Surf. Processes Landforms*, *12*, 95–104, 1987.

Finney, D. J., On the distribution of a variate whose logarithm is normally distributed, *J. R. Stat. Soc. Suppl.*, *7*, 155–161, 1941.

Koch, R. W., and G. M. Smillie, Bias in hydrologic prediction using log-transformed regression models, *Water Resour. Bull.*, *22*(5), 717–723, 1986.

Landwehr, J. M., Some properties of the geometric mean and its use in water quality standards, *Water Resour. Res.*, *14*(3), 467–473, 1978.

Lane, W. L., Extraction of information on inorganic water quality, *Hydrol. Pap. 73*, Colo. State Univ., Ft. Collins, 1975.

Lee, C. Y., Comparison of two correction methods for the bias due to the logarithmic transformation in the estimation of biomass, *Can. J. Forest Res.*, *12*, 326–331, 1982.

Likeš, J., Variance of the MVUE for lognormal variance, *Technometrics*, *22*(2), 253–258, 1980.

Richards, R. P., and J. Holloway, Monte Carlo studies of sampling strategies for estimating tributary loads, *Water Resour. Res.*, *23*(10), 1939–1948, 1987.

Rukhin, A. L., Improved estimation in lognormal models, *J. Am. Stat. Assoc.*, *81*(396), 1046–1049, 1986.

Sichel, H. S., New methods in the statistical evaluation of mine sampling data, *Inst. Min. Metall. Trans.*, *61*, 261–288, 1952.

Thomas, R. B., Estimating total suspended sediment yield with probability sampling, *Water Resour. Res.*, *21*(9), 1381–1388, 1985.

Young, T. C., J. V. DePinto, and T. M. Heidtke, Factors affecting the efficiency of some estimators of fluvial total phosphorus load, *Water Resour. Res.*, *24*(9), 1535–1540, 1988.

T. A. Cohn, L. L. DeLong, E. J. Gilroy, R. M. Hirsch, and D. K. Wells, U.S. Geological Survey, WGS Mail Stop 410, Reston, VA 22092.

(Received April 4, 1988;
revised September 27, 1988;
accepted October 10, 1988.)