Auckland
Regional Council
TE RAUHĪTANGA TAIAO

Redwoods Stream Sediment Yield

June TR 2009/062

Auckland Regional Council
Technical Report No.062 June 2009
ISSN 1179-0504 (Print)
ISSN 1179-0512 (Online)
ISBN 978-1-877528-74-3

This report is part of a series of reports that were commissioned during the period 1993-1999 that were used to support the establishment of the Mahurangi Action Plan. They are being made available following a review of technical information.

Reviewed by:

Name: Amy Taylor

Position: Project Leader Land

Organisation: ARC

Date: 1/06/09

Approved for ARC Publication by:

Name: Grant Barnes

Position: Group Manager Monitoring & Research

Organisation: ARC

Date: 1/06/09

Recommended Citation:

Hicks, D. M.; McKerchar, A. I. (2000). Redwoods Stream sediment yield. Prepared by NIWA for Auckland Regional Council. Auckland Regional Council Technical Report 2009/062.

Redwoods Stream sediment yield

D. M. Hicks
A. I. McKerchar

Prepared for
Auckland Regional Council

NIWA Client Report: CHC00139
Project No.: ARC00501
June 2000

National Institute of Water & Atmospheric Research Ltd
PO Box 8602, Riccarton, Christchurch
New Zealand
Tel: 03 348 8987
Fax: 03 348 5548

CONTENTS

Reviewed by:

Charles Pearson

Approved for release by:

Derek Goring

Executive Summary

Sediment yields from Redwoods Stream are compared before and during timber harvesting. The analysis uses Locally Weighted Scatterplot Smoothing (LOWESS) to fit sediment rating curves and uses the conditional variance of residuals to correct for logarithmic transformation bias and to determine confidence intervals on yield estimates. When the sediment rating curves for the pre-harvest and harvesting periods are applied to the same standard discharge record, no significant change in sediment yield is indicated. Misleading results would arise from using simple linear regression to model the sediment rating relationships. This is because the data violate fundamental assumptions of the regression procedure.

1 INTRODUCTION

This study was commissioned by Mr B. Handysides acting for the Auckland Regional Council. The council has been monitoring the effects of the disturbance caused by roading and then logging on sediment loads carried by Redwoods Stream, a tributary of the Mahurangi River that drains into the Hauraki Gulf at Warkworth, north of Auckland.

The council has monitored flow and sediment load for a catchment area of 0.60 km^2. Flow was monitored with conventional stream gauging equipment, and suspended sediment data were measured using an automatic sampler operated on a flow-proportional basis. An almost continuous series of 505 samples were collected from May 1994 to December 1995, while the catchment was undisturbed (hereafter termed the "pre-harvest" period). Roading (and salvage logging) of the basin commenced on 20 January 1997. General log harvesting commencing on 9 April 1997 and was completed on 11 February 1998. Cable logging was used over the entire area. Further measurements of sediment load and streamflow were made during the harvesting period, beginning on 13 February 1997 and finishing on 31 March 1998. Three hundred and seventy sediment samples were collected during this period (hereafter termed the "harvesting" period).

Preliminary analysis of the data by Auckland Regional Council yielded results that were not conclusive, and the aim of this study is to undertake a thorough analysis of the data to determine whether the disturbance due to harvesting has affected the sediment yields compared with the pre-harvest period. In presenting the analysis, the intention is to also demonstrate the appropriate use of bias correction methods, which are necessary for this type of analysis where logarithm-transformed data are used.

2 ESTIMATION OF SEDIMENT YIELD DURING PRE-HARVEST PERIOD

There was an intensive data gathering effort in the pre-harvest period, specifically for 21 May 1994 to 25 December 1995. During this period streamflows were monitored and sediment load rates were monitored by almost continuous sampling. Automated samples of the suspended sediment were drawn from the stream and composited in a bottle which was filled by eight samples. After the streamflow volume reached 1300 m^3, a new bottle was used. There were 505 bottles of composite samples for this period. The bottles thus contain samples of sediment collected over differing periods of time, depending on the streamflow, and can be described as flow proportional sediment samples. The sample from each bottle provides a weighted estimate of the mean sediment concentration for each sampling interval: they are not estimates of the instantaneous load at any time. An estimate of the sediment load is given by the concentration multiplied by the average flow rate over each time interval.

The sequence of data is continuous over the 18-month pre-harvest period, except for several occasions when it appears that samples were either not collected (because the auto-sampler had filled all its bottles and had shut down before it could be serviced), not analysed, or otherwise had some problems. The missing data were identified by computing the apparent volume of runoff between samples, equal to the mean discharge associated with the sample times the time interval between samples. Gaps are flagged where this volume exceeds the trigger volume for the flow-proportional sampling (Fig. 1). These missing samples span a total of 71 days.

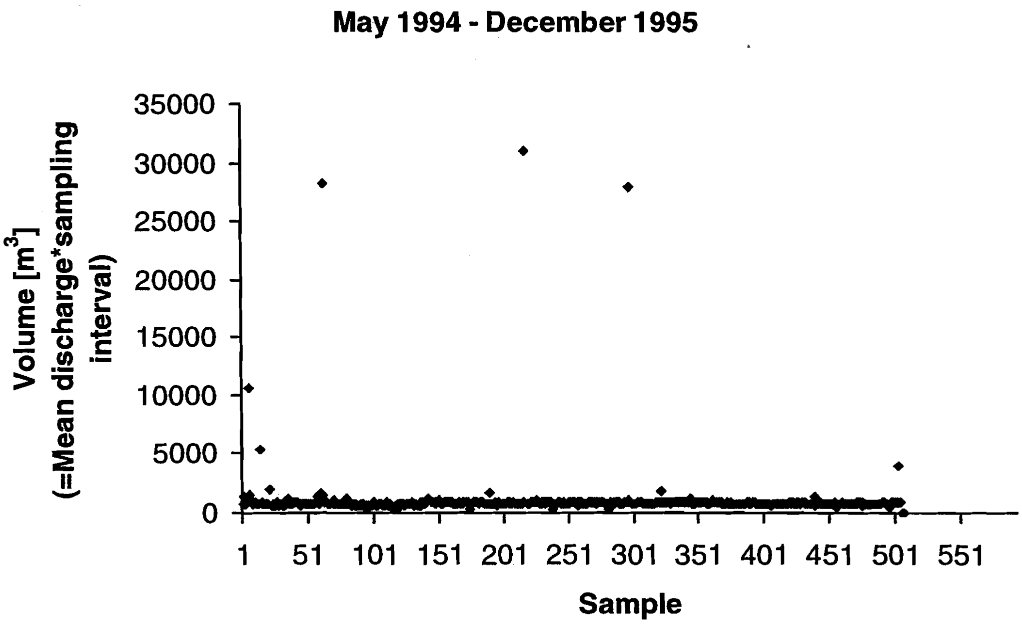

Figure 1. *Sampling trigger volume indicated by product of mean discharge over sampling duration and time since previous sample, for period of near-continuous sampling between May 1994 and December 1995. The trigger volume should be quasi-constant for a continuous record. The outliers flag short periods of missing/unanalysed samples.*

This period of missing record aside, the pre-harvest data can be used to compute the sediment yield directly without the need to construct a sediment load vs. flow rate relationship, or rating curve (simply as the sum of the products of the sample concentration, mean water discharge between samples, and time interval between samples). The resulting load (Table 1), based on 4040 samples (i.e., 8 flow-proportional sub-samples composited within each of 505 sample bottles), provides a measurement of the sediment load over the periods of continuous record, and is termed the "direct" load.

Table 1. *Estimates of load for pre-harvest period (21.05.94 to 25.12.95).*

Method	Load (t)
Direct	169
Rating by standard linear regression	72
Rating by standard linear regression with bias correction	145
LOWESS	95
LOWESS with bias correction using local variance (± signifies 1 standard error)	152 ± 27

The estimate derived in this way provides a standard for assessment of the reliability of procedures that use sediment rating curves, and we discuss this aspect below. In addition, the sample is useful for establishing the shape of the rating relationship.

Error in the direct estimate arises through measurement errors in the sediment load and flow rate, and the assumption of quasi-steady state over each interval. Further error occurs when sediment rating curves are used with flow records to obtain load estimates. These additional errors are quantified and expressed as standard errors.

The 505 data points for the pre-harvest period are plotted in Fig. 2. Superimposed on this plot are the band-averaged sediment loads: these are the average loads for each of 48 "bins" for streamflows. These band averages show some scatter, but indicate the shape required for the sediment/flow rating curve. This sediment rating should model the conditional mean load over the range of flow encountered for the period of record. Normally one would have a much smaller sample of data points than those given in this figure.

Figure 3 is a plot of the percentage of the cumulative load compared with flow magnitude. From this plot it is evident that, for example, 80 percent of the load is carried by flows with natural logarithms exceeding 9.9, i.e. flows exceeding 20 l/s, and 60 percent of the load is carried by flows with natural logarithms exceeding 10.6, i.e. flows exceeding 40 l/s. From this figure it can be concluded that the important region of the rating curve for correct estimation of sediment load is the region of the higher flows.

Figure 4 shows a standard linear regression fitted to the log-transformed data. Although the coefficient of determination (i.e. R^2) for this regression is 0.772, the regression generally under estimates the loads at higher flows. The failure of the linear regression to fit the higher flows well is thus expected to lead to substantial under-estimation of the sediment load, as is evident in Table 1. The poor fit at higher discharges occurs because the vast bulk of the data occur at lower discharges, and it is these that are driving the linear regression results.

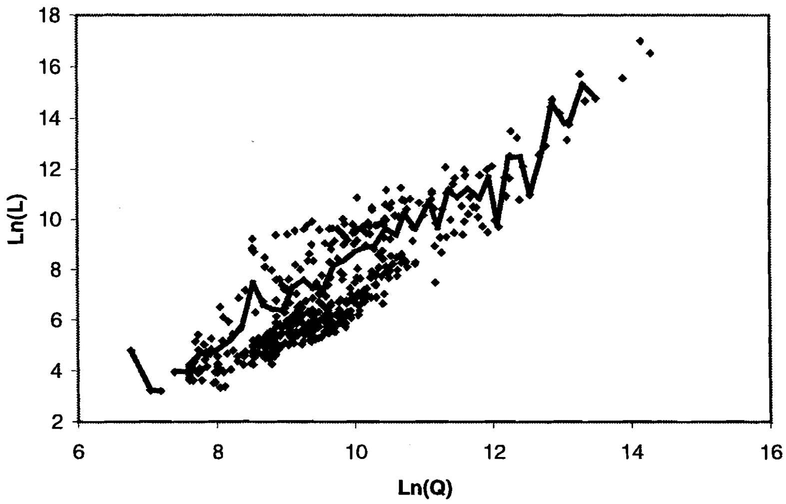

Figure 2. *Natural logarithms of sediment load (mg/s) plotted against streamflow (ml/s for Redwoods Stream for the measurement period 12 May 1994 to 25 December 1995. The line overplotted is the average sediment load when the data are grouped into 50 streamflow bands. This line is discontinuous because of a band at low flows where there were no data points.*

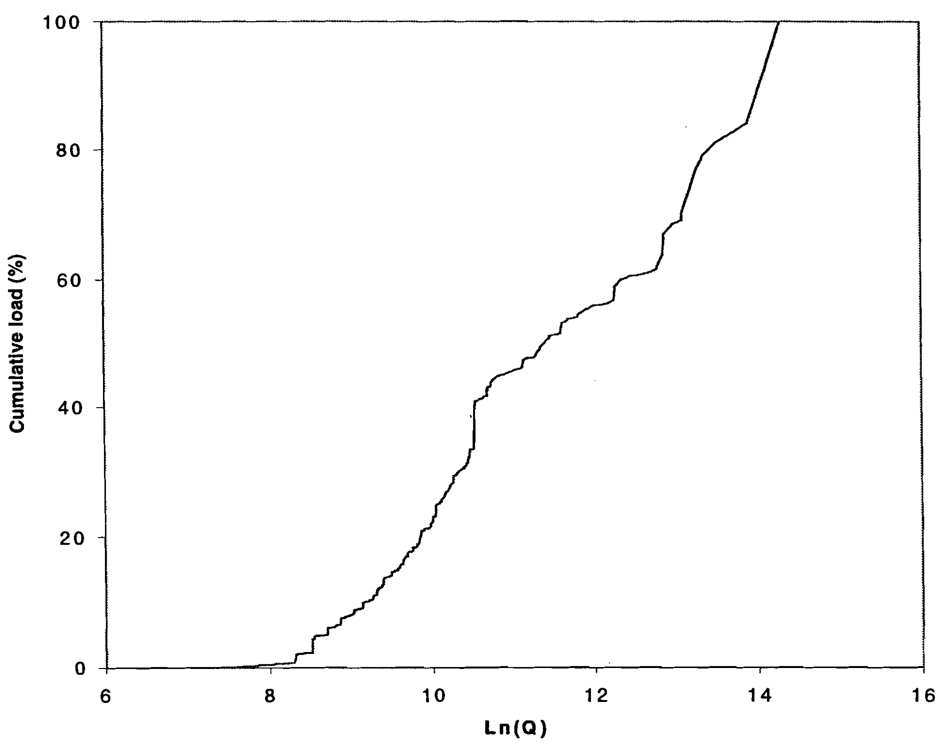

Figure 3. *Percentage of total load plotted against natural logarithm of flow (ml/s).*

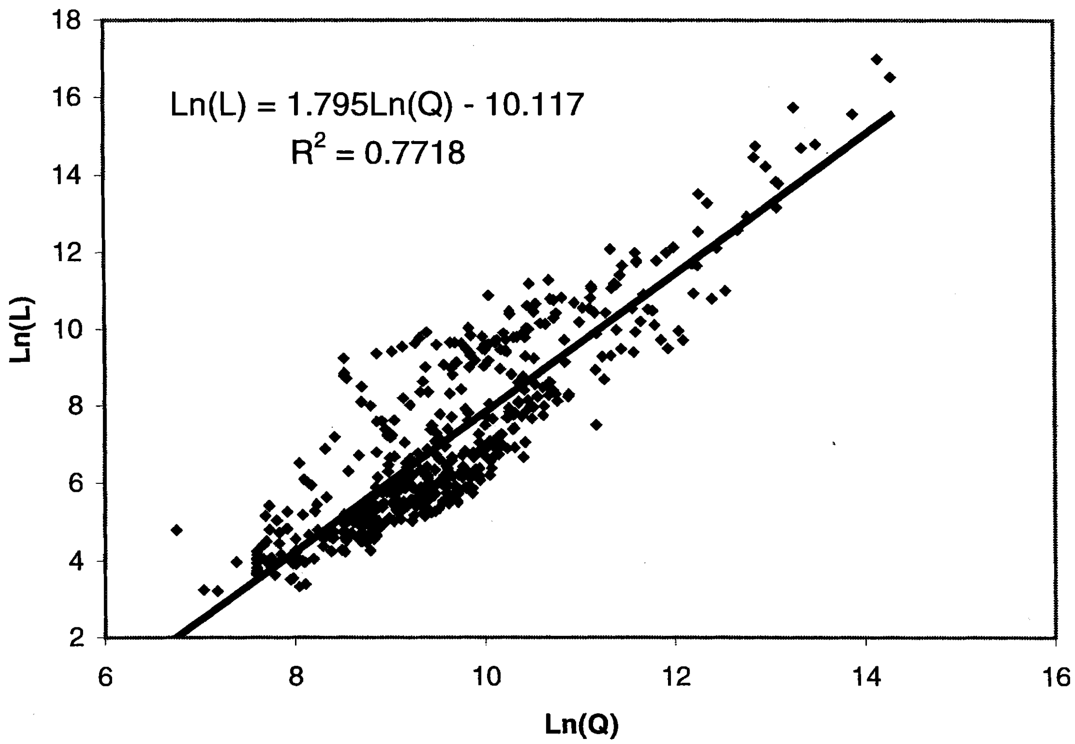

Figure 4. *1994 - 1995 load vs flow data with a simple linear regression line fitted.*

Figure 5 shows the scatter of the residuals from the linear regression model (i.e. $Ln(L_{obs}) - Ln(L_{pred})$) against the values of flow. Not only are the residuals for high and low flows generally positive, but also the scatter of the residuals is higher for low flows than for high flows. That is, the variance of the residuals is not constant, but varies with flow magnitude. In statistical terms, the residual variance is described as hetroscedastic. This feature does not disqualify the use of linear regression, but it does limit what can be said about the errors of estimate and it does affect the correction for log induced bias, which will be discussed below.

Figure 6 is a probability plot showing the distribution of the residuals compared with a normal probability line. At the lower tail, there are some departure from the straight line, but otherwise over the bulk of the data there is a reasonable fit, indicating that the residuals data are approximately log-normally distributed.

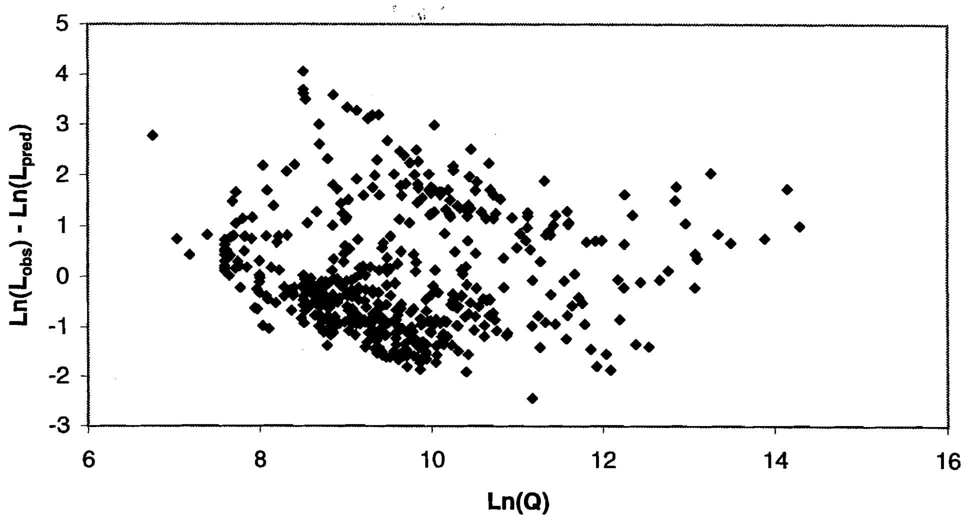

Figure 5. *Residuals ($Ln(L_{obs})$-$Ln(L_{pred})$) for the linear regression in Fig.2.*

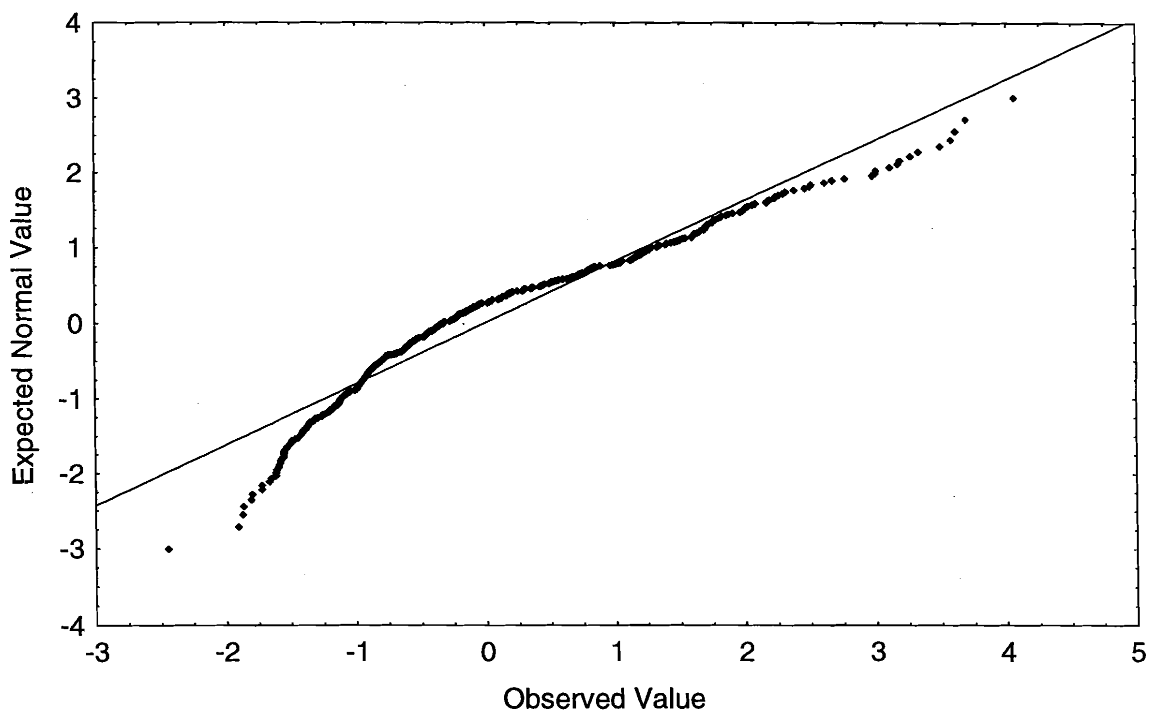

Figure 6. *Probability plot of the residuals for standard linear regression.*

Given the inadequacy of the linear regression, several options are available including non-linear regression. We choose to use the modern approach of Locally Weighted Scatterplot Smoothing (LOWESS). First, however, we outline the issue of bias correction.

2.1 Bias correction

The use of logarithms of data for constructing sediment rating curves means that the regression lines are fitted to the geometric means of the data. As these are less than the arithmetic means, a bias correction factor is necessary to achieve an unbiased load. This is given by:

$$BCF = \exp(s^2/2)$$

where s^2 is the variance of the residuals. (This result derives from the method of moments equations for fitting a log-normal distribution, and its application for sediment load estimation is detailed in Ferguson (1987)). In the form given above, there are the implicit assumptions that the residuals are log-normally distributed (which they are in this case, as shown above) and that the variance is independent of flow magnitude, which, as shown above, is not the case.

2.2 LOWESS estimates

LOWESS provides an objective empirical approach to curve fitting which requires no a priori assumptions about the form of the relationship. It can be superior to standard least squares regression in cases such as that shown in Figure 4 where lnL – lnQ relation exhibits curvature.

Locally weighted scatterplot smoothing was applied to the data, with a range of "stiffness" estimates which control the extent of the smoothing. A reasonable fit of a LOWESS curve is shown in Figure 7. Conventional LOWESS analysis uses a constant "stiffness" factor f over the entire range of the independent variable, but in this case we changed the value of f to improve the fit because there are fewer data points at the high-discharge tail of the rating relationship. The resulting compound LOWESS curve uses two values of the stiffness factor f: f = 0.02 for low flows (flows less than 353 l/s) and f = 0.1 for higher flows (flows greater than 353 l/s).

Because of the variation of the variance of the residuals (Fig. 5), the BCF is calculated for a localised estimate of the variance of the (log-spaced) residuals, s_i^2. The localised variance is calculated over a window of data consistent with the data window made in the curve fitting. Figure 7 includes a second LOWESS curve which is the compound LOWESS curve corrected by the BCF, where the BCF is calculated using local values of variance. The variance is calculated across a window of values matching the LOWESS curve stiffness factor, f, as described by Hicks *et al.* (2000).

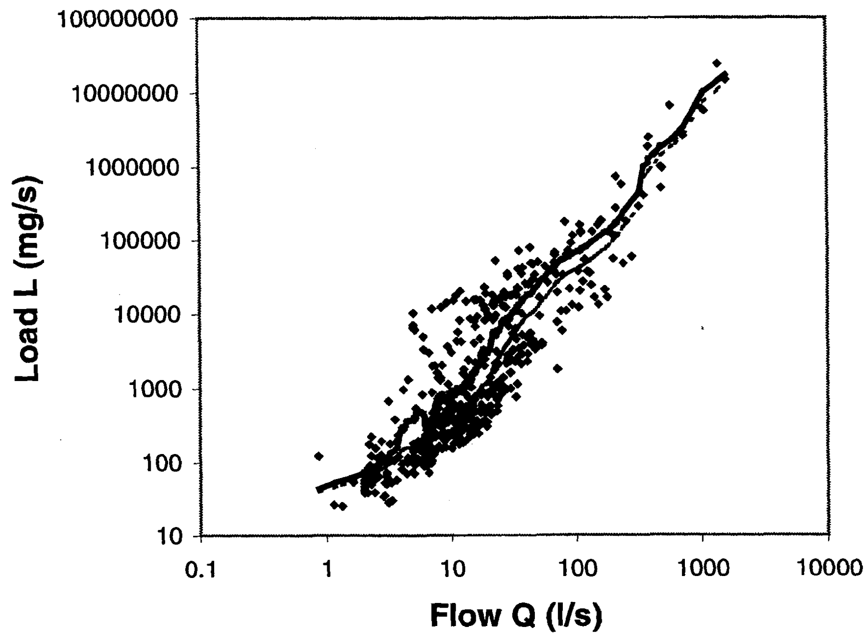

Figure 7. *Pre-harvest data with a compound Lowess curve fitted, without and with bias correction. The dashed line is the compound Lowess line fitted with two stiffness values, f=0.1 and f=0.02. The solid line is the same Lowess fit with bias correction applied, using local variance estimate as explained in the text.*

2.3 Error estimates

We used the method of Verhoff *et al.* (1980) to estimate a confidence interval on the sediment yield estimate. This sums the conditional variance in the load estimate, weighted by the proportion of time that each given discharge occurs. We estimated the conditional variance, S_i^2, as

$$S_i^2 = (\exp(s_i)-1)*L_i)^2$$

Where s_i is the conditional standard error of the residuals (using log values) described above and L_i is the load estimated for the i-th discharge value using the bias corrected rating relationship. The confidence interval on the yield estimate, Y, becomes

$$Y \pm t * \sqrt{(\Sigma S_i^2)}$$

Where t is the appropriate Students-t statistic for given degrees of freedom, ν, and significance level. $\nu = f*N - 2$, where f is the LOWESS stiffness factor and N the total number of data values in the rating relationship.

2.4 Summary of estimates for the pre-harvest period

In Table 1 the alternative estimates of the sediment load for the pre-harvest period are listed. These results show that the bias correction increases the estimates substantially. Even with the simple bias correction applied, the standard linear regression estimate is low. The compound LOWESS estimate with the local bias correction factor applied is closest to the direct estimate, and it agrees with the direct estimate to within one standard error. This, and the preceding comments about the adequacy of the simple linear regression approach, justifies our using the LOWESS method for estimating the sediment rating and sediment yield for the subsequent periods of catchment disturbance by roading and logging.

3 DIFFERENCES BETWEEN THE PRE-HARVEST AND HARVESTING PERIOD YIELDS

One difficulty with comparing sediment yields between two relatively short periods (of the order of 1-2 years in this case) is that the sediment yield may be affected as much by differences in runoff as by differences in erosion processes or sediment sources. For example, Hicks (1994) showed that for Manukau Basin, near Auckland, the potential sediment yield could vary by as much as a factor of 10 from one year to the next due to variations in runoff. To avoid this complication at Redwoods Stream, we chose to compare the potential sediment yields for the same reference time period and series of water discharges for which we computed the pre-harvest yield (i.e. from 21 May 1994 until 25 December 1995). Thus for the harvesting phase, we derived a sediment rating from sediment load and water discharge data collected during the harvesting period, but we applied this rating to the pre-harvest discharge record to compute the *potential* harvesting yield.

We used the 370 data samples collected over the period 13 February 1997 until 31 March 1998 to derive a sediment rating representative of the harvesting phase. With these data, we applied the same LOWESS procedure as above to fit a rating relationship, and used the conditional variance to correct for logarithmic bias and to estimate confidence intervals on the resulting sediment yield. The only difference was that this time, after inspecting the data, we chose a constant stiffness factor of f = 0.15 to fit the LOWESS curve. The basis for this choice was the relatively smooth curve smooth curve that it gave, as shown in Figure 8. The bias-corrected rating relationship for the harvesting phase is shown in Figure 8. We extrapolated this rating to cover the full range of water discharge over the reference period.

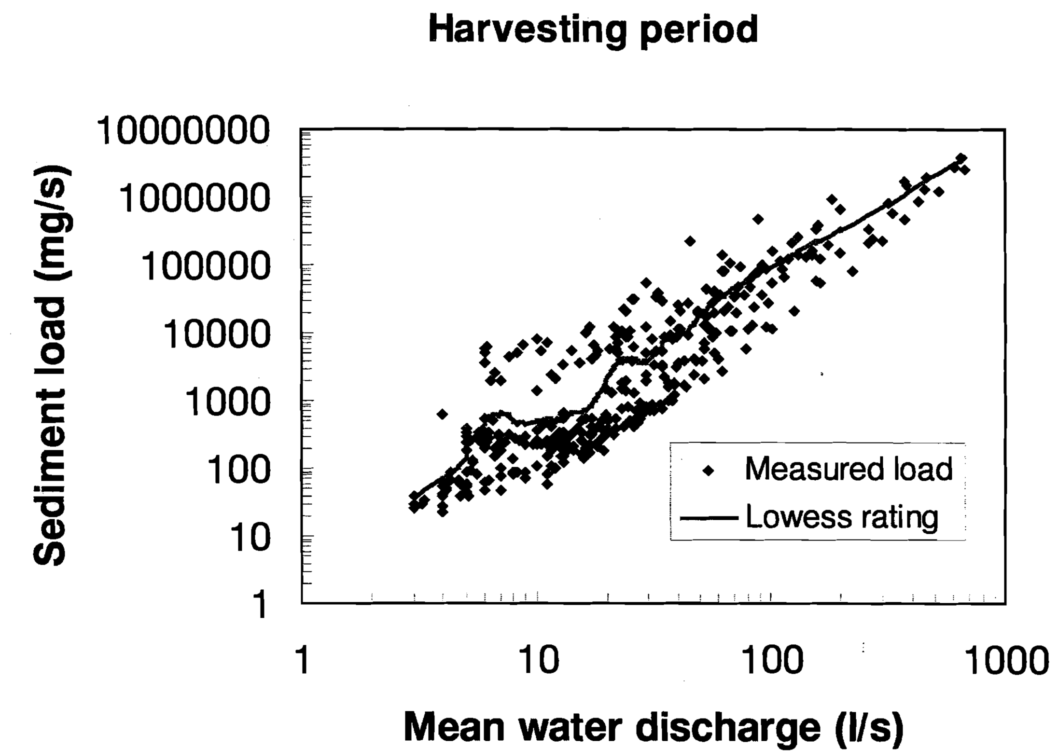

Figure 8. *Data and bias-corrected LOWESS rating for the harvesting period.*

The potential harvesting yield so estimated (Table 2) is 1.09 times the LOWESS-rated yield for the pre-harvest period. However, since the yield for the pre-harvest period lies well inside the 95% confidence interval of the harvesting yield, we conclude that the harvesting activity has led to no significant change in sediment yield.

Table 2. *Comparison of sediment yield estimates for the period May 1994 – December 1995 using rating relationships derived for the pre-harvest and harvesting periods.*

Period and method	Load (t)	Lower bound, 95% confidence limit	Upper bound, 95% confidence limit
Pre-harvest, direct	169	-	-
Pre-harvest, LOWESS rating	152	97	206
Harvesting, LOWESS rating	166	90	242
Pre-harvest, standard regression	72	-	-
Harvesting, standard regression	92	-	-

Confidence limits for the linear regression estimates in Table 2 are not given as the requirements for the residuals to be free of bias and have constant variance are not met.

4 DISCUSSION

Modern computationally intensive statistical methods were used in this analysis in order to avoid the limitations of standard linear regression for defining sediment rating curves. One of the key limitations of the simple regression approach in this instance is that the residuals of the regression had statistics that varied with the magnitude of the flows. Failure to allow for this, and also to correct for logarithmic transformation bias, would have produced misleading results. As shown in Table 2, the pre-harvest period yield estimated using a linear regression to model the pre-harvest rating (uncorrected for logarithm bias) is 72 t, and compares with a yield of 92 t using a linear regression model of the harvesting period rating. Both yields are low compared to the more accurate estimates, and might be used to (incorrectly) interpret a significantg increase in sediment yield due to harvesting operations. Figure 9 shows the rating data for both the pre-harvest and harvesting periods, plus the linear regression models. Notice how the data for the two periods overlie the same area, but the regression line for the harvesting period is steeper, suggesting (incorrectly) significantly greater loads.

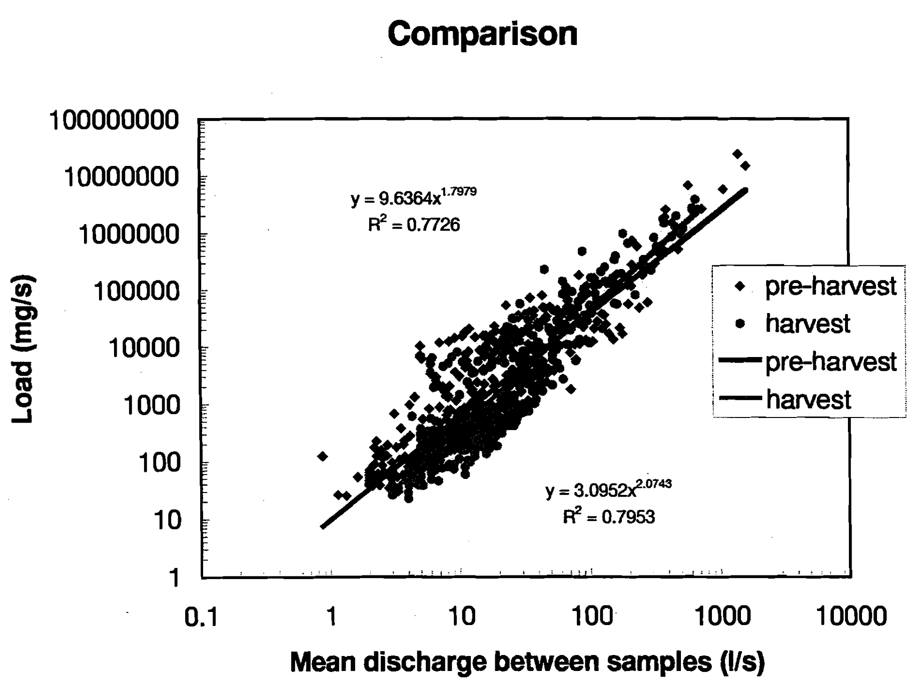

Figure 9. *Relationships between sampled load and water discharge for the pre-harvest and harvesting periods. The lines and equations are for linear regression fits to the data.*

5 CONCLUSION

The timber harvesting activity at Redwoods Stream, from February 1997 until March 1998, resulted in no significant change in sediment yield compared to the pre-harvest period.

The analysis reported in this study uses Locally Weighted Scatterplot Smoothing (LOWESS) to fit sediment rating curves and uses the conditional variance of residuals to correct for logarithmic transformation bias and to determine confidence intervals on yield estimates. Misleading results would arise from using simple linear regression to model sediment rating relationships.

6 REFERENCES

Ferguson, R.I. 1987. Accuracy and precision of methods for estimating river loads. *Earth SurfaceProcesses and Landforms 12*: 95-104.

Hicks, D.M.; Gomez, B.; Trustrum, N.A. 2000. Erosion thresholds and sediment yields, Waipaoa River basin, New Zealand. *Water Resources Research 36(4):* 1129-1142.

Hicks, D.M. 1994. Storm sediment yields from basins with various landuses in Auckland area. Report prepared for the Auckland Regional Council, July 1994. *NIWA Consultancy Report ARC802/1.*

Verhoff, F.H.; Yaksich, S.M.; Melfi, D.A. 1980. River nutrient and chemical transport estimation. *Journal of Environmental Engineering Division, ASCE, Vol.106, No. EE3*: 591-608.