

Sediment Attributes Stage 1

Prepared for Ministry for the Environment

June 2016

Prepared by:

D M Hicks
M Greenwood
J Clapcott (Cawthron Institute)
R Davies-Colley
J Dymond (Landcare Research)
A Hughes
U Shankar
K Walter

For any information regarding this report please contact:

Murray Hicks
+64-3-343 7872
m.hicks@niwa.co.nz

National Institute of Water & Atmospheric Research Ltd
PO Box 8602
Riccarton
Christchurch 8011

Phone +64 3 348 8987

NIWA CLIENT REPORT No: CHC2016-058
Report date: June 2016
NIWA Project: MFE16502

Quality Assurance Statement		
	Reviewed by:	Les Basher (Landcare Research)
	Formatting checked by:	Tracy Webster
	Approved for release by:	Jo Hoyle

© All rights reserved. This publication may not be reproduced or copied in any form without the permission of the copyright owner(s). Such permission is only to be given in accordance with the terms of the client's contract with NIWA. This copyright extends to all forms of copying and any storage of material in any kind of information retrieval system.

Whilst NIWA has used all reasonable endeavours to ensure that the information contained in this document is accurate, NIWA does not give any express or implied warranty as to the completeness of the information contained herein, or that it will be suitable for any purpose(s) other than those specifically contemplated during the Project or agreed by NIWA and the Client.

Contents

- Executive summary 11**

- 1 Introduction 16**
 - 1.1 Background 16
 - 1.2 Aim and objectives..... 16
 - 1.3 Report structure and interim reports 17

- 2 Data 18**
 - 2.1 Data requirements and sources..... 18
 - 2.2 Collation, processing, and conventions 20
 - 2.3 Database 20

- 3 Sediment rating curve parameter estimation models..... 23**
 - 3.1 Introduction 23
 - 3.2 SRC Data..... 24
 - 3.3 Sediment rating parameter fitting 27
 - 3.4 Predicting SRC parameters 29
 - 3.5 Discussion 50

- 4 How sediment rating curves change when catchment sediment load changes 51**
 - 4.1 Introduction 51
 - 4.2 Results from SRC datasets that capture a change in sediment load 51
 - 4.3 Sensitivity of SRC b-parameter to sediment load indicated by modelling 54
 - 4.4 Discussion 54

- 5 Mutual relationships between visual water clarity, turbidity and suspended matter in rivers..... 58**
 - 5.1 Introduction 58
 - 5.2 Data..... 60
 - 5.3 ESV inter-relationships and parameters..... 63
 - 5.4 c^* 69
 - 5.5 Predicting power law coefficients and c^* 70
 - 5.6 Discussion 79

- 6 Suspended sediment particle size distributions and relationship to sediment load 80**
 - 6.1 Introduction 80

6.2	Spatial variation in SS particle size and its relationship with sediment load and other catchment characteristics	80
6.3	Change in SS particle size associated with changes in catchment sediment load .	90
7	Relationships between streambed deposited sediment and sediment load	93
7.1	Introduction	93
7.2	Summary of data available	93
7.3	Regression analyses	94
7.4	Discussion	108
8	Analytical frameworks to link catchment sediment loads and sediment environmental state variables	110
8.1	Introduction	110
8.2	Flow duration curve	111
8.3	Analytical framework 1	112
8.4	Analytical framework 2	122
9	Accuracy of analytical frameworks and their fitness-for-purpose	130
9.1	Introduction	130
9.2	Robustness of analytical frameworks	130
9.3	Utility of frameworks for national-scale assessment of change in sediment load to meet sediment ESV targets.....	133
10	Framework for further work.....	135
10.1	Introduction	135
10.2	Further research needs by topic.....	135
10.3	Research needs by motivation/type.....	141
10.4	Overarching needs.....	142
10.5	Value of research	142
11	Conclusions	150
12	Acknowledgements	155
13	Glossary of abbreviations and terms	156
14	References.....	161
Appendix A	Statement of work – Description of services	166
Appendix B	Data-seek letter	168

Appendix C	Database conventions	171
Appendix D	MS-Access database tables.....	185
Appendix E	Random Forest regression approach.....	191
Appendix F	Inter-relationships between visual clarity, turbidity, and total suspended sediment	192
Appendix G	Suspended sediment particle size variation during floods: Mararoa River example	196
Appendix H	Sediment Assessment Methods.....	197
Appendix I	Standard error of regression line	198
Appendix J	Derivation of suspended sediment percentiles from flow percentiles	199

Tables

Table 3-1:	Distribution of sites across climate, topographic source of flow, and Strahler stream order classifications.	27
Table 3-2:	Predictor variables used in Random Forest models.	32
Table 3-3:	Descriptions of four Random Forest models used to predict rating curve coefficients.	33
Table 3-4:	Out-Of-Bag (OOB) R^2 results for Random Forests predicting intercept and slope values for the three rating curve types.	34
Table 3-5:	Results from individual one-way ANOVAs testing the effect of source of flow category (SOF) on coefficients from the three rating curve methods.	37
Table 3-6:	Mean (\pm 1 SD) coefficient values for SRCs generated using three different rating curve methods for sites in different source of flow (SOF) categories.	38
Table 3-7:	Summary of methods for calculating SSC.	41
Table 3-8:	Model performance metrics for different methods of predicting intercept and slope coefficients.	43
Table 3-9:	Mean model performance metrics for different methods of predicting log SSC.	46
Table 5-1:	Linear regression statistics for relationships between VC, TSS, and turbidity for pooled NRWQN dataset for cases of all sites and sub-set with unaltered flow regimes.	65
Table 5-2:	Summary statistics for the at-a-site SMA regression intercepts and slopes for the mutual relationships of VC, TSS and turbidity and for the site-median c^* .	66
Table 5-3:	Linear regression results for models predicting SMA regression intercepts (a) from slopes (b) for the mutual relationships of VC, TSS and turbidity.	69
Table 5-4:	Predictors variables used in Random Forest models for predicting c^* and the intercepts and slopes of the mutual relationships of VC, TSS and turbidity.	71

Table 5-5:	Performance results for Random Forest models predicting the intercepts (<i>a</i>) and slopes (<i>b</i>) for the mutual relationships of VC (visual clarity), TSS, and turbidity as well as site-median <i>c</i> *.	73
Table 6-1:	Distribution of sites across climate, topographic source-of-flow, geological, and Strahler stream order classifications as extracted from the REC.	83
Table 6-2:	Predictors variables used in Random Forest models for predicting Ave%<16 and Ave%<63.	84
Table 6-3:	Distribution of sites across alternative lithology grouping.	85
Table 6-4:	Model performance metrics for Random Forest models of Ave%<16 and Ave%<63.	86
Table 6-5:	Specific turbidity at Motueka at Gorge, Motueka at Woodmans Bend, and Motupiko at Christies before and after Easter 2005 storm.	92
Table 7-1:	Number of deposited fine sediment data for each sample method.	94
Table 7-2:	Mean and range of variables used in regression analyses.	95
Table 7-3:	Statistics of <i>SIS</i> data and <i>SSC</i> at one quarter the mean annual flood and at the mean annual flood at sites with both data types available.	106
Table 9-1:	Definitions for robustness of prediction of median <i>ESV</i> value associated with a 50% reduction in catchment sediment load.	131
Table 9-2:	Robustness of methods used to predict sediment <i>ESVs</i> for frameworks 1 and 2 for the Wairua at Purua example where sediment attributes are measured regularly.	131
Table 9-3:	Uncertainties in component relationships for analytical frameworks, as derived in previous sections.	132
Table 9-4:	Expected robustness of methods used to predict sediment <i>ESVs</i> for frameworks 1 and 2 for sites where sediment attributes are not measured.	133
Table 10-1:	Summary of further research needs to improve understanding, fill gaps in data coverage, test hypotheses, and/or develop more sophisticated models within the analytical chain linking <i>ESVs</i> to catchment sediment load.	143
Table 10-2:	Value definitions for research topics in terms of relative cost and scientific importance.	149

Figures

Figure 2-1:	Relationships between tables in the MS-Access database.	22
Figure 3-1:	Location of sampling sites.	26
Figure 3-2:	Unstandardised and standardised relationships between <i>SSC</i> (<i>C</i>) and water discharge (<i>Q</i>) at 10 sites in Central Otago.	28
Figure 3-3:	Intercept and slope parameters for linear regressions fitted to individual sites (<i>n</i> = 271) using three rating curve variations.	31
Figure 3-4:	The importance of individual predictors to each of the four RF models.	35
Figure 3-5:	Univariate partial dependence plots showing the shape of relationships between the two most important predictors for the four different Random Forest models: Slope, Int_O/E, Int_WRENZ, and Int_unstand.	36
Figure 3-6:	Univariate partial dependence plot showing the shape of relationships between SRC slope and catchment sediment yield.	37
Figure 3-7:	Intercept and slope coefficients for linear regressions fitted to individual sites (<i>n</i> = 271) coded by source of flow (SOF) categories for three rating curve	

	methods: SSC standardised by C_{mean} calculated from O/E sediment loads (top), SSC standardised by C_{mean} calculated from sediment load from WRENZ (middle), and unstandardized SSC with points coded by SOF category (bottom).	39
Figure 3-8:	Predicted versus observed SRC parameters generated using the methods in Table 3-7.	42
Figure 3-9:	Predicted versus observed SSC for all site visits generated using the methods in Table 3-7.	45
Figure 3-10:	Boxplot of the ratio of the root mean square error to the standard deviations of the observed data (RSR) for observed versus predicted suspended sediment concentrations generated using eight different methods across 271 sites.	47
Figure 3-11:	Boxplot of Nash Sutcliffe Efficiency (NSE) for observed versus predicted suspended sediment concentrations generated using eight different methods across 271 sites.	48
Figure 3-12:	Boxplot of root mean square error (RMSE) for observed versus predicted suspended sediment concentrations generated using eight different methods across 271 sites.	49
Figure 4-1:	SRCs for: (a) Motueka at Gorge, (b) Motupiko at Christies, (c) Motueka at Woodmans Bend before and after the Easter 2005 rainstorm; (d) Waipaoa at Kanakanaia and (e) Mangatu at Omapere before and after Cyclone Bola (1988); and (f) Okura at Weiti Forest during and after forest harvesting.	52
Figure 4-2:	Simple examples of change in catchment-outlet SRCs arising from relative change in sediment loads from tributaries.	56
Figure 4-3:	Conceptual trends in the relationships between river discharge (Q) and suspended-sediment concentration (C) as expressed by changes in the sediment rating curve slope and vertical offset (from Warrick 2015, Figure 2).	57
Figure 5-1:	Mutual relationships of VC (black disc visibility), turbidity and TSS at a monitoring site (PW3; 49 ha) at the Whatawhata Research Station (1995-2014).	62
Figure 5-2:	Relationship between VC (as measured by the black disc method) and total suspended solids concentration at the un-impacted NRWQN sites.	64
Figure 5-3:	Histograms of power function parameters (from SMA regression) for VC versus TSS plots.	67
Figure 5-4:	Relationships between visual clarity and turbidity (left) and between TSS and turbidity (right) using pooled data from the 64 un-impacted NRWQN sites.	68
Figure 5-5:	Relationship between beam attenuation coefficient (BAC) and TSS at the 77 NRWQN sites.	69
Figure 5-6:	Distribution of c^* for (a) all NRWQN data (2011-2015) and for (b) the median values for each site.	70
Figure 5-7:	Plots of c^* versus four REC2 derived catchment characteristics for the 64 NRWQN sites with un-impacted flow regimes.	71
Figure 5-8:	Predicted versus observed coefficient values for the full NRWQN site dataset (n = 77) generated using RF models.	74
Figure 5-9:	Predicted versus observed coefficient values for the non-flow impacted dataset (n = 64) generated using RF models.	75
Figure 5-10:	The importance of individual predictors for the three RF models with OOB $R^2 > 0.2$ for the full dataset (n = 77).	76

Figure 5-11:	The importance of individual predictors for the three RF models with OOB $R^2 > 0.2$ for the non- flow impacted dataset (n = 64).	76
Figure 5-12:	Univariate partial dependence plots showing the shape of relationships between the three most important predictors for the 3 RF models run on the full data set.	77
Figure 5-13:	Univariate partial dependence plots showing the shape of relationships between the three most important predictors for the 3 RF models run on the dataset excluding flow impacted sites (n = 64).	78
Figure 6-1:	Locations of sites with SS PSD data.	82
Figure 6-2:	Relationship between median grainsize (D50) and % mud (Ave%<63) for average PSDs from 59 sites.	82
Figure 6-3:	Averaged PSDs by lithology groups.	86
Figure 6-4:	Observed against hold-one-out cross-validation (CV) predicted values for the Ave%<16 and Ave%<63 PSD variables.	87
Figure 6-5:	The importance of individual predictors to each of the two particle size RF models.	88
Figure 6-6:	Univariate partial dependence plots showing the shape of relationships between the four most important predictors for the Random Forest models for Ave%<16 and Ave%<63.	89
Figure 6-7:	Specific turbidity vs particle size for two turbidity sensor types, from Foster et al. (1992).	90
Figure 6-8:	Turbidity vs SSC trends before and after the Easter 2005 storm at three sites in the Motueka catchment.	91
Figure 7-1:	Distribution of deposited sediment samples.	94
Figure 7-2:	Correlation between $\log(\text{Catchment sediment yield})$ and deposited sediment metrics.	96
Figure 7-3:	Correlation between $\log(\text{Segment sediment yield/stream power})$ and deposited sediment metrics.	97
Figure 7-4:	Partial plots for the main effects of a simple general regression model of <i>Fines</i> as a function <i>Segment sediment yield/stream power</i> , <i>Catchment sediment yield</i> , <i>method</i> , <i>mesohabitat</i> , and <i>CSOF</i> . Solid lines show line-of-best-fit and dashed lines show 95% confidence intervals.	98
Figure 7-5:	Univariate partial dependence plots (smoothed fitted functions) of the relationships between <i>Fines</i> and <i>Catchment sediment yield</i> , <i>Segment sediment load/stream power</i> and environmental descriptors, and sampling <i>method</i> and location identifiers.	99
Figure 7-6:	Univariate partial dependence plots (smoothed fitted functions) of the relationships between <i>Fines</i> collected in run habitat and <i>Catchment sediment yield</i> , <i>Segment sediment load/stream power</i> and environmental descriptors, and sampling <i>method</i> and location identifiers.	100
Figure 7-7:	Scatter plots of observed versus predicted values from a) all data and b) run habitat-restricted data model for <i>Fines</i> .	101
Figure 7-8:	Frequency distribution of predicted proportional reduction in % fine sediment cover (<i>Fines</i>) in response to a 50% reduction in <i>Segment sediment load</i> from the BRT model fitted with run-habitat data.	101
Figure 7-9:	Partial plots for the main effects of a general linear regression model of <i>SIS</i> as a function of <i>Segment sediment load/stream power</i> , <i>Catchment sediment yield</i> ,	

	and <i>CSOF</i> . Solid lines show line-of-best-fit and dashed lines show 95% confidence intervals.	102
Figure 7-10:	Univariate partial dependence plots (smoothed fitted functions) from a BRT model of the relationships between $\log(SIS)$ and sediment load and other environmental descriptors.	103
Figure 7-11:	Partial plots for the main effects of a simple linear regression model of <i>Shuffle</i> as a function of <i>Segment sediment load/stream power</i> , <i>Catchment sediment yield</i> , and <i>CSOF</i> . Solid lines show line-of-best-fit and dashed lines show 95% confidence intervals.	103
Figure 7-12:	Univariate partial dependence plots (smoothed fitted functions) from a BRT model of the relationships between <i>Shuffle</i> and sediment load and other environmental descriptors.	104
Figure 7-13:	Sites with both <i>SIS</i> measurements and reliable SRC useful for testing the “frozen bedload” hypothesis.	105
Figure 7-14:	Relationship between <i>SIS</i> and suspended sediment concentration (SSC) at quarter of the mean annual flood discharge (Q_{maf}).	106
Figure 7-15:	Scatterplot of temporal variation of <i>Fines</i> at (left) one segment and (right) all segments where more than one replicate was measured in relation to native forest cover.	107
Figure 8-1:	Schematic diagram indicating the analytical chain that links VC, LP, SSC, and DS to the management of catchment sediment sources (i.e., framework 1).	111
Figure 8-2:	Flow duration curve of the Wairua River at Purua.	112
Figure 8-3:	Log-log plot of SSC versus discharge for the Wairua River at Purua.	113
Figure 8-4:	Log-log plot of VC versus SSC for the Wairua River at Purua.	114
Figure 8-5:	Median VC as a function of % reduction in sediment load for the Wairua River at Purua.	116
Figure 8-6:	Increase in median VC as a function of % reduction in sediment load for the Wairua River at Purua.	117
Figure 8-7:	Log-log plot of ED versus SSC for the Wairua River at Purua.	118
Figure 8-8:	Median ED as a function of % reduction in sediment load for the Wairua River at Purua.	119
Figure 8-9:	Increase in ED as a function of % reduction in sediment load for the Wairua River at Purua.	120
Figure 8-10:	Decrease in embeddedness as a function of % reduction in sediment load for the Wairua River at Purua.	121
Figure 8-11:	Framework 2, linking catchment load of optical cross-section (LOCS) with visual clarity and light penetration via relationship between beam attenuation coefficient (BAC) and water discharge.	122
Figure 8-12:	Log-log plot of BAC versus discharge for the Wairua River at Purua.	124
Figure 8-13:	Log-log plot of BAC versus SSC for the Wairua River at Purua.	125
Figure 8-14:	Median VC as a function of % reduction in sediment load for the Wairua River at Purua (in framework 2).	126
Figure 8-15:	Increase in median VC as a function of % reduction in sediment load for the Wairua River at Purua (under framework 2).	126
Figure 8-16:	Log-log plot of ED versus BAC for the Wairua River at Purua.	127
Figure 8-17:	Median ED as a function of % reduction in sediment load for the Wairua River at Purua.	128

Figure 8-18: Increase in median ED as a function of % reduction in sediment load for the Wairua River at Purua (under framework 2). 128

Executive summary

Background

Sediment is an important contaminant in freshwaters (and downstream coastal waters) in New Zealand. It affects ecosystem health through various modes of impact which can be quantified by four environment state variables (ESVs): suspended sediment concentration (SSC), visual water clarity (VC), light penetration (LP), and deposited fine sediment (DS). The NPS-FM does not currently define attributes for sediment; however, sediment has been identified as a priority area for the development of attributes and bottom lines for future revisions of the NPS-FM. One key challenge to implementing this concerns how to transform catchment sediment loads into ESVs. In the first instance, this transformation is required to make a national-scale assessment of how much sediment loads may need to change to achieve national ESV targets. Also, at the implementation stage, if regional councils choose to set sediment related objectives then being able to transform changes in sediment loads to changes in ESVs will help them to put in place justifiable actions and limits. This report presents results from research using existing data that addresses this challenge.

The basic analytical framework to transform sediment load to ESVs hinges on the relationship between SSC and water discharge, termed the sediment rating curve (SRC). When combined with a flow duration curve, the SRC enables the catchment sediment load to be determined and it also enables SSC exceedance percentiles to be determined. Relationships between SSC and VC, LP, and DS then enable these latter three ESVs to be linked back to sediment load, so that, for example, a given change in sediment load can be converted to an ESV value associated with a given exceedance percentile. This study explores these relationships and the assumptions involved in linking them together. Since the ESV inter-relationships are strongly influenced by sediment characteristics, notably the particle size distribution (PSD), this too is examined.

Objectives

The key study objectives were to:

- Collate all nationally available data held by research institutes and territorial authorities from which sediment rating curves (SRCs) can be defined and to support the development of methods to link SSC to VC and LP, turbidity, DS, and suspended load PSD.
- Develop models to estimate the parameters defining SRCs for locations without data.
- Determine how the parameters of SRCs change in response to changes in catchment sediment loads.
- Identify and characterise the relationship between turbidity, VC, SSC and LP and develop methods for predicting turbidity, VC, and LP as functions of SSC.
- Examine the extent to which sediment PSD changes with change in sediment load and in what circumstances, using PSD data if available or using specific turbidity as a proxy measure of PSD.
- Analyse relationships between sediment loads and measures of streambed DS, conclude if an empirical approach can be developed and, if so, scope what new data is required to deliver functional relationships.

- Provide analytical frameworks for the use of these methods to determine catchment sediment load limits to achieve objectives that are enumerated in terms of the ESVs.
- Estimate and describe the sensitivity of the each step in the analytical chain and indicate the steps that most limit the accuracy of the analysis.
- Provide a framework for further work to develop more sophisticated and accurate methods for relating ESVs to catchment loads.

Predicting sediment rating curve parameters at any river site

Sediment rating curves (SRCs) were fitted to 271 sites broadly scattered around New Zealand with a model of the general form $C/C_{\text{mean}} = a(Q/Q_{\text{mean}})^b$ (where C is SSC, Q is water discharge, and C_{mean} and Q_{mean} are their respective annual mean values). Random Forest (RF) regression models were used to relate the **a** and **b** parameters to catchment/site characteristics.

The main factor determining the **a**-parameter was found to be catchment sediment yield, with lesser influences from land-cover and soil texture. The main factors determining the SRC **b**-parameter were those linked to catchment slope (i.e., unit stream power, elevation, steepness, stream order). These findings generally align with those observed in international datasets.

The SRC **a**-parameter at any site can be predicted to a factor of $\times/\div 2.29$, which aligns with the accuracy of sediment yield estimators developed previously from a similar dataset.

The RF models can be used to predict SRC parameters for all stream segments in the national network of channels that underpins the River Environment Classification (REC, version 2).

How do sediment rating curves change if the catchment load changes?

A fundamental assumption of the proposed analytical framework is that only the **a**-parameter of the SRC changes as catchment sediment load changes. However, analysis of SRC changes in catchments experiencing floods or land-cover change that changed the catchment sediment supply showed mixed effects on the SRC parameters. While all catchments experienced significant change in the **a**-parameter (by factors ranging from 1.6 to around 4), two also experienced significant changes in the **b**-parameter.

A common factor at sites where the SRC **b**-parameter did not change was that the event's impact was reasonably uniform over the catchment. Of the two sites where the SRC **b**-parameter did change, one had experienced changes in sediment supply over only part of its catchment, whereas the other likely experienced an increase in runoff as a result of forest harvesting. These observations, corroborated by overseas observations and a simple modelling exercise, showed that the assumption of a stable SRC **b**-parameter under changing catchment sediment load does not hold-up where tributary SRCs and sediment load changes are not uniform within a catchment and/or load change is accompanied by a change in runoff.

This is not expected to be an important factor for national-scale assessment of the effects of sediment load change on ESVs, since at that scale it is reasonable to assume that sediment load change is uniformly distributed across catchments. However, it may be important when implementing regional sediment management policy (for example, where erosion control may be focussed in priority areas such as eroding stream banks). In such cases, a potential way forward would be to add into the analytical framework a simple water and sediment routing model that

calculated downstream changes in both **a** and **b** parameters following localised changes in sediment load.

Relationships between suspended sediment concentration, visual clarity, and turbidity and their prediction

Strong inter-relationships were found between SSC (represented by laboratory measurements of total suspended solids, TSS), VC, and turbidity in rivers, consistent with previous work. However there was appreciable variability, with about a 10-fold range in VC at a given TSS observed across diverse rivers. VC and TSS are inversely related but not perfectly, such that a halving of TSS typically does not double VC but increases it only about 65%.

RF models had only weak ability to explain the observed regional variation in the parameters defining these various relationships – indicating that catchment variables in readily available databases do not well capture the particle size, shape, and composition that theoretically together control light attenuation by sediment and thus VC. Nonetheless, while the amount of variance in the source data explained by the regressions models was small, the end-result in regard to parameter prediction accuracy still appears reasonable. For example, the standard error on the *a*-parameter in the relationship between VC and TSS (i.e., $VC = aTSS^b$) was $\pm 17\%$.

Suspended sediment particle size distribution and its controls

Suspended sediment particle size distribution (PSD) is important because it is the main factor controlling the relationships between SSC and VC and LP. Thus understanding the factors that control PSD underpins understanding of (i) regional variation in relationships between VC, SSC, and turbidity (which can be calibrated as a proxy for both), and (ii) whether a change in the sediment load from a catchment (e.g., after erosion mitigation work) might also change these relationships by changing the PSD of the catchment sediment load.

RF models were developed from sampled suspended sediment PSD data to predict both the fine silt and clay component (i.e., finer than 16 μm) of the suspended load and the mud component (i.e., finer than 63 μm). The mud component was also found to strongly and inversely correlate with suspended sediment median particle size. The main factor influencing the regional variation of mud content was catchment lithology, but with some weaker control also exerted by land-cover, sediment supply, elevation and rainfall. After lithology and land-cover, the % of fine silt and clay was also influenced by temperature, elevation, and rainfall. The RF models explained 39% of the observed variation in mud content and 34% of the variation in fine silt and clay content. Unfortunately, a lack of concurrent data on PSD, SSC, and VC limits the extent that these PSD predictors can be used to improve prediction of the SSC – VC relationship parameters.

Does the sediment particle size distribution change if the sediment load changes?

No PSD data were available to directly assess changes in PSD accompanying changes in sediment load, but we used changes in specific turbidity (the ratio of turbidity to SSC) to infer changes in PSD from two catchments following extreme hydrological events, one of which (the Motueka) had a well-recorded increase in sediment loads during and following the event. We conclude that changes in sediment load can cause changes in the PSD and so changes in the relationships between SSC and optical properties. Whether these changes are significant will depend on (i) the PSD of the sediment delivered from the affected sources compared with the catchment-average PSD, and (ii) the importance of that source to the total sediment load.

Controls on deposited sediment

Our analysis of DS data showed very weak dependence on catchment sediment load – certainly not enough to justify any functional relationship. While not initially expected, this is perhaps not that surprising given issues with data collection and because the bulk of river sediment delivery is likely to occur at periods of high flow that would effectively flush the sediment through the stream network to receiving environments. The importance of low elevation and low channel slope in explaining variance in DS metrics in our models supports this hypothesis. A measure of antecedent flow is likely to improve our ability to link DS to catchment loads because we could then account for time since last bed disturbance.

Our results suggest that the local delivery of sediment and how it is ‘captured’ by the local stream morphology is more informative of DS than sediment load from the upstream catchment. The implications for management could be a focus on local habitat to minimise the chronic delivery of fine sediment that occurs during stable flows, sourced, for example, from eroding stream banks.

The addition of other environmental variables (describing elevation, slope, geology and flow) improved our ability to predict deposited sediment; however, the explanatory power of any of the models is modest. Similarly modest results were observed in a recent United Kingdom study, which also showed that stream power (the product of channel slope and flow) was the most significant explanatory variable of DS and that the influence of sediment load was small.

The “frozen bedload” hypothesis predicts a direct relationship between DS and the SSC at $\frac{1}{4}$ the mean annual flood discharge, which is when bedload motion is assumed to stop on flood recessions and thereby trap suspended sediment within substrate pore spaces. There were insufficient data to robustly test this hypothesis; however, the available data showed no such relationship. Indeed, if anything the data suggests that fine sediment may accumulate progressively at flows below this discharge.

Analytical frameworks linking environmental state variables to catchment sediment loads

Two analytical frameworks were developed that enable ESV targets (e.g., VC threshold not exceeded more than a certain % of time) to be quantitatively related to catchment sediment load.

Framework 1 services all four ESVs and links VC, LP, and DS to sediment load via SSC. Framework 2 only services VC and LP and links these to the “load of optical cross-section” (LOCS) via the beam attenuation coefficient (BAC), which is an optical equivalent to SSC and is effectively the inverse of VC. A drawback with framework 2 is that it remains unclear how the LOCS relates to the actual load of sediment and to sediment sources. In both frameworks, flow duration curves are combined with rating curves (sediment vs discharge for framework 1; BAC vs discharge for framework 2) to link the ESVs with loads and to assign them exceedance percentiles.

Using relationships based directly on measurements at a case-example site, the robustness (as defined by standard error) of predictions of VC and DS associated with a given change in sediment load was weak while the robustness of predictions of LP (represented by the euphotic depth) and SSC was moderate. The robustness improved if only the *change* in ESV was predicted from a given % change in sediment load. In the case of VC, framework 2 was more robust in predicting VC than was framework 1. The main reason for this, though, was that the site studied had more measurements of VC than SSC, which may not be the case at other measurement sites.

Using the national regression models developed in this study to predict the relationships, the larger error terms rendered both frameworks only weakly robust when predicting absolute values of the ESVs associated with given changes in sediment load. However, the predictions of *change* in ESV (from an unknown initial state) were more robust.

Utility of analytical frameworks for national-scale assessment of change in sediment load to meet sediment environmental state variable targets

Framework 1, which links the regression relationships developed in this study, provides a workable approach for making national-scale assessments of the implications to sediment loads of setting national targets for SSC, VC, and LP. However, its accuracy will be weak if targets for these ESVs are set in terms of absolute values. This is largely due to the high uncertainty in estimating SRC parameters. Results will be more robust if targets are set in terms of *change* in ESV associated with a change in sediment load. Framework 1 cannot be used with any confidence at all to assess the implications of DS targets because of the lack of any reliable relationships between sediment load and DS.

Framework 2 remains work-in-progress, has more limited utility, and is therefore not recommended at this stage.

Framework for further research

Numerous areas were identified where further research and development is required to enable management of fine sediment in waters under the NPS-FM. These have been tabulated, classified, detailed by site and methodology, and prioritised in regard to scientific importance, relative cost, and net value (i.e., importance/cost ratio).

Further data collection is required to fill gaps in geographic coverage and/or to provide a stronger basis for refinement of predictor models, notably for SRC parameters and particle size characteristics. Research to improve understanding includes detailed monitoring of the processes controlling fine sediment deposition. Several simplifying assumptions require further testing with existing or new data, including: (i) that the concentration duration curve (CDC) links directly to the flow duration curve through the SRC; and (ii) that changes in catchment sediment load do not change PSD and ESV inter-relationships. More sophisticated model development includes improvements to sub-models in the analytical frameworks to deal with: (i) the general case where a change in catchment sediment load causes both SRC parameters to change downstream; and (ii) cases where the above listed assumptions may fail (e.g., an improved CDC predictor is required).

The research topics/questions of greatest value are generally those that are high in scientific importance and low in cost (which typically means use of existing datasets rather than collecting new ones). Examples include testing whether the CDC estimated using the sediment rating and flow duration curves matches the observed CDC, and adding complexity to sub-models in the analytical framework.

Nonetheless, there is an overarching need for further field data that is collected in an integrated way, with concurrent measurements (of high quality) of SSC, VC and particle characterisation (PSD, shape and composition) conducted at base flow and over high flow events. This could be done in a relatively small number of dedicated 'sediment' sites in diverse experimental catchments. Also, at these or other sites, sampling of DS should be extended to experimental studies of sedimentation processes to resolve whether there is any functional relationship to be found between sediment load (or SSC) and DS (or not).

1 Introduction

1.1 Background

The National Policy Statement for Freshwater Management (NPS-FM) requires regional councils, through their regional plans, to set freshwater objectives that provide for freshwater values, and to set limits and management actions to achieve those objectives. The NPS-FM contains attributes that assist regional councils to define freshwater (i.e., numeric) objectives and justifiable policies (including limits) for achieving these.

Sediment is an important contaminant in freshwaters (and downstream coastal waters) in New Zealand. It affects ecosystem health through various modes of impact which can be quantified by four environment state variables (ESVs): suspended sediment concentration (SSC), visual water clarity (VC), light penetration (LP), and deposited fine sediment (DS)¹. The NPS-FM does not currently define attributes for sediment; however it has been identified as a priority area for the development of attributes and bottom lines for future revisions of the NPS-FM.

The Ministry for the Environment (MfE) and the Ministry for Primary Industries (MPI) convened a process that considered how attributes for sediment can be developed. The process concluded (MfE 2015) that research and development is required to solve two key problems associated with sediment attributes: (i) the transformation of catchment sediment loads into ESVs; and (ii) defining numerical thresholds for sediment-related ESVs that relate to effects on ecology and other environmental values. MfE then commissioned a literature review that established the current state of knowledge internationally around these two problems (Davies-Colley et al. 2015).

MfE subsequently issued separate requests for research into these two problems. This report presents the results of research around the first problem: transformation of catchment sediment loads into ESVs.

1.2 Aim and objectives

As detailed in the Statement of Work (Appendix A), the research aim is to use existing data to develop methods that link catchment sediment loads to the ESVs at a level of accuracy and precision that enables regional councils to set sediment related objectives and to put in place justifiable actions and limits (for example, to achieve objectives on SSC exceedance percentiles in impacted waterways by limiting catchment sediment exports).

The research objectives were to:

- A. Collate all the nationally available data (up to 30 June 2015) held by research institutes and territorial authorities from which sediment rating curves (SRCs) can be defined, including data previously collated as part of the Ministry-funded Auckland Council project “Integrating three regional council sediment monitoring datasets for the purposes of calibrating a sediment yield predictive model for freshwater catchments”.
- B. Collate all the nationally available data held by research institutes and territorial authorities (up to 30 June 2015) to support the development of methods to link sediment concentrations to the ESVs, including flow data and flow duration curves, measured visual

¹ This report considers deposited sediment (DS) of mud and sand grades (finer than 2 mm). Hereafter, whenever deposited sediment is referred to it concerns deposited fine sediment.

water clarity and light penetration data, measured turbidity data, measured deposited sediment data, and particle size distribution (PSD) data.

- C. Assess the quality of the above data and identify a set of sites that maximises the geographic and environmental coverage of New Zealand's river catchments for which the SRC information is of sufficient quality for analysis.
- D. Develop models to estimate the parameters defining the SRCs for locations without data.
- E. Determine how the parameters of SRCs change in response to changes in catchment sediment loads.
- F. Identify and characterise the relationship between turbidity, VC, SSC and LP, and develop regionalisation(s) to provide methods for predicting turbidity and clarity and light penetration as functions of SSC.
- G. Examine the extent to which PSD changes with change in load and in what circumstances, using PSD data if available or using specific turbidity as a proxy measure of PSD.
- H. Analyse relationships between sediment loads and measures of streambed sediment deposition, conclude if an empirical approach can be developed and, if so, scope what new data is required to deliver functional relationships.
- I. Where appropriate, and based on the outputs of the above studies, provide analytical frameworks for the use of these methods to determine catchment sediment load limits to achieve objectives that are enumerated in terms of the ESVs.
- J. Estimate and describe the sensitivity of the each step in the analytical chain and indicate the steps that most limit the accuracy of the analysis.
- K. Provide a framework for further work to develop more sophisticated and accurate methods for relating ESVs to catchment loads. The framework should provide estimates of the amount of new data needed including the number of sites, sampling methods and frequencies and considerations for site location. The framework should describe how the data will be used to develop new methods or conduct research to test new ideas and assumptions.

1.3 Report structure and interim reports

This report generally addresses Objectives A-K above in sequence, but with data collation (Objectives A and B) combined in the one section and data quality and fitness-for purpose assessment (Objective C) split across the sections covering the various analysis topics.

This report assimilates material presented in two Interim Reports:

- Hicks, M. (2016) Interim Report on Data Assessment for Sediment Attribute Stage 1 Study. NIWA Christchurch Memo to Ministry for the Environment, 30 March 2016.
- Clapcott, J., Hicks, M. (2016) Interim Report on Deposited Sediment. NIWA Christchurch Memo to Ministry for the Environment, 15 April 2016.

2 Data

2.1 Data requirements and sources

2.1.1 Sediment data

Sediment related data were required for this study in three general areas:

- Data to create parameterised relationships between suspended sediment concentration (SSC) and discharge, visual water clarity, and turbidity.
- Data on suspended sediment composition, including information on particle size and organic content.
- Data on deposited sediment cover.

Associated metadata were also required on site information and data-collection purpose, field sampling/measurement method, instruments and laboratory methods used and associated standards/protocols.

Data were compiled from numerous sources. These included:

- An existing NIWA database with data from regional councils and the National River Water Quality Network (NRWQN) sites. This database (the “UnwinMfEDB”) was compiled for a previous NIWA contract for MfE (project MFE15503) and included:
 - previous regional council data compiled for a NEMAR project (1990 to “ragged end” dates between 2011 and 2013)
 - updated regional council data downloaded from the LAWA website (1st Jan 2004 to end of 2013, but only for 7 water quality variables)
 - the 77 NRWQN sites (1989 to end of 2011), which includes SSC, clarity and turbidity data.
- An existing database held by NIWA of suspended sediment gauging data. This was collated largely from data collected up to 2000 by NIWA and regional/district councils and precursor organisations (Ministry of Works and Development, Catchment Boards), and comprised data from concurrent suspended sediment and water discharge gaugings during freshes and floods. The sediment gaugings used depth-integrating samplers at multiple verticals, and provided the discharge-weighted, cross-section averaged SSC (using methods as described in, or similar to, Hicks and Fenwick 1993).
- A linked but smaller database of suspended sediment particle size analyses collected by NIWA through the 1990s, generally from depth-integrated multi-vertical gaugings but also from auto-sampler collected bank-side samples.
- The New Zealand Freshwater Fish Database (NZFFD), from which visual estimates of deposited sediment cover for fished reaches were extracted for all available cards (>8000). For sites in which the relative proportions of morphological habitat units (runs, riffles, pools, etc.) were reported, we compiled a ‘dominant’ morphological unit category if coverage of any one particular unit type was >50%.

- Data compiled for the generation of the deposited Sediment Assessment Methods (SAM) protocols by Clapcott et al. (2011).
- Data supplied for this study by regional/district councils that updates/adds to that provided by previous data requests. The data-seek letter sent to regional/district councils is shown in Appendix B.
- Miscellaneous other relevant datasets held by research organisations, including the authors of this report (listed in the relevant sections of this report).

Universities were excluded from the data search because of the time that would have been required to search through dissertations. Data was also not sought from consulting organisations for similar reasons and also because of potential proprietary issues. Scion Research declined to supply any data, citing proprietary concerns associated with the forestry industry.

The request to regional/district councils sought data up to 30 June 2015. In practice, the data supplied terminated at varying dates.

Sites were excluded that were not on mainland New Zealand (i.e., sites on Stewart Island or the Chatham Islands were not included).

2.1.2 Site/catchment data

For many of the analyses, information was also required on site or upstream catchment characteristics.

Generally, this information was extracted from the River Environment Classification version 2 (REC2). REC2 is a database of spatial attributes summarised for every segment in New Zealand's river network. River segments are defined based on a 30 metre digital elevation model of New Zealand (DEM). The database variables have been previously derived from analyses of the DEM (distances and slopes), maps (geology), remote sensing (land cover), and modelling and extrapolation of climate station and other observations (runoff and potential evaporation). REC2 is an update from the previous REC1, where major updates include a refining of the river network and segment locations and overlays of information from a more recent land cover database (LCDB3). Catchment/site parameters pre-calculated for the REC2 network for use in this study included mean annual suspended sediment load (Hicks et al. 2011) and channel width at mean flow (Booker and Hicks 2013).

In order to extract catchment information using the REC2, sites had to be assigned to river segments (NZSegments) within the REC2 database. This required that all sites had site location information. Eastings and Northings in the NZTM projection were used for this project.

NZSegment assignment involved a multi-stage procedure. Initially, sites were matched to the nearest NZSegment using the site coordinates and NZSegment centroid coordinates. However, this automated approach provided incorrect assignments in a minority of cases, typically where main-stem sites were matched to first order tributaries and where sites at junctions were placed on the wrong tributary or the wrong end of the junction. Such errors were checked for and were manually corrected in a workflow that included personal knowledge of sites, flagging all order 1 segments as suspect, manual inspection using ARCGIS layers, cross-checking against REC1 reach numbers, and cross-checking derived parameters such as catchment area and mean flow against independent data. In such cases, additional site naming/location information (e.g., "North Branch of Mistake Creek upstream from confluence") often proved invaluable in locating the correct NZSegment. Sites

that have had their NZSegment assignment checked are indicated in the MS-Access database SiteInfo table (see details in section 2.3.1). All NZFFD NZSegments are checked as the data is entered into the database (S Crowe, NIWA, pers. comm.) and so these sites were not rechecked. Sites with unconfirmed locations (in regard to segment number) were flagged and removed from the datasets used for analysis.

2.2 Collation, processing, and conventions

Relevant data from existing databases were extracted and compiled into a new MS-Access database (detailed in Section 2.3).

New data provided from councils was processed in a two stage process. First, relevant data for this study was identified and extracted² into 'compilation' Excel files that covered data useful for analysis of sediment rating curves (SRCs), SSC-VC-turbidity-LP inter-relationships, sediment composition, and DS, and metadata relating to site information and methods. Second, the compilation files were exported into the MS-Access database.

In order to make data comparable between the different sources we developed conventions for assigning project-specific and consistent site names and parameter identification codes. These are detailed in Appendix C.

2.3 Database

2.3.1 MS-Access database

A Microsoft Access database was created to store the data used in this project. The database was designed to have minimal replication of information across tables. Queries can be created to extract the data required for particular tasks.

The following tables were constructed (table details are provided in Appendix C and Appendix D):

1. A main data table that contains all the sample measurement information (DataTable; Table D-1).
2. A table with all relevant site location information, linked to the DataTable through the SSSiteID parameter (SiteInfo, Table D-2).
3. A table explaining the source database codes used in SiteInfo (SourceDB, Table C-1).
4. A table explaining the source organisation codes used in SiteInfo (Source, Table C-2).
5. A table with any catchment information extracted for an NZSegment during the course of analysis (Catchment Info, Table D-3). This table was linked to the SiteInfo table via the NZSegment parameter.
6. A table with any measured catchment information as provided by the source organisation (MeasureCatch, Table D-4). This table was linked to the SiteInfo table via the SSSiteID parameter.

² Often, the data supplied included parameters not needed for this study (e.g., nutrient concentrations) or unsuited for establishing quantitative relationships (e.g., qualitative descriptions of deposited sediment cover).

7. The MHPParameters table with the description of consistent parameter codes and names assigned to different parameter methods (Table C-3).
8. The FullOriginalParameters table, which has information on the parameter descriptions as provided by the councils (Table D-5). This table is largely only complete for records extracted from the UnwinMfEDB. For other data sources it is necessary to view the original returned files.

The relationships between these tables are shown in Figure 2-1.

Only data used or processed for the project was included in the database. Note that not all data extracted for an analysis task was used, with additional data selection steps occurring within each task. For example, some sites were excluded from the sediment rating analysis in Section 3 because of issues with data quantity or relationship quality. See individual task descriptions in Sections 3 to 7 for data so excluded.

2.3.2 Data extraction from database

Data can be extracted from the MS-Access database using queries, generally indexing off site (SSSiteID), Date, Time, and parameter(s) of interest (MHpID).

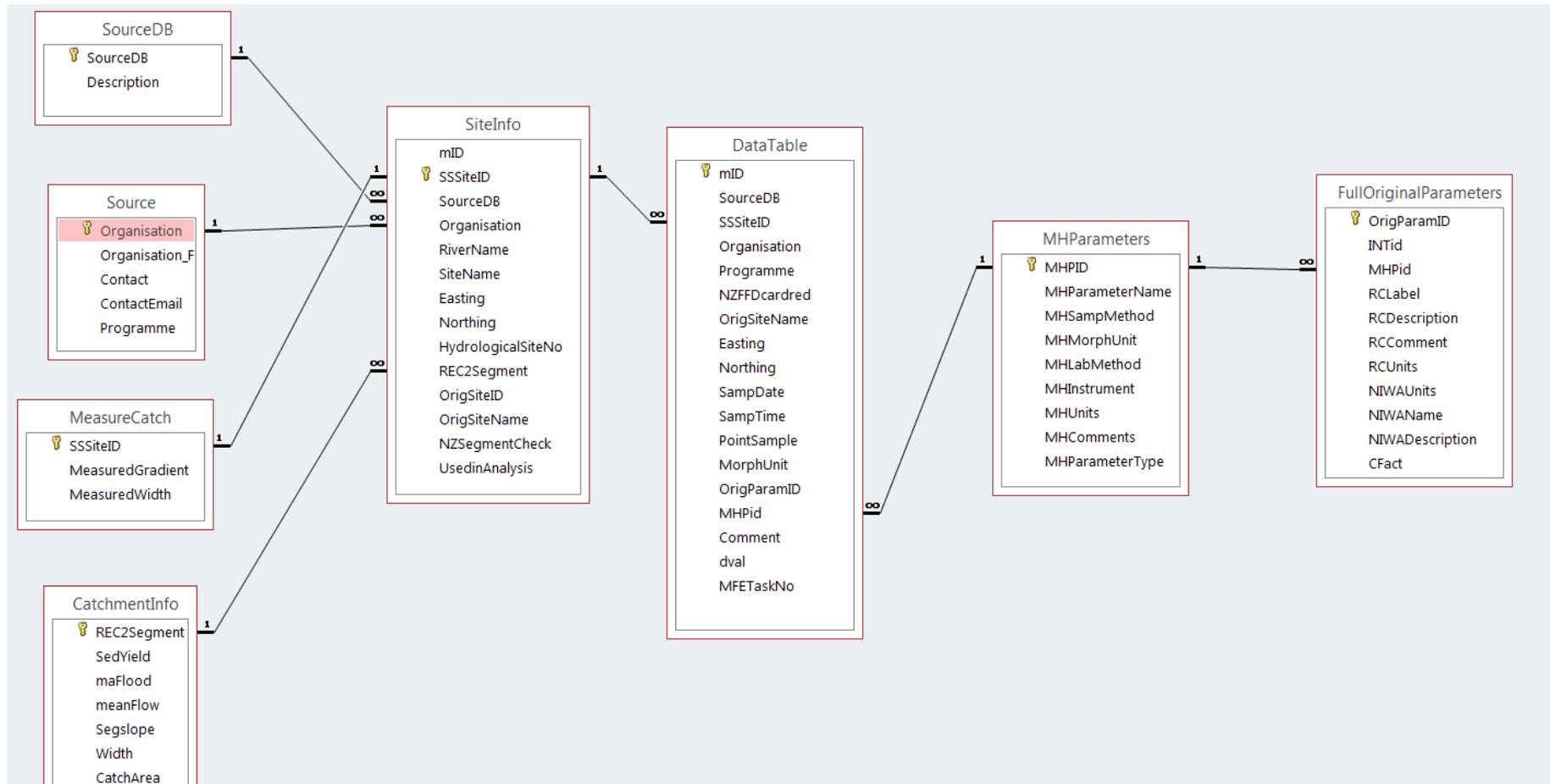


Figure 2-1: Relationships between tables in the MS-Access database.

3 Sediment rating curve parameter estimation models

3.1 Introduction

Sediment rating curves, representing the relationships between SSC (C) and water discharge (Q), are a key component of the ‘working’ analytical framework that links the ESVs to catchment sediment load. This is because the SRCs, when coupled with flow duration curves (FDCs), provide concentration duration curves (CDCs) which can be used to assess how often a SSC target would be exceeded. The purpose of this section is, therefore, to develop regional models (regionalisations) of the parameters of the SRC so that these can be estimated for locations without data.

The basic approach involved two steps: (i) fitting SRC parameters to datasets with concurrent measurements of C and Q, and (ii) relating these SRC parameters to catchment characteristics to derive predictive relationships for each parameter.

In keeping with most international literature (e.g., see review by Hicks and Gomez 2016), we adopt a simple two-parameter “power law” SRC model: $C = aQ^b$. This is traditionally fitted to data using linear regression of the log-transformed C and Q data, i.e., $\log(C) = \log a + b \log(Q)$. While there are alternative, often more accurate methods for fitting SRCs at individual sites (e.g., LOWESS, Hicks and Gomez 2016), these require empirical determination of many parameters and so are not suited to a generalised, regionally-applicable model.

Yang et al. (2007) report on spatial and temporal variations in SRCs from Chinese rivers based on analysis of several decades of manual suspended sediment gaugings. They found that the **a** and **b** SRC parameters tended to be inversely correlated with distinct spatial patterns. High **b** (and low **a**) values tended to occur in relatively steep, rock-confined upper reaches (with high unit stream power), whereas the reverse tended to occur in lowland reaches with meandering planform and low gradients (low unit stream power). Temporally, periods after large floods were typically characterised by lower **b** values and higher **a** values than pre-flood, while trends for progressively increasing loads over time were associated with increased **b** values.

Tran (2014) found similar results from North American rivers but also noted that the inverse relationship between **a** and **b** was more prominent in larger rivers (mean flow > 218 m³/s), which excludes most New Zealand rivers. Tran (2014) also observed that larger rivers (as indexed by mean discharge) tended to have a smaller **a** and larger **b**, so in the downstream direction (as mean discharge increases but unit stream power decreases with decreasing channel gradient) the factors controlling the parameters will change in inverse ways and so tend to balance each other - thus net downstream trends may be weak.

Kettner and Syvitski (2008) showed that for North American rivers **a** was largely determined by the sediment load and mean flow, while **b** was related to basin average temperature (T), maximum relief (R), and mean annual sediment yield (Y). Their multiple-regression-derived relation for **b** equates to:

$$b = 0.4 - 0.025T + 0.00013R + 0.145 \log(Q_{ST}) \quad (1)$$

where T is basin mean temperature (°C), R is basin relief (m), and Q_{ST} is time-averaged basin suspended sediment load (kg/s). The log-transform of the sediment discharge term indicates that it exerts only weak influence on **b**.

Morehead et al. (2003) found that rivers were spread across **a-b** space according to river type (distinguished as glacier/snow-melt dominated, mountainous, maritime-small, tectonically active, and storm dominated).

Based on this brief review, our expectation was that:

- **a** should show some inverse correlation with **b** but this should be weak given the small size of most New Zealand rivers (in a global context)
- **a** should be strongly determined by sediment yield
- **b** should relate to channel steepness, catchment relief, and/or unit stream power (all of which can be expected to be well correlated in the New Zealand context)
- **a** and **b** might be spread by a source-of-flow classification.

3.2 SRC Data

Data from more than 1300 sites with concurrent discharge³ and SSC measurements were collated from multiple sources, including regional councils, historic NIWA databases and the NRWQN. Discharge at a site was measured using either rated stage records or manual gaugings made during site visits. SSC was determined in a variety of ways. At most sites (from the NIWA suspended sediment database), SSC was measured by full sediment gaugings using depth-integrating samplers at multiple verticals (Hicks and Fenwick 1993), with laboratory analyses undertaken on the whole collected sample. At other sites, samples were more often point samples (manually or auto-sampled) and analysed in the laboratory using the total suspended solids (TSS) method.

While point-sampling tends to misrepresent the true cross-section averaged SSC, this aspect was set aside in order to increase the number of sites included in the analysis. There are few datasets from New Zealand rivers that have concurrent all-of-cross-section SSC gaugings and bankside point-sampling, but those that do indicate that the ratio of all-of-cross-section/point SSC typically varies between 1 and 2 but may be less than 1. The ratio depends on the particle size distribution of the suspended load and the turbulence intensity, which together control the extent of lateral and vertical sediment mixing and depend on local hydraulic conditions, and also on the proximity of the point-sampling to the streambed. For example: Basher et al. (2011) found that this ratio varied between 1 and 2 and averaged 1.5 across four sites in the Motueka Catchment; Hicks (2008) found that it ranged between 1.1 and 2.1 and averaged 1.3 in the Amethyst River in South Westland; while Curran-Cournane et al. (2013) found that it ranged between 0.82 and 1.08 for streams in the Auckland region.

Use of the TSS laboratory approach also compromises the accuracy of the true sampled SSC, since it involves extracting a small aliquot from the original field sample and this extraction process may poorly and erratically represent the sand fractions in suspension (Guo 2006). For example, the dataset supplied by Horizons Regional Council for 13 sites across the Manawatu-Wanganui region included matched samples analysed by the TSS and all-of-sample approaches. This showed erratic average variations among the sites, ranging from 33% underestimation of the true SSC to +40% overestimation (Hicks and Hoyle 2012). Similarly, data collected by Auckland Council ranged from 44% overestimation to 17% underestimation of the true SSC by using TSS, with an average overestimation by 11% (Curran-Cournane et al. 2013). Again, for the purposes of this study this

³ We only used data from sites where discharge was either gauged concurrently with the SSC sampling or else was taken from a reliably rated nearby discharge record.

uncertainty was regarded as acceptable and was set aside in the interest of making data available for analysis from more sites.

For the purpose of extracting catchment characteristics from GIS layers, each site was indexed by its NZSegment within the New Zealand REC2 river network⁴. Sites were assigned to NZSegments as described in section 2.1.2. Using this method 104 sites were checked manually, resulting in 32 corrections of NZSegment numbers. Seven sites could not be placed on a confirmed NZSegment and were excluded from further analyses.

Further data quality checks involved checking:

- that there were no duplicate sites or data (sites were combined if rating curves for the individual sites in close proximity [<100 m apart] overlapped sufficiently; duplicate data that had the same measurement value and were collected on the same date at the same time were removed from the combined sites) - we combined 57 pairs, 18 triplicates, and one quadruplicate of sites in this manner
- the trend of the dataset - sites were excluded from the analysis if sediment rating curves (after log transformation) were not generally linear and positive, had inadequate R^2 values (sites with $R^2 < 0.45$ were removed), or were too steep (sites with slopes exceeding 3 were discarded)
- for clusters of SSC values at lab-analysis detection limits (often, SSC values are assigned to the detection limit (e.g., 1 mg/l) even if their true value is less, and with log-transformed data this can result in a cluster of misplaced points that biases the curve-fitting) – such points were removed from the datasets
- for an adequate number of data points – sites with fewer than 7 data points were discarded⁵.

271 sites remained after these checks. The number of data points ranged from 7 to 1947, and had a median value of 36 points. The sites were generally well spread around the country (Figure 3-1), although Marlborough, Northland, and Fiordland had few sites. The sites covered a broad range of channel sizes as indicated by Strahler stream order, catchment types as indicated by REC topographic source of flow categories, and climates as indicated by REC climate categories (Table 3-1).

⁴ The New Zealand river network and associated databases describe the spatial configuration of New Zealand's rivers along with an associated hierarchical classification called the River Environment Classification (REC; Snelder and Biggs 2002). A recent update of the REC (REC2) was used in this project, with the channel network revised and reaches referred to as "segments".

⁵ While this may seem a small number, sites with this few samples were retained provided the samples both covered a broad range in discharge and showed acceptable R^2 values.

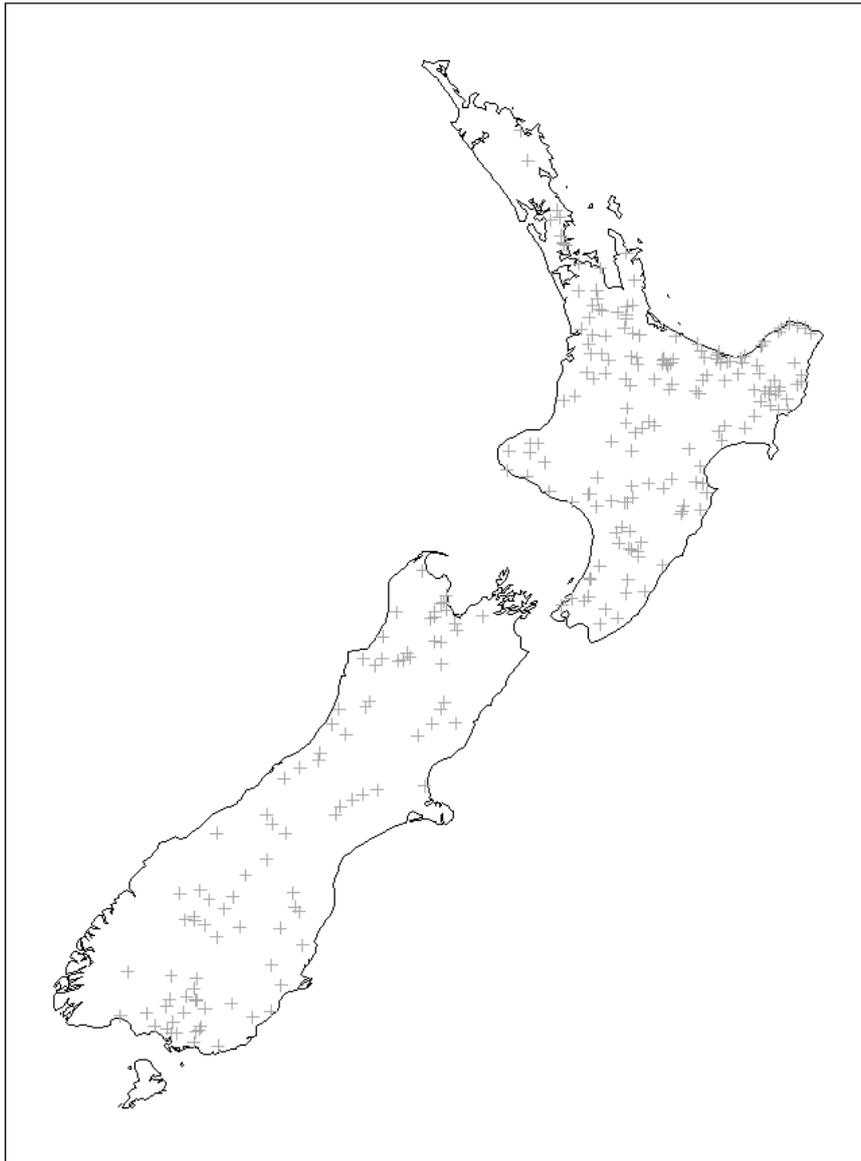


Figure 3-1: Location of sampling sites. n = 271.

Table 3-1: Distribution of sites across climate, topographic source of flow, and Strahler stream order classifications. Refer to Snelder and Biggs (2002) for full descriptions of codes.

Description	Number of sites
Strahler Stream Order	
1	3
2	6
3	19
4	37
5	96
6	78
7	30
8	2
Topographic source of flow	
Glacial mountain	7
Mountain	18
Hill	101
Low elevation	141
Lake-fed	4
Climate	
Cool-dry	40
Cool-wet	116
Cool-extremely wet	30
Warm-dry	10
Warm-wet	69
Warm extremely wet	6

3.3 Sediment rating parameter fitting

3.3.1 Data standardisation

A preliminary step in the SRC analysis was to standardise the SSC and discharge values. Standardising discharge accounts for variation in river size. For example, 100 m³/s would be a base flow in a large river (with associated low SSC) but a flood in a small one (with high SSC). Standardising SSC accounts for differences in catchment sediment supply which varies widely within New Zealand (Hicks et al. 2011). In the context of this study, where it is intended that SRCs are to be used to link changes in mean annual sediment load to changes in SSC, the SSC standardisation explicitly includes mean annual sediment load in the predictive model.

This involved dividing the SSC by the discharge-weighted mean SSC (C_{mean} , equal to the mean annual SS load [L, g/yr] divided by the mean annual water discharge [Q_{mean} , m³/s], multiplied by the number

of seconds in one year) and dividing the discharge by the mean discharge⁶. Thus the SRC model becomes:

$$C/C_{\text{mean}} = C/(L/Q_{\text{mean}}) = a (Q/Q_{\text{mean}})^b \quad (2)$$

On plots of log(C) vs log(Q), the discharge standardisation shifts the SRC left or right, while the SSC standardisation shifts the SRC up or down. Thus the standardisation alters the **a** parameter (rating offset) with a multiplier but does not alter the **b** parameter (rating slope).

As an example, Figure 3-2 contrasts standardised and unstandardised SRC data from sites in Central Otago. Note how the data from the various sites converge into the same “space” after standardisation.

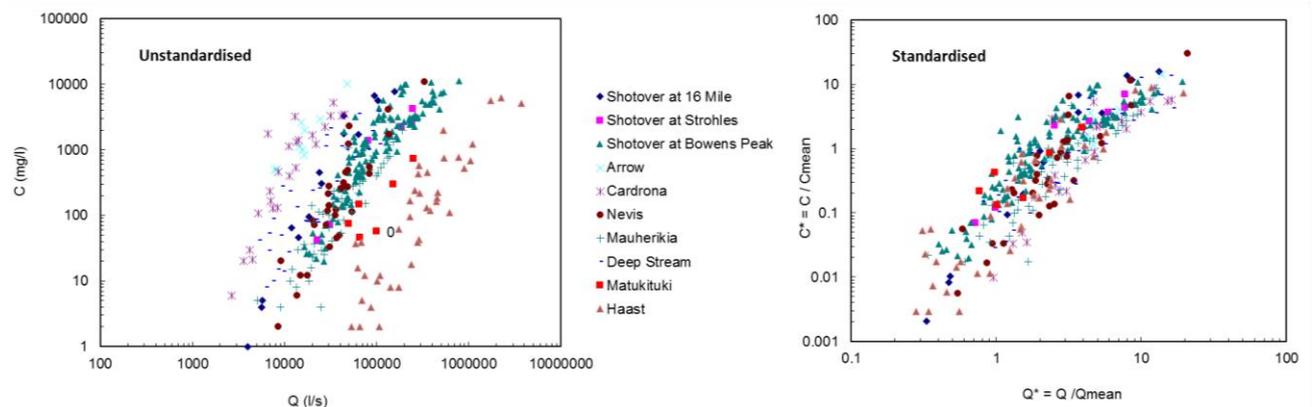


Figure 3-2: Unstandardised and standardised relationships between SSC (C) and water discharge (Q) at 10 sites in Central Otago. C is standardised by the discharge weighted mean SSC, Q by the mean discharge.

Where flow records were available, mean discharge was taken directly from those flow records. For sites without a flow recorder (44 of the 271 sites), mean discharge was estimated from the REC database (where mean flow is estimated using the approach of Woods et al. 2006).

Mean annual sediment load was estimated in two ways. The first approach derived the sediment load by direct integration of the SRC with the flow record (where a flow record existed). For most sites, the mean annual loads had previously been calculated using this method by Hicks et al. (2011) and these values were used. For other sites with flow records available the load was calculated for this study using the same method. For the remaining (44) sites, the load was estimated using NIWA’s national empirical sediment load predictive model as detailed in Hicks et al. (2011), which we term the WRENZ model. This relates sediment load to mean annual rainfall and an Erosion Terrain classification developed by Landcare Research. A sediment load from this model has been calculated for every segment across the REC2 network.

The first approach uses the best available estimates of sediment load at each site. The second approach was to simply use the WRENZ model estimates for all sites – while more uncertain overall, this has the advantage of including a load estimator in the model derivation that is also available to be used in the model application.

⁶ The equivalent standardisation approach was adopted by Kettner and Syvitski (2008).

We extracted rating curve parameters and developed predictive models for three variations:

- SSC standardised by C_{mean} calculated using the observed rating curves and available flow records, standardised discharge (termed “O/E standardised”)
- SSC standardised by C_{mean} calculated using the WRENZ sediment load, standardised discharge (termed “WRENZ standardised”)
- unstandardised SSC, standardised discharge (termed “Unstandardized”).

We included the third, unstandardised SSC variation to check on the efficacy of the standardisation approach.

3.3.2 Parameter fitting

Two-parameter linear regressions were fitted to the log-transformed data at each site for each of the three SRC variations described above, and the intercept (parameter **a**) and slope values (parameter **b**) of the fitted lines were extracted⁷. Regression slope values generated for all three variations are the same (since the standardisation applies the same multiplier to all data points at a site), leaving us with four different coefficient values to model (Table 3-3):

- Intercept O/E standardised (Int_O/E)
- Intercept WRENZ standardised (Int_WRENZ)
- Intercept unstandardized (Int_unstand)
- Slope (same for all variations).

3.3.3 Relationship between SRC slope and intercept

The slope and intercept values for the three rating curve variations showed no obvious relationship (Figure 3-3), indicating that slope and intercept needed to be predicted separately⁸.

3.4 Predicting SRC parameters

3.4.1 Variables potentially influencing SRC parameters

Advancing from the review in Section 3.1 of factors influencing the SRC parameters, we chose the potential predictor variables as listed in Table 3-2. As well as measures of catchment sediment yield, unit stream power⁹, stream order, runoff (as indexed by rain), and temperature, we added proportional land-cover, source-of-flow, climate-type, and catchment soil particle size (sourced from Leathwick et al. 2002 LENZ Database).

3.4.2 Predicting SRC parameters by regression

Approach

Random Forest (RF) regression models (Breiman 2001) were used to separately generate models of the four regression parameters listed in Section 3.3.2 using the 15 environmental predictors in

⁷ The **a** and **b** parameters derived are included with the MS-Access database.

⁸ If a tight relationship has been observed between slope and intercept then the prediction of either slope or intercept would give a good estimate of the other and so only one would need to be predicted off catchment characteristics.

⁹ Unit stream power depends on channel segment slope and mean water discharge per unit width. This was found to correlate highly with other measures of slope (basin-average slope, average channel slope upstream) and thus these were removed from the final RF model.

Table 3-2. The four models are named in Table 3-3.

RF models are an ensemble of regression trees from which a final prediction is based on the predictions averaged over all trees. They were chosen as the method to model regression coefficients for this study because they have several benefits over standard linear regression techniques. Because RFs are a non-parametric method that can handle non-linear relationships explicitly, these benefits include fewer assumptions about data structure and the shape of relationships between predictors and responses than parametric methods. RFs also have inbuilt cross-validation with models tested against data held out of the set used to create the predictions (Breiman 2001, Ellis et al. 2012). Details of the RF approach are provided in Appendix E.

The 15 environmental predictor variables had pairwise correlations of <0.7 . To check for nonlinear effects in the standardisation process, the two variants of C_{mean} were also individually included as predictors in the relevant RF models even though they are correlated with sediment yield ($r = 0.64$ and 0.86 , respectively).

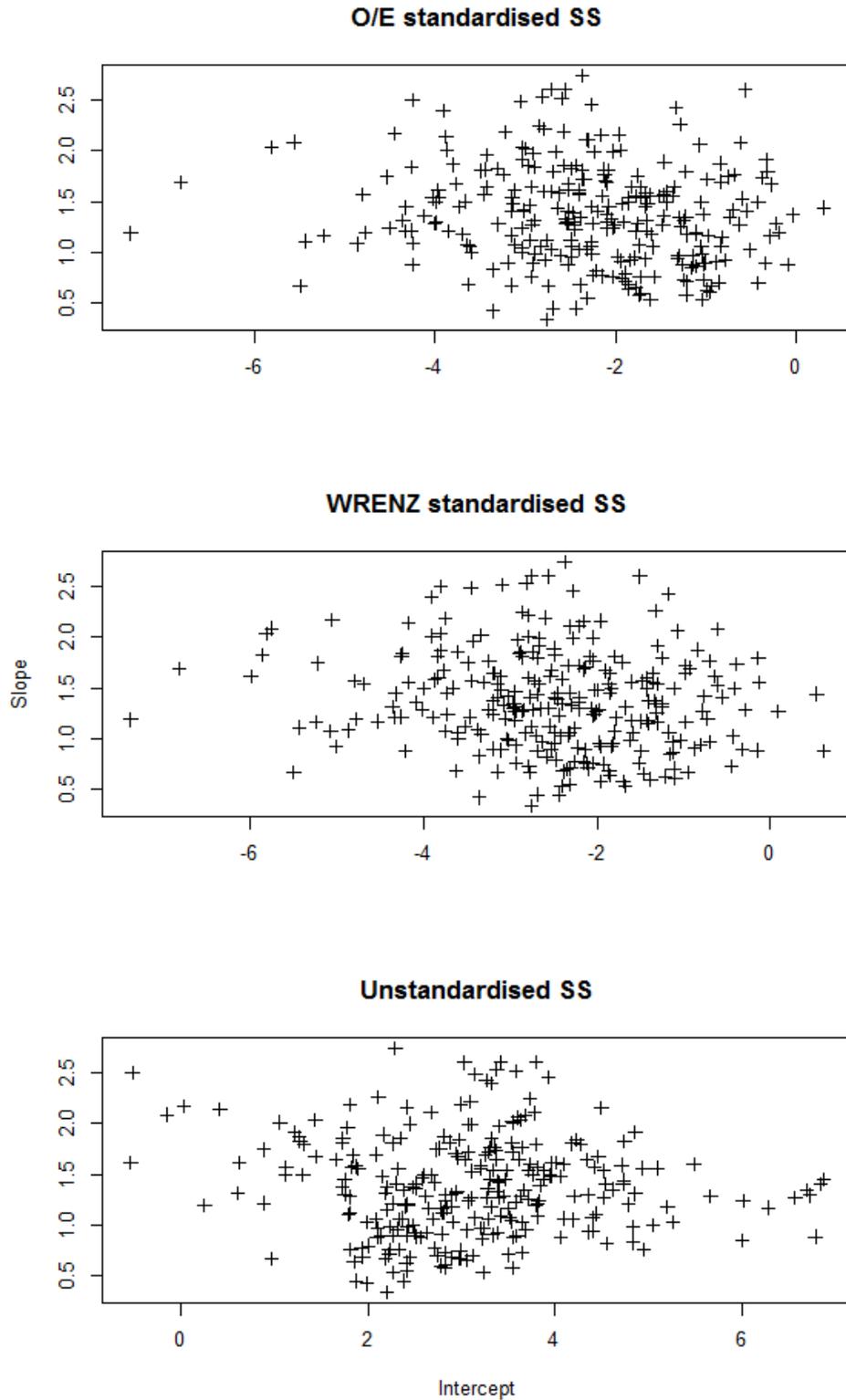


Figure 3-3: Intercept and slope parameters for linear regressions fitted to individual sites (n = 271) using three rating curve variations. a) SSC standardised by C_{mean} calculated from observed/estimated (O/E) sediment yields, b) SSC standardised by C_{mean} calculated from sediment load from WRENZ model, and c) unstandardized SSC.

Table 3-2: Predictor variables used in Random Forest models. Wetted width for unit stream power calculations was estimated as wetted width at mean flow generated following the method of Booker and Hicks (2013). C_{mean} was standardised using observed or estimated mean annual sediment load values for the Int_O/E and Slope models and by WRENZ generated sediment load for the Int_WRENZ model.

Predictor	Description	Units	Notes
C_{mean}	Mean annual sediment load / mean annual river flow	g/m^3	Used sediment yield from WRENZ or O/E as appropriate to coefficient being tested
Unit stream power	REC2_segment slope * mean flow / width at mean flow	$\text{t}/\text{s}/\text{m}$	REC2_segment slope equals segment mean slope in degrees. Mean flow estimated using Woods et al. (2006) model. Width at mean flow estimated using Booker and Hicks (2013) model.
Sediment yield	Sediment load / REC2_us_Catarea	$\text{t}/\text{y}/\text{m}^2$	REC2_us_Catarea = upstream catchment area in m^2 .
REC2_headw_dist	Distance to the furthest headwater reach	m	
REC2_us_tmin	Upstream mean minimum winter air temperature	$^{\circ}\text{C}$	
REC2_StreamOrder	Strahler stream order	Ordinal scale	
REC2_REC1_CLIMATE	Climate categories from Snelder and Biggs (2002)		Generated for REC1 and mapped to REC2: Cool dry, Cool wet, Cool extremely wet; Warm dry, Warm wet, Warm extremely wet.
REC2_REC1_SRC_OF_FLW	Source of flow categories from Snelder and Biggs (2002)		Generated for REC1 and mapped to REC2. Glacial-mountain, mountain, hill, lowland, and lake-fed categories.
REC2_us_rain	Mean annual rainfall in catchment upstream	mm	
REC2_us_psize	Upstream catchment average soil particle size	Ordinal scale ¹⁰	Sourced from the LENZ Database and based on soil data, as reported by Leathwick et al. (2002)
REC2_us_elev	Mean elevation of the catchment upstream	m (above mean sea level)	
REC2_us_lakePerc	Upstream area of the catchment covered by lakes	%	

¹⁰ Particle size classes include silt and clay, sand, gravel, coarse to very coarse gravel, boulder to massive (from Leathwick et al. 2002)

Predictor	Description	Units	Notes
Prop_us_Agland	Proportion of upstream catchment in agricultural land	Proportion	Combination of upstream cover of LCDB ¹¹ categories: 30, 33, 40
Prop_us_Forest	Proportion of upstream catchment in forest	proportion	Combination of upstream cover of LCDB categories: 64, 68, 69, 71
Prop_us_Scrubland	Proportion of upstream catchment in scrubland	proportion	Combination of upstream cover of LCDB categories: 51, 52, 55, 56, 58

Table 3-3: Descriptions of four Random Forest models used to predict rating curve coefficients.

Method	Model name	Rating curve type	Coefficient type predicted
Random Forest	RF_Int_O/E	Observed/Estimated standardised	Intercept
Random Forest	RF_Int_WREnz	WREnz standardised	Intercept
Random Forest	RF_Int_unstand	Unstandardised	Intercept
Random Forest	RF_Slope		Slope

Results¹²

Out-Of-Bag R^2 for the four RF models was greater than 0.40 in all cases (Table 3-4), with the predictive model for intercept values for the unstandardized SSC regression having the greatest predictive power, followed by the intercept for regressions standardised using O/E yield values (Table 3-4). We consider this a reasonable result based on similar studies and when considering the accuracy of the predicted SSCs (later section).

The most important predictor for all intercept models was the C_{mean} predictor relevant to the intercept being tested (Figure 3-4). Sediment yield per unit area (SedperArea) was the second most important predictor for the O/E and unstandardized intercept models, while upstream particle size (REC2_us_psize) was slightly more important in the Int_WREnz model. The elevation of the upstream catchment was the most important predictor of the SRC slope and was almost twice as important as any other predictors (Figure 3-4).

Sediment yield showed a positive relationship with the intercept from the unstandardized rating curves but a negative relationship with curves standardised by O/E sediment yield (Figure 3-5). This suggests a non-linear relationship between the intercept and sediment yield with a power-law exponent less than 1.¹³ The intercept for the WREnz standardised rating curves declined as upstream particle size increased. This is logical since higher sediment production is expected when soil particle size becomes finer-grained, all other things being equal.

SRC slope showed a positive relationship with both upstream elevation and unit stream power, indicating that the slopes of rating curves were generally steeper in reaches at higher elevation and

¹¹ LCDB = Land Cover Database. Version 3 was used.

¹² Refer to Appendix E for explanations of Random Forest terminology and performance measures.

¹³ If with standardisation $a \propto \text{Yield}^k$ and $0 < k < 1$, then after standardisation (which divides SSC by C_{mean} which correlates with sediment yield) $a \propto \text{Yield}^{k-1}$.

with higher unit stream power (Figure 3-5) – which is consistent with Yang et al. (2007) and Tran (2014). Sediment yield per catchment area (SedperArea) was of moderate importance (8th most important predictor out of 16), with an increase in SRC slope at relatively low levels of sediment yield (Figure 3-6)¹⁴.

Table 3-4: Out-Of-Bag (OOB) R² results for Random Forests predicting intercept and slope values for the three rating curve types. OOB R² provides an estimate of the predictive power of the model for new cases, with a maximum value of 1.

Model name	Rating curve type	Parameters predicted	OOB R ²
RF_Int_O/E	Observed/Estimated standardised	Intercept	0.51
RF_Int_WRENZ	WRENZ standardised	Intercept	0.46
RF_Int_unstand	Unstandardised	Intercept	0.55
RF_Slope		Slope	0.41

3.4.3 Alternative approach for predicting rating parameters using one environmental predictor

An alternative to the Random Forest models, which used 15 environmental predictors, is to use a simplified model that has only one predictor variable. Results from this method can be compared with the more intensive Random Forest methods to assess any additional explanatory or predictive power the more complex models have over using a single environmental predictor. The predictor we used was the Source of Flow (SOF) category from the REC, which was developed to aid in describing patterns of seasonality of flow regimes, frequency of high flows, and sediment transport regimes (Snelder and Biggs 2002). The source of flow categories are: Glacial mountain (GM), mountain (M), hill (H), low elevation (L) and lake-fed (Lk).

Results

We investigated whether the SRC parameters varied between SOF categories using one-way analyses of variance (ANOVA). These analyses showed that the slope coefficient differed significantly between the SOF categories, with rating curves having steeper slopes in higher elevation locations (mountain and hill) than in lower elevation locations (Figure 3-7, Table 3-5, Table 3-6). The three intercept coefficients showed no significant relationships with SOF category, apart from the O/E standardised intercept which showed a trend towards lake sites having higher intercepts than the other categories (Figure 3-7, Table 3-5, Table 3-6).

Given the significant differences in the slope of rating curves between SOF categories but marginal effect on the intercept of rating curves, we used the REC SOF category as our single predictor of coefficients. To do this we extracted the mean values of the intercepts and slopes of sites within each SOF category (SOF.Mean).

¹⁴ It is of note that partial dependence plots (such as in Figure 3-5) show the marginal contribution of a predictor to the response (i.e., the response as a function of the predictor when the other predictors are held at their mean value). They are not a perfect representation of the relationship between each predictor and response, particularly if there are interactions between predictors or predictors are strongly correlated; however, they provide useful information for interpreting the model (Friedman and Meulman 2003).

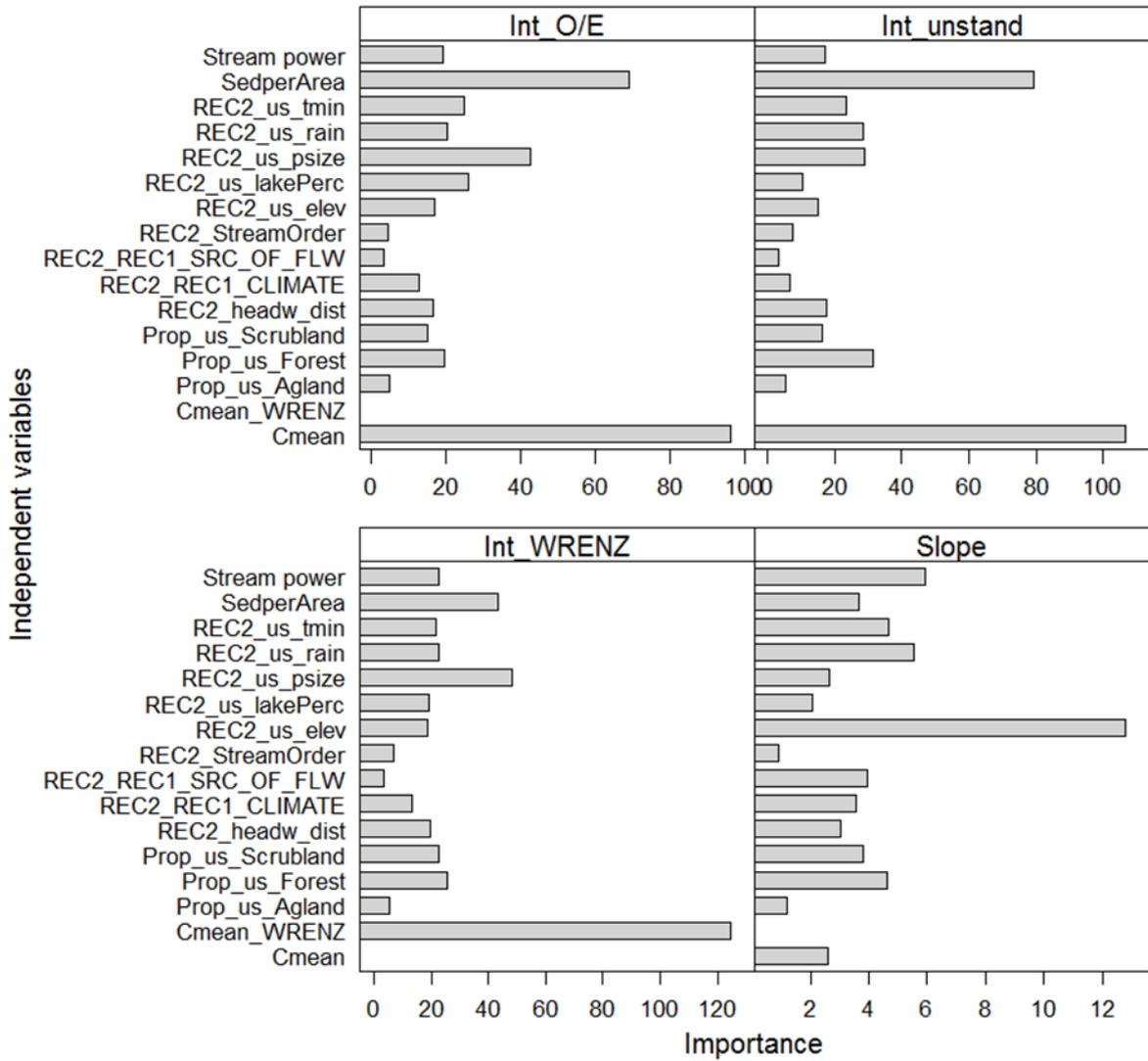


Figure 3-4: The importance of individual predictors to each of the four RF models. Importance is measured as increasing node purity from splitting on the selected variable averaged over all trees in the Random Forest. C_{mean_WRENZ} was only included in the model for Int_WRENZ, while C_{mean} was excluded.

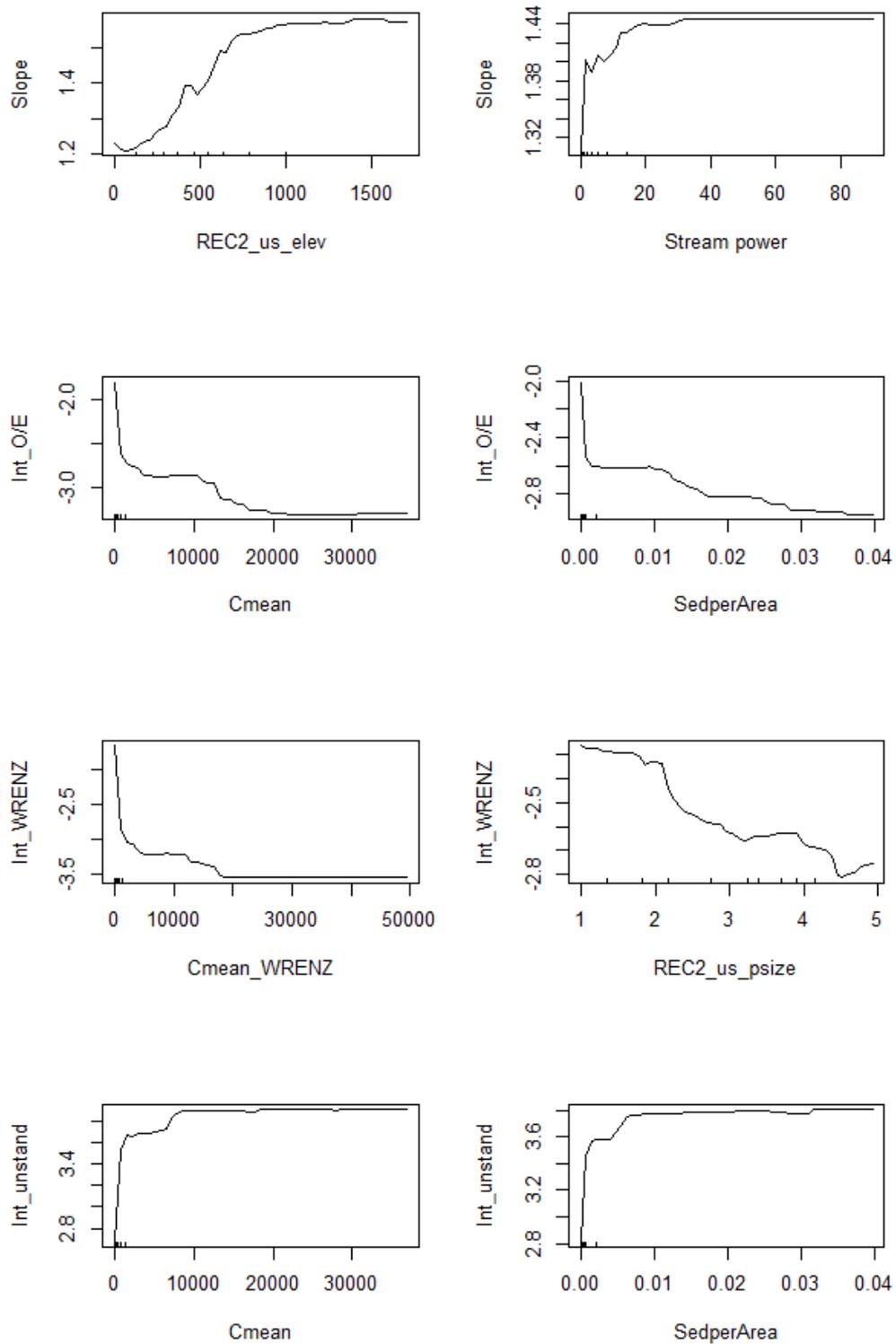


Figure 3-5: Univariate partial dependence plots showing the shape of relationships between the two most important predictors for the four different Random Forest models: Slope, Int_O/E, Int_WRENZ, and Int_unstand. The spread of predictor values across the gradient is shown by the “rug” on the x-axis (each tick is a decile).

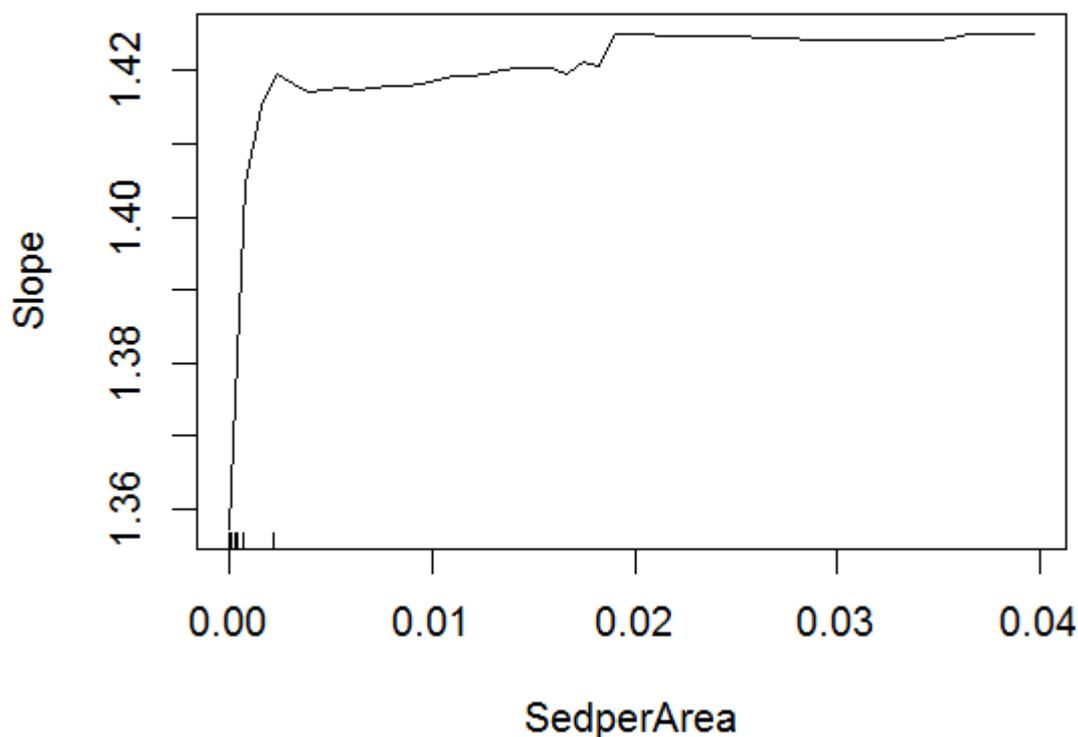


Figure 3-6: Univariate partial dependence plot showing the shape of relationships between SRC slope and catchment sediment yield. The spread of predictor values across the gradient is shown by the rug on the x-axis (each tick is a decile).

Table 3-5: Results from individual one-way ANOVAs testing the effect of source of flow category (SOF) on coefficients from the three rating curve methods. A significant result ($P < 0.05$) indicates that the coefficient values differ between SOF categories. See Table 4-6 for the mean coefficient value (and its standard deviations) for each SOF category. R^2 shows how much of the variance in the observed coefficients is explained by using SOF categories as the predictors. There are 4 degrees of freedom in the SOF model, leaving 266 residual degrees of freedom.

Coefficients	degrees of freedom	F	P	R^2
Intercept O/E standardised	4, 266	2.3	0.06	0.03
Intercept WRENZ standardised	4, 266	0.6	0.70	0.01
Intercept unstandardized	4, 266	1.3	0.27	0.02
Slope	4, 266	18.8	<0.001	0.22

Table 3-6: Mean (± 1 SD) coefficient values for SRCs generated using three different rating curve methods for sites in different source of flow (SOF) categories. The number of sites in each SOF category is indicated.

SOF	Slope	Int_O/E	Int_WRENZ	Int_unstand
Glacial mountain (n = 7)	1.61 \pm 0.38	-2.10 \pm 0.69	-2.11 \pm 0.68	3.95 \pm 0.76
Mountain (n = 18)	1.82 \pm 0.55	-2.43 \pm 0.94	-2.62 \pm 1.17	3.15 \pm 1.50
Hill (n = 101)	1.57 \pm 0.43	-2.43 \pm 1.19	-2.60 \pm 1.31	2.95 \pm 1.36
Lowland (n = 141)	1.16 \pm 0.41	-2.29 \pm 1.29	-2.47 \pm 1.28	3.16 \pm 1.14
Lake (n = 4)	1.52 \pm 0.25	-0.61 \pm 0.17	-1.91 \pm 1.55	2.97 \pm 0.49

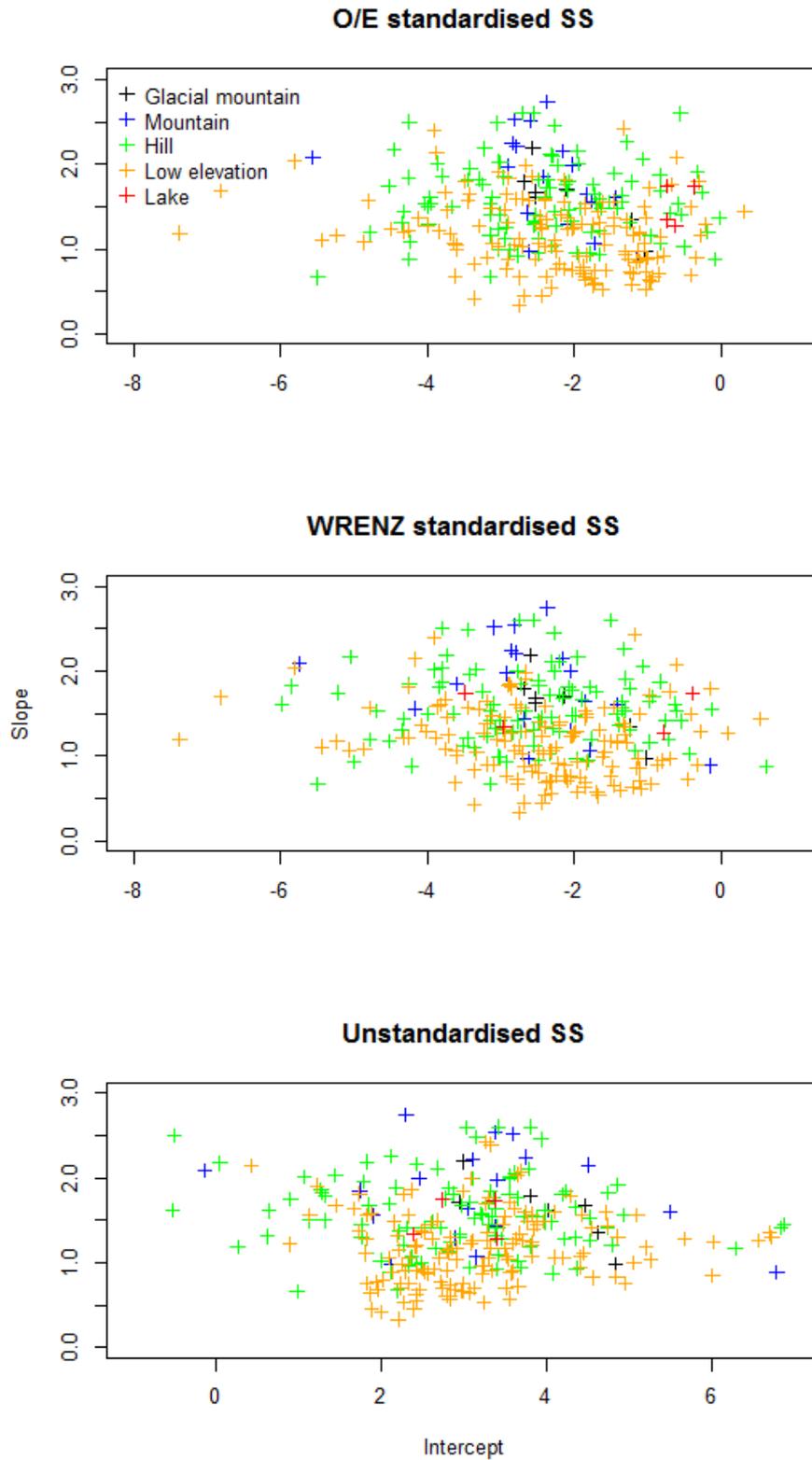


Figure 3-7: Intercept and slope coefficients for linear regressions fitted to individual sites (n = 271) coded by source of flow (SOF) categories for three rating curve methods: SSC standardised by C_{mean} calculated from O/E sediment loads (top), SSC standardised by C_{mean} calculated from sediment load from WRENZ (middle), and unstandardized SSC with points coded by SOF category (bottom).

3.4.4 Comparison of model performance

Approach

To provide best-case and worst-case benchmarks against which the RF and SOF models could be compared, several other sets of predicted coefficients were generated (Table 3-7). The worst-case scenarios included:

- a dataset in which substitute intercept and slope values from the regressions on SSC standardised using C_{mean} calculated with O/E sediment yields were selected from a random site (random_obs/est standardised)
- a dataset in which substitute intercept and slope values from the regressions on unstandardized SS concentrations were selected from a random site (random_unstandardised)
- a dataset in which the mean intercept and slope values were used for all sites (mean).

The best-case scenario used the actual intercept and slope values calculated from the regression for that site (fitted). This scenario is the best that could be achieved if we can predict intercepts and slopes perfectly for a site.

We tested the ability of the different methods to predict (i) the parameters of the relevant rating curve types, and (ii) the SSC at each observed discharge using the appropriate combinations of model coefficients. Three model performance metrics were calculated for each test: Nash–Sutcliffe efficiency (NSE), root mean square error (RMSE), and the ratio of the root mean square error to the standard deviation of observed data (RSR)¹⁵. The NSE ranges from $-\infty$ to 1, with 1 indicating a perfect match to predictions, 0 indicating that predictions are as accurate as the mean of the observed data, and negative values indicating that the observed mean is a better predictor than the model.

Results for prediction of SRC parameters

The regression **a** and **b** parameters predicted using the six different methods (Table 3-7) varied in their relationship with observed parameter values (Figure 3-8, Table 3-8). The fitted method uses the exact regression coefficients for each site and thus has a 1:1 relationship between observed and predicted coefficients for both the slope and intercept. As expected, both randomised methods (generated based on standardised and unstandardized regressions) showed poor relationships between predicted and observed coefficient values. Using the mean coefficients of (i) all sites or (ii) each SOF category creates “tower-like” patterns in the observed versus expected parameter plots where the same value is used for all, or for many sites, respectively (Figure 3-8). The four parameters generated using the RF models showed degrees of relationship between observed and predicted values that were somewhere between the extremes of best-case (fitted method) and worst case (random methods).

¹⁵ Refer to Moriasi et al. (2007), Nash and Sutcliffe (1970), and Pineiro et al. (2008) for test details.

Table 3-7: Summary of methods for calculating SSC.

Method Name	Rating curve type	Predicted variables	Description	What it represents
Random_O/E	O/E standardised	Int_random Slope_random	Observed values of intercept and slope taken from a random site. Observed values are from regressions of SSC standardised by O/E yield C_{mean})	A worst case prediction for standardised regressions
Random_unstand	unstandardised	Int_random_unstand Slope_random_unstand	Observed values of intercept and slope taken from a random site. Observed values are from regressions of unstandardized SSC	A worst case prediction for unstandardised regressions
Mean	O/E standardised	Int_Mean Slope_Mean	Observed values taken as mean fitted intercept and slope for all sites	A situation where the rating curves for all sites are the same
Fitted	O/E standardised	Int_Fitted Slope_Fitted	Fitted intercept and slope values from the original regressions	The best that could be achieved if we can predict intercepts and slopes perfectly
SOF.Mean	O/E standardised	Int_SOF.mean Slope_SOF.mean	Predicted coefficients are the mean coefficients from source of flow (SOF) categories	Predictions for intercept and slope for the standardised by O/E rating curve type based only on source of flow categories
Random Forests	O/E standardised WRENZ standardised unstandardised	Int_WRENZ Int_O/E Int_unstand Slope_RF	Predicted coefficients are out-of-bag estimates predicted using 15 environmental variables	Predictions for intercept and slope for each rating curve type for each site as if that site had not been visited

Negative NSE values for the randomised methods indicated that the observed mean was a better predictor than the randomised values (Table 3-8, Figure 3-8). NSE values of 0 for the mean methods indicated that predictions were as accurate as the mean of the observed values, as expected. The methods using mean values from SOF categories for predicted intercepts and slopes had positive NSE values, indicating that they performed better than the randomised or mean methods. However, the four Random Forest methods all had NSE values that were at least twice as high as the best SOF model, indicating that they perform better than any of the other methods apart from the fitted method (Figure 3-8).

Note that the RMSE errors on the intercept (Table 3-8) are for log values, so the actual error in the SRC **a** parameter is a factor equal to e^{RMSE} . For example, the factorial RMSE on the **a** parameter predicted by the RF model using unstandardized SSC data is $\times/\div e^{0.83} = 2.29$. It is of note that this is similar to the RMSE factorial error associated with sediment yield predictors (e.g., Hicks et al. 2011).

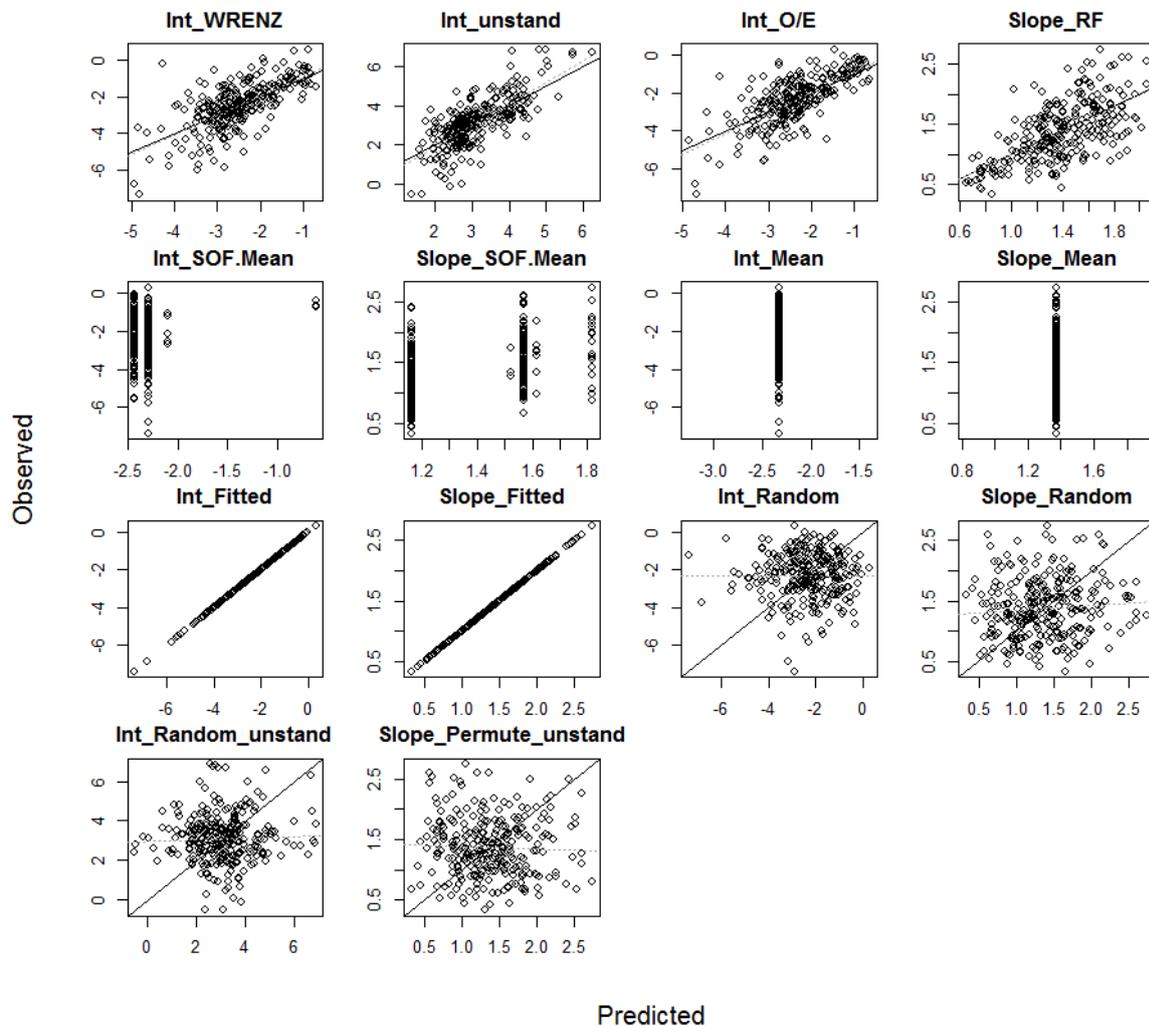


Figure 3-8: Predicted versus observed SRC parameters generated using the methods in Table 3-7 Error! reference source not found. Solid lines are 1:1, dashed grey lines are fitted regression lines between observed and predicted values.

Table 3-8: Model performance metrics for different methods of predicting intercept and slope coefficients. Metrics were generated using a hold-one-out cross-validation procedure (i.e., each point was systematically held out of model development and was then compared to the value predicted from the model that it was held out of). R^2 is a measure of how good the relationships between predicted and observed values are. NSE is Nash-Sutcliffe efficiency, RMSE is root mean square error, RSR is ratio of the root mean square error to the standard deviation of the observed data. Note that the errors for the intercept are for log values, so the actual error of the SRC a parameter is a \times/\div factor equal to e^{RMSE} .

	Model representation	R^2	NSE	RMSE	RSR
Int_WRENZ	RF predictions of intercept	0.45	0.45	0.95	0.74
Int_unstand	RF predictions of intercept	0.55	0.55	0.83	0.67
Int_O/E	RF predictions of intercept	0.51	0.51	0.86	0.70
Slope_RF	RF predictions of slope	0.41	0.41	0.38	0.71
Int_SOF.Mean	Sites within SOF categories have same intercept	0.03	0.03	1.21	0.98
Slope_SOF.Mean	Sites within SOF categories have same slope	0.22	0.22	0.43	0.88
Int_Mean	All sites have same intercept	<0.01	0	1.30	1
Slope_Mean	All sites have same slope	<0.01	0	0.49	1
Int_Fitted	Fitted intercept values, best case scenario	1	1	0	0
Slope_Fitted	Fitted slope values, best case scenario	1	1	0	0
Int_Random	Randomised intercept, worst case scenario	<0.01	-1.00	1.74	1.42
Slope_Random	Randomised slope, worst case scenario	<0.01	-0.84	0.66	1.36
Int_Random_unstand	Randomised intercept, worst case scenario	<0.01	-0.91	1.71	1.38
Slope_Random_unstand	Randomised slope, worst case scenario	<0.01	-1.10	0.70	1.44

Results for prediction of SSC at a given discharge

There were differences in the performance of the models used to generate predictions of SSC at a given discharge (Figure 3-9, Table 3-9).

As expected, using parameter values from the original fitted regressions for each site generated the best predictions of SSC, with all sites having RSR values under 1 (Figure 3-10), positive NSE values (Figure 3-11), and low RMSE values (Figure 3-12) - all indicating good model performance.

Poor performance of the randomised predictions (Random_O/E and Random_unstand) indicated that there were considerable between-site differences in the slopes and intercepts of the SRCs. This poor performance was indicated by relatively high scatter in the observed versus predicted SSC plots (Figure 3-9) and in the high RSR and RMSE and largely negative NSE values for many sites (Table 3-9, Figure 3-10 through Figure 3-12).

Methods using coefficients generated from either the overall mean or the mean of SOF categories performed better than the random methods but not as well as any of the RF or fitted methods (Figure 3-9 through Figure 3-12, Table 3-9).

The three RF methods for predicting regression coefficients performed the best out of the non-fitted methods. These three methods showed relatively tight relationships between observed and predicted values that were also close to the 1: 1 line (Figure 3-9), as well as generally low RSR and RMSE and largely positive NSE values (Table 3-9, Figure 3-10 through Figure 3-12). Similar model

performance was seen between the methods using unstandardized SSC values compared to values standardised by O/E sediment yields (Table 3-9). Predictive performance was reduced slightly by the use of WRENZ standardised SSCs (Table 3-9).

Note that the RMSE errors on the predicted SSC values in Table 3-9 are for log values, so the actual error in the predicted SSC is a factor equal to e^{RMSE} . For example, the factorial RMSE on the SSC predicted by the RF models using both the standardized and unstandardized SSC data is $\times/\div e^{1.16} = 3.19$. By comparison, the factorial error associated with the fitted SRCs is $e^{0.81} = 2.25$. In other words, using these RF models to predict the SSC at a given discharge at any site (instead of using a measured SRC) increases the uncertainty on the SSC prediction by a factor of 1.42 on average.

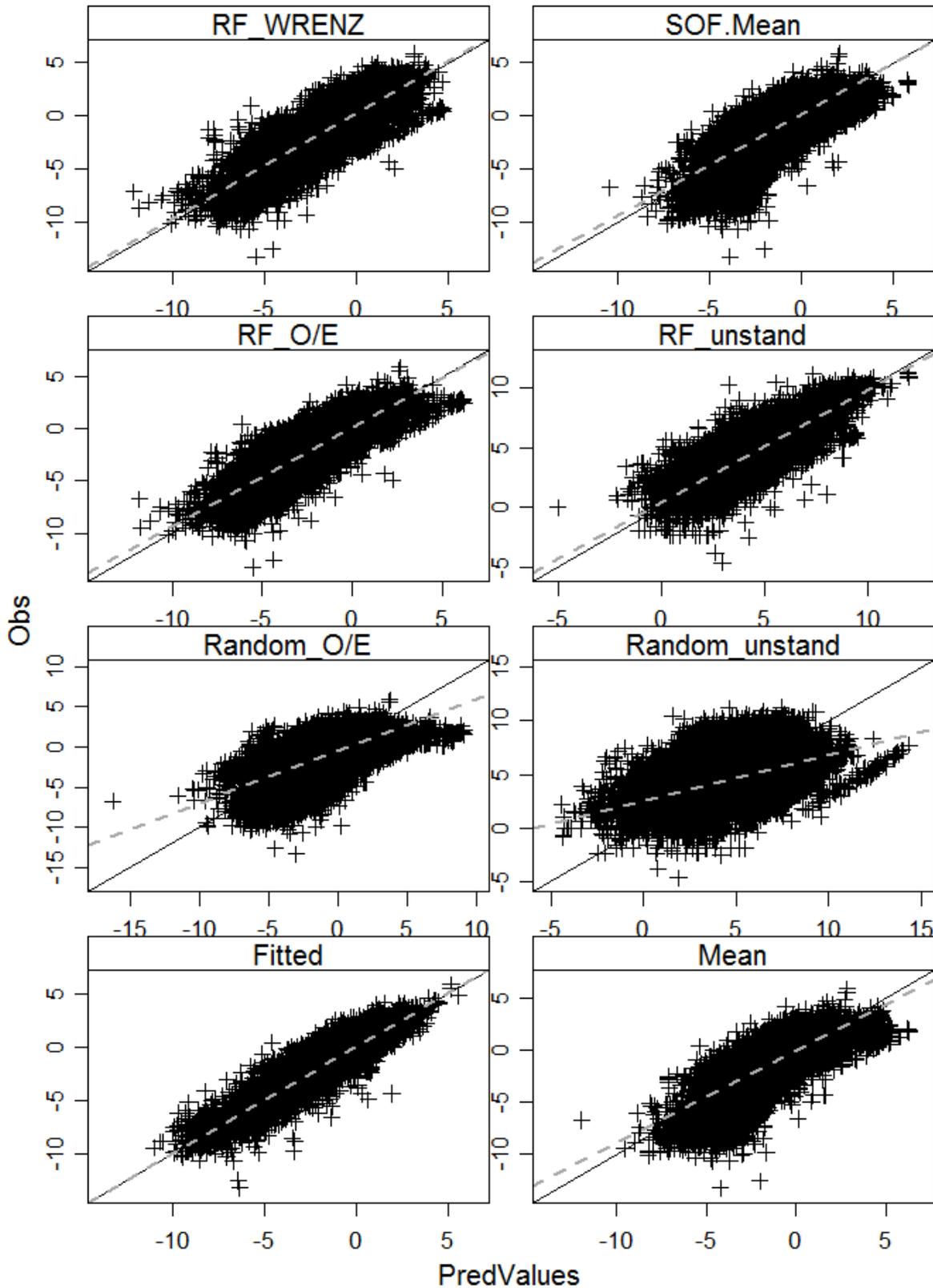


Figure 3-9: Predicted versus observed SSC for all site visits generated using the methods in Table 3-7. Solid lines are 1:1, dashed grey lines are fitted regression lines between observed and predicted values. n = 23793.

Table 3-9: Mean model performance metrics for different methods of predicting log SSC. ± values indicate the standard error on the mean. NSE is Nash-Sutcliffe efficiency, RMSE is root mean square error, and RSR is ratio of the root mean square error to the standard deviation of the observed data. Note that errors are for log SSC, so the actual error in SSC is a \times/\div factor equal to e^{RMSE} . n = 271 sites.

	Model representation	NSE	RMSE	RSR
RF_WRENZ	Random Forest: SS concentration standardised by WRENZ sediment yield	0.16 ± 0.06	1.21 ± 0.03	0.82 ± 0.02
RF_unstand	Random Forest: SS concentration unstandardized	0.26 ± 0.05	1.16 ± 0.03	0.79 ± 0.02
RF_O/E	Random Forest: SS concentration standardised by O/E yield	0.25 ± 0.05	1.16 ± 0.03	0.79 ± 0.02
SOF.Mean	Sites within SOF categories have same coefficients	-0.28 ± 0.10	1.40 ± 0.04	0.98 ± 0.03
Mean	All sites have same coefficients	-0.38 ± 0.11	1.45 ± 0.04	1.02 ± 0.04
Fitted	Fitted coefficient values, best case scenario	0.71 ± 0.01	0.81 ± 0.02	0.51 ± 0.01
Random_O/E	Randomised coefficients, O/E standardisation, worst case scenario	-1.70 ± 0.33	1.93 ± 0.06	1.35 ± 0.06
Random_unstand	Randomised coefficients, unstandardized, worst case scenario	-0.93 ± 0.16	1.71 ± 0.06	1.18 ± 0.05

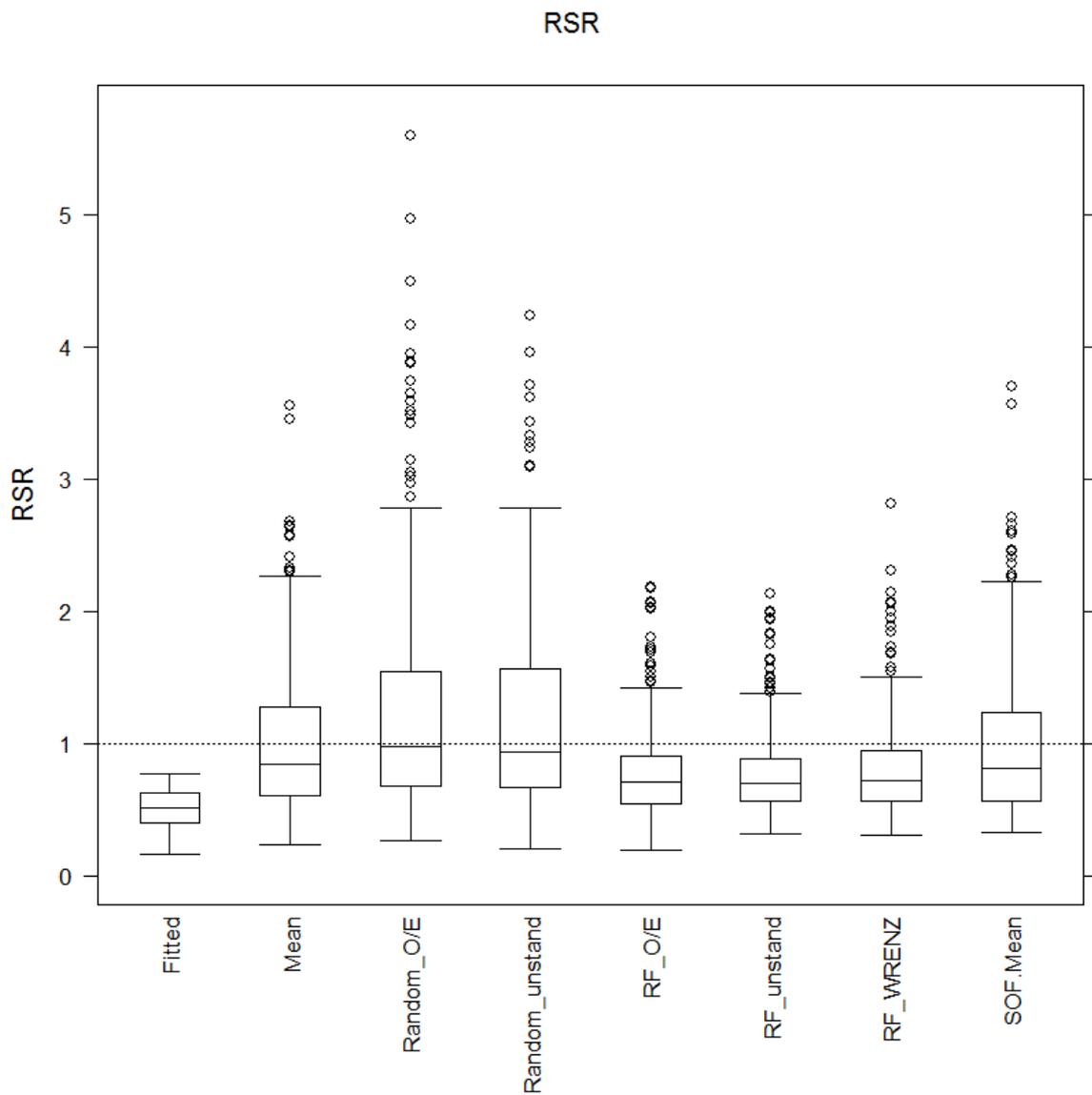


Figure 3-10: Boxplot of the ratio of the root mean square error to the standard deviations of the observed data (RSR) for observed versus predicted suspended sediment concentrations generated using eight different methods across 271 sites. See Table 3-7 for method details. A lower RSR indicates a better fit between predicted and observed values.

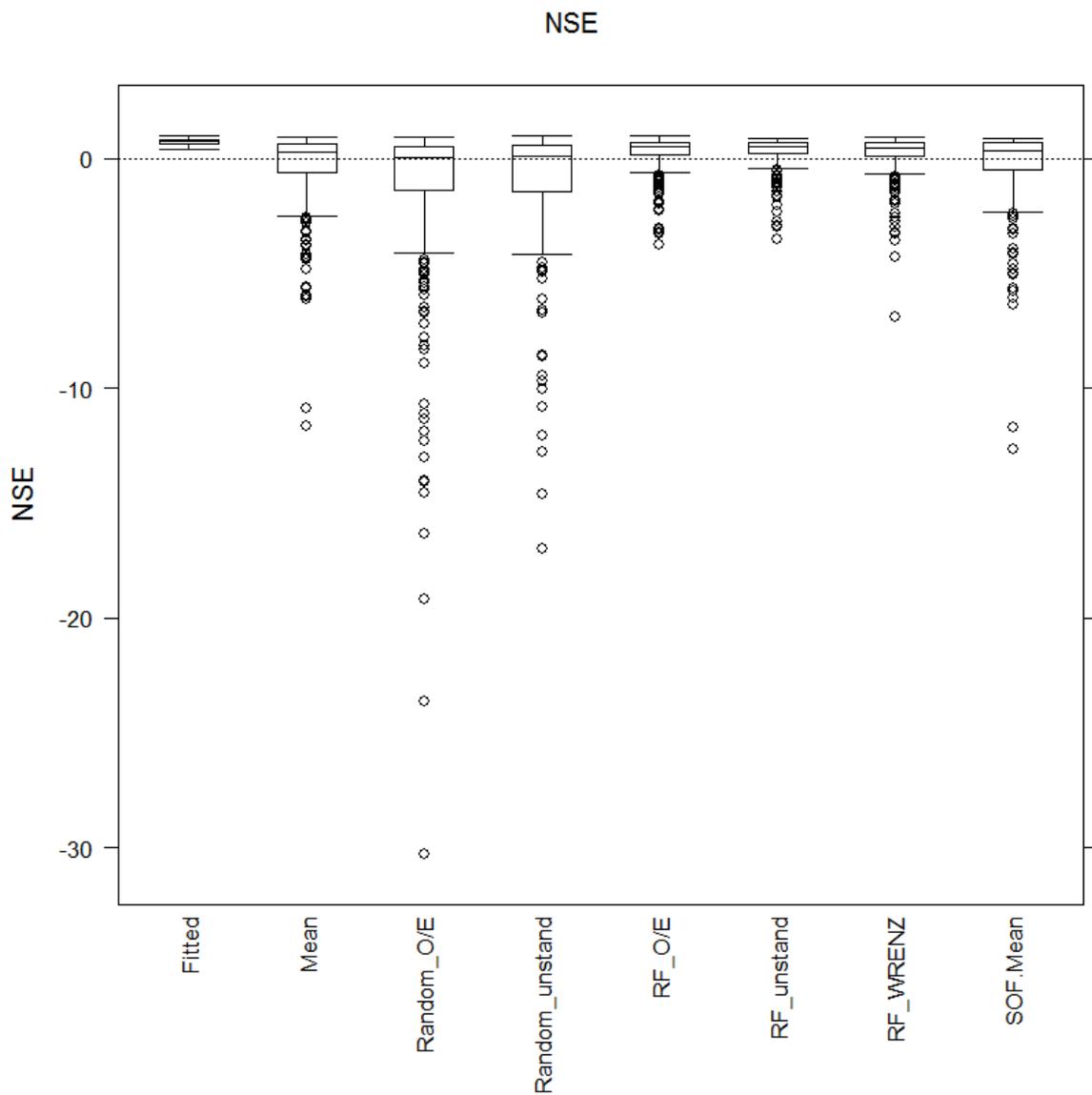


Figure 3-11: Boxplot of Nash Sutcliffe Efficiency (NSE) for observed versus predicted suspended sediment concentrations generated using eight different methods across 271 sites. See Table 3-7 for method details. NSE ranges from $-\infty$ to 1, with 1 indicating a perfect match to predictions, 0 indicating that predictions are as accurate as the mean of the observed data, and negative values indicating that the observed mean is a better predictor of the model.

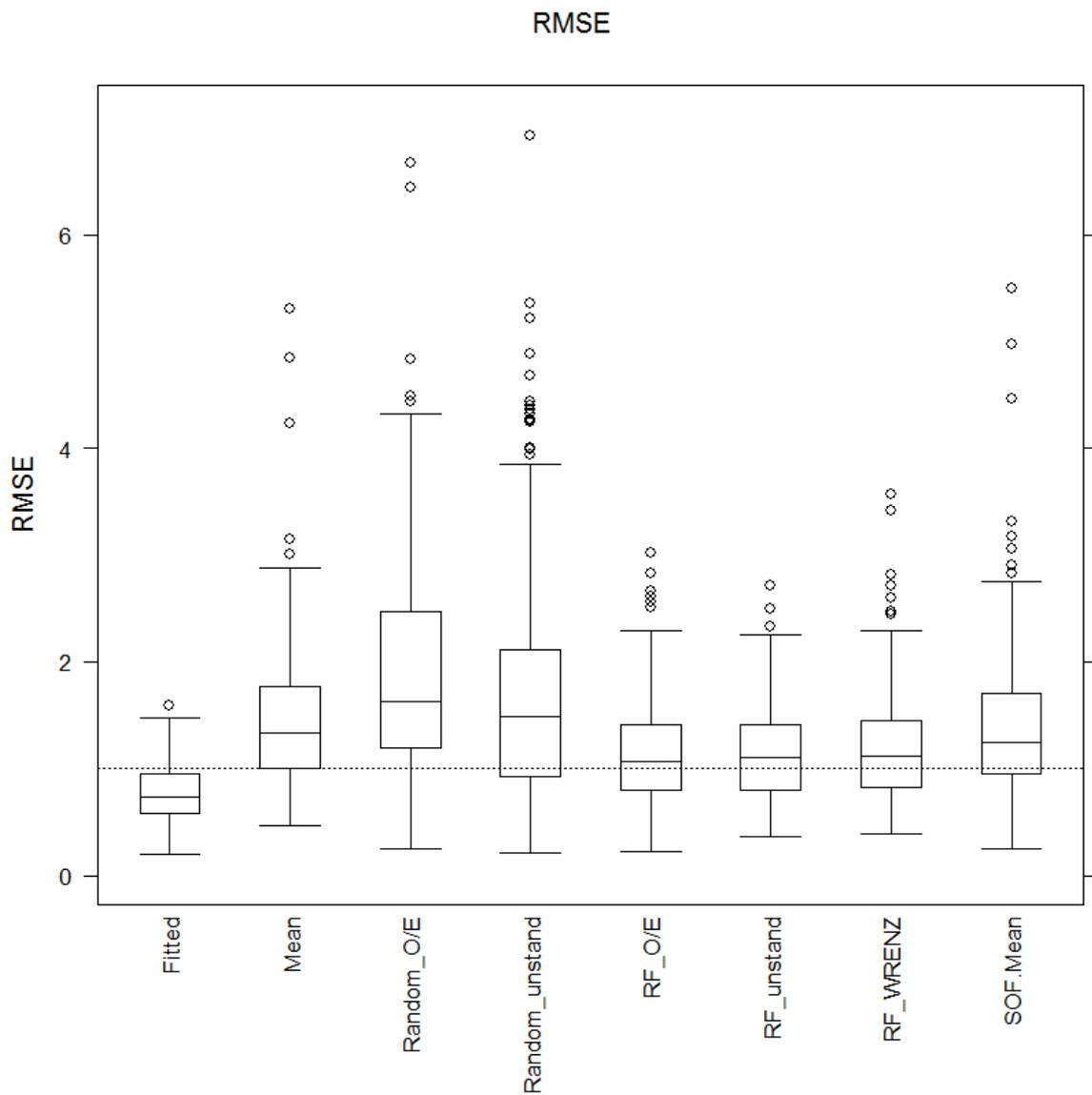


Figure 3-12: Boxplot of root mean square error (RMSE) for observed versus predicted suspended sediment concentrations generated using eight different methods across 271 sites. See Table 3-7 for method details. A lower RMSE indicates a better fit between observed and predicted values.

3.5 Discussion

Summarising the above results in regard to predicting SRC parameters from catchment characteristics:

- The RF models explained approximately 50% of the variance in the observed SRC parameters.
- The RF model predicting unstandardized SSC performed marginally better than the RF model predicting SSC standardised with discharge-weighted mean SSC.
- All the RF models performed substantially better than did a simple alternative model that related the SRC parameters to source of flow, and they performed much better than models that used only the mean of the observed SRC parameters.
- The catchment parameters exerting the strongest influence in the RF models on the SRC parameters (sediment yield for the **a**-parameter; factors linked to catchment slope for the **b**-parameter, e.g., elevation, steepness, stream order, unit stream power) generally align with those observed in international datasets.
- The factorial RMSE on the **a**-parameter is $\times/\div 2.29$, which aligns with the accuracy of previous sediment yield estimators. The RMSE on the **b**-parameter is ± 0.38 .
- The factorial RMSE on SSC predicted at any given discharge at any site by the RF models developed using both the standardized and unstandardized SSC data (along with the observed mean annual sediment loads) is $\times/\div 3.19$. This combines the uncertainty in the predicted **a** and **b** parameters but is only 1.42 times greater, on average, than the factorial RMSE associated with the observed fitted SRCs.

We regard these as reasonable results given the substantial natural data-scatter observed on SRCs. Considering that there is only a small difference in performance between the RF models using unstandardized and standardised SSC, we recommend using the standardised SSC as a matter of convenience for the sediment framework. Either RF model can be implemented by performing a one-off prediction of the SRC **a** and **b** parameters for every New Zealand REC2 segment within the R development environment, then simply accessing these results via “lookup” functions.

4 How sediment rating curves change when catchment sediment load changes

4.1 Introduction

A fundamental assumption of the ‘working’ analytical framework that links the ESVs to catchment sediment load is that a change in catchment sediment load (e.g., as a result of erosion mitigation works) will only alter the **a**-parameter of the SRC, not the **b**-parameter (and thus the SRC will simply shift vertically in log-log space). This means, conveniently, that the change in **a** will be proportional to the change in sediment load.

In this section we check this assumption in two ways. Firstly by examining several SRC datasets where the catchment sediment load is known to have changed due either to extreme hydrological events or to land-use change. In each case we apply statistical tests to evaluate if the rating curve **a** and **b** parameters change following the catchment disturbance. Secondly by considering, more generally, the dependence of **b** on catchment sediment load as informed by the results presented in Section 3.

4.2 Results from SRC datasets that capture a change in sediment load

4.2.1 Motueka

As reported in Hicks and Basher (2008) and Basher et al. (2011), suspended sediment loads were continuously monitored for 5-7 years (2002-2009) at four flow-recorder sites in the Motueka Catchment using turbidity as a proxy for SSC. These sites included the Motueka main-stem at Woodman’s Bend near the coast, the upper Motueka at Gorge, and the Motupiko at Christies. Automatic samplers were used throughout the monitoring campaign to characterise the relationship between SSC and turbidity, and the auto-samples thus also provided a record of the relationship between SSC and discharge. An intense rainstorm occurred across the upper Motueka and Motupiko sub-catchments during Easter 2005, and this reactivated gullies that, for a period of several years, provided elevated sediment loads from these tributaries. The storm had little impact in the western tributaries of the Motueka.

Figure 4-1a and Figure 4-1b plot the SRCs for the upper Motueka and Motupiko sites. Data are separated before and after the Easter 2005 storm. At both sites the log-transformed SRCs are straight so we fitted a first order polynomial to the before and after datasets (i.e., $\ln C = \mathbf{a} + \mathbf{b} \ln Q$), and then used a t-test to assess if there were statistically significant differences (at the 5% level) in their **a** and **b** parameters. The results showed a significant increase in **a** at both sites following the storm (equating to factor of 3.81 and 3.93 increases in SSC at Motueka at Gorge and Motupiko at Christies, respectively). Neither site showed a significant difference in **b** (2.08 ± 0.07 vs 2.13 ± 0.10 at Motueka Gorge; 1.70 ± 0.11 vs 1.88 ± 0.11 at Motupiko at Christies). Thus the effect of the Easter 2005 storm was to vertically shift the SRCs at both sites but not change their slopes.

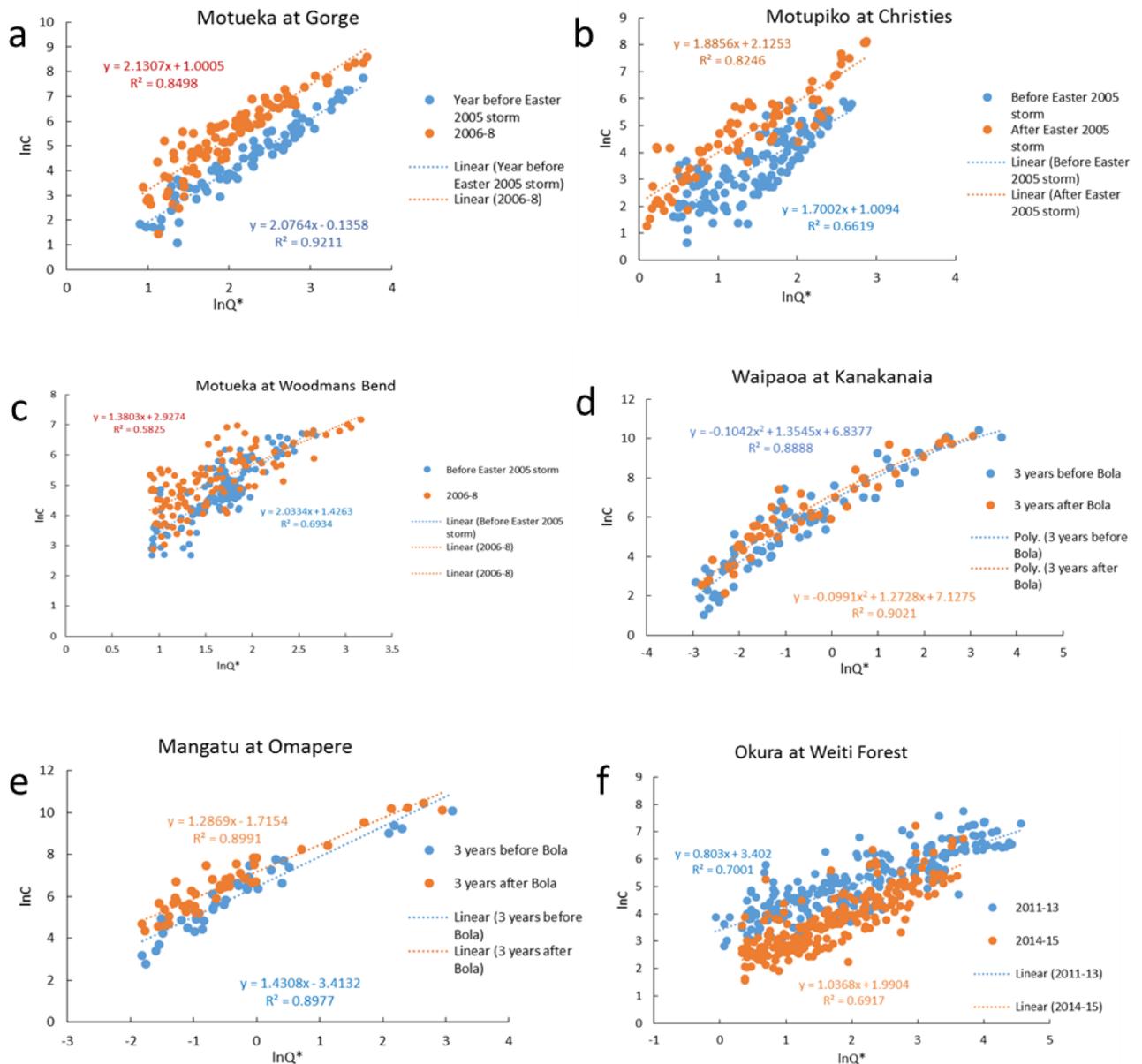


Figure 4-1: SRCs for: (a) Motueka at Gorge, (b) Motupiko at Christies, (c) Motueka at Woodman's Bend before and after the Easter 2005 rainstorm; (d) Waipaoa at Kanakanaia and (e) Mangatu at Omapere before and after Cyclone Bola (1988); and (f) Okura at Weiti Forest during and after forest harvesting. Data have been transformed to log values, discharge (Q) has been normalised by mean discharge.

Figure 4-1c plots the SRC for the Motueka mainstem at Woodman's Bend before the Easter 2005 storm and then through 2006-8. Again, first order polynomial SRCs are fitted. Our analysis showed a significant increase in **a** following the storm (equating to factor of 1.58) and also a significant reduction in **b** (from 2.03 ± 0.11 to 1.38 ± 0.11). Thus the effect of the Easter 2005 storm was to lift and clockwise-rotate the SRC at Woodman's Bend. This indicates that after the storm SSC was relatively higher at lower flows but tended to be little different at higher flows. Not shown here (but as reported by Hicks and Basher, 2008), the Wangapeka River, which is a western tributary of the Motueka and did not experience the extreme Easter 2005 rainfall, showed no change in its SRC. We

interpret that the change in both parameters of the Woodman's Bend SRC is because the load changed non-uniformly across its catchment. We re-visit this point in the discussion.

It is of note that Yang et al. (2007) also observed a reduction in **b** and increase in **a** associated with load increases in large Chinese rivers (Section 3.1).

4.2.2 Waipaoa

Suspended sediment rating data has been collected from the Waipaoa River at Kanakanaia Bridge and from its tributary the Mangatu at Omapere since the 1950s. As reported in Hicks et al. (2000), in March 1988 Cyclone Bola caused widespread land-sliding and gullyng across the East Cape area, and this was accompanied by an upshift in the SRC that lasted for several years.

Figure 4-1d plots the Waipaoa at Kanakanaia Bridge SRC data for the three years prior to Cyclone Bola and the three years after. The log-transformed SRC is curved¹⁶, so we fitted a second order polynomial to the before-Bola dataset and then used a t-test¹⁷ to assess if there were statistically significant differences (at the 5% level) in the means and trends (with discharge) of the residuals from this rating for the before-Bola and after-Bola data. The results showed a significant increase in the mean of residuals post-Bola (equating to a factor of 1.43 increase in SSC) but neither the before- or post-Bola residuals show a non-zero trend with discharge. In other words, the effect of Cyclone Bola was to upshift the SRC by a factor of 1.43 but it did not significantly change its shape.

Figure 4-1e plots the Mangatu at Omapere SRC data for the three years prior to Cyclone Bola and the three years after. The log-transformed SRC is straight so in this case we fitted a first order polynomial to the before-Bola and after-Bola datasets (i.e., $\ln C = a + b \ln Q$), and then used a t-test to assess if there were statistically significant differences (at the 5% level) in the before-Bola and after-Bola **a** and **b** parameters. The results showed a significant increase in **a** post-Bola (equating to a factor of 2.15 increase in SSC) but no significant difference in **b** (1.43 ± 0.08 vs 1.29 ± 0.07). Thus, as on the Waipaoa mainstem, the effect of Cyclone Bola was to upshift the Mangatu River's SRC but it did not significantly change its slope.

4.2.3 Weiti Forest

Auto-sampled SRC data was collected by Auckland Council from the Okura at Weiti Forest site prior to, during, and following full harvesting of the exotic forest in the catchment (84% of the catchment was in plantation forest, the remainder in pasture). Figure 4-1f compares the SRC data over the harvesting period (May 2011 – December 2013, which includes the pre-harvesting phase involving roading and landing-area formation/upgrading) with the post-harvesting phase (2014 – 2015) when it is expected that erosion scars associated with harvesting activities would commence to stabilise and/or become revegetated. Our analysis showed a significant decrease in **a** after harvesting (by a factor of 0.36) and also a significant increase in **b** (from 0.80 ± 0.03 to 1.04 ± 0.05); thus the SRC was lowered and rotated counter-clockwise. This indicates that after harvesting concentrations were relatively lower at lower flows but tended to be little different at higher flows. We suspect that this change was also influenced by an increase in runoff following tree harvesting.

¹⁶ Curved log-transformed SRCs are sometimes observed in the New Zealand context, particularly in rivers with high SSCs such as occur in the highly erodible East Cape region. Such curvature results in inaccuracies when the SRC is fitted with a 2-parameter model, as we have used in Section 3; however, use of a 3-parameter model and the attendant requirement to also predict the third parameter adds another level of complication that we chose to avoid for this study.

¹⁷ T-testing involves the assumption that the data being tested are normally distributed. The K-S test was used to evaluate this assumption, and in all cases reported in this section the log-transformed residuals from the SRCs were concluded to be normally distributed (at the 5% significance level).

4.2.4 Other sites

We also examined SRCs from six sites in the Manawatu and Rangitikei Catchments¹⁸ for temporal shifts associated with the extreme flood of February 2004 (which peaked at 3650 m³/s at the Manawatu at Teachers College site on 16/2/2004). Using t-tests as above, none of these showed any significant changes in SRC parameters following this event. This was surprising given the widespread hillslope erosion observed after this event; however, the SRC data scatter was large and so any signal of that event in regard to a transient increase in sediment supply would appear to have been no larger than normal variability. Similarly, examination of the SRC from Waitomo at Aranui Caves Bridge through the 1990s (following a phase of riparian management during the 1980s) showed no significant change in SRC parameters between the 1990-1995 and 1996-1999 periods.

4.3 Sensitivity of SRC b-parameter to sediment load indicated by modelling

An alternative way of assessing the sensitivity of SRC parameters to changes in catchment sediment loads is through the relationships observed in the modelling analysis in Section 3. Firstly, these show strong dependence of the **a** parameter (SRC intercept) on catchment mean annual sediment load (either directly or with load normalised by the mean discharge-weighted SSC) – as is expected. More importantly however (in the context of this section), the **b** parameter (SRC slope) appears to be only weakly related to sediment load (with strongest control exerted by catchment relief). In other words, this indicates that the **b** parameter should be relatively insensitive to catchment load.

4.4 Discussion

The site-by-site analysis results from Section 4.2 showed that temporal changes in catchment sediment load relating to sediment availability had mixed effects on the SRC slope. In the Waipaoa, Mangatu, Upper Motueka, and Motupiko cases the SRC slope was not significantly changed whilst the SRC shifted vertically by factors of around 1.5 – 4. The common feature of these sites was a widespread storm and associated increase in sediment supply. Of the two sites where the SRC slope also changed, the Motueka at Woodman's Bend had experienced changes in sediment supply over only part of its catchment, whereas the Okura at Weiti Forest also likely experienced an increase in runoff as a result of forest harvesting. Thus based on this limited set of observations, assuming a stable SRC slope at any site experiencing a change in sediment supply from upstream may only be reasonable if the sediment supply is changed uniformly across the catchment and/or the factors causing the change in sediment supply do not also change the runoff regime. For example, if a catchment sediment supply was artificially reduced by mitigating bank erosion along one key reach, or if widespread hillslope erosion was mitigated by re-forestation, these may well lead to changes in both the slope and offset of the SRC at the catchment outlet.

A simple modelling example illustrates the effect of non-uniform sediment supply. Consider three tributaries (1, 2, and 3 in Figure 4-2) all producing equal discharge (Q) but potentially different sediment loads and SRC parameters. The SRC below their confluence is derived by summing, at given discharges, the products of SSC (C) and Q from each tributary to get the total load and then dividing this by the total discharge. C at each tributary is defined by a rating curve $C = aQ^b$. Figure 4-2 shows three trials. In Trial 1, each tributary has the same **b** values and initially the same **a** values; then, Tributary 1 is treated to reduce its sediment load by 90% (**a** reduces from 0.1 to 0.01). In this case, the **a** at the catchment outlet reduces after treatment but the **b** remains unchanged, thus the SRC rating simply shifts down vertically. In Trial 2, each tributary has different **b** values but all are treated

¹⁸ Manawatu at Teachers College, Manawatu at Hopelands, Pohangina at Mais Reach, Makuri at Tuscan Hills, Rangitikei at Mangaweka, Rangitikei at Pukeokahu.

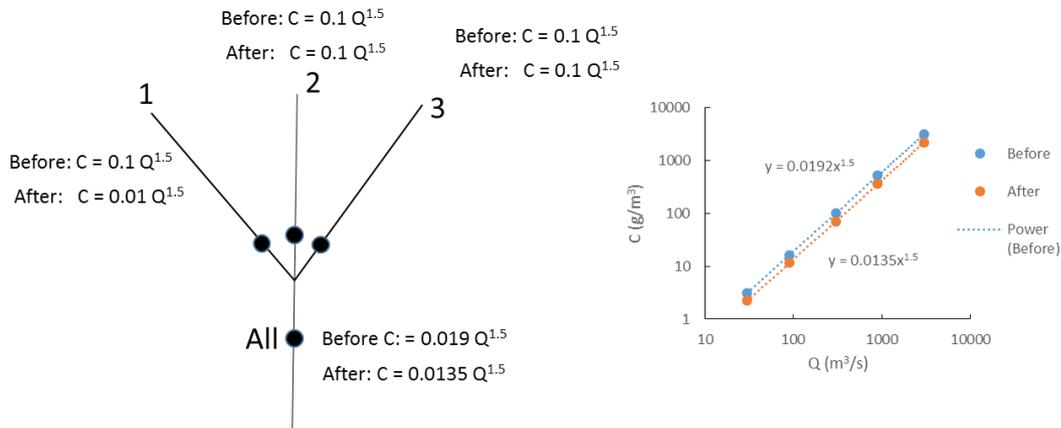
to reduce their loads by 90%. In this case, after treatment only the **a** changes at the catchment outlet as the rating shifts down. In Trial 3, the tributaries have different **b** values but only Tributary 1 is treated. In this case, after treatment both **a** and **b** change at the catchment outlet so that the rating lowers and flattens. Further iterations of this simple model demonstrate that the extent of change in **b** after treatment depends on the spread of tributary **b** values, the **a** values, and the relative discharges from the tributaries. In reality, the response will be further complicated by phase lags between water and sediment arrival from various tributaries at the catchment outlet.

Warrick (2015) noted that **a** values related to sediment supply while **b** values reflected the relative rate of supply of water and sediment to the river channel. He showed nine possible outcomes on rating curve slope and offset arising from changes in sediment load and/or water discharge (Figure 4-3, cases a – f). Vertical offsets with no slope change (cases d and f) can occur due to either a change in load or a change in discharge (which can either dilute or increase SSC). Combinations of changes in relative sediment load and runoff lead to the other outcomes. For the New Zealand cases discussed in Section 4.2: the Waipaoa, Motueka Gorge, and Motupiko examples align with Figure 4-3d; the Motueka at Woodman’s Bend aligns with Figure 4-3a; and the Okura at Weiti Forest aligns with Figure 4-3i.

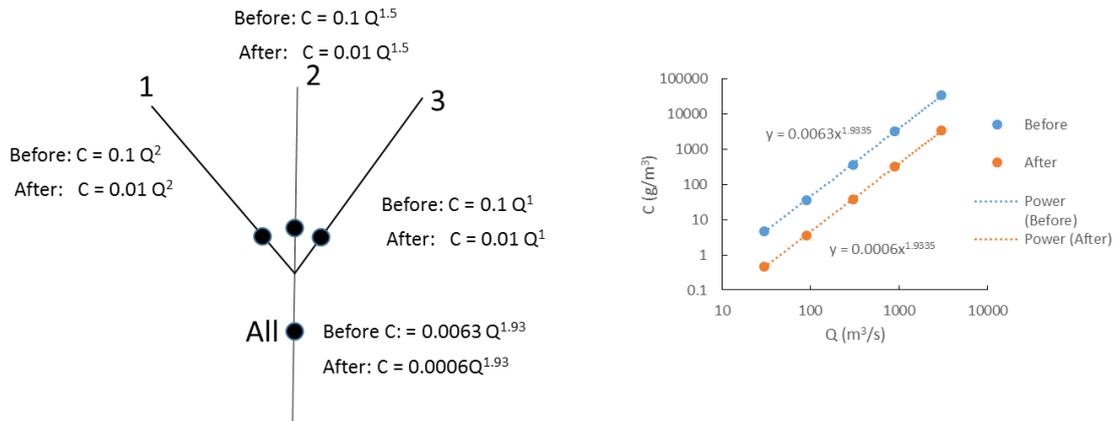
Warrick (2015) also illustrated case examples of changes in SRC parameters from US and Chinese rivers, and undertook Monte Carlo simulations of multiple realisations of changes in water and sediment supply. A key finding from this exercise was that **b** values changed if the sediment load increased more (or less) during larger (or smaller) events. In practice, this is influenced by the various erosion processes that come into play during different sized events. An example would be where bank erosion tended to dominate the sediment load during base flows and flood recessions: bank stabilisation would therefore drop the SRC more over the low-mid flow range, steepening the rating overall and increasing the **b** value.

Thus, as illustrated both by the New Zealand field data and our simple model, and as corroborated by Warrick (2015), the assumption of a stable SRC slope under changing catchment sediment load does not hold-up where tributary ratings and sediment load changes are non-uniform and/or the load change is accompanied by a change in runoff regime. However, a potential way forward with the ‘framework’ would be to route SRC parameters from tributary sources downstream in a manner similar to that illustrated in Figure 4-2. This approach and its validation are discussed further in Section 10.

Trial 1: Uniform b, non uniform treatment



Trial 2: Non-uniform b, uniform treatment



Trial 3: Non-uniform b, non-uniform treatment

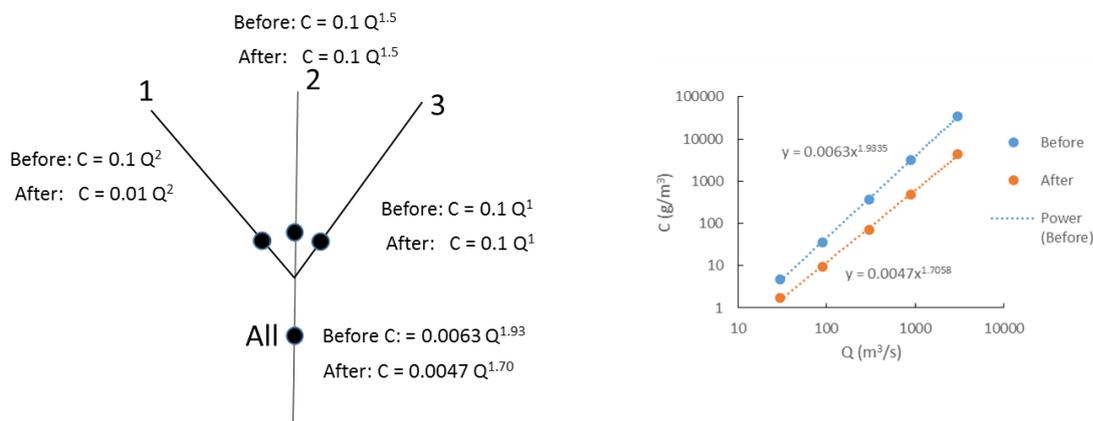


Figure 4-2: Simple examples of change in catchment-outlet SRCs arising from relative change in sediment loads from tributaries. Trial 1 has uniform a and b values for the SRCs across all three tributaries, then a is reduced by 90% in Tributary 1. Trial 2 has different b values among the tributaries but a is reduced by 90% in all tributaries. Trial 3 has different b values among the tributaries and a is reduced by 90% only at Tributary 1. At the catchment outlet, only a changes after treatments in Trials 1 and 2 but both a and b change in Trial 3.

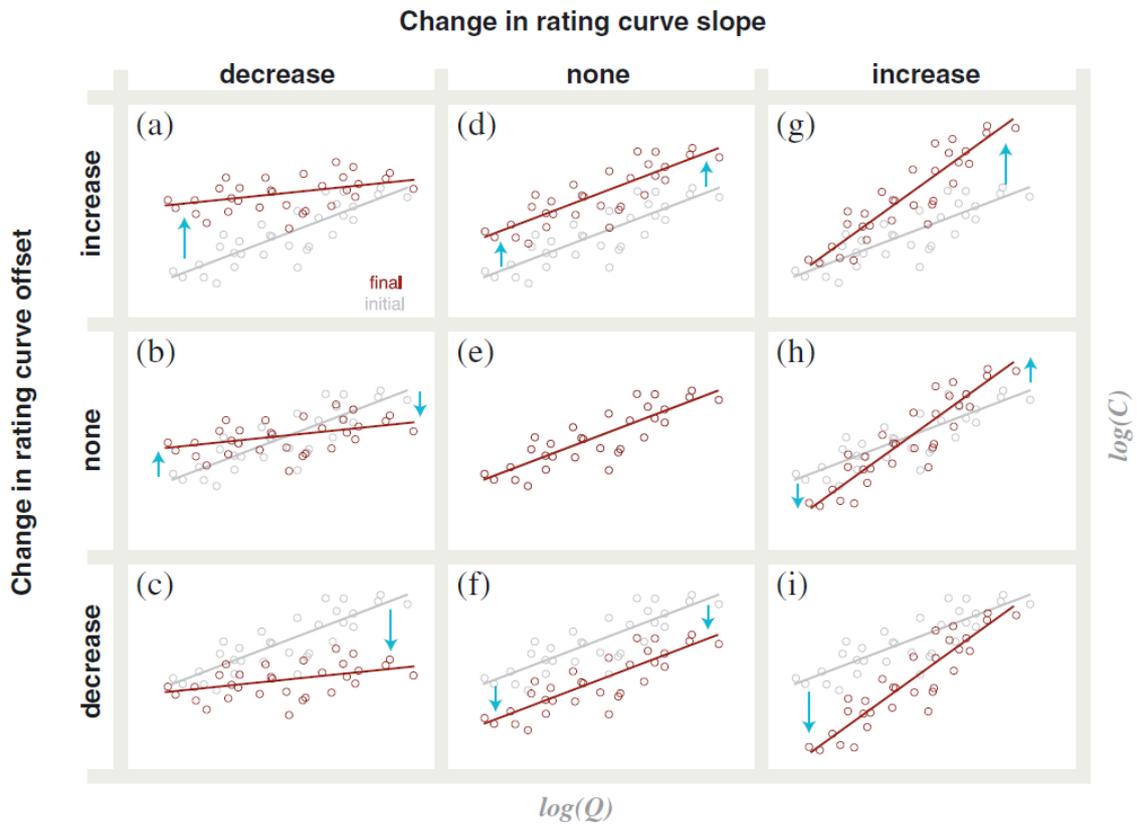


Figure 4-3: Conceptual trends in the relationships between river discharge (Q) and suspended-sediment concentration (C) as expressed by changes in the sediment rating curve slope and vertical offset (from Warrick 2015, Figure 2). Each plot shows log-transformed synthetic Q and C data from two time intervals, initial (light colour) and final (dark colour). Arrows show the trends in the least-squares power function rating curves (lines).

5 Mutual relationships between visual water clarity, turbidity and suspended matter in rivers

5.1 Introduction

This section examines and develops predictive models for the relationships between suspended sediment concentration (SSC), turbidity, visual water clarity (VC), and light penetration (LP) for the purpose of inter-converting between the water column sediment ESVs. We include turbidity because it is an often-used proxy for both SSC and VC and it is likely that it will be important, when measured continuously, for implementation of NOF-sediment.

We expected to be able to relate these variables to each other at-a-site with simple models and then to relate the at-a-site model parameters to catchment characteristics available on the REC2 database.

Since both VC and LP depend on how light is attenuated in water by suspended and dissolved constituents, the following brief review sets out our theoretical expectations of these relationships.

Light attenuation is caused by two main optical processes: *light scattering* in which light photons change direction but not energy, and *light absorption* in which photon energy is converted to another form (ultimately heat) (Kirk 2011). Suspended particulate matter (SPM) is the dominant light scattering material in all but the clearest natural waters, and typically SPM also contributes strongly to light absorption. Total light beam attenuation (per unit length of light path) is the sum of absorption and scattering:

$$c = a + b \quad (3)$$

where c is the beam attenuation coefficient (BAC, units 1/m), a is the absorption coefficient, and b is the scattering coefficient.

VC is inversely related to the (photopic) beam attenuation coefficient (which is very close to the beam attenuation coefficient (c) at 550 nm near the peak sensitivity of the human eye – Zanevald and Pegau 2003):

$$c(550) = 4.8/VC \quad (4)$$

The beam attenuation coefficient (c) can be partitioned into contributions from the three light attenuating constituents: SPM (p), CDOM (g) and water itself (w), thus:

$$c(550) = c_p(550) + c_g(550) + c_w(550) \quad (5)$$

SPM is usually the dominant light-attenuating constituent controlling VC and LP, and the value of c_p depends on particle concentration, composition, size, and shape. Thus for a given type or mixture of particles there tends to be a linear relationship between c and SSC – the higher the value of SSC, the higher the value of c .

Light beam attenuation by water itself, $c_w(550)$, is $\sim 0.058/m$ (Pope and Fry 1997) - which is very low and can be safely ignored in rivers and, in fact, in all but the *very* clearest natural waters (Gall 2013). However, the contribution of CDOM, $c_g(550)$, might sometimes be important in restricting visual water clarity in strongly coloured but low turbidity waters. For example, the 95 percentile CDOM content in the NRWQN, as indicated by $a_g(440)$, is about 2.5/m. The corresponding value of $c_g(550) =$

$a_g(550) \sim 0.43/\text{m}$ (accounting for the typical absorption spectral shape of CDOM), which would imply a visibility of 11 m with no SPM present. In Frosty Creek (Westland), which is very strongly humic-coloured but of low turbidity, Davies-Colley and Nagels (2008) reported $a_g(440) = 12.6/\text{m}$ and $\text{VC} = 0.47$ m, when expected VC due to CDOM attenuation alone would be about 2.2 m. So even in this very coloured and non-turbid stream, particles still dominated light beam attenuation and controlled VC. So it seems that CDOM is seldom very important as regards VC in New Zealand rivers.

CDOM is more likely to be important to LP because of its strong light absorption – which is the process that actually removes light from the water column. The main optical effect of SPM is light scattering. So, as first pointed out by Kirk (1985), the main mechanism by which suspended fine sediment affects LP is by forcing photons to take a tortuous path down through the water column – increasing average path length of photons over a given depth interval and increasing their probability of absorption. The stronger effect of light absorption than scattering (by particles) to attenuation of (diffuse) light with depth can be seen in a semi-empirical equation given by Equation (6) (Kirk 2011). The attenuation of diffuse light with depth (symbol K_d , units: $1/\text{m}$) is directly proportional to the light absorption coefficient, a , but to the square root of the light scattering coefficient, b :

$$K_d = [a^2 + Gab]^{0.5} \quad (6)$$

where G is a factor depending on the structure of the light field in water.

The overall effect of sediment particles on LP depends on whether they merely scatter light or also absorb light – typically due to chemically adsorbed (coloured) organic matter. Typically, sediment in waters both scatters and absorbs light, but there are counter examples. For example, the glacial flour (dominated by mica) in Lake Pukaki lacks organic matter and has an extremely high ratio of scattering to absorption, with the result that light penetrates much more deeply into that lake than would otherwise be expected and a much higher fraction of the incident light back-scatters out of it (Gallegos et al. 2008). Lake Pukaki may indeed be one of the ‘brightest’ lakes in the world because its light absorption is so low compared to its light scattering by sediment (Davies-Colley et al. 2003).

CDOM is not routinely measured by regional authorities, so accounting for this constituent would require a model predicting unmeasured CDOM with sufficient accuracy, for example from REC parameters of catchments. This could be done in principle (using the CDOM data in the NRWQN), but would be appreciably onerous for present purposes and was judged not worthwhile. In the NRWQN, although there are numerous coloured rivers these are also turbid rivers, so VC remains dominated by SPM. LP is slightly affected by CDOM, and a correction for the CDOM effect should be possible in future when CDOM is more commonly measured or after a national model is developed for its estimation.

Thus assuming that suspended particulate matter is the dominant light-attenuating constituent, we expect that the parameters of the relationship between VC (and LP) and SSC should depend primarily on the suspended particle characteristics (size, shape, and composition¹⁹).

Moreover, if we further assume a linear relationship between c and SSC then the parameter $c^* = c/\text{SSC}$ should also index particle characteristics (in similar fashion to specific turbidity, equal to turbidity/SSC, as discussed in Section 6.3.1). The ratio c^* has units m^2/g and is expected to be inversely related to particle size above about $1 \mu\text{m}$ in diameter (e.g., Davies-Colley and Smith 2001).

¹⁹ Particle composition (particularly whether inorganic or organic) determines the particle density, and thus exerts an important control on the relationship between optical cross-section and mass concentration (as measured by SSC).

Fine clays of around 1 μm have near-maximal c^* values $\sim 1 \text{ m}^2/\text{g}$ (and plate-shaped clay minerals can have substantially higher values), whereas coarse silt particles of 50 μm have a $c^* \sim 0.02 \text{ m}^2/\text{g}$. The underlying concept here is that if c^* can be 'regionalised', then this provides a simple means for converting VC to TSS or vice versa²⁰.

Elliott et al. (2013) introduced the concept that the beam attenuation coefficient, c , can be considered as an optical cross section (OCS) per volume of water (analogous to a sediment concentration), with the load of optical cross-section (LOCS) and yield of optical cross-section (YOCS) the optical equivalent to catchment sediment load and yield. Elliott et al. (2013) found that catchment slope, rainfall, rock induration and proportion of pasture cover were important factors in determining YOCS, thus we might expect that these similarly influence regional variation in c^* .

5.2 Data

5.2.1 Sources

To address this objective we sought data where, ideally, all three water column ESVs (SSC, VC and LP) plus turbidity were measured on the same site visits (or in the same water samples). We expected there would be very little data on LP into rivers, the only important exception being NIWA's own data collected for the purpose of developing statistical models of LP as a function of optical water quality (Davies-Colley and Nagels 2008). However we anticipated that several regional authorities would have data for SSC as well as turbidity and VC from SoE monitoring networks.

We also revisited three NIWA datasets in which SSC, VC and LP have been measured simultaneously:

- The NRWQN (77 river sites) – to which suspended matter assays (TSS²¹, also VSS, POC and PON) were added in 2011-2015 (as reported by Davies-Colley et al. (2014); and Ballantine et al. (2014))
- Whatawhata Research Station stream sites (8 sites) related to the integrated catchment management (ICM) experiment (riparian planting, pine afforestation, and native planting – Quinn et al. (2009) as reported by Hughes and Quinn (2014)
- Kaipara tributary rivers (6 sites) – to which VC and bench turbidity measurements were added in 2013- 2015 'piggy-backed' on sediment load monitoring (which, uniquely, included auto-samples of stormflows – sometimes with sub-centimetre visibility) (Hughes et al. 2014).

No new LP data was found in our data search, but we consider that the simple semi-empirical model of LP as a function of VC and CDOM that was developed by Davies-Colley and Nagels (2008) can probably be used to estimate LP sufficiently accurately in NZ rivers where required. Therefore, we concentrate hereafter in this section on VC data.

We note also the dearth of CDOM data and the need for more measurements (as recommended by Davies-Colley et al. 2012) or a spatial model for its prediction.

²⁰ It must be acknowledged that c^* is not an ideal parameter for indexing the relationship of VC and TSS. This is because in rivers c^* typically trends systematically as TSS increases mainly with increases in flow, throwing increasingly coarser material into suspension.

²¹ The bulk of SSC data analysed for this section was analysed in the laboratory using the TSS method, thus henceforth in this section SSC is referred to as TSS and $c^* = c/\text{TSS}$.

5.2.2 Dataset screening

Analysis commenced with scatterplots of (three combinations of two) variables: TSS, turbidity, and VC. Initially the plots were inspected as an indicator of data quality, on the basis that these three variables are known from previous work to be reasonably closely inter-related (Davies-Colley and Smith 2001; Davies-Colley et al. 2014).

This plotting immediately revealed some data quality concerns. For example, Figure 5-1 for Whatawhata Research Station stream sites showed more data scatter than expected (and some apparent outliers). The suspect data was traced to inaccurate (biased) VC on some occasions on the basis that the VC was also biased low when plotted against turbidity, while turbidity and TSS were more closely related. We think that the VC data from Whatawhata may be biased on occasion due to measurements being made under riparian shade on clear days when partial shadowing of the path of sight is expected to affect the sighting range. Another possible source of bias is measurement without sufficient water behind the black disc, such that the river banks affect the sighting range. This kind of problem is probably common in small streams and rivers with (a) discontinuous riparian cover and (b) restricted viewing range in small channels, but far less problematic in the mostly large rivers of the NRWQN.

Similarly, when we plotted scatterplots of regional SoE data for VC, TSS, and turbidity there was often more data scatter than expected and appreciable numbers of apparent outliers. Regional data for small rivers may be affected by biased VC in the same way as at Whatawhata – arising from readings with partial shadowing of the path of sight or insufficient water path behind the disc. Another potential problem with visibility measurement, particularly in small streams, is where insufficient care is taken to avoid disturbance plumes created by wading in the channel.

Some of the regional TSS data was also of insufficient quality, with numerous non-detects (often with $TSS < 3 \text{ mg/m}^3$) and 'striping' of TSS data due to numerical rounding to 1 significant figure when TSS was low (approaching, but still above, the detection limit). TSS data in the add-on to the NRWQN, in contrast, were measured with large volume samples – up to 5 litres in the very clear Motueka River – so as to avoid approaching the detection limit even in very clear water. In any case, as noted above, CDOM is not routinely measured by councils so correction for (diffuse) light attenuation by this constituent with resulting effects on LP would require predictions of CDOM from a model that has not yet been developed.

Based on the above, we elected not to use the Whatawhata Research Station data or regional SoE data, particularly in view of quality concerns. That left the NRWQN (sediment add-on) data for exploring mutual relationships of TSS/turbidity/VC. The data at three NRWQN sites (Hoteo at Gubbs (AK1), Mangakahia at Titoki Bridge (WH3) and Wairua at Purua (WH4)) were supplemented by measurements of TSS/turbidity/VC on samples collected during storm events by NIWA, with the assistance of Auckland Council and Northland Regional Council, as part of NIWA's Cumulative Effects of Stressors on Aquatic Ecosystems Research Programme.

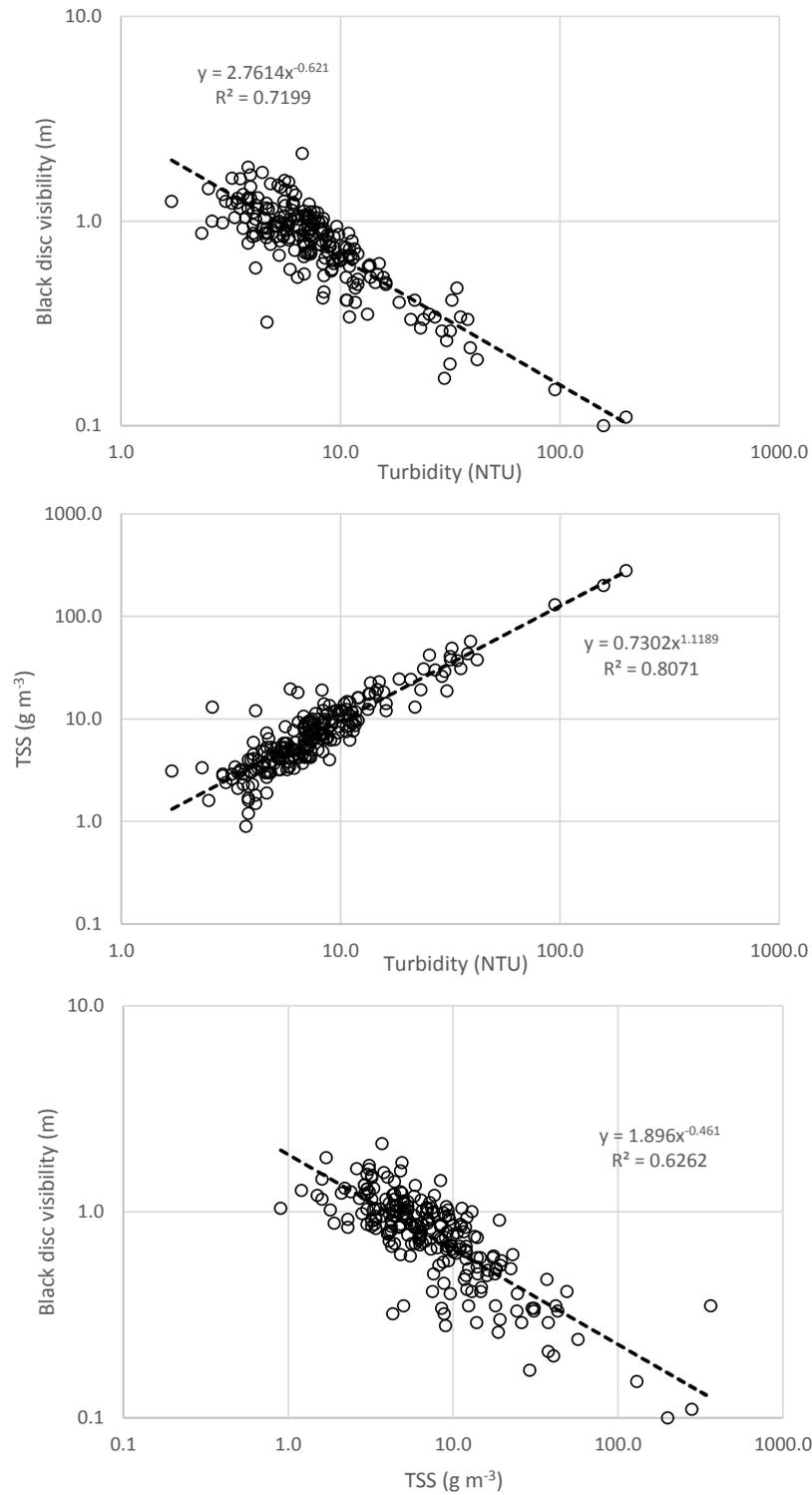


Figure 5-1: Mutual relationships of VC (black disc visibility), turbidity and TSS at a monitoring site (PW3; 49 ha) at the Whatawhata Research Station (1995-2014). Regression lines were fitted as power functions.

5.2.3 Data analysis

Data analysis was carried out in four steps using a combination of MS Excel and R statistical software:

- exploring the mutual relationships of TSS, VC and turbidity
- fitting power-law (two parameter) regression models to the at-a-site relationships between VC, TSS and turbidity, using standard major axis (SMA)²² regression on the \log_{10} transformed data with the 'lmodel2' package in R
- exploring the statistics of the c^* values at-a-site, using this as an alternative, single-parameter estimator of VC off TSS
- developing Random Forest (RF) models in R to predict: (i) the two parameters of the at-a-site relationships between TSS, VC and turbidity off catchment-characteristics; (ii) the at-a-site median c^* off catchment-characteristics.

These analyses were run separately on two datasets:

- the full set of NRWQN sites ($n = 77$)
- a reduced dataset where flow impacted sites (those influenced by upstream lakes or reservoirs) were removed ($n = 64$).

We ran these two analyses to assess if lake outflows had any significant impact on optical signature.

We used a similar approach to the prediction of SRC parameters in Section 4. Further details of the Random Forest methodology and terminology can be found there and in Appendix E.

5.3 ESV inter-relationships and parameters

5.3.1 VC vs TSS

Figure 5-2 shows VC plotted against TSS for the pooled data from the NRWQN²³. There are a few outliers that were judged 'improbable' and removed for the purpose of subsequent analyses.

Highlighted on Figure 5-2 are the points for the clearest river site in the NRWQN (Motueka at Gorge, NN2) and the least clear river (Waipaoa at Kanakania, GS1). Interestingly, although these two rivers differ greatly in characteristic VC and SPM, their plots overlap, suggesting that the particles present in their waters are similar in character when they have similar TSS and VC.

²² SMA regression minimises unexplained variance in both the X and Y directions and is preferred over ordinary least squares regression when the regression models are to be applied to predict X off Y as well as vice-versa.

²³ This plot updates a similar plot by Davies-Colley et al. (2014) that was affected by non-detects and numerical rounding causing 'striping'. Both issues have since been addressed for the NRWQN data used in the present study, using the original laboratory worksheets to replace non-detects by best estimates and increasing numerical precision to avoid rounding error.

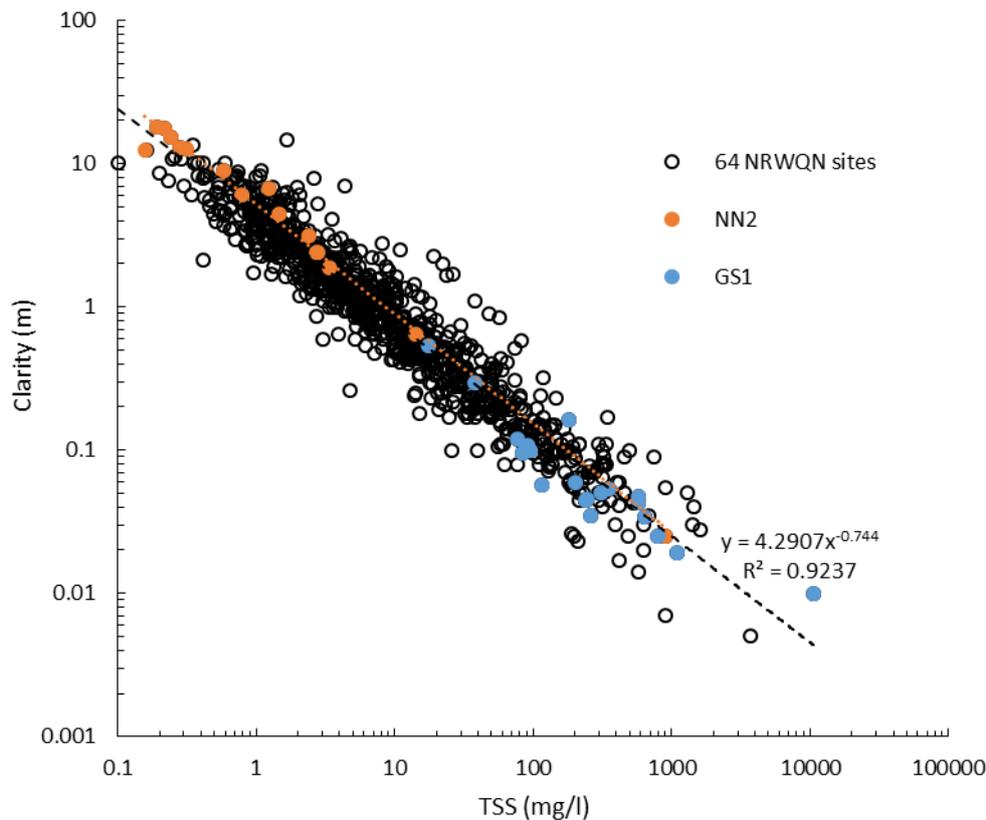


Figure 5-2: Relationship between VC (as measured by the black disc method) and total suspended solids concentration at the un-impacted NRWQN sites. Points for the Motueka at Gorge (blue) and Waipaoa at Kanakania (orange) are highlighted. Data are derived from 12 months (2011) of monthly sampling and four years (2012-2015) of selective (flow dependent) sampling during monthly visits.

As a preliminary exploration, a power-law regression model of the form

$$VC = a' \times TSS^b \quad (7)$$

where a' (coefficient)²⁴ and b (exponent) are fitting parameters, was fitted to the pooled data for both the full set of NRWQN sites and the sub-set with un-impacted flow regimes (Table 6-1). For the un-impacted sites this gave $a' = 4.28$ and $b = -0.743$, explained 92% of the variance in the logs of the pooled observed VC values (i.e., $R^2 = 0.92$), and had a factorial standard error of the estimated VC of $\times/\div 1.49$ (which equates to a percentage error of $\pm 49\%$). The high R^2 suggests a very strong relationship, but this is over a very wide data range (e.g., 10,000 for VC) and the predictive utility of this pooled-data model is limited by its high standard error. It is of note that the exponent is smaller in magnitude than 1, indicating that a proportional change in TSS will result in a proportionally smaller change in VC (e.g., doubling TSS will only reduce VC by a factor of 1.7). Very similar results were obtained with the full dataset (Table 5-1), suggesting little is gained by removing the impacted sites.

²⁴ The log transformed version of Equation (7) is $\log_{10}VC = a + b \log_{10}(TSS)$, where the intercept $a = \log_{10}a'$.

Table 5-1: Linear regression statistics for relationships between VC, TSS, and turbidity for pooled NRWQN dataset for cases of all sites and sub-set with unaltered flow regimes. Regression model of form $y = a'x^b$. FSE is the factorial standard error on the estimate of y .

Regression Coefficients	Excluding flow impacted (n = 64)				Including flow impacted (n = 77)			
	a'	b	R^2	FSE on a'	a'	b	R^2	FSE on a'
VC-TSS	4.28	-0.743	0.92	$\times/\div 1.49$	4.48	-0.748	0.91	$\times/\div 1.52$
VC-turbidity	3.19	-0.781	0.94	$\times/\div 1.44$	3.14	-0.776	0.93	$\times/\div 1.44$
TSS-turbidity	1.65	0.99	0.90	$\times/\div 1.80$	1.81	0.960	0.88	$\times/\div 1.87$

The same power-law model (i.e., Equation (7)) was fitted by SMA regression to the log transformed TSS and VC data for each of the NRWQN sites. Appendix F details the regression results while Table 5-2 provides summary regression statistics across sites for the two datasets (i.e., all 77 sites, sites with un-impacted flow regimes). There is little difference between these two dataset cases, so we focus on the sub-set of sites with un-impacted flow regimes.

Figure 5-3 shows the distributions of the regression parameters for these sites. The intercept a for the VC vs TSS relationships ranged between 0.38 and 1.16, averaged 0.66, and had a standard deviation of 0.16. Untransformed, these figures indicate a factor of 6.3 range in a' and a factorial standard deviation of $\times/\div 1.44$ (equating to a % error of $\pm 44\%$). The exponent (b) ranged from -0.50 to -1.07 (standard deviation of ± 0.12) with a median value of about -0.7. As noted above for the pooled dataset, this indicates that typically a proportional increase in TSS in a particular river induces a lesser proportional reduction in VC. The reason may be that an increase in TSS at-a-site is typically accompanied by a coarsening of particle size (as flow increases), and coarser particles attenuate light less than fine particles so VC is not proportionately affected.

Table 5-2: Summary statistics for the at-a-site SMA regression intercepts and slopes for the mutual relationships of VC, TSS and turbidity and for the site-median c^* . Results are for two data sets: (i) excluding flow impacted sites, n=64; (ii) all sites, n = 77. The regressions were fitted to log-transformed data. The means of the intercepts (a) are log values; the numbers beside in brackets give the re-transformed mean coefficient, i.e., a' in Equation (7). Similarly for the intercept standard deviation; the numbers beside in brackets give the factorial standard deviation for the re-transformed coefficient (equal to 10 raised to the power of the standard deviation of the intercept derived in log-space). Ranges are very similar for all sites case, so are not shown.

Coefficient	Excluding flow impacted sites			Including flow impacted sites	
	mean	SD	Range	mean	SD
VC-TSS intercept (a)	0.66 (4.57)	0.16 (1.44)	0.38 to 1.16	0.68 (4.79)	0.16 (1.44)
VC-TSS slope (b)	-0.76	0.12	-1.07 to -0.05	-0.76	0.13
VC-TSS R^2	0.88	0.15	0.15 to 0.99	0.82	0.23
VC-turbidity intercept (a)	0.50 (3.16)	0.12 (1.32)	0.25 to 0.82	0.50 (3.16)	0.11 (1.29)
VC-turbidity slope (b)	-0.79	0.09	-1.01 to -0.57	-0.78	0.10
VC-turbidity R^2	0.88	0.13	0.17 to 0.99	0.85	0.17
Turbidity-TSS intercept (a)	0.20 (1.58)	0.24 (1.74)	-0.49 to 0.86	0.23 (1.70)	0.26 (1.82)
Turbidity-TSS slope (b)	1.05	0.19	0.59 to 1.45	1.05	0.21
Turbidity-TSS R^2	0.84	0.19	0.13 to 0.99	0.78	0.29
Median c^*	0.70	0.25	0.24 to 1.41	0.70	0.25

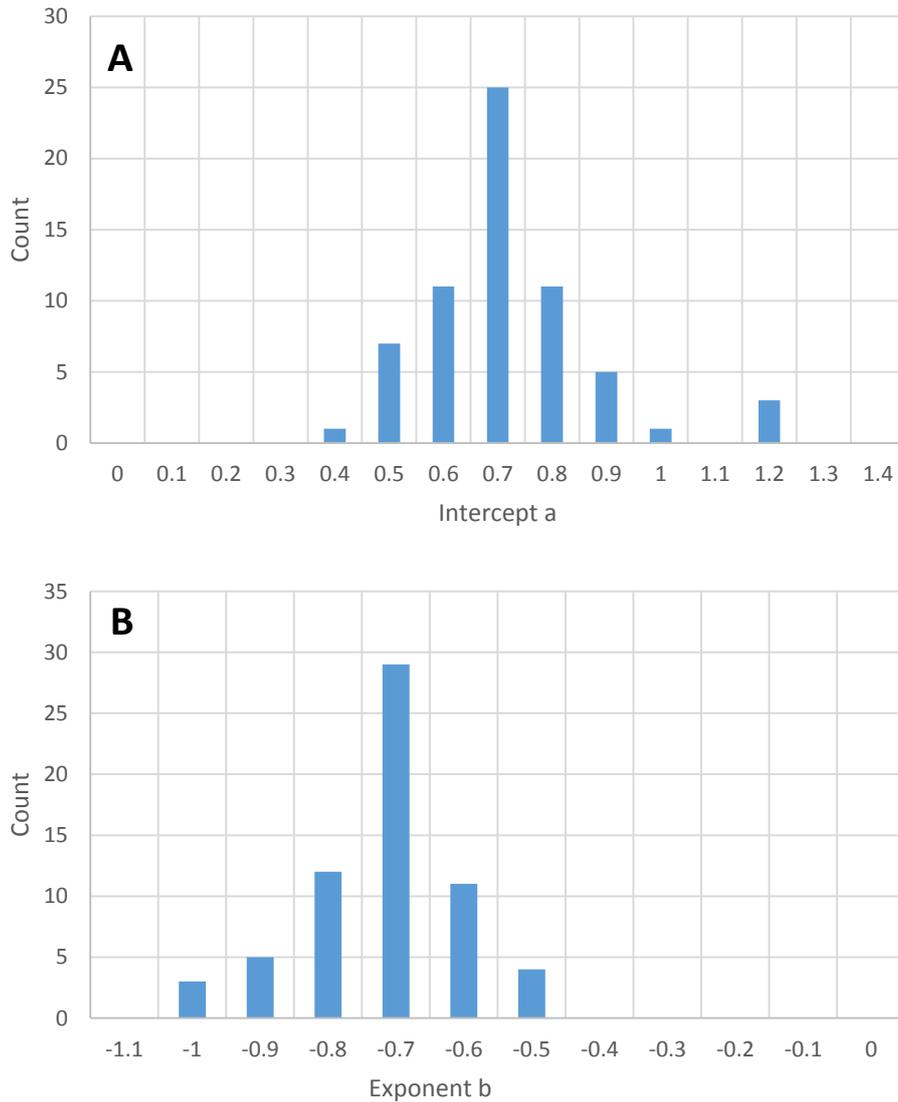


Figure 5-3: Histograms of power function parameters (from SMA regression) for VC versus TSS plots. A: intercepts (log transformed data), B: exponents.

5.3.2 TSS and VC vs turbidity

Figure 5-4 shows the relationships between VC and turbidity and between TSS and turbidity for the pooled dataset, while Table 5-1 includes statistics for the associated regression models. Again, there was little difference in results for the two dataset cases. The VC vs turbidity relationship shows very similar parameters to the VC vs TSS relationship but with slightly better factorial standard errors for the former ($\times/\div 1.44$) compared to the latter ($\times/\div 1.49-1.52$). This suggests that turbidity is a slightly better predictor of VC than is TSS, which is to be expected since both VC and turbidity are optical measures (i.e., of light attenuation and absorption and light back-scattering, respectively). Interestingly, turbidity is more linearly related to TSS ($b = 0.99$) but the TSS vs turbidity relationship shows more data scatter and the largest factorial standard error ($\times/\div 1.80$). Thus these data confirm that, for a given TSS, both VC and turbidity respond to suspended particle physical characteristics but in different ways.

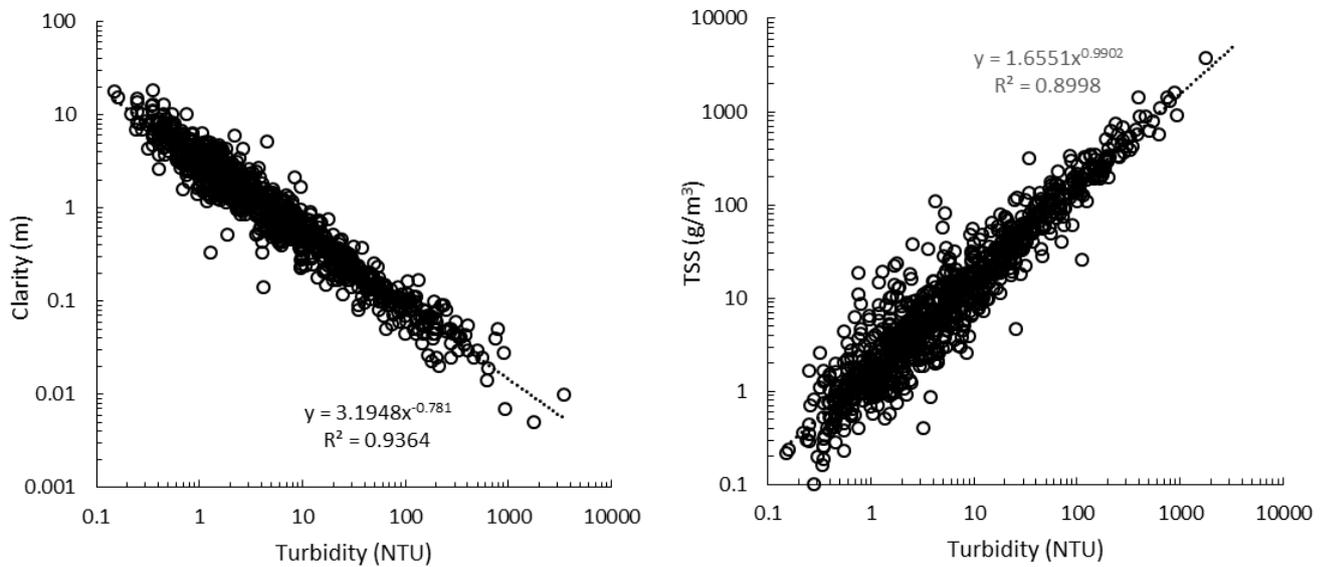


Figure 5-4: Relationships between visual clarity and turbidity (left) and between TSS and turbidity (right) using pooled data from the 64 un-impacted NRWQN sites.

Table 5-2 includes summary statistics for the SMA regression slopes and intercepts for the VC vs turbidity and TSS vs turbidity relationships at individual NRWQN sites. Again, there was little difference in results for the two dataset cases, and the results were very similar to those obtained using the pooled dataset. The main difference was that the factorial standard deviations associated with the individual site intercepts were lower than the factorial standard errors associated with the pooled datasets (e.g., $\times/\div 1.74$ compared with $\times/\div 1.80$ for the TSS-turbidity relationships for the un-impacted sites).

5.3.3 Regression model parameter inter-dependence

We found that the a and b parameters for the at-a-site VC vs TSS relationships were not independent (Table 5-3), with a low slope value (b) tending to be associated with a low intercept (a). On the one hand, this means that the intercept is a poor index of the expected VC at a given TSS. However, it also potentially simplifies ‘regionalisation’ since by being able to predict one parameter from the other it is only necessary to predict one parameter from catchment characteristics. We found similarly that the equivalent intercepts and slopes for the turbidity vs VC and turbidity vs TSS relationships were also not independent (Table 5-3). The regression parameters (α , β) for slope vs intercept models (i.e., $a = \alpha + \beta b$) for all three inter-relationships for the un-impacted sites are included in Table 5-3.

Table 5-3: Linear regression results for models predicting SMA regression intercepts (a) from slopes (b) for the mutual relationships of VC, TSS and turbidity. Flow regulated sites excluded; $n = 64$. Significance levels (p) on R^2 are < 0.001 in all cases. Predictive model of the form $a = \alpha + \beta b$. Figures in brackets after standard error of estimate give factorial standard error on untransformed a (a') and equals $10^{\text{Std error on estimate of } a}$.

Regression parameters	R^2	α	β	Standard error on estimate of a
TSS-VC	0.57	-0.104	-1.006	± 0.102 ($\times/\div 1.26$)
VC-turbidity	0.57	-0.271	-0.984	± 0.080 ($\times/\div 1.20$)
Turbidity-TSS	0.23	0.864	-0.630	± 0.210 ($\times/\div 1.62$)

5.4 c^*

As outlined in Section 5.1, VC measurements can be converted to beam attenuation coefficient (c), and the ratio $c^* = c/\text{TSS}$ thus provides the potential to represent the VC vs TSS relationship by a single parameter. Figure 5-5 shows the relationship between beam attenuation (c at 550 nm, as estimated from VC via Equation (4)) and TSS for the pooled NRWQN dataset. With an exponent of 0.75 it is essentially the inverse of the VC vs TSS relationship shown in Figure 5-2, thus confirming this potential.

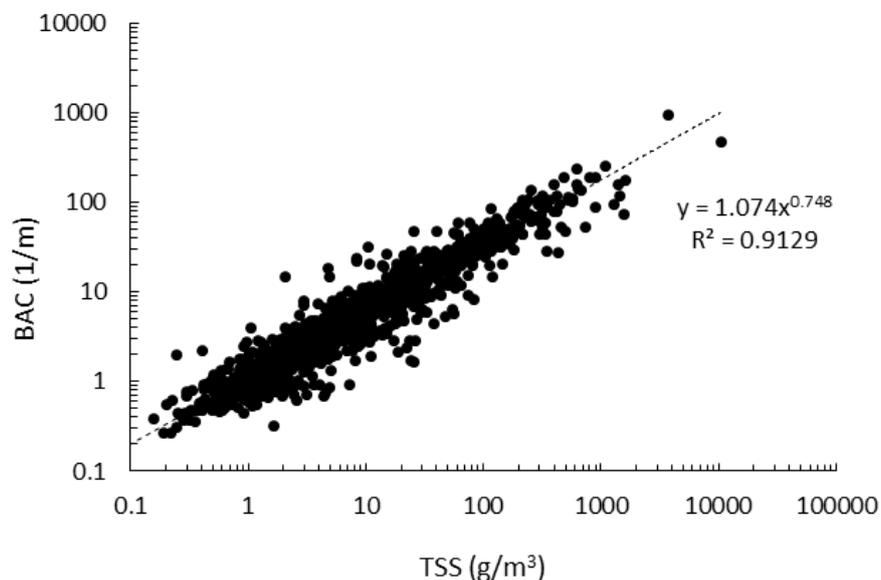


Figure 5-5: Relationship between beam attenuation coefficient (BAC) and TSS at the 77 NRWQN sites. Data are derived from 12 months (2011) of monthly sampling and four years (2012-2015) of selective (flow dependent) sampling during monthly visits.

Table 5-2 lists the mean and standard deviations of the site-median c^* values for the two dataset cases. Both cases show mean values of $0.70 \text{ m}^2/\text{g}$ with standard deviations of $\pm 0.25 \text{ m}^2/\text{g}$.

Figure 5-6 shows the distribution of c^* values at the sub-set of sites with un-impacted flow regimes, both for all data and the site-median values. The site-median c^* values are more tightly distributed (standard deviation of ± 0.25) than are the pooled data (standard deviation of ± 0.49). This reflects

the influence of site-specific factors, and shows that the conversion of TSS to VC or vice-versa can be done more precisely if some empirical paired data is available.

In the following section we explore how the site-median c^* values relate to catchment characteristics and the potential for improving the simple predictive model (based on the mean of site-median c^*).

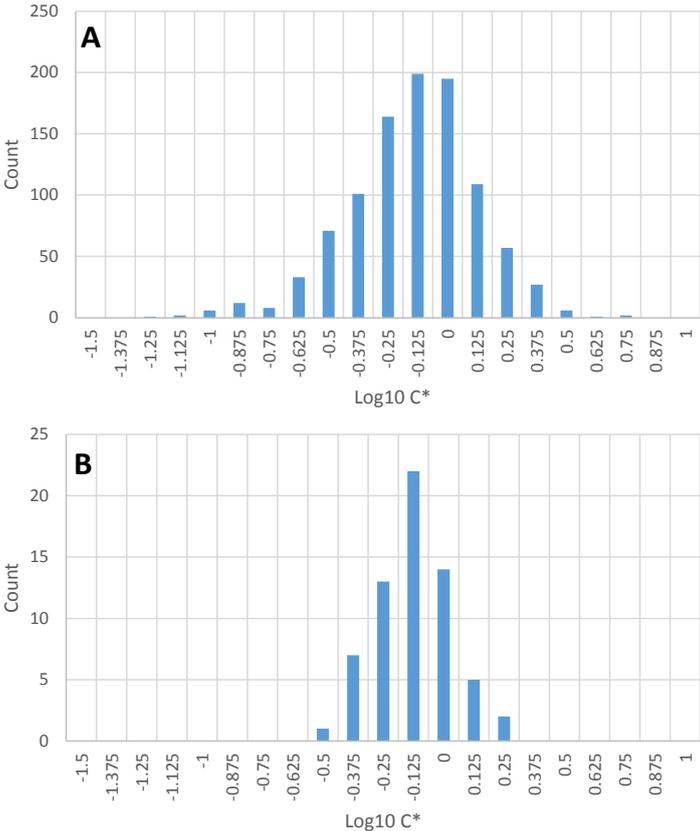


Figure 5-6: Distribution of c^* for (a) all NRWQN data (2011-2015) and for (b) the median values for each site. Note that the distribution is plotted for the \log_{10} of the c^* values.

5.5 Predicting power law coefficients and c^*

In a similar approach to that taken to predict SRC parameters by regression (Section 3), we used the Random Forests method to determine the combination of catchment characteristics that best predict c^* and the intercepts (a) and slopes (b) for the (power law) regression fits for the relationships between VC-TSS, VC-turbidity, and turbidity-TSS.

We applied the RF models to 18 possible explanatory variables available on the REC2 database (Table 5-4). These included measures of catchment slope, rainfall, rock induration and proportion of pasture cover, which is what Elliott et al. (2013) found to influence loads and yields of optical cross-section (See Section 6.1). For a preliminary exploration, we plotted these four variables against c^* (Figure 5-7) but this showed a disappointing lack of pattern - which likely reflects a lack of variables on the REC that adequately predict particle character (size, shape, composition). We also included the percentage of the catchment dominated by silt/clay type soils which, although not included in the

REC2 database, is available from Landcare Research. The 18 explanatory variables used all had pairwise correlations of < 0.7.

The RF models to predict the three pairs of power-law coefficients and c^* were run on the full NRWQN dataset and for the sub-set of 65 sites with un-impacted flow regimes.

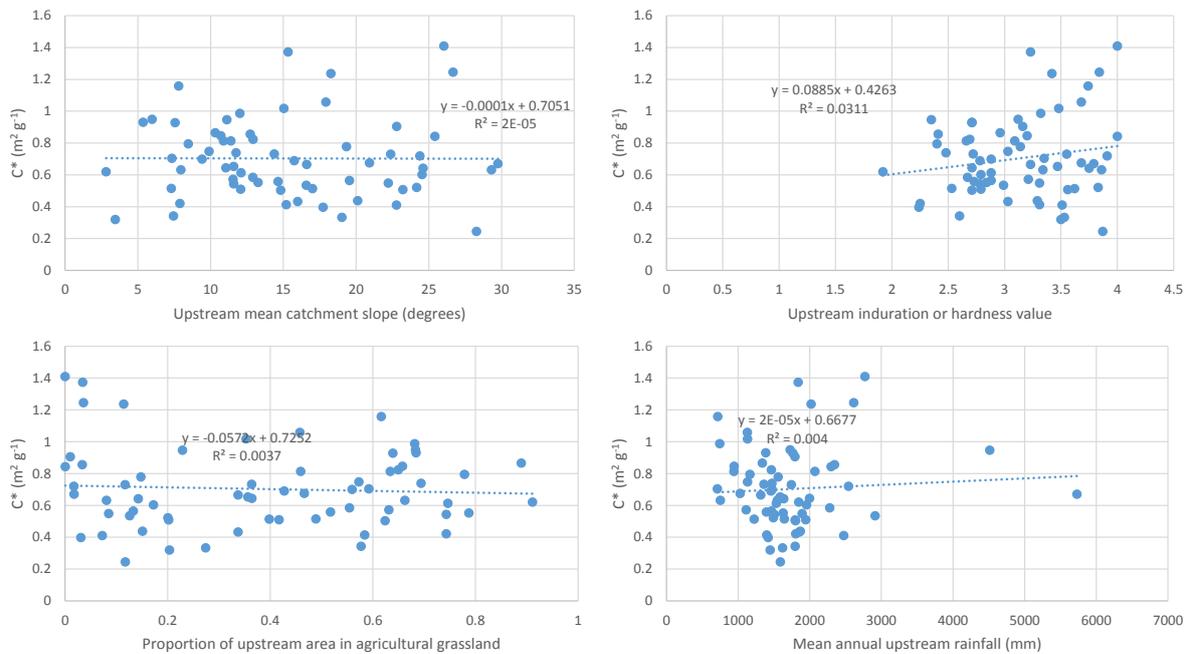


Figure 5-7: Plots of c^* versus four REC2 derived catchment characteristics for the 64 NRWQN sites with un-impacted flow regimes. The characteristics include mean slope, mean rainfall, proportion of area in pasture, and soil induration (or hardness).

Table 5-4: Predictors variables used in Random Forest models for predicting c^* and the intercepts and slopes of the mutual relationships of VC, TSS and turbidity.

Predictor	Description	Units	Range	Notes
WRENZ_upstream_sedtot	Sediment load - from WRENZ	t/y		
REC2_USCalcium	Average calcium concentration of underlying rocks	1= very low to 5 = very high	1-5	
REC2_US_RockPhos	Average phosphorous concentration of underlying rocks	1= very low to 5 = very high	1-5	
REC2_us_rain	Mean annual upstream rain	mm	684 – 6183	
REC2_us_mat	Upstream mean maximum air temperature	°C	3.8 – 13.5	
REC2_us_ind_forest	upstream catchment area	m ²		

Predictor	Description	Units	Range	Notes
REC2_us_hard	Upstream induration or hardness value	Ordinal scale	2 – 5	Highly correlated with REC2_us_particle size
REC2_us_catarea	Upstream mean elevation above sea level of the watershed or basin	m		
REC2_StreamOrder	Strahler stream order		3 – 8	
REC2_seg_elev	Segment mean elevation above sea level	m	15 - 950	
REC2_REC1_SRC_OF_FLW	Source of flow categories from Snelder and Biggs (2002)		See Table 6-2	
REC2_REC1_GEOLOGY	Geology categories from Snelder and Biggs (2002)		See Table 6-2	
REC2_REC1_CLIMATE	Climate categories from Snelder and Biggs (2002)		See Table 6-2	
REC2_MeanFlowCumecs	Mean flow for a segment as in the REC	m ³ /s		
Prop_us_Scrubland	Proportion of upstream catchment in scrubland	Proportion	0.01 – 0.84	Combination of upstream cover of LCDB3 categories: 51, 52, 55, 56, 58
Prop_us_Grassland	Proportion of upstream catchment in grassland	Proportion	0 – 0.85	Combination of upstream cover of LCDB3 categories: 40, 41, 44
Prop_us_Forest	Proportion of upstream catchment in forest	Proportion	0 – 0.93	Combination of upstream cover of LCDB3 categories: 64, 68, 69, 71
us_perc_SiltClay	Percentage of the catchment dominated by silt/clay type soils	Percentage	0 – 98	from the 1:50,000 Fundamental Soils Layers (FSL), maintained by Landcare Research

Model performance

The RF model performance results are listed in Table 5-5. The exclusion or inclusion of the flow impacted sites had little effect on the performance of any of the RF models, with out-of-bag R^2 (OOB R^2 , see Appendix E and Section 13 for explanation) for individual coefficients and c^* being very similar between the two datasets. Model performance did, however, vary between the different coefficients. The three coefficients with OOB $R^2 > 0.20$ were the intercept values for the VC-turbidity and turbidity-TSS relationships as well as for median c^* , which provided the best performance of any model. For all other models the OOB R^2 was close to 0 or negative, indicating that the RF model explained no more, and in some cases less, variance in these coefficients than did their mean values. A similar pattern was observed with Nash Sutcliffe Efficiency (NSE), with only the same three models

indicating that predictions from the RF models were better than simply using the average value (indicated by a NSE of 0). These performance levels are also indicated by plots of observed vs predicted parameters (Figure 5-8 and Figure 5-9).

It is of particular note that the RF analysis provided no regionalisation of the VC vs TSS power-law model (beyond using the national average parameter values given in Table 5-2).

Henceforth we only discuss results from the three models with OOB R² and NSE values >0.20.

Table 5-5: Performance results for Random Forest models predicting the intercepts (a) and slopes (b) for the mutual relationships of VC (visual clarity), TSS, and turbidity as well as site-median c*. Results provided for two datasets: all sites (n=77) and flow impacted sites excluded (n = 64). OOB R² = out-of-bag R²; NSE = Nash Sutcliffe Efficiency; RMSE = Root Mean Square Error. Numbers in brackets in RMSE columns show factorial standard error (equal to 10^{RMSE}).

Regression Coefficients	Excluding flow impacted (n = 64)			Including flow impacted (n = 77)		
	OOB R ²	NSE	RMSE	OOB R ²	NSE	RMSE
TSS-VC intercept	-0.05	-0.14	0.17 (1.48)	-0.14	-0.04	0.17 (1.48)
TSS-VC slope	0.05	0.02	0.12 (1.32)	0.04	0.07	0.13 (1.35)
VC-turbidity intercept	0.31	0.28	0.10 (1.26)	0.36	0.30	0.10 (1.26)
VC-turbidity slope	0.04	0.10	0.09 (1.23)	0.13	0.02	0.09 (1.23)
Turbidity-TSS intercept	0.28	0.21	0.22 (1.66)	0.22	0.28	0.22 (1.66)
Turbidity-TSS slope	0.05	0.07	0.18 (1.51)	0.06	0.05	0.20 (1.58)
Median c*	0.34	0.40	0.19 (1.55)	0.38	0.33	0.20 (1.58)

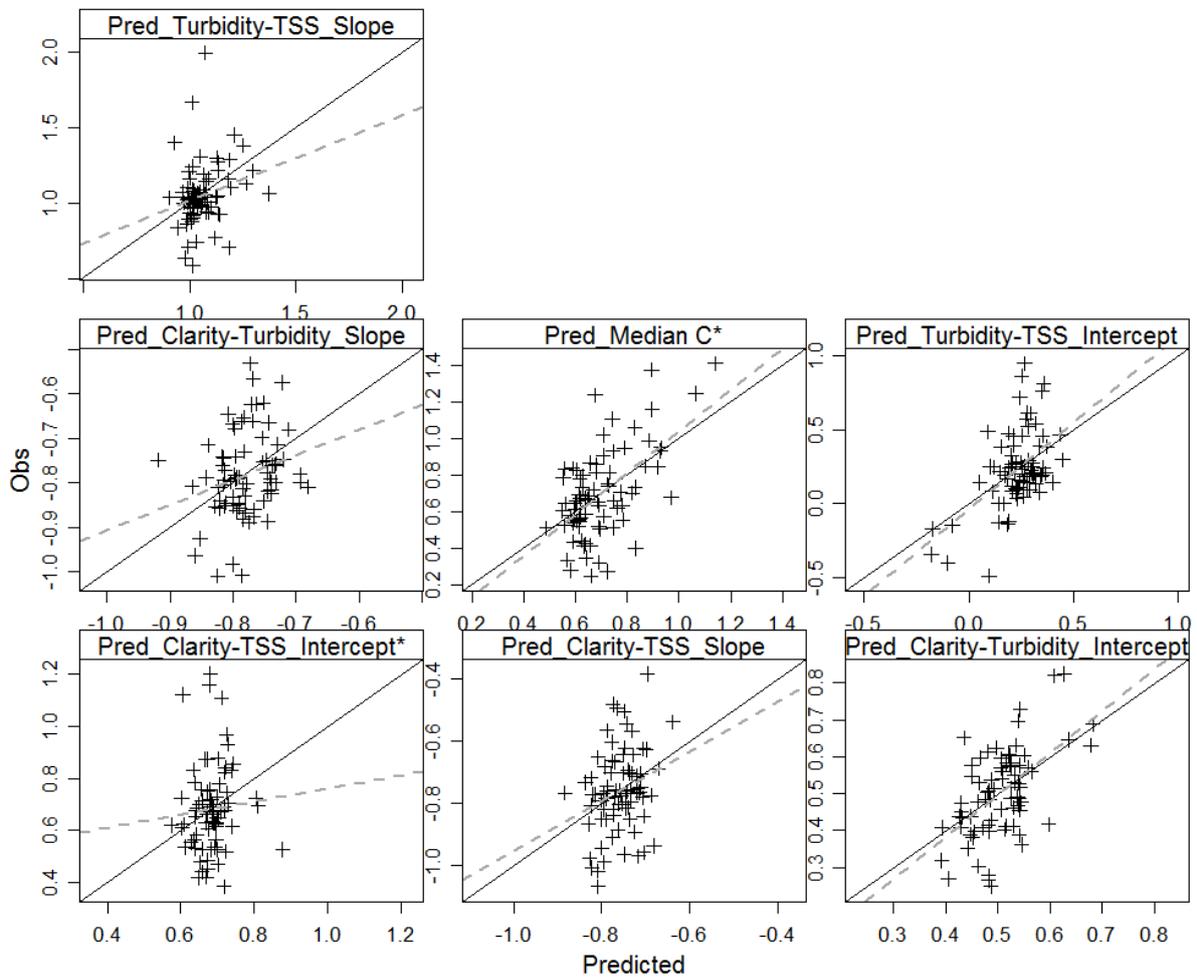


Figure 5-8: Predicted versus observed coefficient values for the full NRWQN site dataset (n = 77) generated using RF models. Solid lines are 1:1 lines, dashed grey lines are fitted regression lines between observed and predicted values. Only the RF models for Turbidity-TSS_intercept, VC-Turbidity_Intercept and Median c* performed reasonably better than just using the average values.

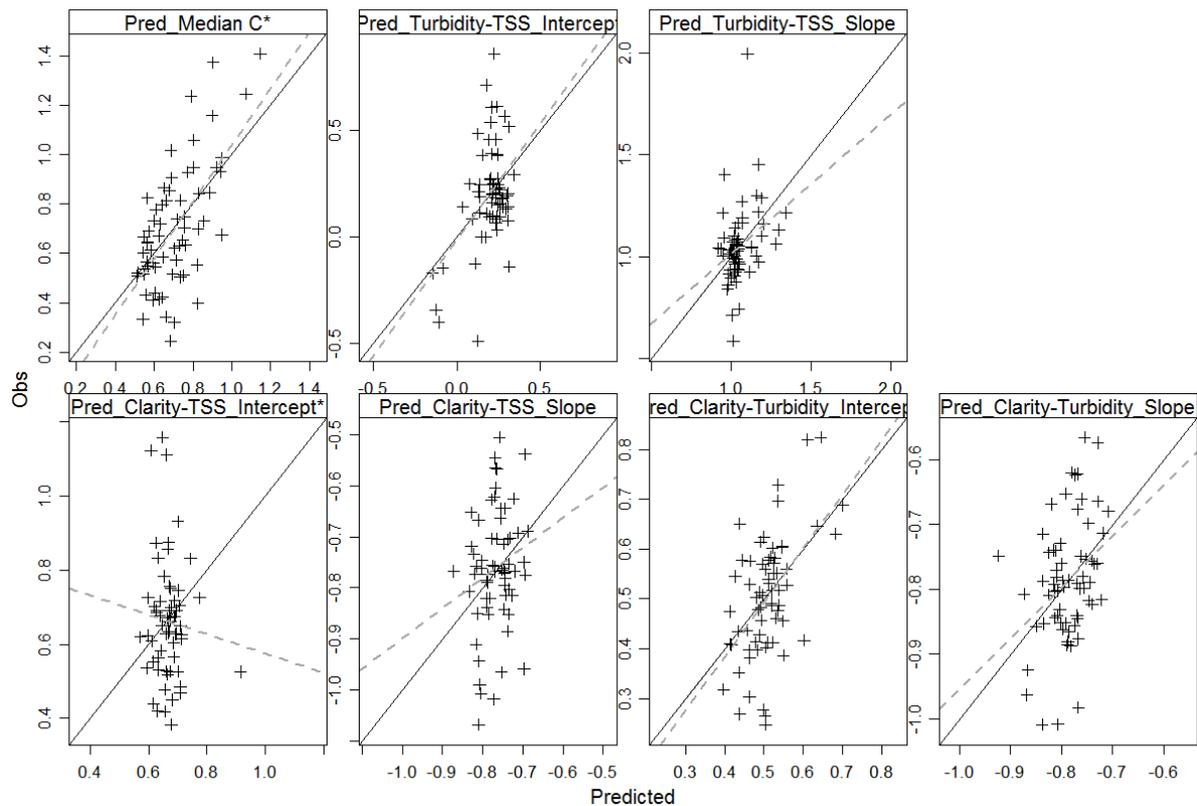


Figure 5-9: Predicted versus observed coefficient values for the non-flow impacted dataset (n = 64) generated using RF models. Solid lines are 1:1, dashed grey lines are fitted regression lines between observed and predicted values. Only the RF models for Turbidity-TSS_intercept, Clarity-Turbidity_Intercept and Median c* performed reasonably better than just using the average values.

VC-turbidity intercept model

The most important predictive variable for the VC-turbidity intercept model, for both the full and the non-flow impacted datasets, was catchment mean maximum air temperature (REC2_us_mat, Figure 5-10; Figure 5-11), with the intercepts being higher as this decreased (Figure 5-12; Figure 5-13). Catchment dominant geology class (REC2_REC1_GEOLOGY) was the second and third most important predictor for the non-flow impacted and full datasets, respectively (Figure 5-10; Figure 5-11), with higher intercepts in sites with catchments falling in the Hard-Sedimentary category (Figure 5-12; Figure 5-13). Sediment load (WRENZ_upstream_sedt0t) was the third most important predictor for the VC-turbidity intercept model run using the non-flow impacted dataset (Figure 5-11), with lower intercepts in sites with higher upstream sediment inputs (Figure 5-13). Mean flow (REC2_MeanFlowCumecs) was the second most important predictor for the full data set (Figure 5-10), with intercepts lower in sites with greater flows (Figure 5-12). We infer that the link with temperature influences the degree of chemical weathering and relative abundance of clay grains (with less clay particles the VC is higher), while catchment lithology should also influence clay content.

TSS-turbidity intercept model

Irrespective of the dataset used (Figure 5-10; Figure 5-11), the three most important predictors for the turbidity-TSS model intercept were catchment air temperature (REC2_us_mat), catchment average slope (REC2_us_slope), and mean flow (REC2_MeanFlowCumecs). The relationships between the intercept and these predictors were negative, negative, and positive, respectively (Figure 5-12; Figure 5-13). These patterns may reflect a trend for more clay minerals in the sediment load as catchment size, flow, and mean temperature all increase.

Median c^* model

The three most important predictors for the median c^* models (Figure 5-10; Figure 5-11) were the extent of silt/clay-type soils (us_perc_Siltclay), mean flow (REC2_MeanFlowCumecs), and catchment size (REC2_us_catarea). These showed positive, negative, and negative relationships with c^* , respectively (Figure 5-12; Figure 5-13). The first is physically meaningful, since more silt/clay should produce higher beam attenuation for a given concentration of sediment.

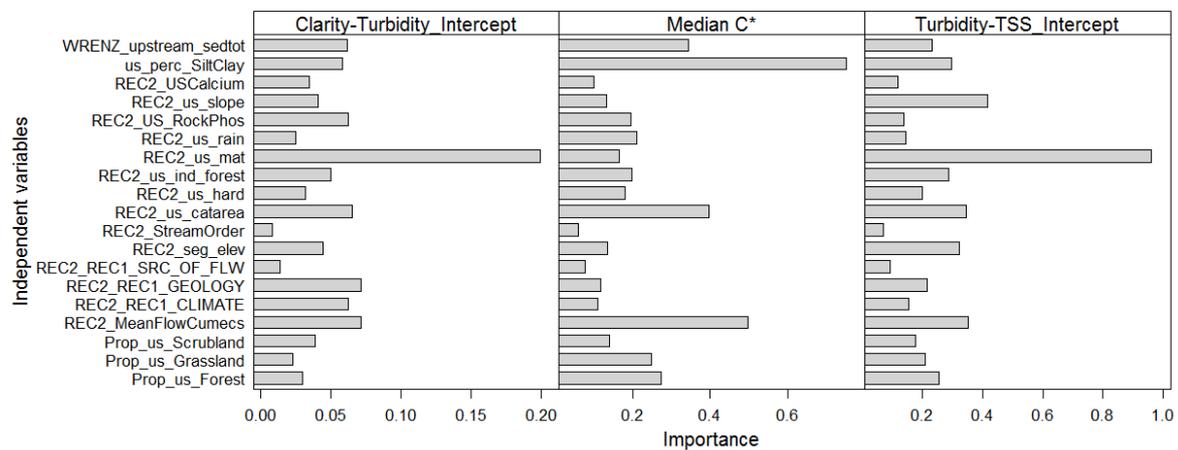


Figure 5-10: The importance of individual predictors for the three RF models with OOB $R^2 > 0.2$ for the full dataset ($n = 77$). Importance is measured as increasing node purity from splitting on the selected variable averaged over all trees in the Random Forest.

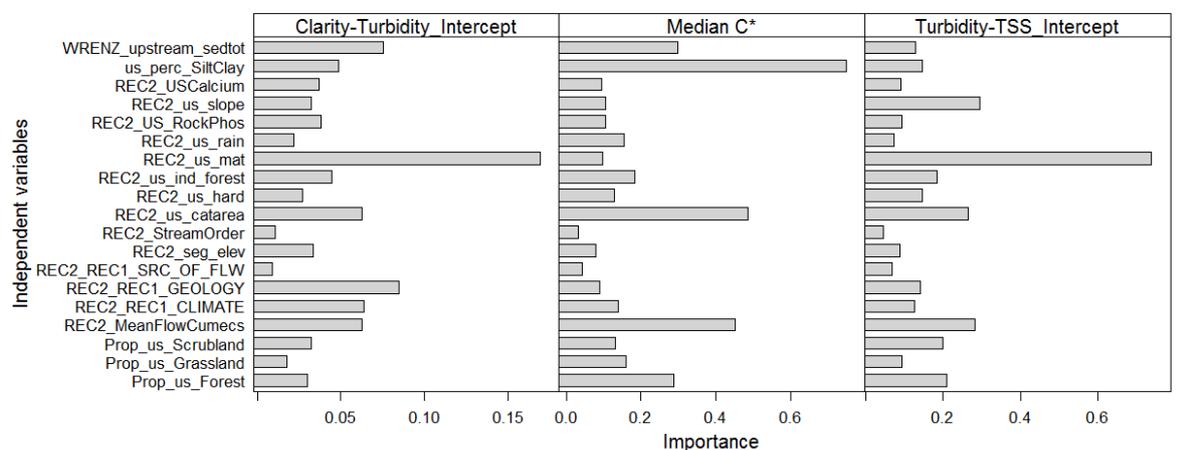


Figure 5-11: The importance of individual predictors for the three RF models with OOB $R^2 > 0.2$ for the non-flow impacted dataset ($n = 64$). Importance is measured as increasing node purity from splitting on the selected variable averaged over all trees in the Random Forest.

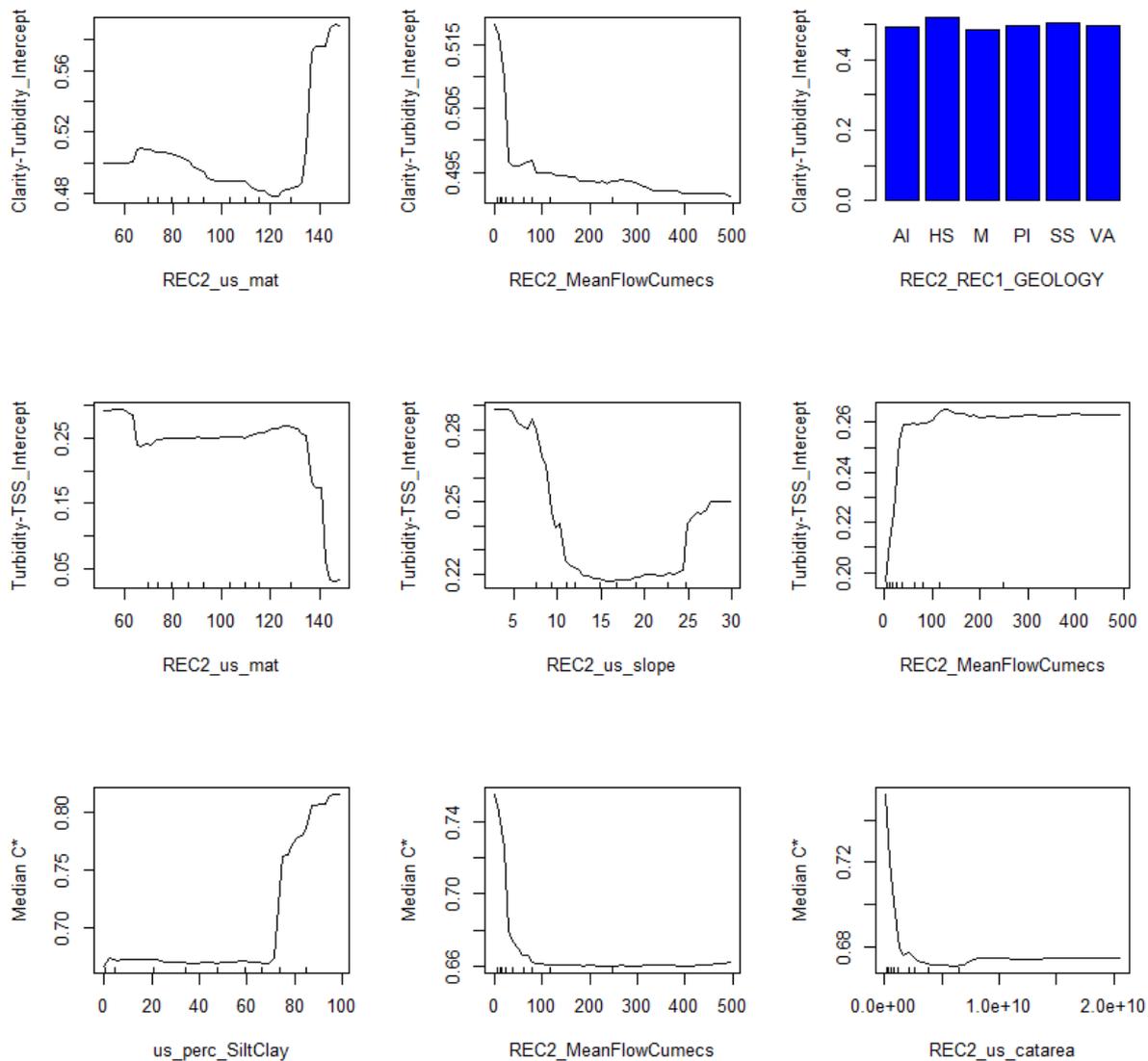


Figure 5-12: Univariate partial dependence plots showing the shape of relationships between the three most important predictors for the 3 RF models run on the full data set. The spread of predictor values across the gradient is shown by the “rug” on the x-axis (each tick is a decile).

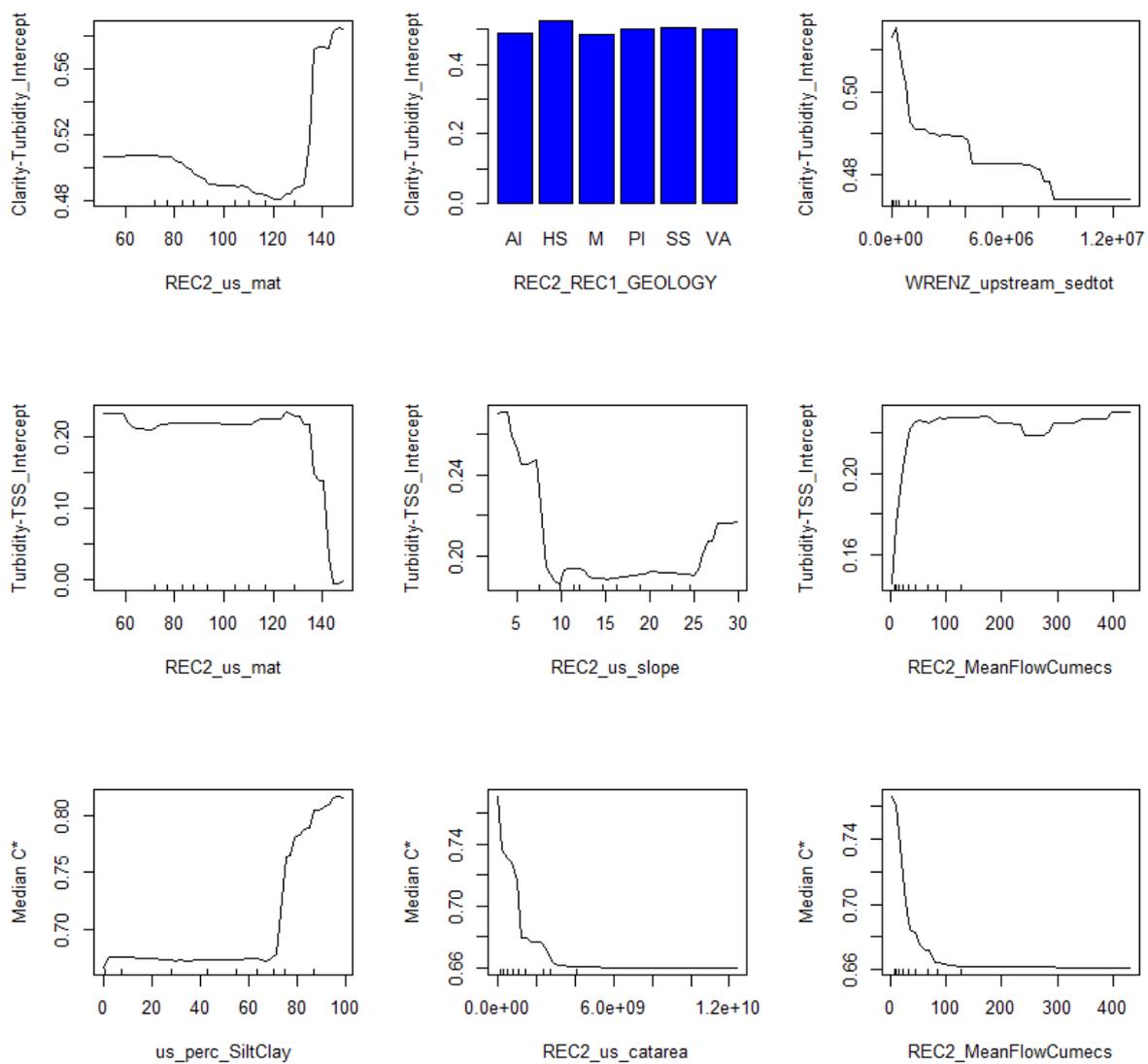


Figure 5-13: Univariate partial dependence plots showing the shape of relationships between the three most important predictors for the 3 RF models run on the dataset excluding flow impacted sites (n = 64). The spread of predictor values across the gradient is shown by the “rug” on the x-axis (each tick is a decile).

5.6 Discussion

It has proven difficult to find a basis for 'regionalising' the relationships between VC and TSS (and turbidity) because the REC does not really 'capture' the particle characteristics that influence optical cross-section. This could change in future, if and when variables are incorporated into the REC that better predict clay content or soil texture.

Nonetheless, while the Random Forest regression models only explain a modest amount of the variance in the observed VC data, the prediction accuracy appears moderate (indeed, this is the case even where the RF models could provide no improvement over the simple national-average values for the VC-TSS relationship).

The best performing model was that predicting site-median c^* (explaining 34% of variance). Since this is a single parameter model, it is perhaps the best to begin with.

To improve on this would require establishing relationships empirically for each river site by measuring both SPM and VC simultaneously on at least a subset of visits/water samples. Because c^* within rivers is appreciably less variable than between rivers, the inter-conversion of TSS and VC can be made with relatively little data. So a relatively brief campaign of paired measurements (e.g., a 12 month add-on to routine monthly SoE data) should suffice to characterise the average c^* value at a site. Typically, c^* may be expected to decrease with increasing TSS as increasingly coarse SPM is thrown into suspension with increasing flow. So, typically, exponent (b) values mostly smaller than -1 (averaging around -0.7) can be expected when fitting power functions to VC versus TSS.

An important finding from this analysis is the low quality of some VC measurements in most of the data available to us. Most of the lower quality data is from sites on small streams such as those around the Whatawhata Research Station. It is important when measuring VC in smaller rivers to take great care to avoid shadowing across the path of sight under clear sun conditions and (reflection from) river banks from affecting the visibility of the black disc. Furthermore, it is important, particularly under clear water (low flow) conditions to avoid measuring VC in disturbance plumes caused by wading. The protocols for measuring VC using the black disc method as originally introduced by Davies-Colley (1988) are currently being formalised in the National Environmental Monitoring Standard (NEMS) for discrete water sampling and testing. It will be important to see these protocols taken up in further data collection for NOF-sediment application (as discussed in Section 10.2.5).

6 Suspended sediment particle size distributions and relationship to sediment load

6.1 Introduction

In theory, suspended sediment (SS) particle size distribution (PSD) exerts strong control on the relationships between SSC and VC and turbidity, since particle diameter determines the “optical cross-section” (Section 5; Davies-Colley et al. 2015). Moreover, SS PSD should also influence deposited sediment (DS), since particle diameter is a primary control on sediment settling and entrainment.

In the context of this study, PSD is important for two reasons. Firstly, understanding the factors that control its spatial variation is important to inform on any spatial (i.e., regional) variation in the relationships between the ESVs. Secondly, temporal change in the PSD presents a potential complication to the proposed sediment NOF frameworks that link changes in ESVs to changes in catchment sediment load (Section 8) – since if the PSD changes as the load changes, then so too will the ESV relationships with load (and also the ESV inter-relationships) change. This could occur, for example, if a large source of clay-rich sediment (such as from stock-trampled riparian margins) was stabilised in a catchment where otherwise erosion processes produced dominantly silt-grade sediment.

In this section, we use available directly-sampled PSD data to assess the factors controlling spatial variation in PSD, including assessing if there is a correlation between PSD and load. We also examine evidence that changes in catchment sediment load cause the PSD to change. In this latter case, no PSD data were available so we investigate changes in the relationships between SSC and turbidity at sites where the sediment load is known to have changed over time.

6.2 Spatial variation in SS particle size and its relationship with sediment load and other catchment characteristics

6.2.1 Data

SS particle size data were available for 59 sites, located in both the North and South islands, in rivers from Strahler stream order 3 to 8, and in a range of catchment types as indicated by REC topographic source of flow categories, climates as indicated by REC climate categories, and catchment geology types as indicated by REC catchment geology categories (Figure 6-1; Table 6-1).

These data were all sourced from NIWA’s database – no more were found in the “data seek” from regional authorities. The NIWA data were largely derived from analysis of composited samples collected by multi-vertical, depth-integrated SS sampling in association with flood gaugings, although some samples were bank-side point-samples collected by auto-sampler. The size grading was determined mainly with manual methods, with settling used for the mud fractions (< 63 µm) and wet-sieving for the sand fractions (> 63 µm), although laser-diffraction or laser time-of-transit instruments were used for some sites. Results are generally available as the cumulative %’s finer by weight at 4, 8, 16, 32, 63, 125, 250, 500, 1000, and 2000 µm.

The number of samples per site varied from one to 45, averaging 4. We represented the PSD at each site by the average of all samples. Sites with multiple samples often showed considerable variation in PSD amongst samples, thus there is significant sampling error for sites with only one to a few

samples. An example of temporal variation in PSD through floods from the Mararoa River is detailed in Appendix G.

For this study, we explored three PSD metrics:

- the % finer than 16 μm (fine silt and clay fractions, labelled “Ave%<16”)
- the % finer than 63 μm (silt and clay fractions, collectively termed mud, labelled “Ave%<63”)
- the median size (transformed to its log value, labelled “logAveD50”).

We chose these because the optical signature of SS is controlled largely by the mud fraction, and particularly the fine silt to clay fraction, while the median size is typically used as a representative “central” statistic when considering sediment deposition and entrainment.

A correlation analysis showed that logAveD50 and Ave%<63 were highly but inversely correlated ($R^2 = 0.91$, Figure 6-2), thus we left the median size out of further analysis. The regression equation is:

$$D50 = 726 e^{-0.047\text{Ave\%<63}} \quad (8)$$

where D50 is in μm (note that this relationship will asymptote to 100% for median sizes finer than 6 μm). This equation can be used to predict D50 given an estimate of the % mud.

Ave%<16 and Ave%<63 were also positively correlated ($R^2 = 0.90$), however, we retained both in subsequent analysis since it remains unclear at this stage of investigation which might be of greater relevance to the optical signature of suspended sediment.

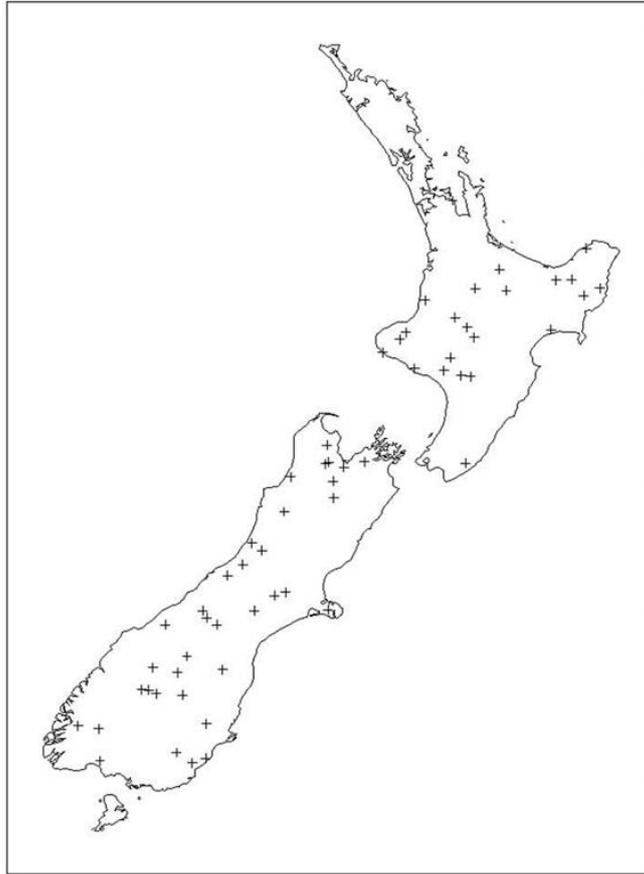


Figure 6-1: Locations of sites with SS PSD data. n=59.

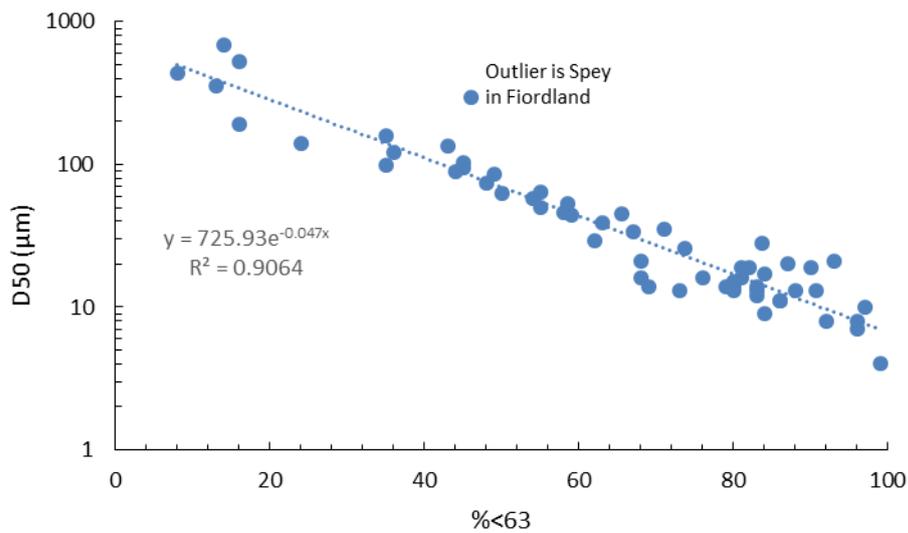


Figure 6-2: Relationship between median grainsize (D50) and % mud (Ave%<63) for average PSDs from 59 sites. The outlier is from the Spey River in Fiordland. It is an outlier because the suspended sand in the Spey is relatively coarse-grade sand (which increases the median size).

Table 6-1: Distribution of sites across climate, topographic source-of-flow, geological, and Strahler stream order classifications as extracted from the REC. Refer Snelder and Biggs (2002) for full descriptions of codes for REC groups.

Description	Number of sites
Strahler Stream Order	
3	5
4	10
5	19
6	18
7	6
8	1
Topographic source of flow	
Glacial mountain	6
Mountain	8
Hill	25
Low elevation	18
Lake-fed	2
Climate	
Cool-dry	10
Cool-wet	23
Cool-extremely wet	16
Warm-dry	0
Warm-wet	8
Warm extremely wet	2
Geology	
Alluvium	4
Hard-sedimentary	28
Plutonic	5
Soft-sedimentary	10
Volcanic-basic	2
Volcanic-acid	10

6.2.2 Analysis approach

We used the Random Forest (RF) regression approach (Breiman 2001) to generate separate models of the Ave%<16 and Ave%<63 particle size classes. Details of the RF methodology, including measures of model-fitting performance and accuracy, are provided in Appendix E.

Predictor variables

The choice of predictor variable was informed by a brief literature review which showed that SS PSD tends largely to be supply-dependent rather than controlled by flow hydraulics (i.e., varying as a direct function of discharge during runoff events; Walling et al. 2000). Thus important controls are

catchment lithology and soil character. Soil character, in turn, is influenced by geomorphic setting (e.g., steepland vs lowland), climate (which influences the balance of physical and chemical weathering and so the proportion of clay grains), and land-cover (e.g., Walling and Woodward 2000; Pavanelli and Selli 2013). Also, samples collected downstream from lakes/reservoirs should be in the fine silt-clay range.

The predictor variables chosen (Table 6-2) included catchment lithology, climate, source-of-flow, stream order, elevation, mean temperature, sediment load, land cover, and soil properties. Sediment load was represented by two variables: the mean annual sediment yield (calculated by combining site SRCs and flow records) and the discharge weighted mean SSC (C_{mean} , based on the mean annual sediment load divided by the mean annual water discharge as determined from flow records). Other catchment characteristics were extracted from the REC2 database, indexing sites by their NZsegment number²⁵. The LENZ soil hardness and particle size ordinal classes (Leathwick et al. 2002) were found to be highly correlated, so the latter was removed from the analysis.

Also, an alternative classification of catchment dominant lithology was based on assessment of the dominant rock-type extracted from the NZ Land Resource Inventory (NZLRI) regrouped into a smaller number of classes (based on constituent mineral grain size and also considering weathering status). The sites sampled for PSD were assigned to one of 13 lithology groups (as listed in Table 6-3, and as reported by Hicks et al. 2004). The averaged PSDs for each lithology class showed a sensible separation (Figure 6-3). For example: the samples from glaciated catchments, catchments with abundant loess drapes, and catchments in soft siltstone and mudstone (e.g., North Island “papa”) were dominated by silt and clay; the granitic/gneissic and volcanoclastic (i.e., ash, tephra) catchments had high proportions of sand and less mud; North Island greywacke catchments had higher mud proportions than the South Island equivalents (indicating more intense weathering in the north).

Table 6-2: Predictors variables used in Random Forest models for predicting Ave%<16 and Ave%<63.

Predictor	Description	Units	Range	Notes
Alternative lithology group	Informal lithology classification based on dominant catchment rock-type.		See Table 6-3	Informed by NZLRI.
C_{mean}	Mean annual sediment load / mean river flow	g/m^3	15 – 11828	Determined for this study
Sediment yield	Mean annual sediment yield	$t/km^2/y$	22.7 – 13228	
REC2_us_hard	Upstream induration or hardness value	Ordinal scale	2 – 5	Highly correlated with REC2_us_particle size
REC2_us_mat	Upstream mean maximum air temperature	$^{\circ}C$	3.8 – 13.5	
REC2_seg_elev	Segment mean elevation above sea level	m	15 - 950	

²⁵ Correct segment number assignment was manually checked for all 59 sites.

Predictor	Description	Units	Range	Notes
REC2_StreamOrder	Strahler stream order		3 – 8	
REC2_REC1_SRC_OF_FLW	Source of flow categories from Snelder and Biggs (2002)		See Table 6-1	
REC2_REC1_CLIMATE	Climate categories from Snelder and Biggs (2002)		See Table 6-1	
REC2_REC1_GEOLOGY	Geology categories from Snelder and Biggs (2002)		See Table 6-1	
REC2_us_rain	Mean annual upstream rain	mm	684 – 6183	
Prop_us_Grassland	Proportion of upstream catchment in grassland	proportion	0 – 0.85	Combination of upstream cover of LCDB3 categories: 40, 41, 44
Prop_us_Scrubland	Proportion of upstream catchment in scrubland	proportion	0.01 – 0.84	Combination of upstream cover of LCDB3 categories: 51, 52, 55, 56, 58
Prop_us_Forest	Proportion of upstream catchment in forest	proportion	0 – 0.93	Combination of upstream cover of LCDB3 categories: 64, 68, 69, 71

Table 6-3: Distribution of sites across alternative lithology grouping.

Label	Dominant lithology	Number of sites
N1	Weaker volcanoclastic	1
N2	Stronger volcanoclastic	3
N5	Nth Island mudstone, siltstone & sandstone	9
N6	Nth Island greywacke & argillite	4
S1	Fiordland gneiss	1
S10	Marble	1
S2	Sth Island foothills greywacke & argillite	7
S3	Sth Island western schist	6
S4	Otago schist	9
S5	Sth Island axial alps greywacke & argillite	8
S6	Glacial	1
S7	Loess	1
S9	Sth Island metasediments & volcanics	8

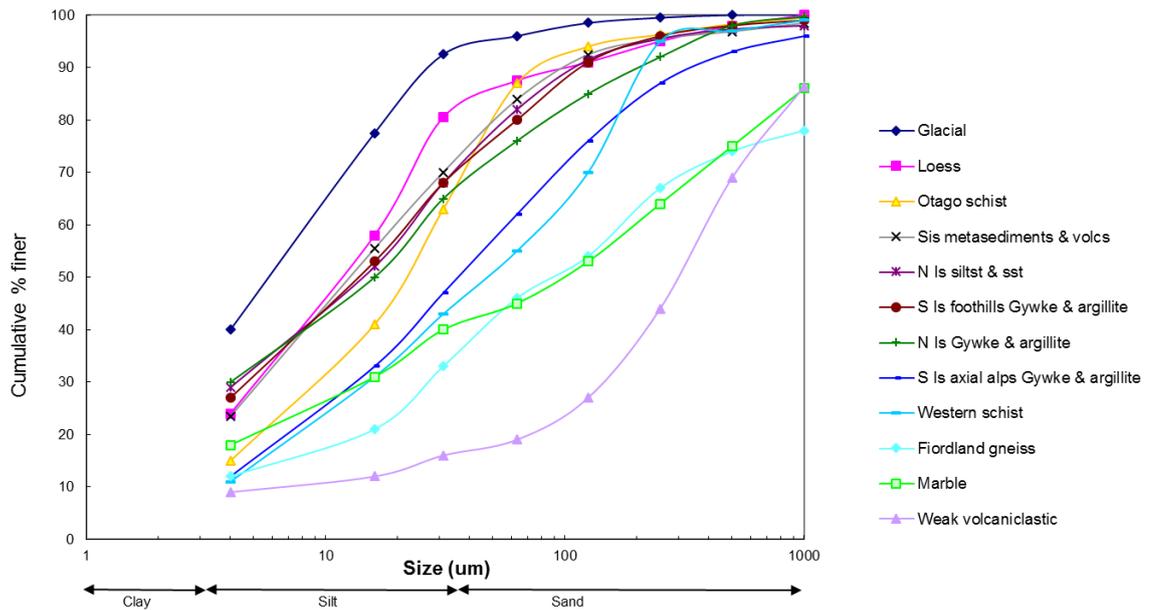


Figure 6-3: Averaged PSDs by lithology groups. Number of samples averaged per group shown in Table 6-3. After Hicks et al. (2004).

RF model-fitting results

Performance results from the RF model-fitting for the Ave%<16 and Ave%<63 response variables are summarised in Table 6-4. OOB R^2 values for Ave%<16 and Ave%<63 were 0.34 and 0.39, respectively. Positive NSE and relatively low RSR values indicate that the models fitted were able to distinguish patterns in percentages of the two particle size categories, even at new sites. There was reasonable correspondence between cross-validation (CV) predictions and observed values for each of these particle size variables (Figure 6-4).

Table 6-4: Model performance metrics for Random Forest models of Ave%<16 and Ave%<63. OOB R^2 is the out-of-bag R^2 and provides an estimate of the predictive performance of the model for new cases. NSE: Nash-Sutcliffe efficiency, RMSE: root mean square error, RSR: ratio of the root mean square error to the standard deviation of the observed data.

RF model	OOB R^2	NSE	RSR	RMSE (%)
Ave%<16	0.34	0.21	0.81	16.3
Ave%<63	0.39	0.24	0.78	18.9

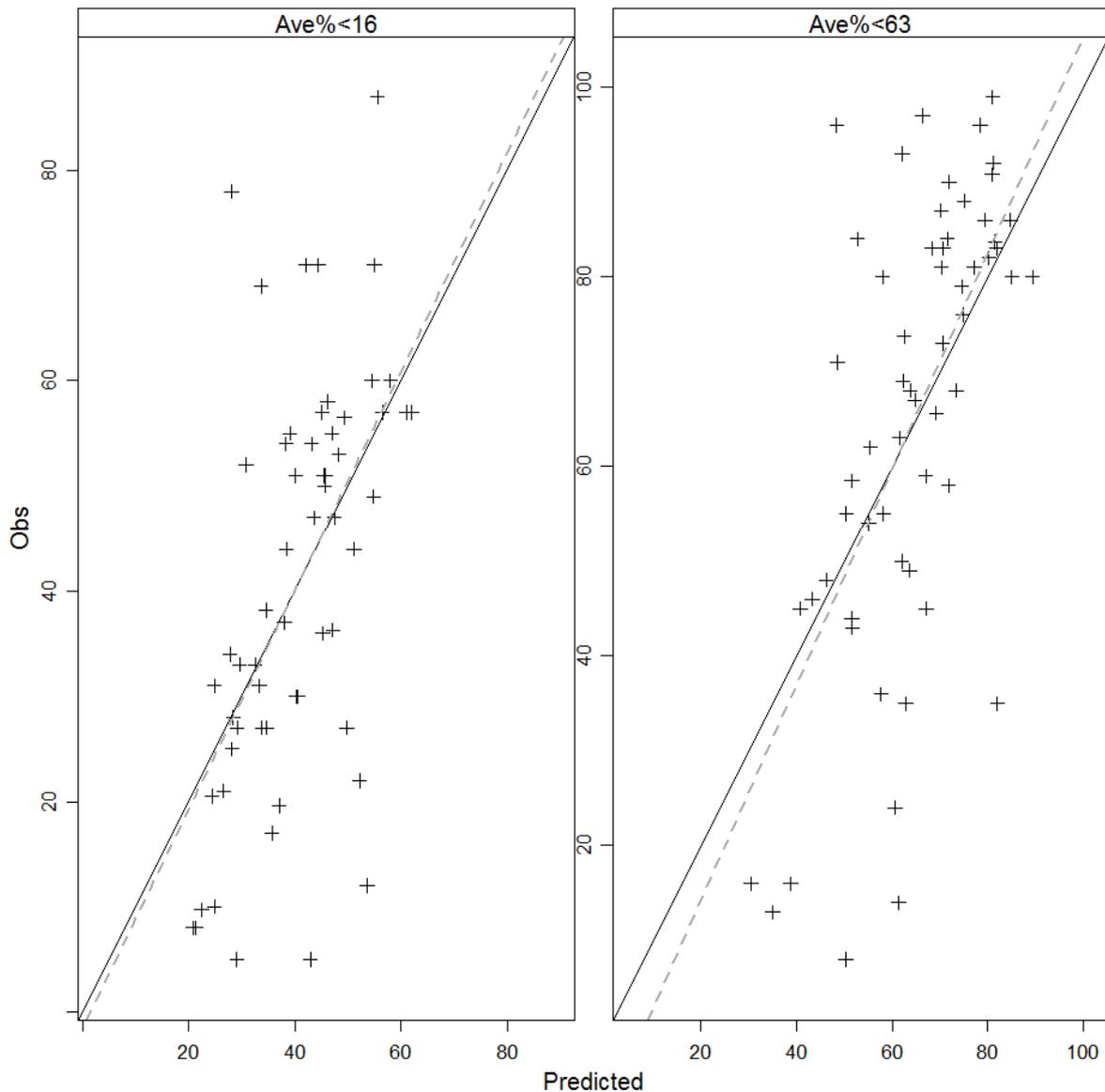


Figure 6-4: Observed against hold-one-out cross-validation (CV) predicted values for the Ave%<16 and Ave%<63 PSD variables. Black line represents 1:1. Grey dashed line is linear regression.

The most important predictors were very similar for both particle size variables (Figure 6-5). Our alternative geology grouping was the most important predictor for both size variables, with the proportion of upstream catchment in scrub or grassland being the next most important. For Ave%<16, segment elevation and upstream temperature were the next most-important predictors, while for Ave%<63 the next most important were C_{mean} and upstream rainfall. C_{mean} was a stronger predictor than was catchment sediment yield²⁶.

Partial dependence plots (Figure 6-6) showed that both particle size classes were proportionally higher in the North Island mud/silt/sandstone and greywacke/argillite groups and in the South Island greywacke/argillite, glacial, and loess lithology groups (N5, N6, S5, S6 and S7), which aligns with the

²⁶ The essential difference between C_{mean} (g/m^3) and sediment yield ($\text{t}/\text{km}^2/\text{y}$) is that the former is the mean annual sediment load divided by annual water discharge while the latter is the load divided by catchment area. Because C_{mean} is discharge-compensated, it arguably provides a better index of sediment availability than does sediment yield.

data summarised in Figure 6-3. These plots also showed both Ave%<16 and Ave%<63 were higher at sites with less upstream scrubland and more grassland in their catchment. Ave%<16 was also higher at lower elevations (suggesting the effect of more chemical weathering compared with physical weathering at lower elevations). Ave%<63 was lower when C_{mean} is lower, suggesting that catchments supplying more sediment with a given amount of runoff also tend to generate higher proportions of mud (which is sensible in regard to expected increased erodibility of finer-grained lithologies).

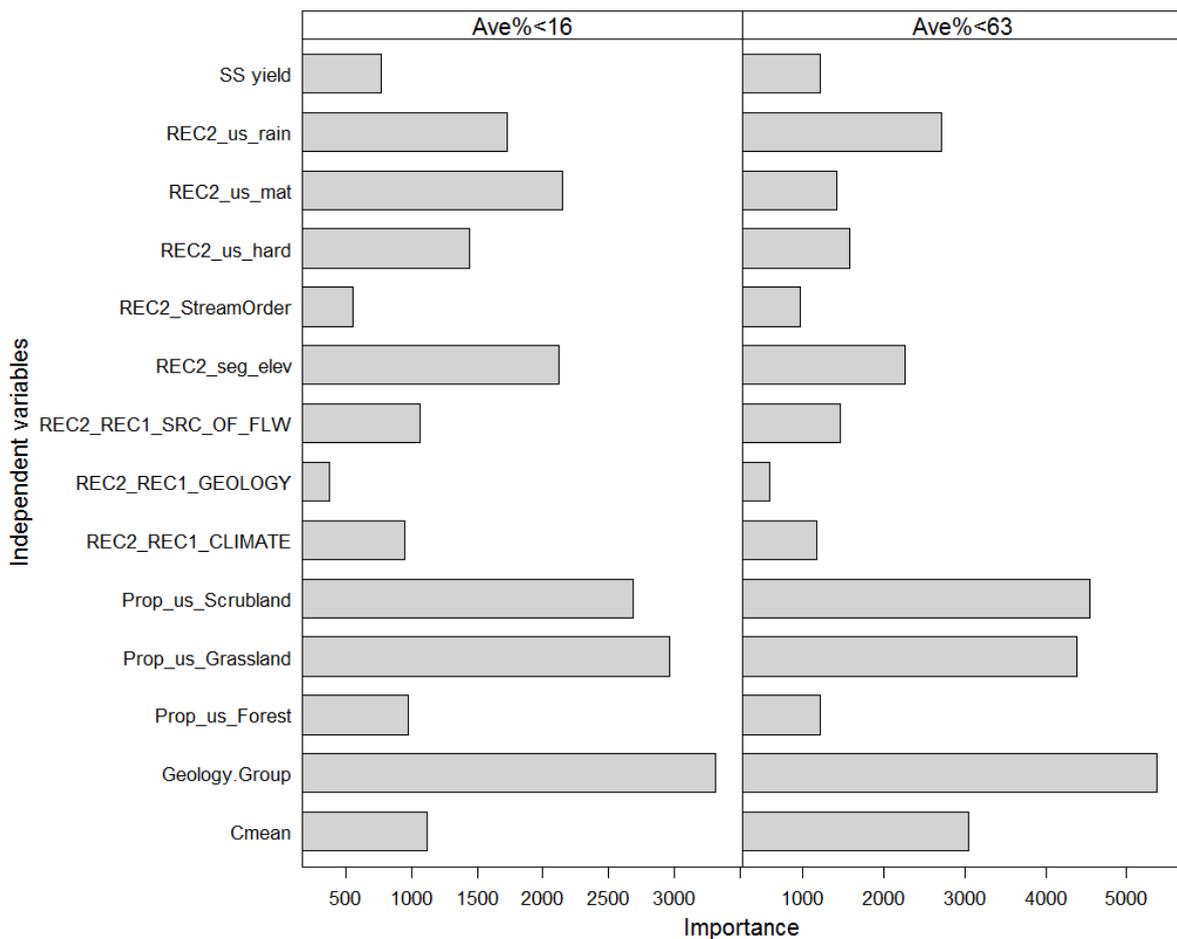


Figure 6-5: The importance of individual predictors to each of the two particle size RF models. Importance is measured as increasing node purity from splitting on the selected variable averaged over all trees in the Random Forest.

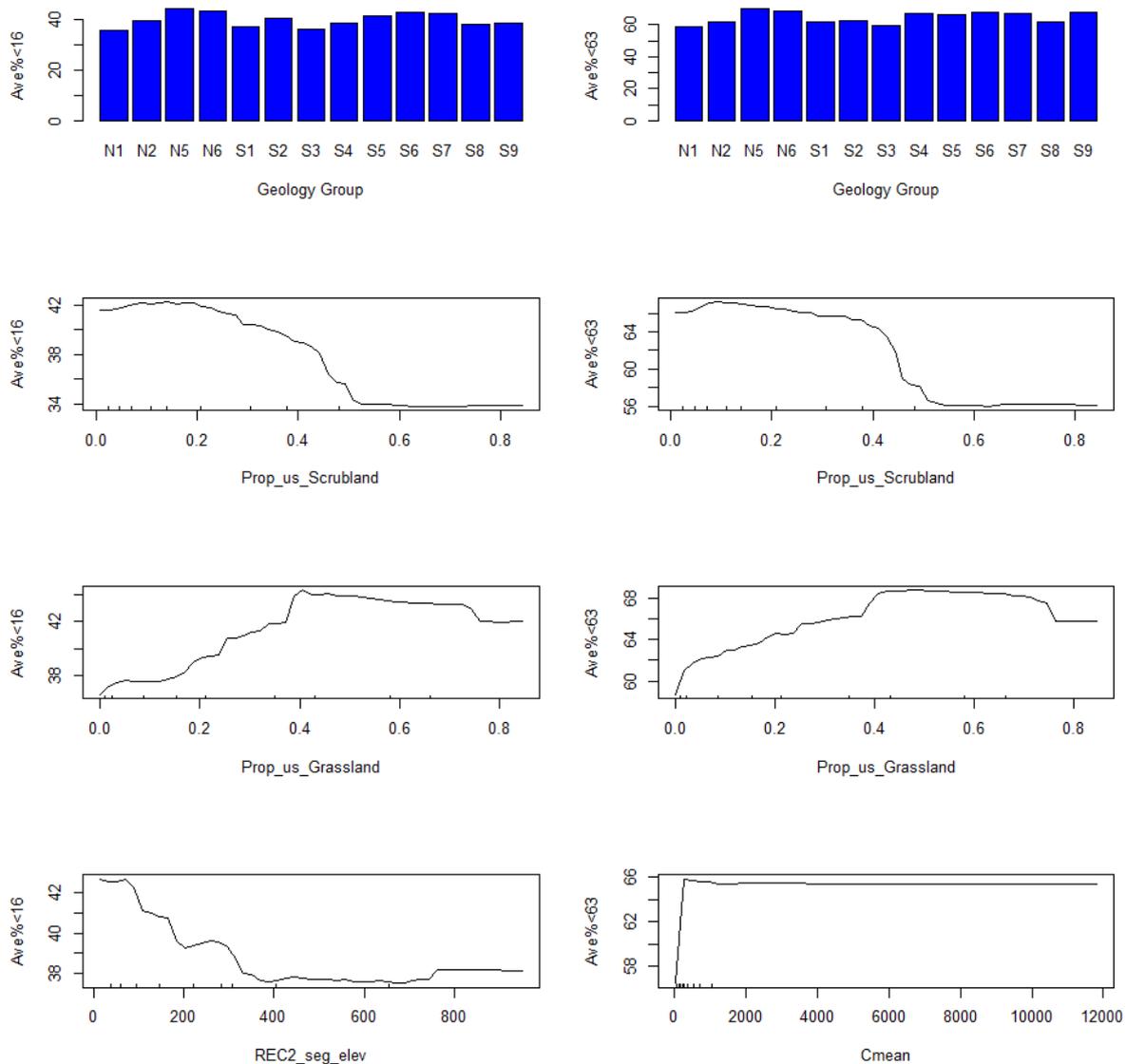


Figure 6-6: Univariate partial dependence plots showing the shape of relationships between the four most important predictors for the Random Forest models for Ave%<16 and Ave%<63. The spread of predictor values across the gradient is shown by the “rug” on the x-axis (each tick is a decile).

6.2.3 Summary from Random Forest modelling of SS particle size

RF models were able to be developed to predict both the % of SS load finer than 16 μm (i.e., fine silt and clay) and the % finer than 63 μm (i.e., the % mud) to an accuracy of $\pm 16\text{-}18\%$. Moreover, the % mud serves as a reasonable proxy for the SS median particle size. The main factor influencing the regional spread of % mud was catchment lithology, but with some control also exerted by land-cover, sediment supply, elevation and rainfall. After lithology and land-cover, the % of fine silt and clay was also influenced by temperature, elevation, and rainfall.

6.3 Change in SS particle size associated with changes in catchment sediment load

6.3.1 Evidence from changes in specific turbidity

Here we examine if changes in catchment sediment load (through changes in sediment supply) cause the PSD to change. No PSD data were available from sites where sediment load has been observed to change, however, we pursue this question indirectly by looking for change in the relationships between SSC and turbidity at sites where the sediment load is known to have changed over time. This is reasonable given that particle size is a strong controller of specific turbidity, which is the ratio of turbidity to SSC (as demonstrated clearly by Foster et al. 1992, Figure 6-7).

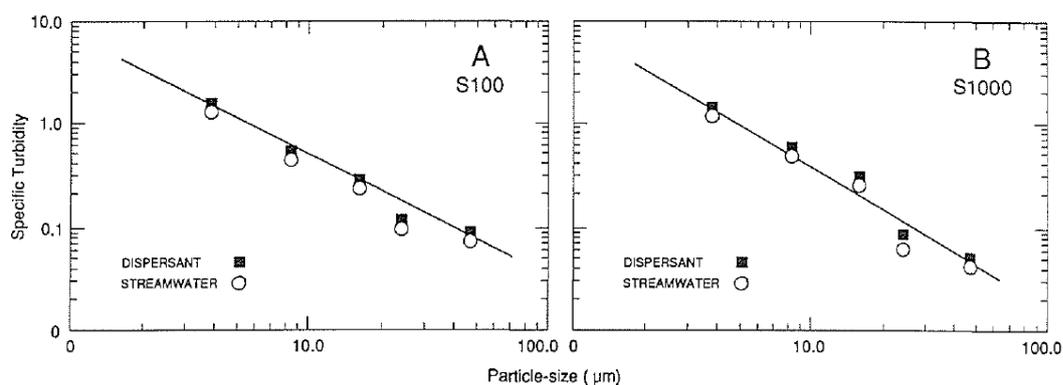


Figure 6-7: Specific turbidity vs particle size for two turbidity sensor types, from Foster et al. (1992).

Specific turbidity is the ratio of turbidity to SSC. The two turbidity instruments were the Partech S100 and S1000. Note more factor-of-20 range in specific turbidity across 4 to 45 µm particle size range.

Data collected from sites in the Motueka Catchment are useful for this purpose. As detailed in Section 4.2.1, a large rainstorm during Easter 2005 activated erosion features that caused a several-year-long phase of increased sediment loading that was detected by turbidity sensors and in auto-samples collected to calibrate the turbidity sensors to SSC. Figure 6-8 shows trends of the SSC vs turbidity relations at three sites before, immediately after, and in the years following this event. The SSC values are from auto-samples, while the turbidity values were all measured from the auto-samples in the laboratory by the same turbidity instrument, which was regularly calibrated with formazin standard solutions. The trend lines pass through zero and so their slopes show the specific turbidity (listed in Table 6-5).

All three sites (Motueka at Gorge, Motupiko at Christies, and Motueka at Woodman's Bend) showed a similar temporal pattern of specific turbidity change: an increase in 2005 immediately after the Easter storm (by 37%, 35%, and 9%, respectively), indicating finer SS grainsize; then lower values (indicating coarser sediment) afterward (Table 6-5)²⁷.

Thus it is clear from this Motueka example that changes in sediment load can alter the size grading of the suspended load and this is manifest as shifts in the relationship between turbidity and SSC.

²⁷ It is of note that the specific turbidity values in Table 6-5 for the Motueka at Gorge (draining old, strongly indurated rocks with coarse-textured soils) and Motupiko at Christies (draining clay rich sediments and soils) sensibly reflect their lithologies, with specific turbidity twice as high at the Motupiko.

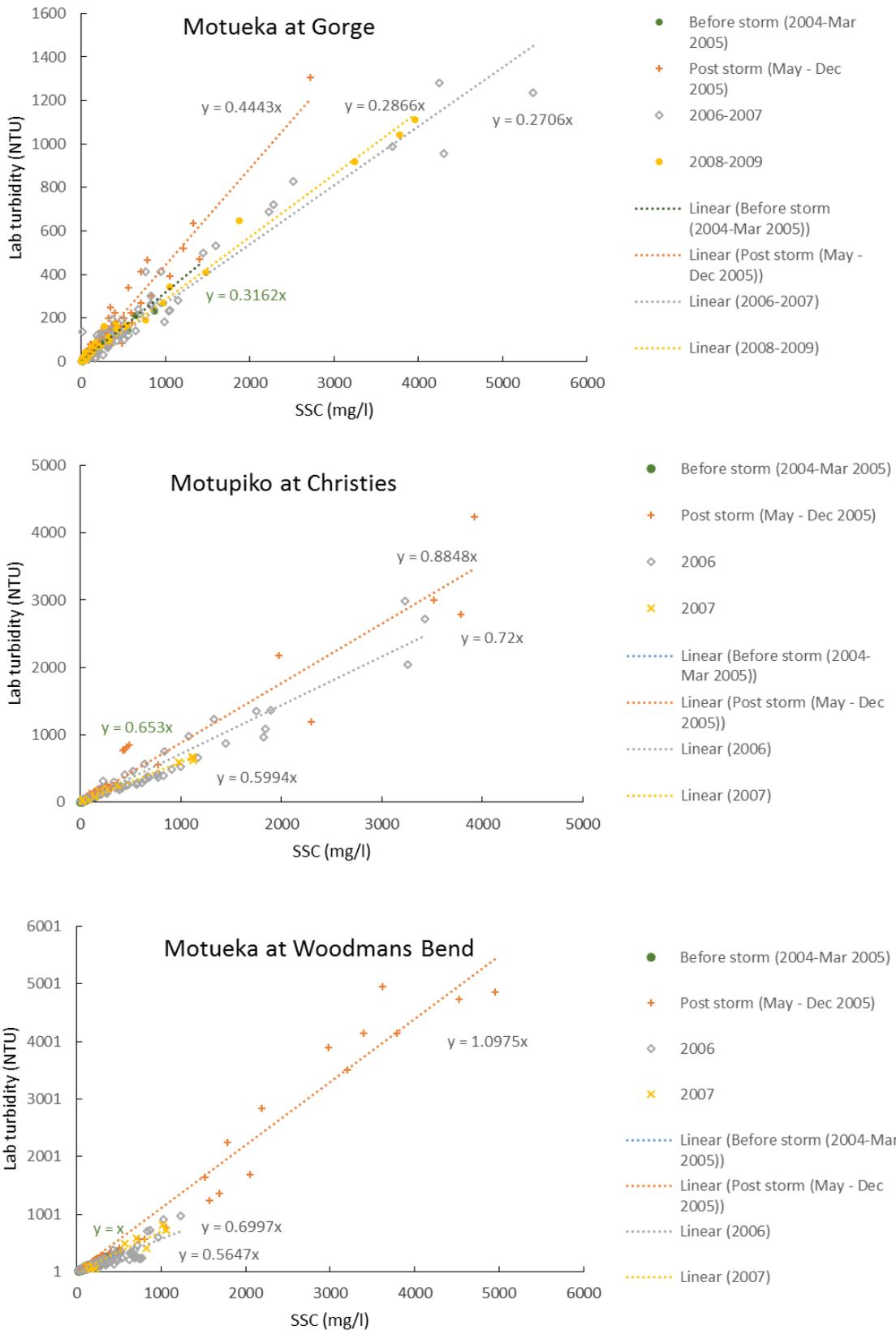


Figure 6-8: Turbidity vs SSC trends before and after the Easter 2005 storm at three sites in the Motueka catchment. Regression trends pass through zero, so their slopes show specific turbidity (NTU/mg/l).

Similar data were analysed for the Manawatu River at Teachers College to search for change in specific turbidity following the large flood of February 2004. Specific turbidity varied from 0.41 before this flood, to 0.40 immediately after in 2004, to 0.37 over 2005-7, but these changes were not significantly different (at the 5% level). In this case, we conclude no significant impact of the Manawatu flood on specific turbidity (despite widespread land-sliding associated with the February 2004 event – L. Basher, Landcare Research, pers. comm.).

Table 6-5: Specific turbidity at Motueka at Gorge, Motueka at Woodman’s Bend, and Motupiko at Christies before and after Easter 2005 storm. Specific turbidity is the trend of the turbidity vs SSC relations shown on Figure 6-8 and is an inverse proxy for particle size.

	Before storm	Post storm, 2005	2006+	2007+
Motueka at Gorge	0.32	0.44	0.27	0.28
Motupiko at Christies	0.65	0.88	0.72	0.6
Motueka at Woodman’s Bend	1	1.09	0.56	0.7

Overall, we conclude that changes in sediment load can cause changes in the SS PSD and so changes in the relationships between SSC and optical properties. Whether changes are significant will depend on the size characteristics of the sediment delivered from the affected sources and the importance of that source to the total sediment load. We note that while the effects of large hydrological events on sediment loads appears to last for several years, and will periodically re-occur, the effects of erosion mitigation works should be permanent.

It remains unclear: (i) why some sites should show significant changes in SRCs and PSDs following extreme hydrological disturbances (e.g., Motueka sites) while others do not (e.g., Manawatu); (ii) how large a storm/flood may be needed to trigger such changes; and (iii) how such naturally-driven changes compare to those potentially caused by erosion mitigation. The need for better understanding of these questions underpins some of the further research that is detailed in Section 10.

6.3.2 Evidence from spatial variation in SS particle size

The analysis of factors influencing spatial variation in SS particle size undertaken in Section 6.2 also informs on whether the SS PSD is likely to change following changes in catchment sediment supply (notably due to erosion mitigation work). This analysis showed that PSD was linked to sediment supply (as indexed by C_{mean}), which hints that if the sediment supply is changed then some change in particle size may follow. However, Section 6.2 showed that the dominant controller of PSD was catchment lithology, thus the sensitivity of PSD change to load change may most likely hinge on the lithological uniformity of the catchment.

Another key factor is the PSD characteristics of the sediment delivered from particular erosion features (e.g., from landslide-prone hillslopes compared with from eroding stream banks). If these are similar then treatment of one should not affect the PSD of the total sediment load, but if they differ then treatment may well alter the total load PSD. Unfortunately there is no PSD data on suspended sediment delivered from specific erosion types.

7 Relationships between streambed deposited sediment and sediment load

7.1 Introduction

This section explores the relationship between fine sediment deposited in and on the stream bed and catchment mean annual sediment loads. The primary hypothesis is that as sediment load increases the amount of fine sediment observed in and on the stream bed will increase.

The aim of the work is to quantify the relationship between modelled sediment load and measures of streambed sediment deposition using empirical models. Three sediment measures will be considered: the surface areal density (g/m^2) of fine, re-suspendable sediment (i.e., mud/silt) trapped within gravelly substrate (i) as measured with a Quorer, and (ii) as measured by a qualitative scoring technique, and (iii) the proportion of the streambed area covered with fine sediment (sand and silt/mud), as assessed visually.

Another aim of this work is to test the hypothesis that the Quorer data represents the amount of suspended sediment captured within the streambed interstices on a flood recession at the flow when bedload stops moving (which can be indexed by a flood statistic such as $\frac{1}{4}$ the mean annual flood discharge). This will be tested by comparing Quorer measurements to SSCs extracted from sediment rating curves at several index discharges.

The strength of the models will be used to determine whether deposited sediment attributes could be confidently linked to management actions. In the absence of robust empirical models, new data required to deliver functional relationships will be identified.

7.2 Summary of data available

Compiled data included 16934 recordings of deposited sediment (DS) using standard protocols (Clapcott et al. 2011; Appendix H, Table H-1) or protocols comparable to standard protocols. These recordings included unique samples and site averages depending on the source of the data. Compiled data were distributed throughout the country and dominated by contributions from the Freshwater Fish Database (NZFFD) (Figure 7-1). Excluded from the compilation were data collected using non-standard protocols, or repeated samples where there were known inconsistencies in sampling method.

Compiled data were labelled by sampling protocol (Table 7-1) and the meso-habitat from which samples were collected. There are five sampling protocols (indicated in brackets below) that result in three deposited sediment measures:

- *Fines* = proportion of streambed area covered by fine sediment, as a percent (SAM1, SAM2 and SAM3),
- *SIS* = surface density of resuspendable inorganic fine sediment embedded in the streambed, in g/m^2 (SAM4 or 'Quorer' method), and
- *Shuffle* = qualitative resuspendable sediment score, a value of 1-5 (SAM5).

Meso-habitats included run, riffle and pool. NZFFD samples were labelled by dominant habitat type at the sample site, but we did not feel this was a true representation of where fine sediment would have accumulated. For example, in a reach with pool and run present all of the fine sediment

observed may have occurred in a pool habitat but would incorrectly be assigned to the run habitat if run was the dominant habitat. Instead we assigned 'reach' to the meso-habitat type for NZFFD samples. All other samples were assigned to a run habitat based on field notes.

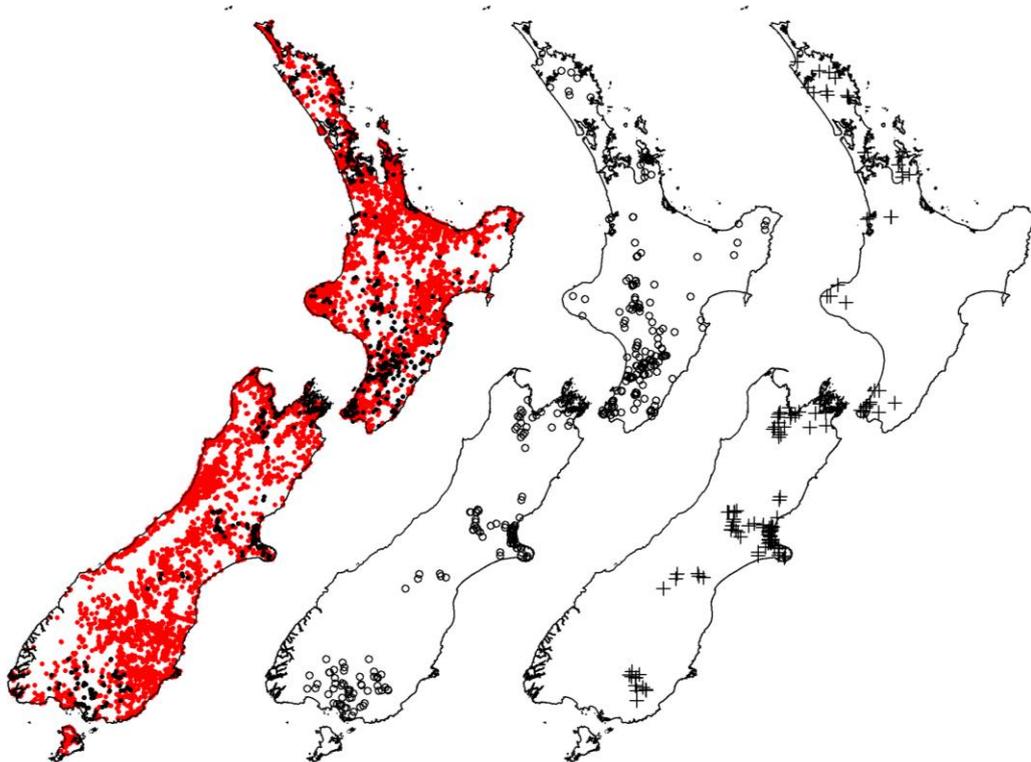


Figure 7-1: Distribution of deposited sediment samples. i) Fines from the New Zealand Freshwater Fish Database (red circles) and other sources (black circles), ii) SIS (open circles) and iii) Shuffle (crosses) data.

Table 7-1: Number of deposited fine sediment data for each sample method. NZFFD = New Zealand Freshwater Fish Database, SAM = Sediment Assessment Methods.

Data source	SAM1 or equivalent	SAM2	SAM3	SAM4	SAM5
NZFFD	10379	0	0	0	0
Regional Council	1388	780	3043	0	0
University	0	31	0	0	0
SAM development	246	239	294	385	149
Total	12012	1050	3337	385	149

7.3 Regression analyses

NZSegment (REC2) and NZReach (REC) identifiers were used to compile environmental data for regression analyses. Primary environment gradients of interest included catchment and segment sediment yield and reach-scale and catchment-scale stream descriptors (Table 7-2).

Table 7-2: Mean and range of variables used in regression analyses.

Variable	Description	Mean (range)	Source
Response variables			
<i>Fines</i>	Percent fine sediment cover on the streambed (%)	20.3 (0-100)	Compiled dataset
<i>SIS</i>	Mass of resuspendable fines embedded in the stream bed (g/m ²)	930 (0-52770)	Compiled dataset
<i>Shuffle</i>	Qualitative score for resuspendable fines in the streambed (0-5)	2.77 (0.5 – 5)	Compiled dataset
Predictor variables			
<i>Catchment sediment load</i>	Predicted sediment load for the total upstream catchment (t/y)	73350 (0 – 10940000)	WRENZ model (Hicks et al. 2011)
<i>Catchment sediment yield</i>	<i>Catchment sediment load</i> divided by catchment area (t/ km ² /y)	507.7 (0 – 68770)	Current project
<i>Segment sediment load</i>	Predicted sediment load at the segment sub-catchment scale (t/y)	409 (0 – 80300)	Hicks et al. 2011, on REC2
<i>stream power</i>	Product of the density of water (1000 kg m ³), acceleration due to gravity (9.8 m/s ²), <i>mean flow</i> , and <i>slope</i>	2.9 (0 – 2453)	Current project
<i>mean flow</i>	Mean annual flow (m ³ /s)	5.3 (0.001 – 495)	Woods et al. 2006, on REC2
<i>slope</i>	Average segment slope (degrees)	1.87 (0 – 26.99)	REC2
<i>width</i>	Average stream width (m)	43.26 (5-130)	Booker and Hicks 2013, on REC2
<i>elevation</i>	Average segment elevation (masl)	283.3 (-34 – 2020)	REC2
<i>mesohabitat</i>	Categorical descriptor of habitat where sediment was sampled ('run' or undefined as 'reach')	NA	Current project
<i>method</i>	Categorical descriptor of sediment sampling method (SAM 1-5)	NA	Clapcott et al. 2011
<i>CSOF</i>	Categorical REC classification at the climate source-of-flow level	NA	REC1
<i>USCalcium</i>	Average calcium concentration of rocks in the catchment, 1 = very low to 4 = very high	1.49 (0 – 4)	FENZ
<i>USHardness</i>	Average hardness of rocks in the catchment, 1 = very low to 5 = very high	3.07 (0 – 5)	FENZ
<i>USDaysRain</i>	Days / year with rainfall in the catchment >25 mm	16.9 (1.2 – 103.4)	FENZ
<i>SegFlowStability</i>	Annual low flow / annual mean flow (ratio)	0.18 (0 – 0.58)	FENZ

We explored the relationship between the three deposited sediment measures and modelled sediment load. Sediment load is the sum of sediment delivery from the contributing landscape at a total upstream catchment scale (*Catchment sediment load*) or local segment sub-catchment scale (*Segment sediment load*) and does not take into account fluvial processes. To account for local hydraulic processes operating on the sediment delivered to the stream channel, we normalised *Segment sediment load* estimates by an index of stream power (the product of mean annual discharge in m³/s and average channel percent slope). We also trialled a second stream power index using mean flood discharge but it provided less explanatory power than the previous; the same was true for an index of unit stream power which takes stream width into account. Finally, we also calculated sediment load normalised by area (i.e., sediment yield) and normalised by mean flow (i.e., sediment concentration).

Variables were transformed where necessary to meet the assumptions of normality for linear regression including log-transformation of *SIS*, *Catchment sediment yield* and *Segment sediment load/stream power*, and logit-transformation of *Fines* because it is a proportional variable.

Deposited sediment (*Fines*, *SIS* and *Shuffle*) had a weak negative correlation with *Catchment sediment yield* at a national scale (Figure 7-2). Catchment sediment concentration had a weaker relationship with DS (data not shown). At the segment scale the normalisation of *Segment sediment load* by *stream power* resulted in a positive correlation with *Fines* and *Shuffle*, but not for *SIS* (Figure 7-3). We did not observe stronger relationships for any other sediment load versions explored, e.g., *Segment sediment yield*. For both catchment and sediment segment measures, univariate relationships suggest that these variables alone explain very little of the variance in the deposited sediment data.

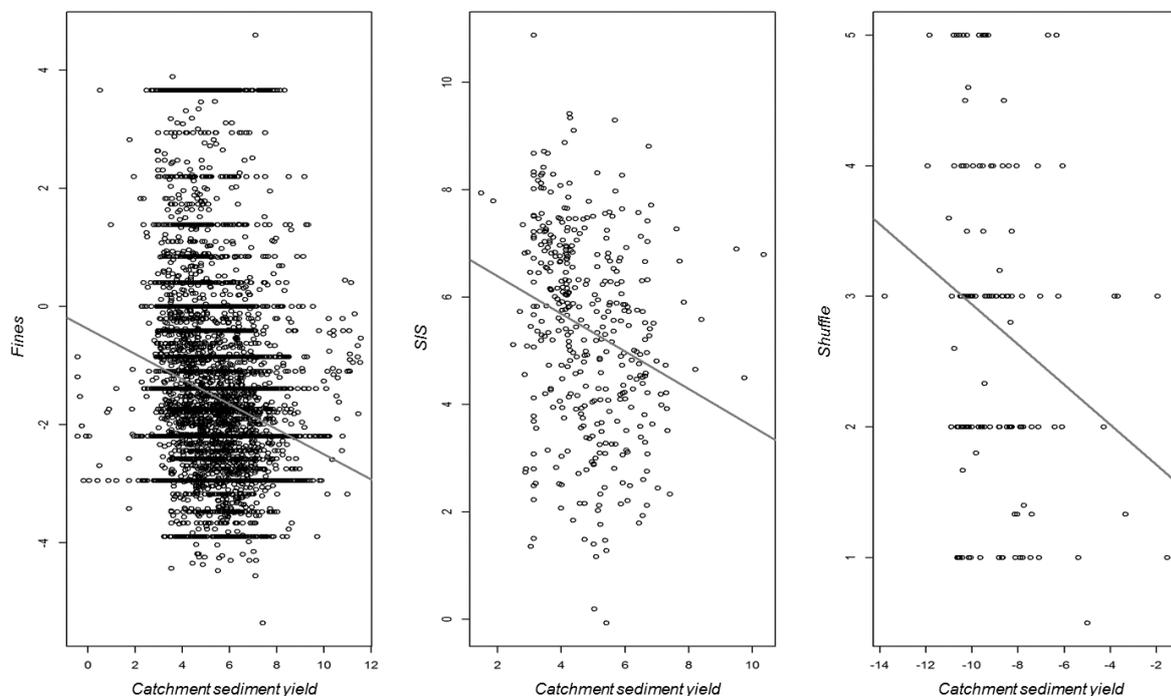


Figure 7-2: Correlation between $\log(\text{Catchment sediment yield})$ and deposited sediment metrics. i) $\text{logit}(\text{Fines})$ ($R^2 = 4\%$), ii) $\log(\text{SIS})$ ($R^2 = 6\%$), and iii) *Shuffle* ($R^2 = 5\%$). Solid lines shows the line-of-best-fit when statistically significant.

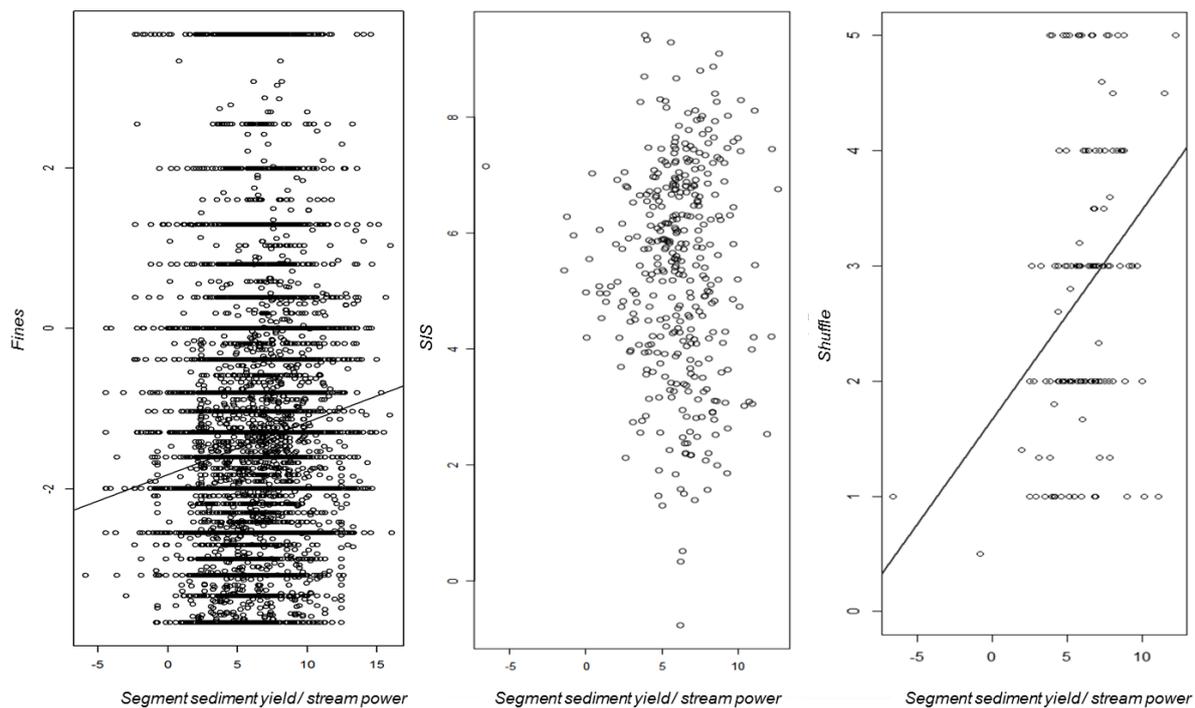


Figure 7-3: Correlation between $\log(\text{Segment sediment yield/stream power})$ and deposited sediment metrics. i) $\text{logit}(\text{Fines})$ ($R^2 = 0.1\%$), ii) $\text{log}(\text{SIS})$ (not significant), and iii) Shuffle ($R^2 = 10\%$). Solid lines shows the line-of-best-fit when statistically significant.

7.3.1 Fines

Sediment load variables (*Catchment sediment yield* and *Segment sediment load/stream power*) were chosen based on the strength of correlations with deposited sediment data. Then we conducted linear regressions to explore the relationship between *Fines* and modelled sediment load and determine the influence of sampling error and other environmental variables in explaining variation in the DS metric. We used weighted regression to accommodate segment replication ($N = 16400$) and weighted each replicate by the inverse of the number of replicates per site, which is equivalent to using a site mean. There were 8239 unique segments in the dataset.

A simple linear regression model for *Fines* as a function of *Catchment sediment yield*, *Segment sediment load/stream power*, *method*, *mesohabitat* and *CSOF* without interactions had $R^2 = 12\%$. We explored the potential for interactions between all variables except *CSOF* (i.e., allowing slopes to vary for all variables but only the mean for *CSOF*) and this partial interaction model had $R^2 = 13.4\%$. Partial plots for the main effects demonstrate an overall positive response of *Fines* to *Segment sediment load/stream power*, a negative response to *Catchment sediment yield*, both with wide scatter in the relationship (Figure 7-4). While *method* and *mesohabitat* were not significant main effects, there was a significant interaction between *mesohabitat* and both sediment load variables. Some *CSOF* classes had significantly higher average *Fines* and some lower average *Fines* (Figure 7-4). We interpreted this model output to mean that all three methods of measuring DS (SAM1, SAM2 and SAM3) provide comparable data and a negative relationship between *Fines* and *Catchment sediment yield* is observed in all cases. However, the relationship was weaker (i.e., lower slope) in runs compared to undefined 'reach' habitats. Furthermore, the relationship between *Fines* and *Segment sediment yield/stream power* was positive in run habitat compared to negative in undefined 'reach' habitat.

This suggests to us that *Fines* data collected from run habitats might provide a more useful response to modelled sediment load than data where the habitat is undefined (i.e., data from the NZFFDB). Finally, the relationship between *Fines* and modelled sediment load varies across the country depending on environmental variability as defined by the REC grouping (i.e., CSOF).

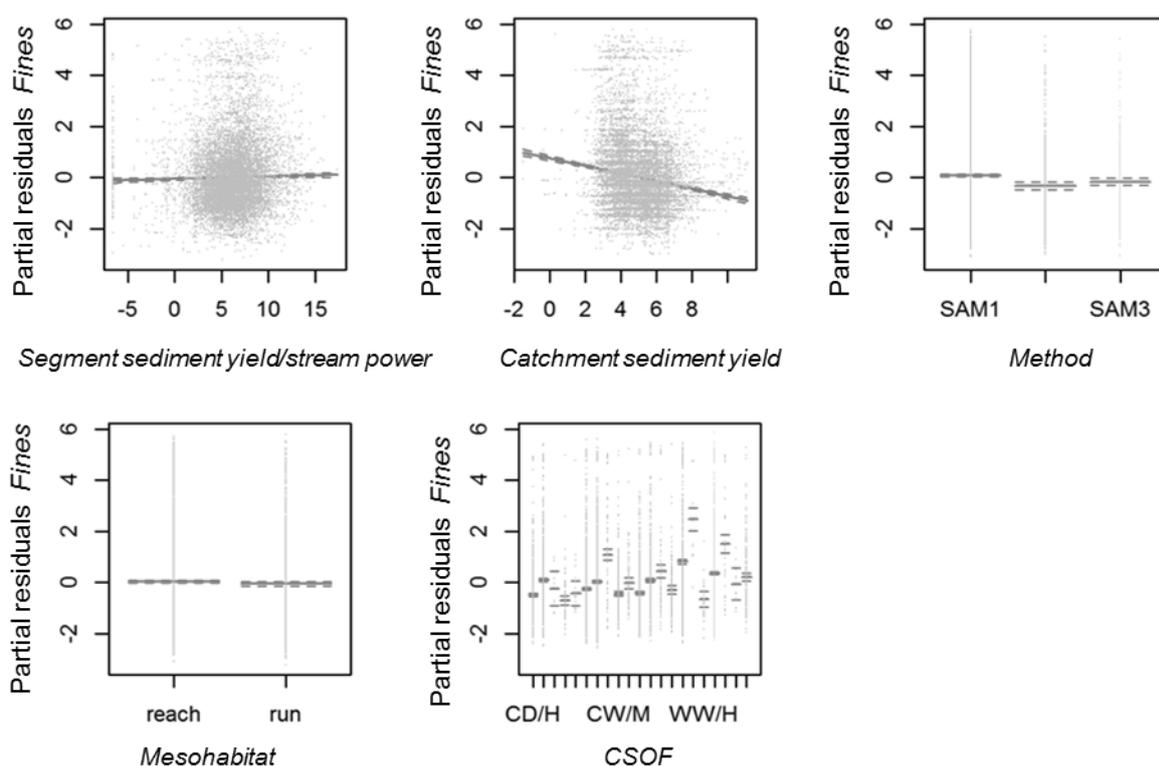


Figure 7-4: Partial plots for the main effects of a simple general regression model of *Fines* as a function *Segment sediment yield/stream power*, *Catchment sediment yield*, *method*, *mesohabitat*, and *CSOF*. Solid lines show line-of-best-fit and dashed lines show 95% confidence intervals.

Next we fitted a boosted regression tree (BRT) model to visualise the shape of response of *Fines* to predictor variables. Logit-transformed *Fines* was modelled as a function of *Catchment sediment yield*, *Segment sediment load/stream power*, geology descriptors (*USCalcium*, *USHardness*), climate descriptors (*USDaysRain*, *SegFlowStability*), and surrogate source of flow descriptors (*USAveSlope*, *Elev*), sampling *method* and *mesohabitat*.

The model diagnostics showed that together these 10 variables explained 22.7% of the deviance in the *Fines* data. *Elevation* was the most informative predictor variable and *mesohabitat* the least. *Catchment sediment yield* and *Segment sediment load/stream power* were the 4th and 6th most explanatory variables and explained 11.2% and 7.2% of the total deviance (i.e., 2.5% and <1% in total).

The response of *Fines* to *Catchment sediment yield* was predominantly negative and for *Segment sediment load/stream power* predominantly positive across the gradient of the predictor variable where the majority of sample data was distributed (as indicated by the rug plots on x axis, Figure

7-5). The fitted functions for *Fines* in response to other variables were mainly intuitive albeit noisy: *Fines* decreased with increasing *Elevation*, *USAvgSlope*, *USDaysRain*, *USCalcium*, and *USHardness*, and *Fines* increased with increasing *SegFlowStability* (Figure 7-5).

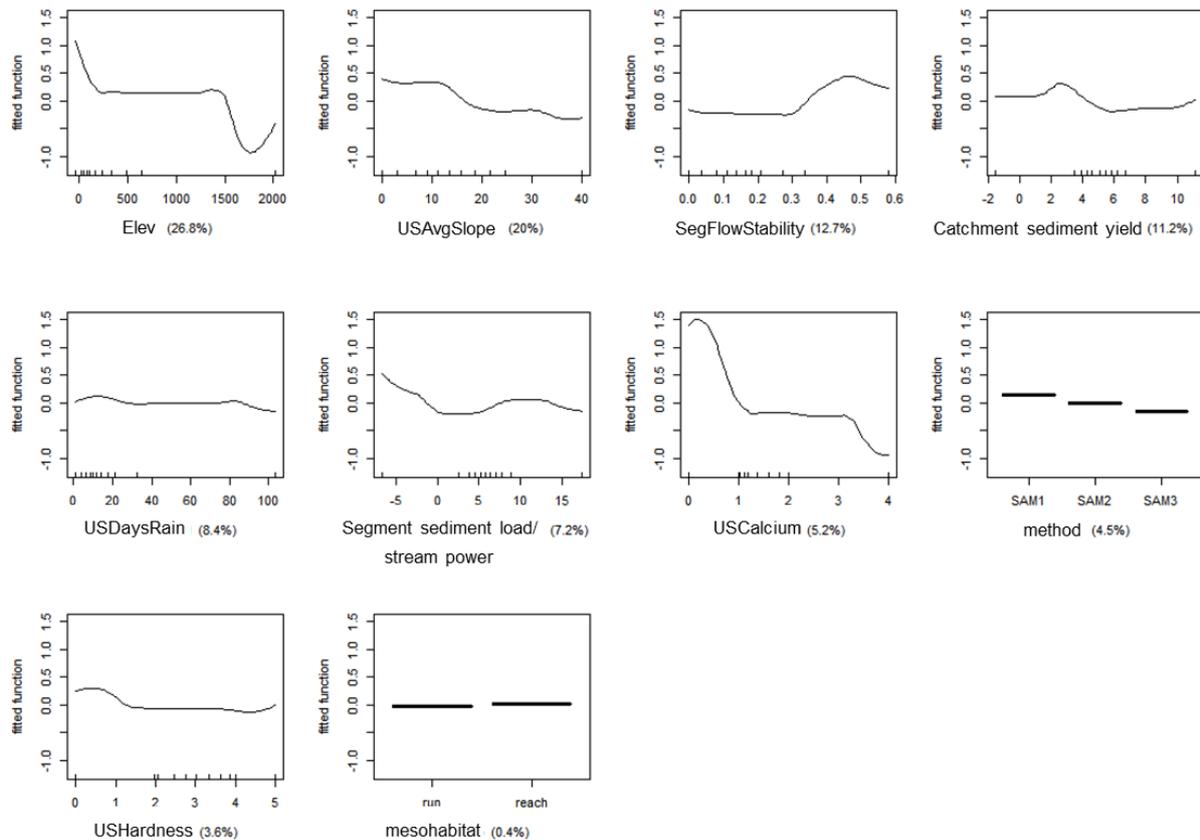


Figure 7-5: Univariate partial dependence plots (smoothed fitted functions) of the relationships between *Fines* and *Catchment sediment yield*, *Segment sediment load/stream power* and environmental descriptors, and sampling *method* and location identifiers. Plots show distribution of data as rug plots on the x axis and the proportion of total deviance explained by each variable in parentheses.

Because simple linear models suggested a more informative predictive relationship might result from examining *Fines* data from run habitat alone, we developed a second BRT model with the same set of predictor variables but excluding *mesohabitat*. Training data included 6036 observations compared to 16360 for the primary model. The run habitat-restricted BRT model explained 55.7% of the variance in the *Fines* data. *Catchment sediment yield* and *Segment sediment load/stream power* were the 1st and 2nd most informative variables explaining 15.9% and 14% of the deviance, respectively (i.e., 8.9% and 7.8% of total deviance). The response of *Fines* to *Catchment sediment yield* remained predominantly negative and the response of *Fines* to *Segment sediment load/stream power* remained predominantly positive (Figure 7-6).

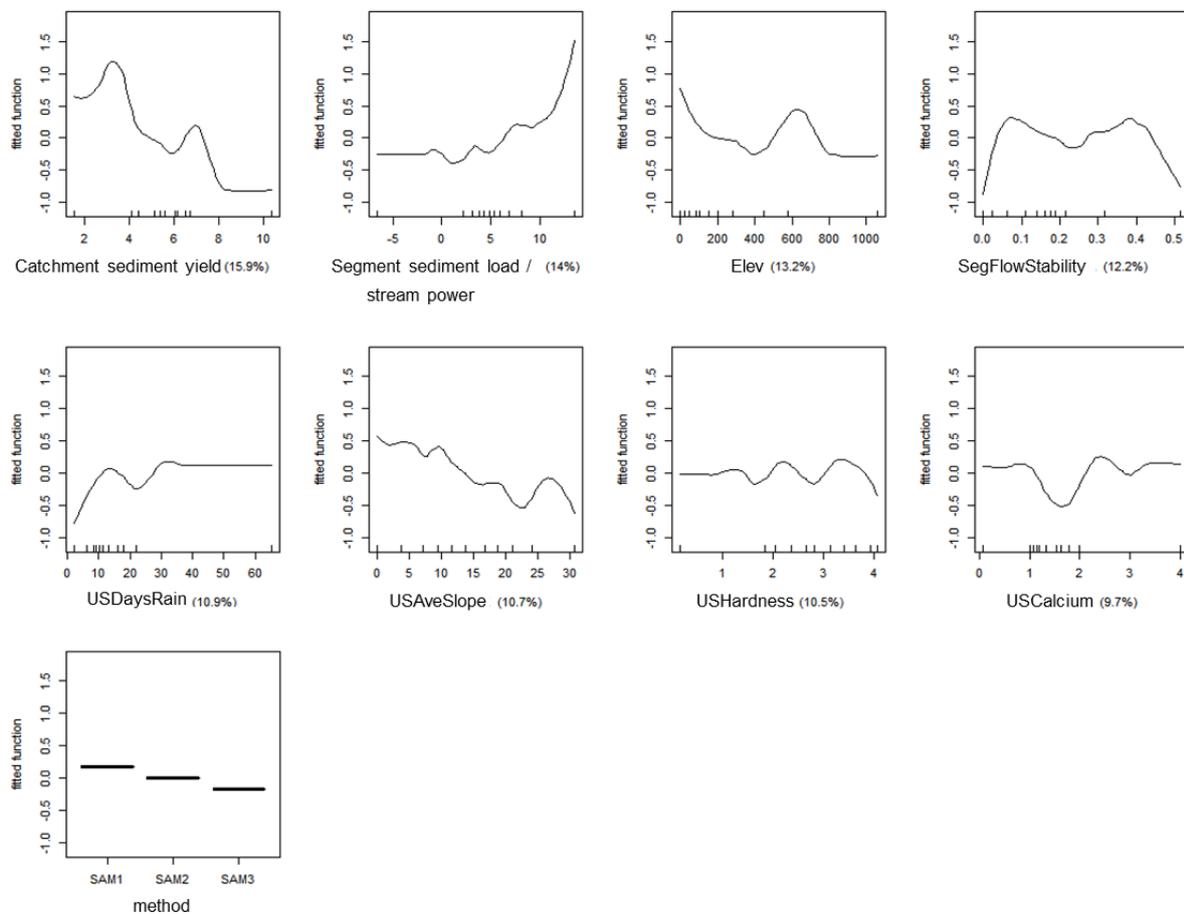


Figure 7-6: Univariate partial dependence plots (smoothed fitted functions) of the relationships between *Fines* collected in run habitat and *Catchment sediment yield*, *Segment sediment load/stream power* and *environmental descriptors*, and *sampling method* and *location identifiers*. Plots show distribution of data as rug plots on the x axis and the proportion of total deviance explained by each variable in parentheses.

Model diagnostic plots for both the ‘all data’ and ‘run habitat-restricted data’ models for *Fines* further illustrate relative model performance (Figure 7-7). The Nash-Sutcliffe Efficiency (NSE) statistic, which indicates how well the plot of observed versus predicted values fits the 1:1 line, shows that the all data model has satisfactory performance (NSE = 0.29), whereas the run habitat-restricted data model has very good model performance (0.73), despite wide scatter in the data. The root mean squared deviation (RMSD) indicates twice the model accuracy in the run habitat-restricted data model (RMSD = 10.2) compared to the all data model (RMSD = 19.9). Both models tend to overestimate *Fines* on average as indicated by negative model bias.

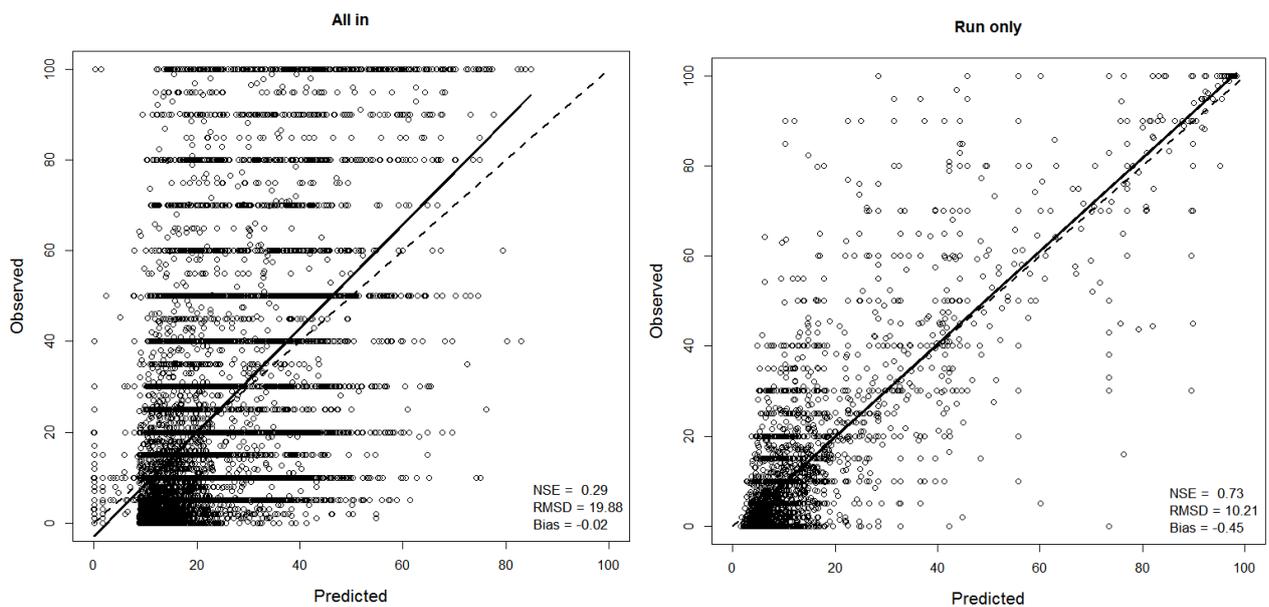


Figure 7-7: Scatter plots of observed versus predicted values from a) all data and b) run habitat-restricted data model for *Fines*. Dashed line is the 1:1 line and the solid line is the line of best fit. Model performance statistics are explained in the text.

We calculated the predicted reduction in *Fines* as a result of a 50% reduction in *Segment sediment load* using the run-habitat restricted data BRT model. The model predicts a resulting 6% reduction in % fine sediment cover on average with a range from 88% reduction to 500% increase (Figure 7-8). For example, a 6% reduction in 20% fine sediment cover equals 18.8% fine sediment cover.

50% reduction in Segment Load

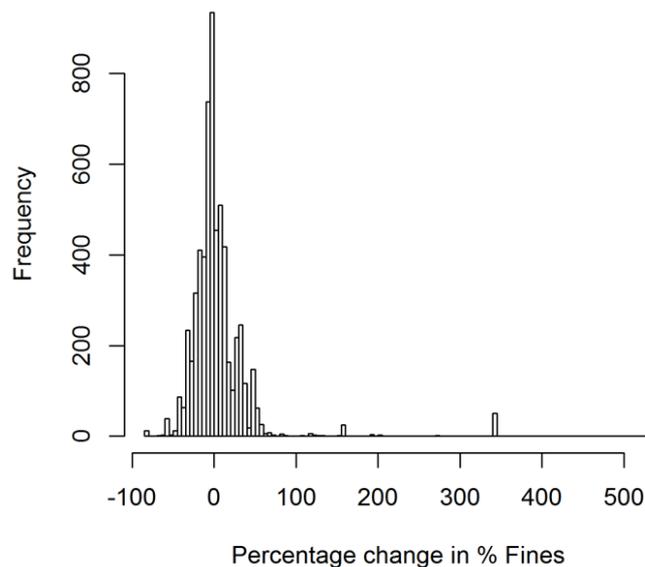


Figure 7-8: Frequency distribution of predicted proportional reduction in % fine sediment cover (*Fines*) in response to a 50% reduction in *Segment sediment load* from the BRT model fitted with run-habitat data. For example, -50 is a halving of % fine sediment cover and +100 is a doubling.

7.3.2 SIS

We conducted regression analyses to explore the relationship between *SIS* ($n = 362$) and modelled sediment load. A general linear model for *SIS* as a function of *Catchment sediment yield*, *Segment sediment load/stream power*, and *CSOF* without interactions had $R^2 = 31\%$. Neither of the sediment load variables were significant in the model and mean *SIS* varied significantly among *CSOF* categories (Figure 7-9).

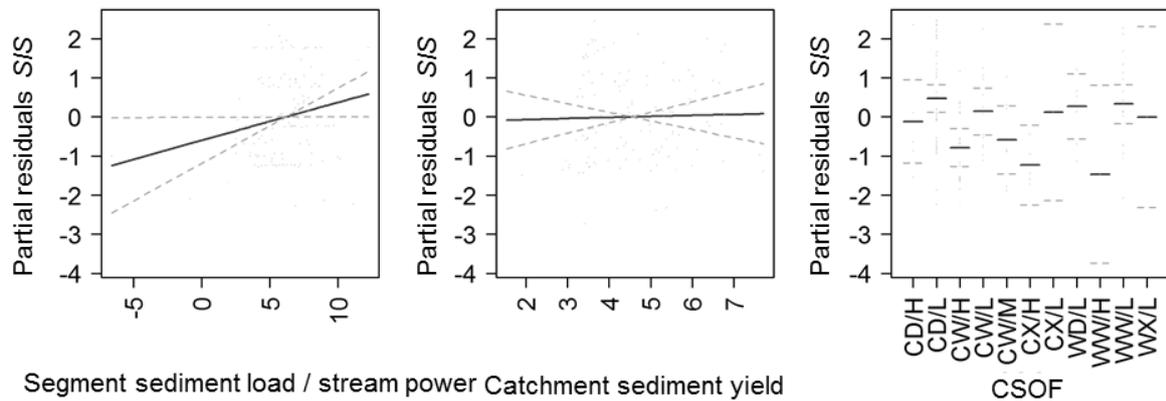


Figure 7-9: Partial plots for the main effects of a general linear regression model of *SIS* as a function of *Segment sediment load/stream power*, *Catchment sediment yield*, and *CSOF*. Solid lines show line-of-best-fit and dashed lines show 95% confidence intervals.

We used a boosted regression tree (BRT) model to further explore the response of *SIS* to modelled sediment load and six continuous environmental variables that are likely to contribute to the distribution of fine sediment on the stream bed. The BRT model explained 39.2% of the deviance in $\log(SIS)$. While *Segment sediment load/stream power* was the most explanatory variable in the model (20.7% deviance), the response of *SIS* to *Segment sediment load/stream power* was predominantly negative, although the somewhat sigmoidal shape suggested the lack of a linear relationship (Figure 7-10). All *SIS* data was collected from run habitats so this was not a factor to explore further.

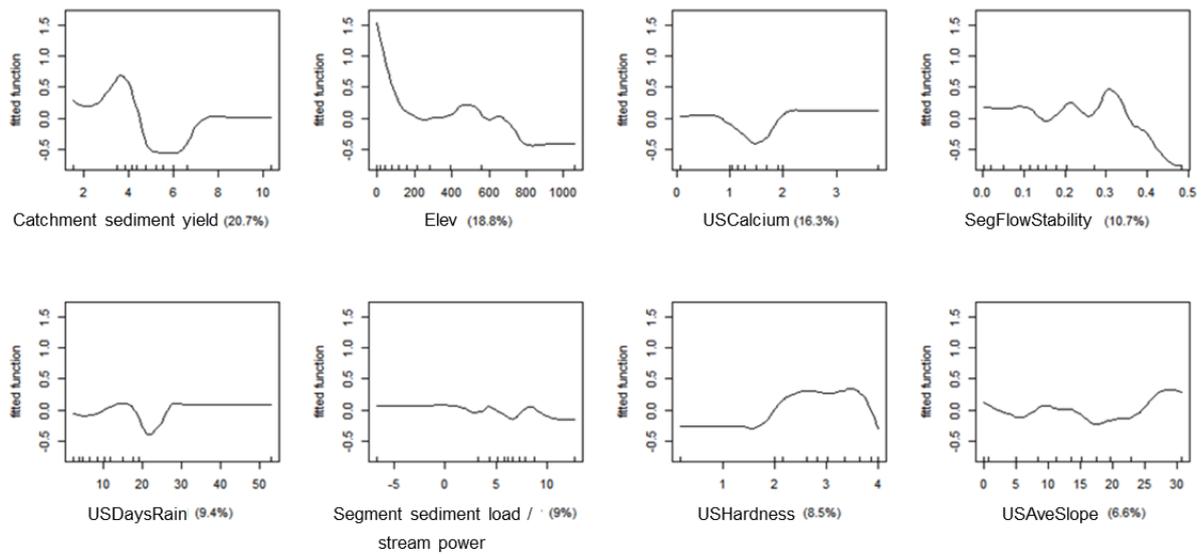


Figure 7-10: Univariate partial dependence plots (smoothed fitted functions) from a BRT model of the relationships between $\log(SIS)$ and sediment load and other environmental descriptors. Plots show distribution of data as rug plots on the x axis and the proportion of total deviance explained by each variable in parentheses.

7.3.3 Shuffle

A simple linear model for *Shuffle* ($n = 130$) as a function of *Catchment sediment yield*, *Segment sediment load/stream power*, and *CSOF* without interactions had $R^2 = 20\%$. (Figure 7-11). The *Shuffle* score significantly increased in response to *Segment sediment load/stream power* and none of the other variables were significant in the model. A BRT model with sediment load and six descriptors of environmental variability explained 11.2% the variance in the *Shuffle* data (Figure 7-12). *Segment sediment load/stream power* was the 2nd most explanatory variable explaining 23.9% of variance (i.e., only 2.7% total deviance). All *Shuffle* data is collected from run habitats so this was not a factor to explore further.

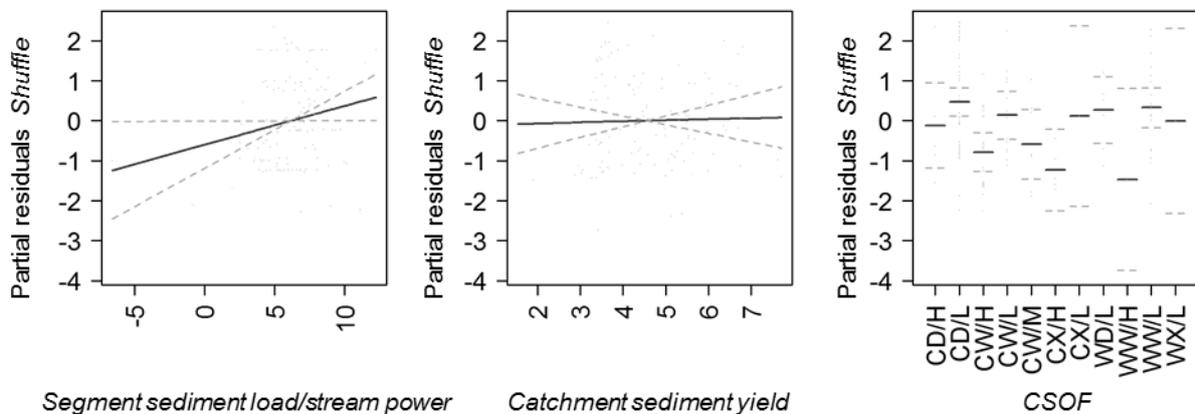


Figure 7-11: Partial plots for the main effects of a simple linear regression model of *Shuffle* as a function of *Segment sediment load/stream power*, *Catchment sediment yield*, and *CSOF*. Solid lines show line-of-best-fit and dashed lines show 95% confidence intervals.

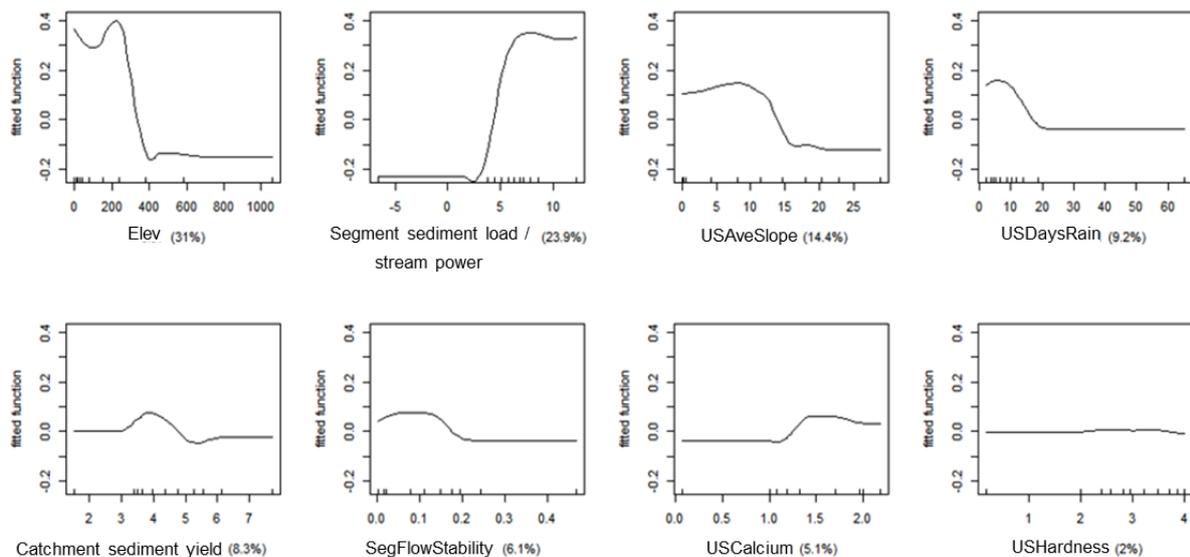


Figure 7-12: Univariate partial dependence plots (smoothed fitted functions) from a BRT model of the relationships between *Shuffle* and sediment load and other environmental descriptors. Plots show distribution of data as rug plots on the x axis and the proportion of total deviance explained by each variable in parentheses.

7.3.4 Analysis of deposited sediment data with matching measured sediment yield data

Only 30 sites were found that had both quantitative measures of DS extent and measurement-based data on mean annual sediment yield (as derived from sediment ratings combined with discharge records) and discharge statistics (i.e., mean flow, mean annual flood). This dataset is too small for developing predictive relations, but is utilised in the following section.

7.3.5 Testing the “frozen bedload” hypothesis

A hypothesis promoted by John Dymond, termed herein the “frozen bedload” hypothesis, is that *SIS* should relate to the *SSC* on a flood recession at the flow when bedload stops moving (which can be indexed by a flood statistic such as $\frac{1}{4}$ the mean annual flood discharge). This assumes that at that time the concentration of sediment in the water trapped in pore spaces between pebbles and cobbles matches that in the flow above. It also assumes that the substrate pore spaces are sealed from further exchanges of fine sediment with the river flow once the bedload stops moving.

The 30 sites where both *SIS* was measured and reliable sediment rating curves were available were located throughout New Zealand (Figure 2-1). They were located mainly in the lower half of the North Island and axial mountains of the upper South Island; stream order ranged from 2 to 7 but was mainly 4-5; catchment source of flow was predominantly either mountain/hill country or lowland; and sediment yield ranged from 6 to 4434 t/km²/yr.

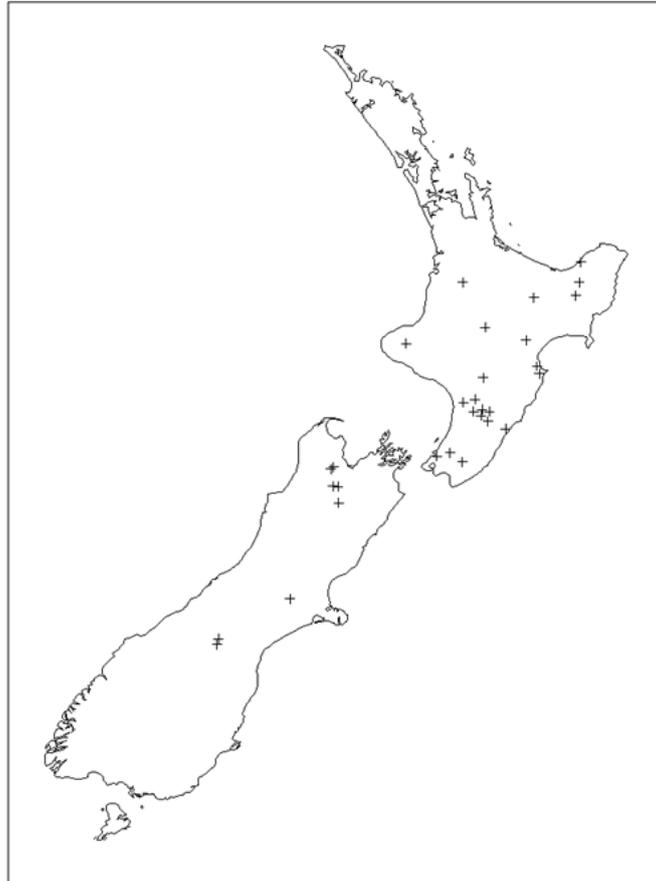


Figure 7-13: Sites with both *SIS* measurements and reliable SRC data useful for testing the “frozen bedload” hypothesis.

There was no relationship observed between *SIS* and SSC at one quarter of the mean annual flood discharge²⁸, as extracted from the rating relationship (Figure 7-14). This does not support the hypothesis.

Statistics of the *SIS* and SSC values at one quarter the mean annual flood discharge are listed in Table 7-3. The *SIS* statistics are typically around 50% of the matching SSC statistics (e.g., the mean *SIS* is 55% of the mean SSC). By the “frozen bedload” hypothesis, the *SIS* values should generally be around 4% of the SSC values²⁹. That they are around 50% on average suggests that if the bedload stops moving at one quarter the mean annual flood discharge then fine sediment must be accumulating into the substrate on flood recessions after the bed has stopped moving.

Table 8-3 also provides equivalent statistics for the SSC at the mean annual flood discharge. In this case, the *SIS* statistics range between 6% and 24% of the matching SSC statistics (with the mean *SIS* being 8% of the mean SSC). This is closer to the hypothesized 4%, and suggests that the discharge when bedload ceases moving may be better indexed by a discharge closer to the mean annual flood

²⁸ Mean annual flood discharge (equal to the average of the series of annual peak discharges) was determined directly from flow records at the observations sites.

²⁹ The areal *SIS* values (g sediment per m² bed area sampled by the Quorer) stem from a nominal bed-disturbance depth of 0.1 m. Thus the mass of fine sediment per unit volume of substrate (g/m³) should equal *SIS*/0.1. Assuming a porosity of 0.4 for the substrate, then the fine sediment mass per unit volume of pore-space should equal *SIS*/(0.1 × 0.4) = 25 *SIS* g/m³. It is this concentration that should align with the SSC in the water trapped in the pores spaces when the bed ceases motion.

discharge. However, indexing off the mean annual flood discharge provides a similarly poor relationship between SSC and SIS ($R^2 = 0.03$; not plotted).

Given the few data points available, however, this analysis cannot be regarded as a robust test of the “frozen bedload” hypothesis. Further data will need to be collected to do this.

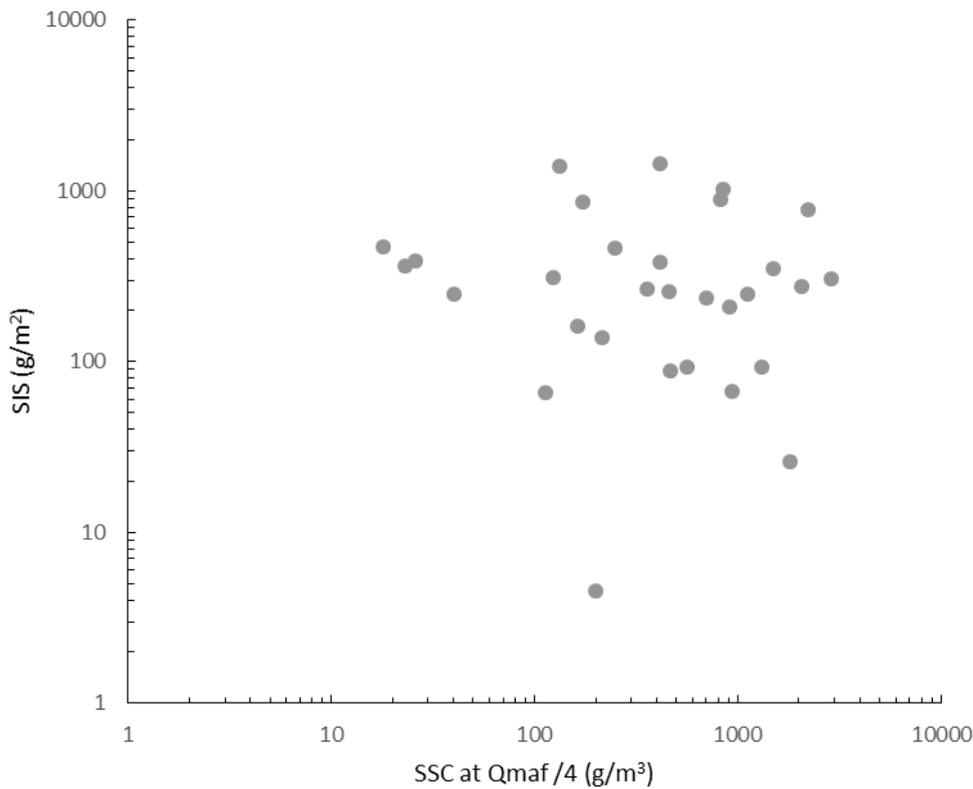


Figure 7-14: Relationship between SIS and suspended sediment concentration (SSC) at quarter of the mean annual flood discharge (Q_{maf}). $R^2 = 0.007$.

Table 7-3: Statistics of SIS data and SSC at one quarter the mean annual flood and at the mean annual flood at sites with both data types available. Q_{maf} is peak discharge of mean annual flood.

	SCC at $Q_{maf}/4$ (g/m^3)	SCC at Q_{maf} (g/m^3)	SIS (g/m^2)
Count	30	30	30
Min	18	19	4.5
Max	2881	26900	1451
Mean	709	5000	396
Median	437	2440	271
Standard deviation	741	6100	380

7.3.6 Further exploration of variation in the deposited fine sediment data

Of the 8239 unique stream segments with associated deposited sediment data, 1957 sites had between 2 and 181 replicate measures. In our regression analyses we weighted the contribution of these replicates during model development so that a 'mean' value is effectively used. This approach may not best represent the state of DS at any given stream site. For example, *Fines* in one segment varied from 0% to 65% (Figure 7-15). The average standard deviation at sites with multiple measures was 8.9% and there was a significant relationship between variation in *Fines* and native forest cover (Figure 7-15). Firstly, high within-segment variability may be due to the fact that there were multiple sites within a stream segment. Segment lengths average 698 m (0.1 – 29,137 m) at a national scale and hence any given segment can easily accommodate multiple sites that may vary from each other due to local geomorphology. Secondly, high within-segment variation may reflect high temporal variation. This suggests that temporal variation in *Fines* will be important to assess when considering the frequency component of any potential sediment attribute and critical for determining temporal trends. It is also a potential reason for unexplainable variance in the *Fines* model in relation to modelled mean annual sediment load measures.

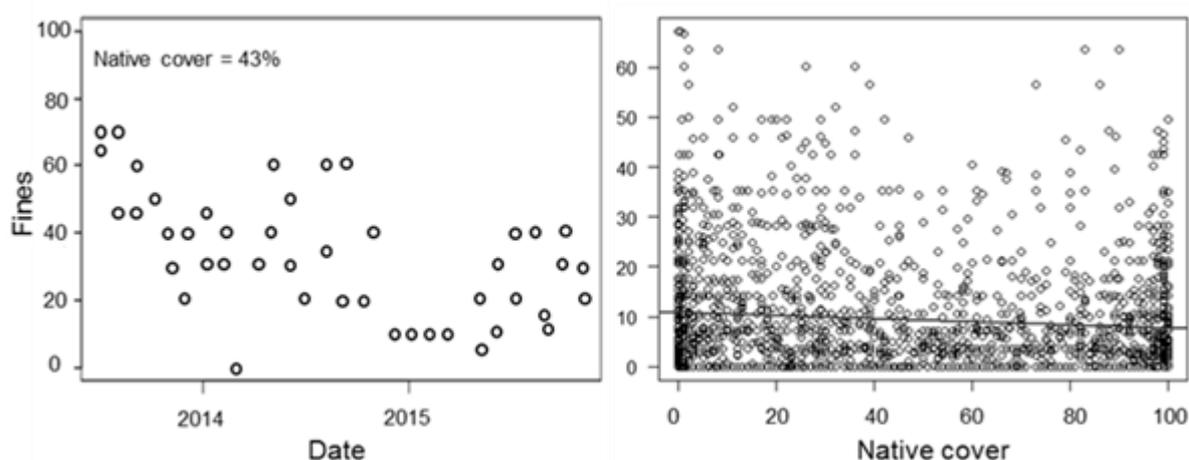


Figure 7-15: Scatterplot of temporal variation of *Fines* at (left) one segment and (right) all segments where more than one replicate was measured in relation to native forest cover. Solid line shows best-fit line.

One way to account for temporal variation in deposited sediment may be to view measures in relation to flow history. The relative importance of *flow stability* in all DS models suggests variation in flow is important. Certainly (e.g., Hicks et al. 2004), the majority of sediment delivery occurs during infrequent, episodic events such as during 'flood' flows (i.e., acute effects) and less sediment delivery occurs during stable flows (i.e., chronic effects). Furthermore, increased bed-movement occurs during flood flows as sediment on the streambed is re-entrained then deposited during receding flows. We have no antecedent flow measures to explore to determine how long a bed moving event occurred prior to measurement of DS. This is another potential source of unexplainable variation in the current data.

It is possible that the model used to predict sediment load introduces a significant source of error to the relationship between sediment load and DS. According to Hicks et al. (2011), the sediment load model is less accurate for catchments <10 km² due to the bias of training data towards larger rivers. We explored this possibility by developing parallel models for all deposited sediment measures

excluding sites with catchments <10 km². This halved the size of our training data and showed no significant improvement in model performance compared to the full data analyses reported above.

7.4 Discussion

The fact that DS decreased in response to increasing catchment sediment loads did not support our primary hypothesis. This is perhaps not that surprising for reasons discussed earlier and also because the majority of sediment delivery is likely to occur at periods of high flow that would effectively flush the sediment through the stream network until a receiving environment is reached (i.e., at low elevation, low slope). The dominant significance of *Elevation* and *USAvgSlope* in explaining variance in DS metrics in our models supports this idea. A measure of antecedent flow is likely to improve our ability to link DS to catchment loads, because we could then account for time since last bedload disturbance.

According to the 'stable channel balance' premise, channel aggradation (i.e., accumulation of sediment) occurs when the product of sediment load and sediment size is greater than the product of slope and discharge (Lane 1955). As such, it is important to take account of the relationship between these variables when quantifying the relationship between sediment load and DS. For example, Herbst et al. (2011) showed that the relationship between predicted sediment yield and DS was mediated by stream power, which is a product of slope and flow. We too found that the local hydraulic conditions measured as stream power were important in demonstrating the relationship between *Segment sediment load* and DS. Our results suggest that the local delivery of sediment and how it is 'captured' by the local stream morphology is much more informative of DS than sediment load from the upstream catchment. The implications for management could be a focus on local habitat to minimise the chronic delivery of fine sediment that occurs during stable flows, e.g., bank erosion. For example, research by Holmes et al. (2016), demonstrated that continuous fencing of a streamside area greater than 300m² resulted in improved stream habitat via decreased fine sediment delivery and retention in a spring fed agricultural stream.

There is a large amount of compiled data available that measures DS in streams. Our analyses suggest that not all of it is useful for defining the sediment load – DS relationship. In particular, estimates of fine sediment cover from the NZFFD appear particularly variable. These data were not collected specifically for assessing stream habitat and, as such, may introduce more noise than explanatory power to models. Focussing on run habitat data provided the most promising relationship between *Segment sediment load* and DS. However, even in this case the model predicted a broad range in the response of DS to a hypothetical 50% reduction in sediment load, providing little evidence of trend data that would inform management.

The addition of environmental variables (describing elevation, slope, geology and flow) substantially improved our ability to model DS as a function of modelled sediment load, especially for the measures of *Fines* (% cover of the streambed) and *SIS* (re-suspended inorganic sediment using the Quorer method). However, the explanatory power of any of the models is modest and similar to that observed in a recent study exploring drivers of DS in UK streams (Naden et al. 2016). Similar to our results, Naden et al. (2016) also showed that stream power was the most significant explanatory variable of DS and the contribution of sediment yield was small (1%) marginally significant ($p < 0.05$). They suggested that stream power be taken into account when setting instream sediment targets. Likewise our results demonstrate high spatial variation in the relationship between sediment yield and DS and highlight the potential need to 'regionalise' relationships when developing sediment limits.

Finally, there were insufficient data to explore the relationship between SSC and DS and the “frozen bedload” hypothesis. Based on the results of current analyses there is no empirical relationship that could robustly be used for predicting the response of DS in streams to land-based management actions. We do not think that any of our current models can be confidently used to describe the relationship between sediment load (as predicted by the Hicks et al. 2011 model or SRCs) and DS. Targeted sampling at sites with established rating curves following periods of variable flow may help improve model performance (as discussed in Section 10). It may further help characterise the ‘normal/expected’ state of DS and determine temporal change in response to natural storm-driven variability versus a change in response to land management.

8 Analytical frameworks to link catchment sediment loads and sediment environmental state variables

8.1 Introduction

An analytical chain for relating catchment sediment loads to the four sediment environmental state variables (ESVs) is shown in Figure 8-1. This chain uses SSC as an intermediate link between catchment sediment load and the other ESVs and is called framework 1. We also explore an alternative analytical chain that uses beam attenuation coefficient as an alternative intermediate link between catchment sediment load and the optical ESVs (i.e., VC, LP). We term this framework 2.

The flow duration curve is a key component of both frameworks since it provides a means of estimating the temporal frequency distributions of the ESVs. In framework 1 this is achieved via the relationship between flow and SSC (i.e., the sediment rating curve). In framework 2 it is achieved via the relationship between flow and the beam attenuation coefficient.

In this section, we detail each of the component relationships that link these frameworks together, beginning with the flow duration curve. We illustrate each relationship with data from an example site in Northland (Wairua River at Purua³⁰), and we demonstrate how the uncertainty in the predictions from each relationship can be estimated. In Section 9 we show how the accumulated uncertainties impact on the robustness of the predicted ESVs.

The prediction objectives across the analytical chain can be of two types:

- predicting the absolute values of the ESVs associated with a given sediment load (or vice-versa)
- predicting the *change* in the ESVs associated with a given change in sediment load (or vice-versa).

In the following sections we will demonstrate how both objectives can be met.

Both frameworks examined here make two key assumptions:

- the flow duration curves can be directly converted to SSC-duration curves and beam-attenuation-duration curves via the respective 'rating' curves
- changes in catchment sediment loads are directly related to changes in the rating curves (i.e., when the rating curve is represented by a power function $C = aQ^b$, the change in the a-parameter is directly proportional to the change in load while the b-parameter does not change).

The generality of the second assumption was examined in Section 4. Assessing the first assumption was not within the scope of this investigation and requires further research (Section 10.2.2).

³⁰ We chose the Wairua at Purua because it is one of the few sites in the country where data on all of the components in the analytical frameworks have been measured or estimated.

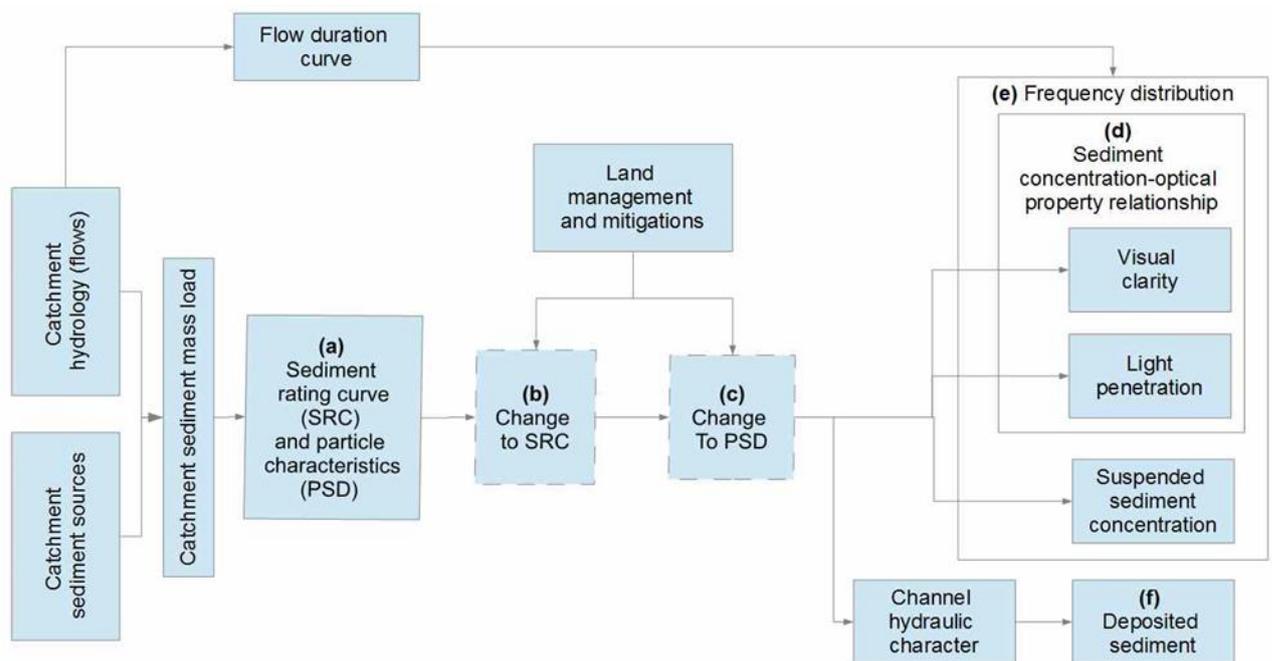


Figure 8-1: Schematic diagram indicating the analytical chain that links VC, LP, SSC, and DS to the management of catchment sediment sources (i.e., framework 1).

8.2 Flow duration curve

The flow duration curve (FDC) gives the percentage of time that the discharge of a river is below a given value. Where the flow duration curve is known, it is possible to read off characteristic flow percentiles, such as the 50% flow percentile, or median flow, which the river is below 50% of the time. Other useful characteristic flow percentiles are 10% (low flow), 80% (high flow), 95% (flood flow). The characteristic flow percentiles may be converted to percentiles of sediment ESVs if there are concurrent measurements of discharge and sediment ESVs. Figure 8-2 shows the flow duration of the Wairua River at Purua as an example. The uncertainties of a measured flow duration curve such as this one are small and may be considered negligible compared to uncertainties of the sediment attributes (Dymond and Christian 1982).

Where there is no flow record available the FDC may be estimated using the predictor developed by Booker and Snelder (2012). Booker and Snelder compared several methods to estimate flow duration curves at ungauged sites across New Zealand and found that the most accurate method was one that represented the FDC by parameters of the generalised extreme value (GEV) probability distribution, with those parameters predicted from catchment/site characteristics using a Random Forests model³¹. The root-mean-square error of their GEV/Random Forests model in predicting standardised flow (i.e., flow divided by mean flow) for a given exceedance percentile, averaged over 379 sites and over all exceedance percentiles, equated to a factor of $\times/\div 1.26$, or approximately $\pm 26\%$ (from Booker and Snelder's Fig. 8). They also found, however, that model predictive performance varied with exceedance percentile.

³¹ The model has since been applied across the REC stream network, so the FDC parameters can simply be found via the stream segment number.

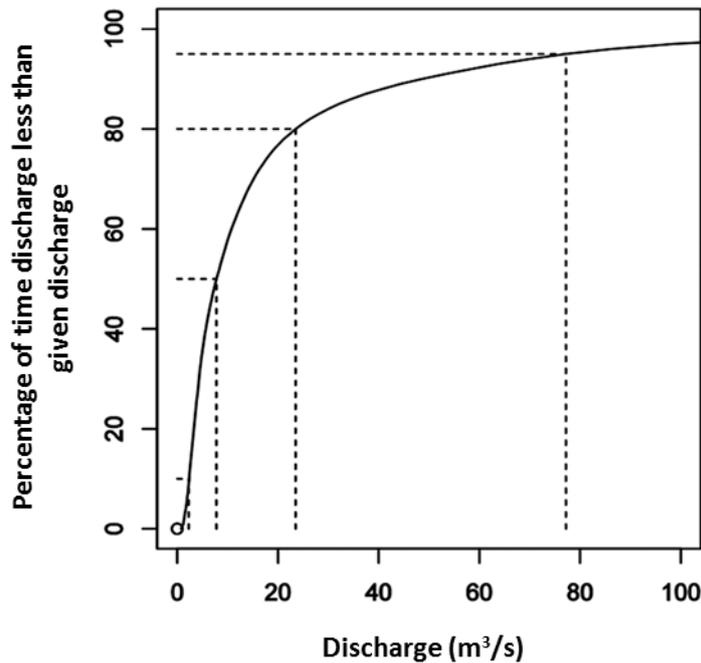


Figure 8-2: Flow duration curve of the Wairua River at Purua. Y-axis gives percentage of time that discharge is less than that shown on the x-axis. The dashed lines show four characteristic flow percentiles, that is, the 10%, 50%, 80%, and 95% flow percentiles (2.3, 7.8, 23.5, and 77.2 m³/s).

8.3 Analytical framework 1

Framework 1 (Figure 8-1) links catchment sediment loads to VC, LP, SSC, and DS through the sediment rating curve (SRC). The component relationships are:

- SSC vs discharge (SRC)
- VC vs SSC
- LP vs SSC
- DS vs SSC.

SSC links to catchment sediment load (L , t/yr) via the relation:

$$L = \sum_i^n L_i = \sum_i^n p_i C_i Q_i \quad (9)$$

where L_i is the sediment load carried in the i th discharge band, Q_i is the mean discharge in each band, p_i is the proportion of time that the water discharge is in the i th band, and C_i is the SSC associated with the given discharge (and is obtained from the SRC).

8.3.1 Sediment rating curve and VC vs SSC relationship

As detailed in Section 3, the SRC is the relationship between SSC and water discharge and is determined from paired measurements of SSC and discharge. Figure 8-3 shows the SRC for the Wairua River at Purua as an example of sites where measurements of SSC have been made. At non-measured sites the SRC would need to be estimated with the SRC parameter estimation models developed in Section 3.

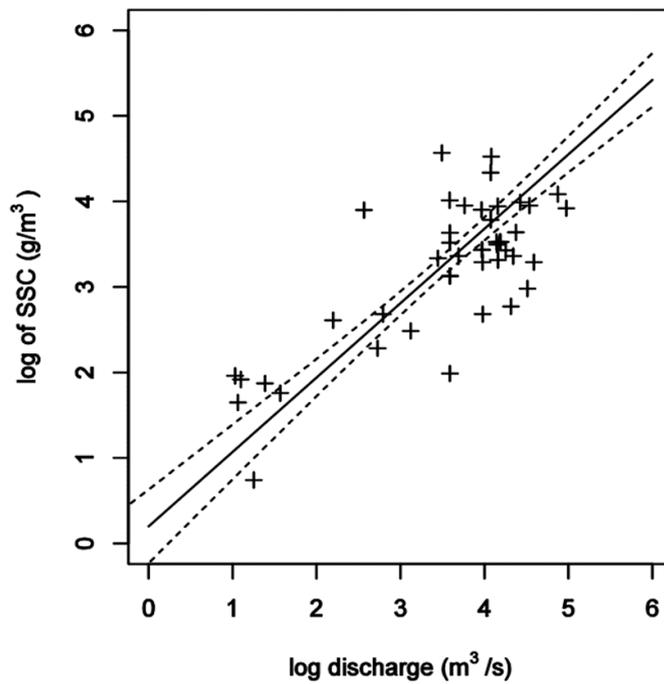


Figure 8-3: Log-log plot of SSC versus discharge for the Wairua River at Purua. Straight line is best-fit linear regression ($y=0.87x+0.2$). Dashed lines show the uncertainty of the regression line (plus or minus one standard error) as derived in Appendix I.

The relationship between VC and SSC is discussed in Section 5 and is typically represented by a power-law function. Figure 8-4 shows the example of the Wairua River at Purua. At non-measured sites the VC vs SSC relationship would need to be estimated with the models developed in Section 5.

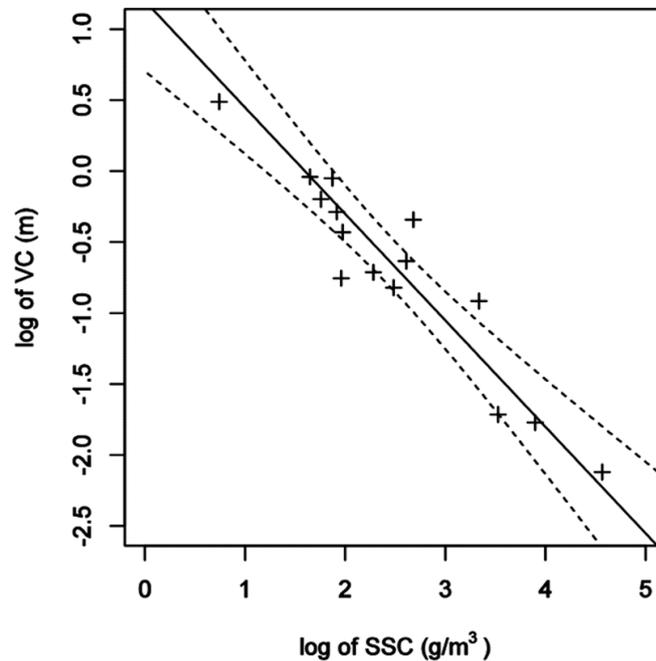


Figure 8-4: Log-log plot of VC versus SSC for the Wairua River at Purua. Straight line is best fit linear regression ($y=-0.75x+1.2$). Dashed lines show the uncertainty of the regression line (plus or minus one standard error).

8.3.2 Characteristic percentiles of SSC and VC

Characteristic exceedance percentiles of SSC are determined from the sediment rating curve and the matching characteristic percentiles of discharge on the flow duration curve. In the Wairua at Purua example, the 10%, 50%, 80%, and 95% percentile flows are extracted from the flow duration curve (Figure 8-2), and the SSC values at those flows (2.5, 7.3, 19.0, and 53.6 g/m³) that are extracted from the sediment rating curve (Figure 8-3) have the same exceedance percentiles. Appendix J shows mathematically why this is so.

The characteristic percentiles of VC are similarly determined from the characteristic percentiles of SSC using the relationship between VC and SSC. In the Wairua at Purua example, from Figure 8-4 the 10%, 50%, 80%, and 95% VC percentiles are 1.6, 0.75, 0.36, and 0.17 m, respectively.

Uncertainty in the VC percentiles derives from two sources. The first is from the uncertainty of the relationship between VC and SSC, shown by the dashed lines in Figure 8-4. The second is from the uncertainty of the underpinning characteristic percentiles of SSC. We can express this mathematically as follows.

Let the relationship between VC and SSC be represented by:

$$WC = F(SC) \tag{10}$$

where WC is log of VC, SC is log of SSC, and F is the function (usually linear) that relates SC to WC .

Dymond and Christian (1982) showed that the error in WC is given by:

$$\Delta(WC) = \Delta F + \frac{\partial(F)}{\partial(SC)} \Delta(SC) \quad (11)$$

where Δ represents a small change.

Assuming the two sources of error are independent then the variance of the VC percentiles may be estimated from:

$$Var[WC] = Var[F] + \frac{\partial(F)}{\partial(SC)}^2 Var[SC] \quad (12)$$

The first term in equation (12) gives % standard errors of 40%, 20%, 20%, and 40% for the four characteristic VC percentiles of the Wairua River (Appendix I). The second term in Equation (12) gives % standard errors of 30%, 15%, 10%, and 10%. $Var[SC]$ is estimated using Appendix I.

8.3.3 Determining absolute VC after a reduction in sediment load

The following approach is used to determine what absolute VC will result from a given reduction in sediment load. It would be applied when a water clarity target was expressed as a specific value of VC with a specific exceedance percentile (for example, a target of 1.25 m for the median VC). The steps are as follows:

- Choose a characteristic percentile of VC. We will choose the 50% percentile, that is, the median, as an example, but any percentile may be used.
- For a range of sediment load reductions, estimate the resulting VC from the relationship between VC and SSC (assuming that the % reduction in SSC is the same as the % reduction in sediment load).

Results for the median VC following this approach are shown in Figure 8-5 for the Wairua at Purua example. To give an idea of possible reductions, as predicted by the SedNetNZ model (Dymond et al. 2016) reforestation of all pasture in the Wairua catchment would achieve a 70% reduction of sediment load in the river, while extensive implementation of farm plans (i.e., soil conservation) would achieve a 50% reduction. These two alternative treatments would increase the median VC from its existing value of 0.75 ± 0.2 m to 1.25 ± 0.5 m and 1.75 ± 0.7 m, respectively.

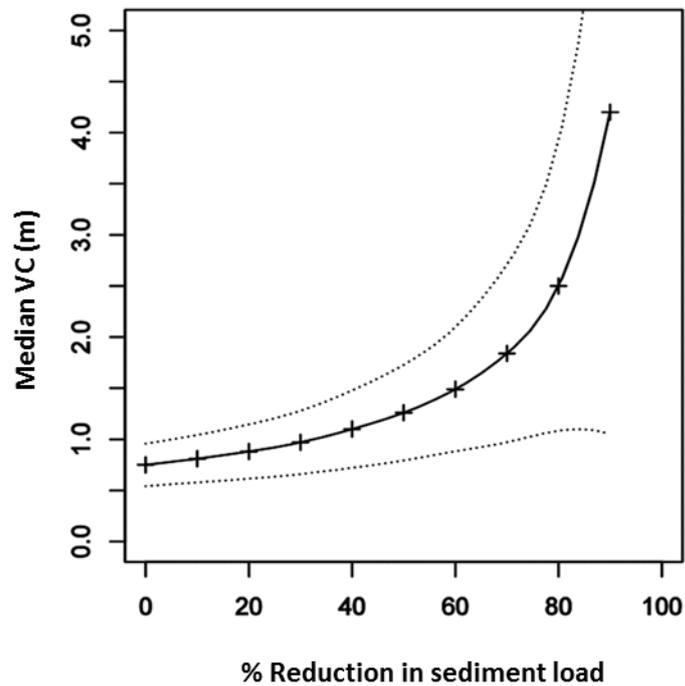


Figure 8-5: Median VC as a function of % reduction in sediment load for the Wairua River at Purua. The dashed lines represent the uncertainty of the estimated median (plus or minus one standard error) as given by equation (12).

8.3.4 Determining the change in VC after a reduction in sediment load

The following approach is used to determine how much VC would change after a reduction in sediment load. It would be applied when a water quality target was expressed as a *change* in VC at a given exceedance percentile rather than as an *absolute* VC target (for example, a target of an increase of 0.5 m for the median VC). In this case the steps are:

- Use the same characteristic percentile of VC as in the previous section (i.e., the median).
- For a range of sediment load reductions estimate the resulting *change in* VC from the relationship between VC and SSC (again assuming that the % reduction in SSC is the same as the % reduction in sediment load).

The results for the median VC following this approach are shown in Figure 8-6 for the Wairua at Purua example. Note that the changes in VC (e.g., a clarity increase of 0.5 ± 0.15 m for a 50% reduction in sediment load) are much the same as indicated from the approach used in Section 8.3.3, but the uncertainties are smaller. This is because only the slope in the linear function (F) of the first term in equation (12) needs to be considered. Thus predicting the sediment load reduction required to induce a given increase in VC can be done more reliably than predicting the load change required to meet a specific clarity target.

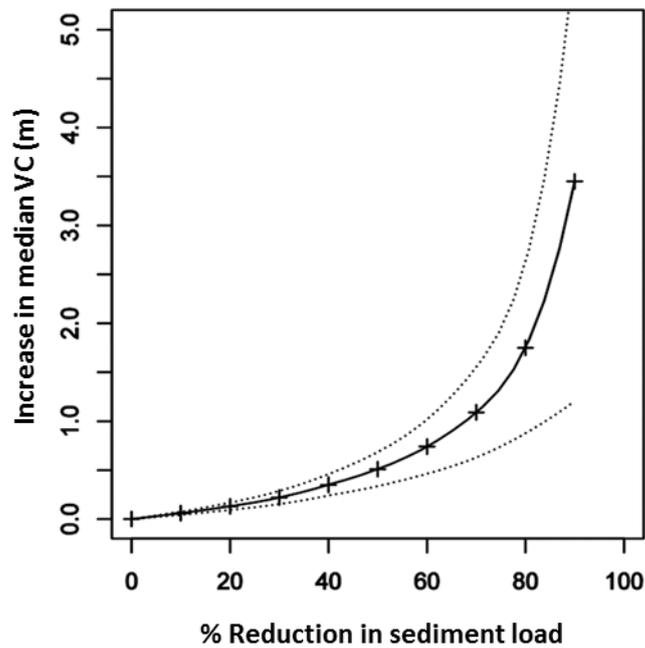


Figure 8-6: Increase in median VC as a function of % reduction in sediment load for the Wairua River at Purua. Dashed lines represent the uncertainty of the estimated median (plus or minus one standard error), which is smaller than that in Figure 8-5.

8.3.5 Characteristic percentiles of light penetration (euphotic depth)

In the following, we index light penetration (LP) by the euphotic depth (ED). The ED is the depth in a water body at which zero net photosynthesis (i.e., carbon dioxide uptake by photosynthesis minus carbon dioxide release by respiration) occurs, and it aligns with the depth where the light intensity is 1 per cent of that at the surface (Allaby 2004).

Characteristic percentiles of ED are determined in much the same way as for VC; the only difference is that the relationship between ED and SSC is used rather than the VC-SSC relationship. In the case-example site at Wairua at Purua, ED was not measured but can be reliably estimated off turbidity³². ED so estimated is related to SSC in Figure 8-7. Using values from this relationship, the 10%, 50%, 80%, and 95% ED percentiles are 2.3, 1.47, 0.98, and 0.64 m, respectively.

³² Euphotic depth may be estimated as 4.6 divided by the square root of turbidity – Davies-Colley and Nagels (2008).

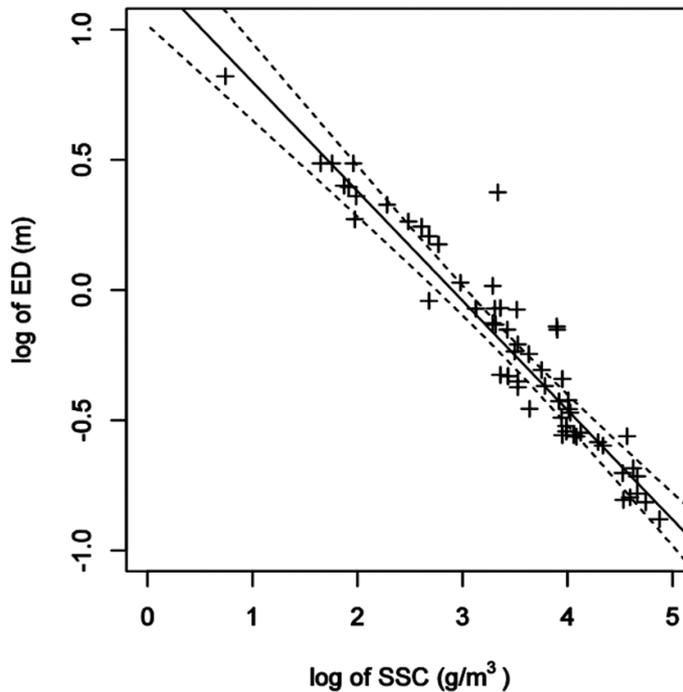


Figure 8-7: Log-log plot of ED versus SSC for the Wairua River at Purua. Straight line is best fit linear regression ($y=-0.42x+1.22$). Dashed lines show the uncertainty of the regression line (plus or minus one standard error).

The uncertainty of ED percentiles derive from two sources. The first is from the uncertainty of the relationship between ED and SSC (shown by the dashed lines in Figure 8-7). The second is from the uncertainty of estimated values of SSC at the characteristic percentiles. We can express this mathematically as follows:

Let the relationship between ED and SSC be represented by:

$$\log ED = G(SC) \quad (13)$$

where $\log ED$ is log of euphotic depth, SC is log of SSC, and G is the function (usually linear) that relates SC to $\log ED$.

The error in $\log ED$ is given by

$$\Delta(\log ED) = \Delta G + \frac{\partial(G)}{\partial(SC)} \Delta(SC) \quad (14)$$

where Δ represents a small change.

Assuming the two sources of error are independent then the variance of the ED percentiles may be estimated from

$$\text{Var}[\log ED] = \text{Var}[G] + \frac{\partial(G)}{\partial(SC)}^2 \text{Var}[SC] \quad (15)$$

The first term in Equation (15) gives % standard errors of 15%, 10%, 5%, and 5% for the four characteristic ED percentiles of the Wairua River example given above (see Appendix I). The second term in Equation (15) gives % standard errors of 15%, 10%, 5%, and 5%. $Var[SC]$ is estimated using Appendix I.

8.3.6 Absolute ED and change in ED after a reduction in sediment load

The same approach as used for VC is used to determine how ED will change after a reduction in sediment load. Figure 8-8 illustrates the *absolute* median ED for the Wairua at Purua as its load changes: for example, a 50% reduction in sediment load will increase ED from 1.5 ± 0.2 m to 2.0 ± 0.4 m. Figure 8-9 shows the *change* in median ED: a 50% load reduction increases ED by 0.50 ± 0.05 m. The uncertainty is less in the latter case because only the slope in the linear function (G) of the first term in Equation (14) needs to be considered.

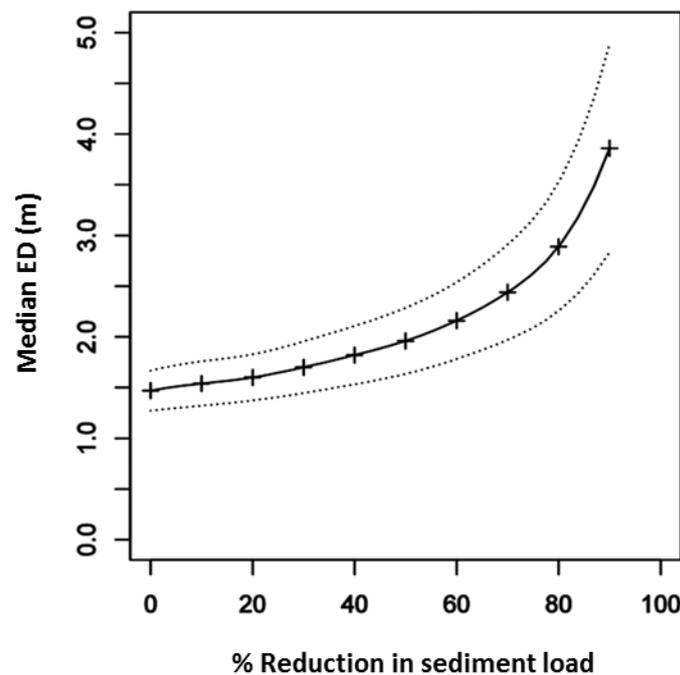


Figure 8-8: Median ED as a function of % reduction in sediment load for the Wairua River at Purua. Dashed lines represent the uncertainty of the estimated median (plus or minus one standard error) as given by Equation (15).

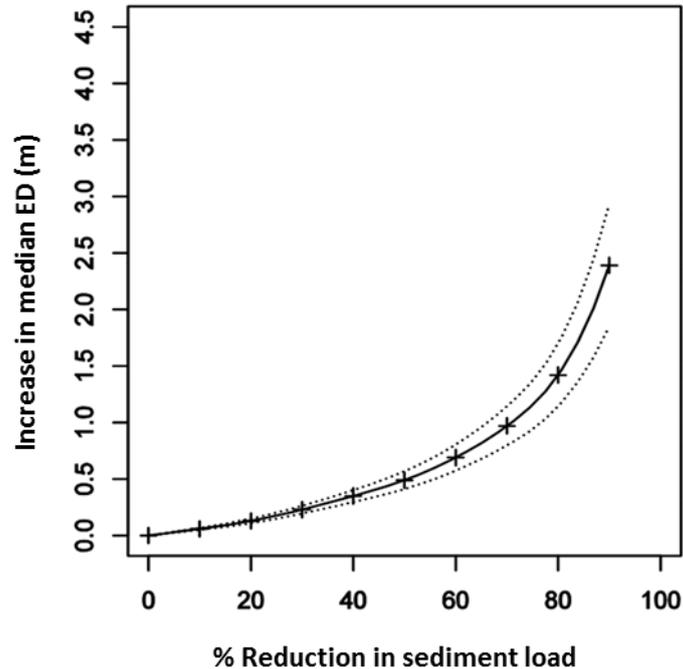


Figure 8-9: Increase in ED as a function of % reduction in sediment load for the Wairua River at Purua. Dashed lines represent the uncertainty of the estimated median (plus or minus one standard error), which is smaller than that in Figure 8-8.

8.3.7 Deposited Sediment (DS)

Fine sediment is often deposited from suspension on stream channel beds, and in the case of gravel beds it is deposited within the bed as well. The sediment deposited on the channel bed is highly variable in space and time and is difficult to characterise and predict. However, sediment deposited in the channel as measured by embeddedness (expressed as fine sediment mass per unit volume of water in channel gravel, g/m^3) has been found to be less variable in river reaches. Moreover, it has been hypothesised that embeddedness is controlled by the flood discharge at which bedload movement ceases. If so, then embeddedness is equal to the SSC of water at the time that bed movement of gravel ceases on the falling limb of a hydrograph (i.e., the “frozen bedload” hypothesis discussed in Section 7.3.5).

The data presented in Figure 7-14, which plots measured *SIS* (g/m^2 of channel) versus estimated embeddedness (g/m^3 of water in channel), does not support the “frozen bedload” hypothesis because there is no relationship observed between *SIS* and embeddedness (i.e., if the “frozen bedload” hypothesis was correct then it is expected that *SIS* is embeddedness times channel porosity times Quorer sampling depth). However, more data is required to exclude the possibility of measurement variation and to perform a thoroughly robust test on the “frozen bedload” hypothesis.

The analysis that follows for DS in Section 8.3.8 assumes that the “frozen bedload” hypothesis is valid. If further research confirms it as invalid, then another approach would be necessary. This is addressed in Section 8.3.9.

8.3.8 “Frozen bedload” hypothesis approach for DS

Clausen and Plew (2004) showed that for New Zealand rivers the flow at which gravel bedload stops moving is approximately equal to one quarter of the mean annual flood. If the relationship between SSC and discharge is given by:

$$C = \varphi(Q) \quad (16)$$

where C is SSC (g/m^3), Q is discharge (m^3/s), and φ is the sediment rating function, then embeddedness is given by:

$$em = \varphi\left(\frac{Q_{maf}}{4}\right) \quad (17)$$

where em is embeddedness and Q_{maf} is the mean annual flood in m^3/s .

The mean annual flood of the Wairua River at Purua is $198 \text{ m}^3/\text{s}$, hence the discharge at which gravel should stop moving on the falling limb of a flood hydrograph is approximately $50 \text{ m}^3/\text{s}$. Figure 8-3 (in natural logarithms) gives a SSC of $36 \text{ g}/\text{m}^3$ at a discharge of $50 \text{ m}^3/\text{s}$, so predicted embeddedness is 36 g of sediment per m^3 of pore water.

If we assume, as previously, that % reduction in SSC at any discharge is linearly related to % reduction in sediment load, then the embeddedness should be inversely related to the % reduction in sediment load. Figure 8-10 illustrates this for the Wairua at Purua example. The uncertainty in this prediction derives from two sources: (i) the uncertainty in estimating the discharge at which bedload ceases, and (ii) uncertainty from the assumptions which relate embeddedness to that discharge. Neither are known so uncertainty cannot be shown on Figure 8-10.

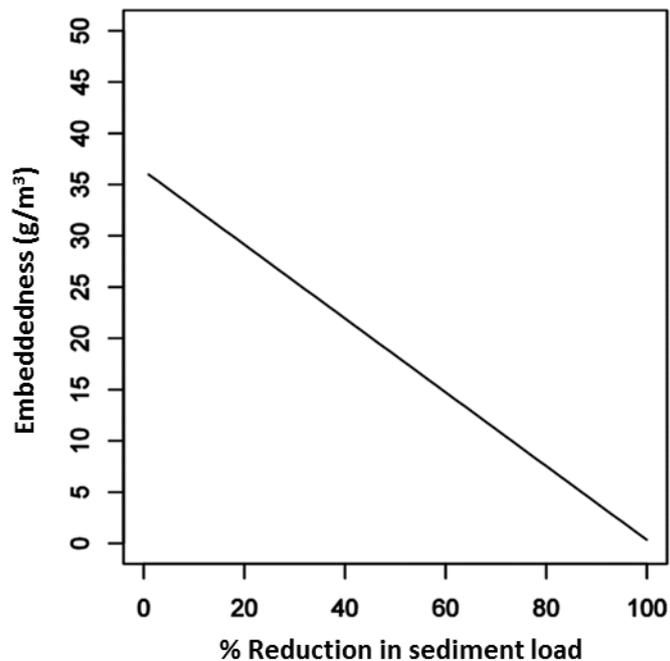


Figure 8-10: Decrease in embeddedness as a function of % reduction in sediment load for the Wairua River at Purua.

8.3.9 Alternative approach for DS

Should the “frozen bedload” hypothesis be confirmed to be invalid by further research, then a possible alternative approach could be to relate DS directly to sediment load and/or other catchment or channel characteristics. Unfortunately, our investigations in Section 7 found no such relationships that were significant. Thus, short of finding adequate relationships from further research, the framework would break down for DS.

8.4 Analytical framework 2

Framework 2 (Figure 8-11) links changes in catchment sediment sources directly to the optical ESVs (i.e., VC and LP) through the beam attenuation coefficient (BAC) and its ‘rating’ relationship with water discharge (BRC). As proposed by Elliott et al. (2013), the “load” of beam attenuation coefficient (or optical cross-section) may be considered analogous to the sediment mass load, and they refer to it as the Load of Optical Cross-section (LOCS). It offers a potentially simpler management tool for fine sediment where the sediment impacts are likely to be mainly on VC or LP.

For SSC and DS, linking through beam attenuation coefficient is not helpful because for these two attributes the impact of reduced sediment load always has to come through reduced SSC. Moreover, it is pointless relating reduced sediment load to SSC indirectly via the beam attenuation coefficient because this introduces additional steps and uncertainty. Thus framework 2 is not suitable where fine sediment effects on SSC and DS are important.

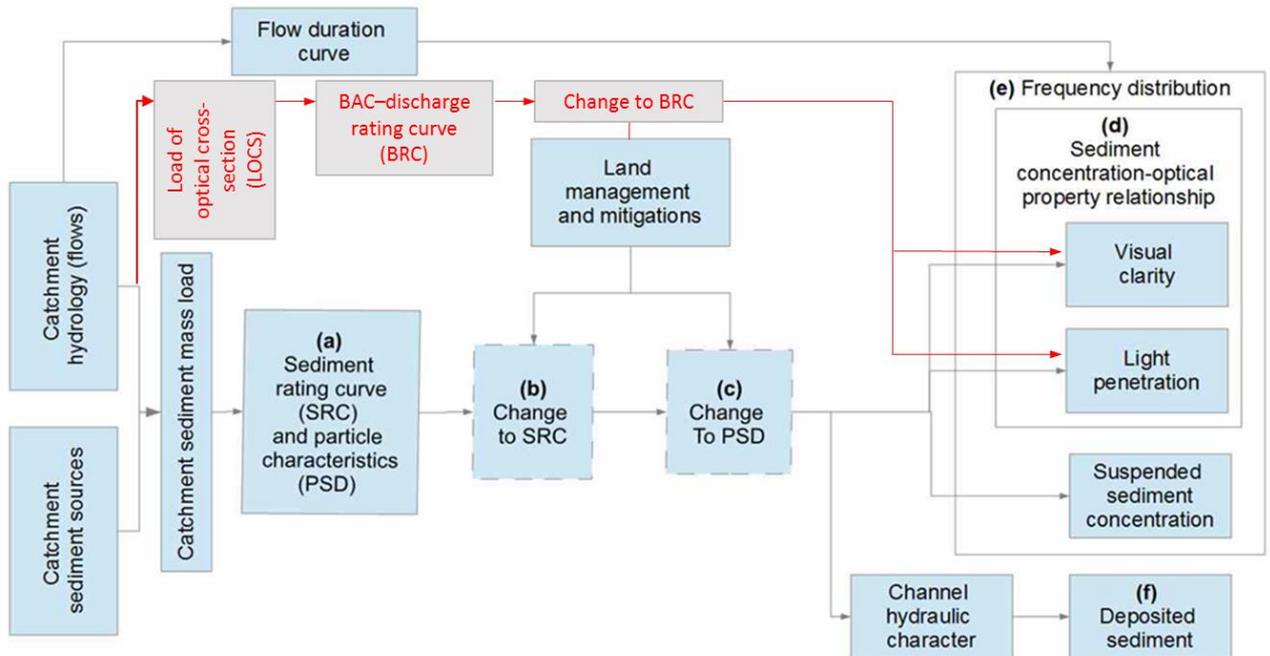


Figure 8-11: Framework 2, linking catchment load of optical cross-section (LOCS) with visual clarity and light penetration via relationship between beam attenuation coefficient (BAC) and water discharge. Framework 2 shown in red overlaid on framework 1.

The component relationships in framework 2 are:

- BAC vs discharge (BRC)
- VC vs BRC
- LP vs BRC.

BAC links to catchment load of optical cross-section (LOCS, t/yr) via the relation:

$$LOCS = \sum_i^n LOCS_i = \sum_i^n p_i B_i Q_i \quad (18)$$

where $LOCS_i$ is the load of optical cross-section carried in the i th discharge band, Q_i is the mean discharge in each band, p_i is the proportion of time that the water discharge is in the i th band, and B_i is the BAC associated with the given discharge (and is obtained from the BRC).

How LOCS relates to the actual catchment sediment load is unclear as yet (e.g., Elliott et al. 2013), thus framework 2 is perhaps best regarded as a potential framework rather than one that is immediately available.

8.4.1 Beam attenuation coefficient and discharge relationship

In the same way as a sediment rating curve is derived, the beam attenuation coefficient rating curve (BRC) is obtained from paired measurements of VC and discharge (with the BAC estimated as 4.8 divided by VC; Davis-Colley and Smith 2001). Figure 8-12 shows the BRC for the Wairua at Purua example site. Like the sediment rating curve, the BRC relationship takes a power form (with the Wairua curve being $BAC = 1.99 Q^{0.65}$). However, unlike SRCs (as in Section 3), so far there has been no work done to develop empirical models so that BRC parameters can be predicted at unmeasured sites.

BAC values read from the BRC can be converted to VC simply via the relation $VC = 4.8/BAC^{33}$.

³³ The equivalent conversion in log space is $\log_e(VC) = 1.57 - \log_e(BAC)$. Clearly, for framework 2 the BAC could be done away with by simply relating VC to discharge. However, the utility of BAC is that it increases with sediment load and is analogous to SSC, whereas VC is inversely related.

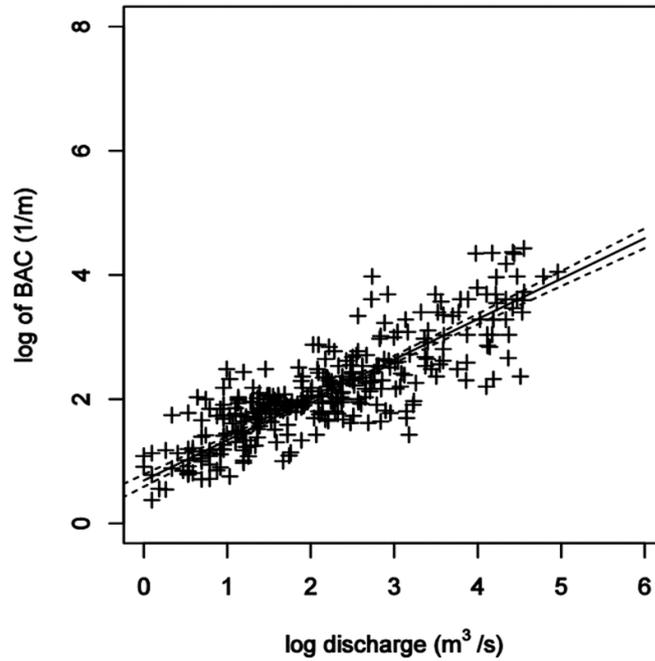


Figure 8-12: Log-log plot of BAC versus discharge for the Wairua River at Purua. Straight line is best fit linear regression ($y=0.65x+0.69$). Dashed lines show the uncertainty of the regression line (plus or minus one standard error).

8.4.2 Characteristic percentiles of VC in framework 2

In framework 2, characteristic exceedance percentiles of BAC (and hence VC) are determined via the BRC by matching characteristic percentiles of discharge on the flow duration curve. Thus for the Wairua at Purua example, Figure 8-2 and Figure 8-12 give the 10%, 50%, 80%, and 95% BAC percentiles as 3.43, 7.58, 15.52, and 33.62 1/m, respectively, and the equivalent VC values are 1.4, 0.63, 0.31, and 0.14 m.

Uncertainty in the VC percentiles derives from two sources. The first is from the uncertainty of the relationship between VC and BAC, which is assumed negligible (Davis-Colley and Smith 2001). The second is from the uncertainty of the characteristic percentiles of BAC. We can express this mathematically as follows.

Let the relationship between VC and BAC be represented by:

$$\log VC = H(\log BAC) \quad (19)$$

where $\log VC$ is log of VC, $\log BAC$ is log of beam attenuation coefficient, and H is the function (usually linear) that relates $\log BAC$ to $\log VC$.

The error in $\log VC$ is given by:

$$\Delta(\log VC) = \Delta H + \frac{\partial(H)}{\partial(\log BAC)} \Delta(\log BAC) \quad (20)$$

where Δ represents a small change.

Assuming the two sources of error are independent then the variance of the VC percentiles may be estimated from:

$$Var[\log VC] = Var[H] + \frac{\partial(H)}{\partial(\log BAC)}^2 Var[\log BAC] \tag{21}$$

The first term in Equation (21) gives % standard errors of 0% for the four characteristic VC percentiles of the Wairua River (this is because there is negligible uncertainty of *H*). The second term gives % standard errors of 7%, 5%, 6%, and 10%.

8.4.3 Determining change in VC after reduction in sediment load in framework 2

Pursuing the LOCS concept to its full intent, framework 2 would operate by linking changes in catchment LOCS to changes in VC. However, since the LOCS remains ill-defined in terms of sediment load constituents, it remains necessary to link it empirically with actual catchment sediment load via the relationship between BAC and SSC. Figure 8-13 shows this relationship for the Wairua at Purua example. This relationship then enables the changes in VC associated with changes in sediment load to be calculated, following the same procedure as detailed in Section 8.3.4.

Figure 8-14 shows absolute median VC as a function of % reduction in sediment load for the Wairua at Purua. This shows, for example, that a 70% reduction of sediment load would increase median VC from 0.65 m to 1.55 ± 0.40 m. Considering only the change in median VC, Figure 8-15 shows that a 70% reduction in sediment load would increase VC by 0.4 ± 0.35 m. The uncertainty is slightly smaller with this because only the error due to estimating the change in BAC corresponding with change in sediment load needs to be considered (not the uncertainty in the initial median VC under the existing sediment load).

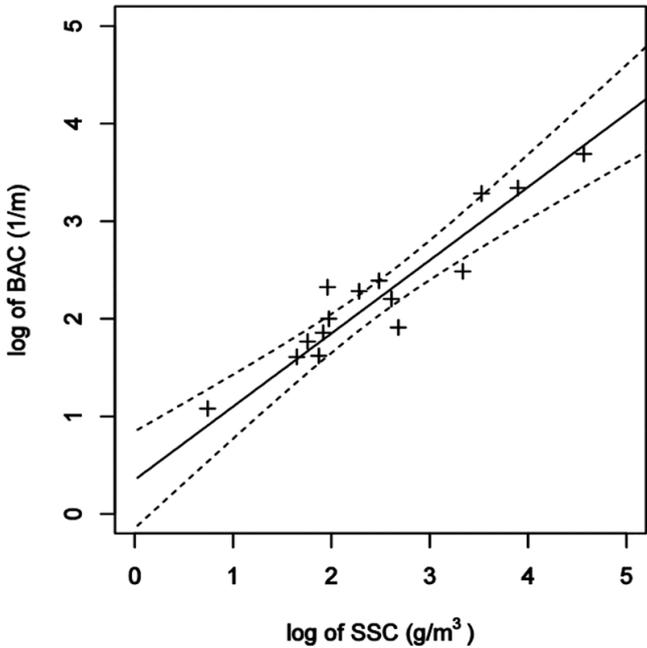


Figure 8-13: Log-log plot of BAC versus SSC for the Wairua River at Purua. Straight line is best fit linear regression ($y=0.35+0.75x$). Dashed lines show the uncertainty of the regression line (plus or minus one standard error).

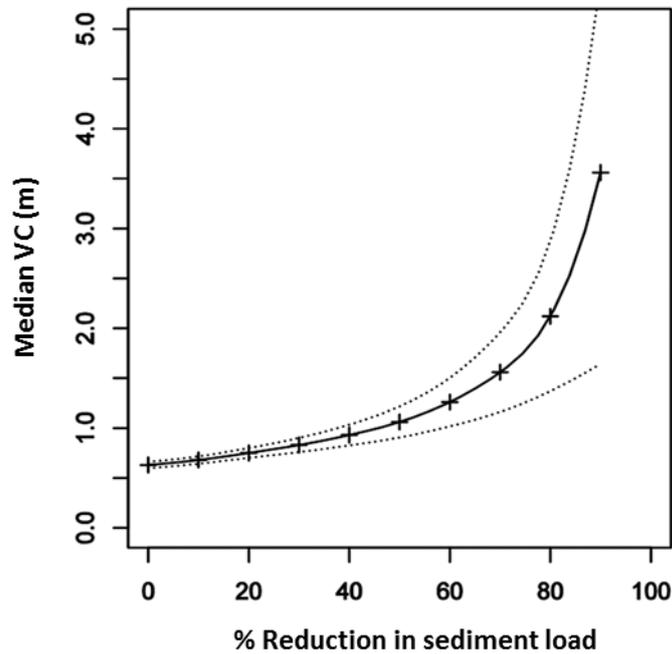


Figure 8-14: Median VC as a function of % reduction in sediment load for the Wairua River at Purua (in framework 2). Dashed lines represent the uncertainty of the estimated median (plus or minus one standard error) as given by equation (21) with the addition of an error due to estimating the change in BAC corresponding with change in sediment load.

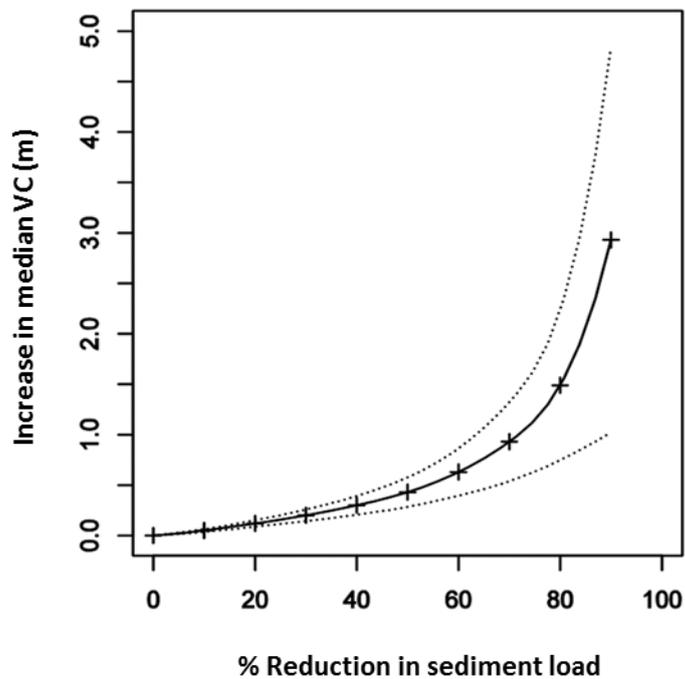


Figure 8-15: Increase in median VC as a function of % reduction in sediment load for the Wairua River at Purua (under framework 2). Dashed lines represent the uncertainty of the estimated median (plus or minus one standard error), which is slightly smaller than that in Figure 8-14.

8.4.4 Characteristic percentiles of ED in framework 2

Characteristic percentiles of ED in framework 2 are determined similarly to those for VC except that an empirical relationship is required between BAC and ED. This relationship is shown for the Wairua at Purua in Figure 8-16 and provides ED values of 2.35, 1.58, 1.10, and 0.75 m corresponding to 10%, 50%, 80%, and 95% percentiles for BAC and discharge.

As previously, the uncertainty of the ED percentiles derives from the uncertainty of the relationship between ED and BAC and from the uncertainty of the characteristic percentiles of BAC. Assuming these two sources of error are independent then the variance of the ED percentiles may be estimated from:

$$Var[\log ED] = Var[U] + \frac{\partial(U)}{\partial(BAC)}^2 Var[BAC] \quad (22)$$

where U is the function (usually linear) that relates logBAC to logED. The first term in Equation (22) gives % standard errors of 3.6%, 2.4%, 3.0%, and 4.8% for the above four characteristic ED percentiles of the Wairua River at Purua. The second term gives % standard errors of 3.4%, 2.5%, 2.9%, and 4.8%.

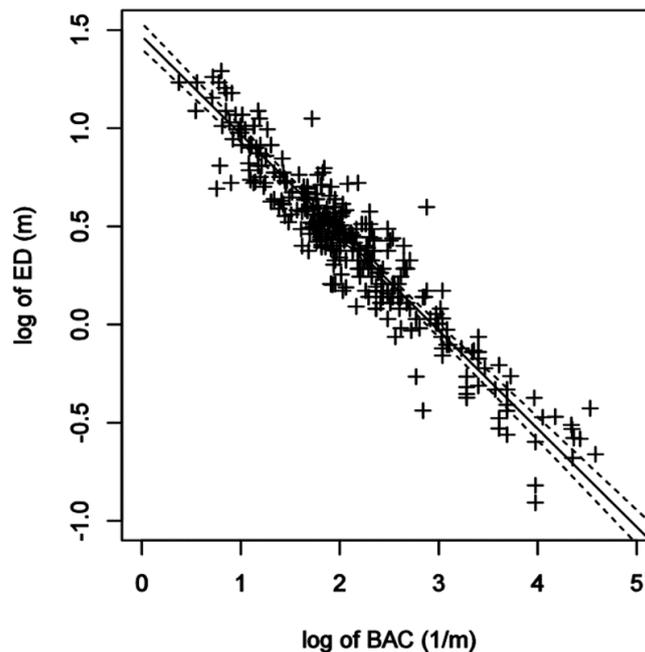


Figure 8-16: Log-log plot of ED versus BAC for the Wairua River at Purua. Straight line is $y=1.47-0.5*x$. Dashed lines show the uncertainty of the regression line (plus or minus one standard error).

8.4.5 Change in ED after reduction in sediment load in framework 2

Changes in ED resulting from sediment load reduction are determined by converting changes in BAC (with load reduction) to changes in ED using the BAC-ED relationship (e.g., Figure 8-13). Following through with the Wairua at Purua case, Figure 8-17 shows, for example, that a 70% reduction in sediment load would increase the ED from 1.55 ± 0.05 m to 2.45 ± 0.20 m. Alternatively, Figure 8-18

shows that a 70% reduction in sediment load would change the ED by 0.90 ± 0.13 m (again, the uncertainty is slightly smaller in the latter case because the initial ED is not required to be estimated).

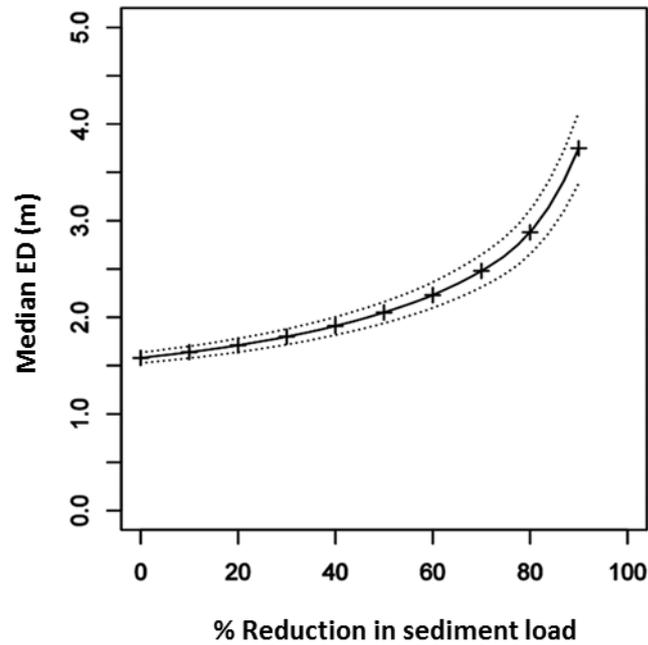


Figure 8-17: Median ED as a function of % reduction in sediment load for the Wairua River at Purua. Dashed lines represent the uncertainty of the estimated median (plus or minus one standard error) as given by equation (22) plus uncertainty due to estimating change in *BAC* from change in sediment load.

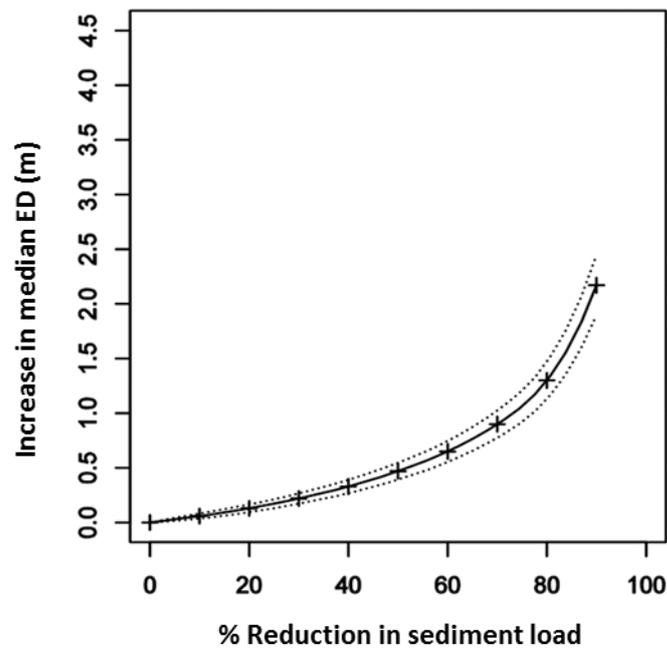


Figure 8-18: Increase in median ED as a function of % reduction in sediment load for the Wairua River at Purua (under framework 2). Dashed lines represent the uncertainty of the estimated median (plus or minus one standard error), which is slightly smaller than that in Figure 8-17.

8.4.6 Discussion on framework 2

The load of optical cross-section (LOCS) that underpins framework 2 is determined by those parts of the suspended and dissolved loads of a stream that have the greatest impact on the optical properties of the water. However, how LOCS relates to the catchment sediment load, and how BAC relates to SSC, remain ill-defined at a national level. Moreover, there is no information at all on what influences BAC-discharge ratings and how they might change when LOCS and/or catchment sediment load changes (as we investigated, for example, for SRCs in Section 4). Thus even though framework 2 might offer the advantage of a potentially more direct link between catchment management and the optical ESVs, the general need to also relate catchment management to SSC and deposited sediment, and the currently poor state of knowledge around what LOCS physically is and what controls it, means that it cannot currently provide a useful tool for national/regional assessment.

9 Accuracy of analytical frameworks and their fitness-for-purpose

9.1 Introduction

As detailed in Section 8, with framework 1 percentiles of flow are converted to percentiles of ESVs using SSC as an intermediate variable. The impacts of reduced sediment loads, through soil conservation / erosion mitigation, on ESVs (SSC, VC, BAC, and DS) at a given percentile are thus able to be determined through the associated reduction in SSC. A methodology for determining the accuracy of each step in this analytical framework has also been presented in Section 8, rather than here in this section, so that the framework and uncertainties are presented together. Propagating the errors associated with each framework step provides the error bars around the plots of median ESV value versus % reduction in sediment load that are given in Section 8.

In the course of this error propagation we observed that plots of *change* in ESV were more accurate than plots of the absolute value of ESV. This is because in the former case it is not necessary to estimate the starting ESV before the sediment load is changed. Thus sediment attribute targets couched as a change (or improvement from a starting state) can therefore be determined more robustly than targets couched in absolute values.

Similarly, with framework 2 percentiles of flow are converted to percentiles of ESVs using the beam attenuation coefficient (BAC, which may be regarded as an optical analogy to SSC). The impacts of reduced sediment loads through soil conservation on ESVs, and their associated uncertainties, were thus also able to be determined through the associated reduction in BAC.

In this section, we firstly compare framework 1 and framework 2 in terms of robustness of prediction of ESV change for a given change in catchment sediment load. We then examine the utility of the two frameworks for national-scale assessment of what level of change in catchment sediment load might be required to meet a given set of targets around the sediment ESVs.

9.2 Robustness of analytical frameworks

We compare framework 1 and framework 2 in terms of robustness of prediction of ESV change for a given change in catchment sediment load for two cases:

- At a typical site where data is available to “populate” the frameworks with locally-measured data, as represented by the Wairua at Purua example.
- At site with no measurements, where we are dependent on estimating the component relationships of the two frameworks off empirical predictors.

In each case, we arbitrarily graded robustness in terms of the uncertainty in median ESV prediction associated with a 50% reduction in sediment load (Table 9-1).

Table 9-1: Definitions for robustness of prediction of median ESV value associated with a 50% reduction in catchment sediment load. The robustness grading is based on the uncertainty (\pm one standard-error) in the predicted ESV. For example, a VC prediction accurate to ± 0.3 m would be graded as “moderately robust”. Limits of robustness scale were defined arbitrarily.

ESV	Highly robust	Moderately robust	Weakly robust
VC (m)	< 0.1	0.1 – 0.4	> 0.4
ED (m)	< 0.1	0.1 – 0.4	> 0.4
DS (g/m ³)	< 0.5	0.5 – 2.0	> 2.0
SSC (g/m ³)	< 0.5	0.5 – 2.0	> 2.0

9.2.1 Robustness of predicted change in ESVs at sites with measurements

Table 9-2 shows the results of this comparison for the Wairua at Purua example. They are based on the observed relationships and results provided in Section 8.

In framework 1, the methods are either moderately or highly robust for VC change, ED and ED change, SSC and SSC change, and could confidently be used for setting sediment attribute targets. For VC and DS, the methods are weakly robust and could not be used confidently for setting targets.

In framework 2, the robustness for methods VC change, ED and ED change, are the same as framework 1, that is, either moderately or highly robust. However, for VC the robustness in framework 2 is moderate rather than weak as in framework 1. The main reason for this is that the Wairua River at Purua site has more measurements of VC than SSC. This may not be the case at other measurement sites. There are no methods in framework 2 for predicting DS, SSC, and SSC change (hence no robustness grading is provided).

Table 9-2: Robustness of methods used to predict sediment ESVs for frameworks 1 and 2 for the Wairua at Purua example where sediment attributes are measured regularly. Robustness is estimated from the uncertainty of prediction and can be either highly robust, moderately robust, or weak. There is no method for DS and SSC in framework 2.

ESV	Framework 1	Framework 2
VC	weak	moderate
VC change	moderate	moderate
ED	moderate	moderate
ED change	high	high
DS	weak	
SSC	moderate	
SSC change	high	

9.2.2 Robustness of predicted change in ESVs at sites with no measurements

At sites where ESVs are not measured and have to be predicted through regional regression relationships, the analytical frameworks will be less robust. For example, the first term in equation (12) will be larger because the function F is estimated from a regional regression rather than from measurements at the site in question. If the Wairua at Purua had to use ESVs estimated from regional regressions, then the % standard error at the median discharge would increase from 20% to 44%. The second term in equation (12) would also be considerably larger because the sediment rating curve would be estimated from a regional regression rather than from the measurements at Wairua at Purua. The % standard error at the median discharge would increase from 15% to over 200%. There would also be an additional term to be considered due to the uncertainty associated with estimating the flow duration curve (i.e., the third term in equation (23):

$$Var[WC] = Var[F] + \frac{\partial(F)^2}{\partial(SC)} Var[SC] + \frac{\partial(F)^2}{\partial(SC)} \frac{\partial(SC)^2}{\partial(Q)} Var[Q] \quad (23)$$

where Q is log of discharge. This third term would add about 25% to the % standard error of WC .

Table 9-3 summarises the uncertainties in predicted values of parameters for the component relationships in the analytical frameworks, as estimated in Sections 3 through 7. The largest uncertainties are the large factorial errors associated with predicting the SRC a -parameter (\times/\div 2.29) and with predicting the embeddedness (S/S) via the SSC predicted at $\frac{1}{4}$ the mean annual flood discharge (\times/\div 3.19; setting aside that the “frozen bedload” hypothesis appears questionable).

Table 9-3: Uncertainties in component relationships for analytical frameworks, as derived in previous sections.

ESV	Parameter	Uncertainty	Reference
Flow duration curve	Q/Q_{mean} at given exceedance percentile	\times/\div 1.25 (\pm 25%)	Section 9.2
SRC ($C = aQ^b$)	a	\times/\div 2.29	Section 3.4.2
“	b	\pm 0.38	“
VC vs SSC ($VC = aC^b$)	a	\times/\div 1.44 (\pm 44%)	Section 5.3.1 Error! Reference ource not found.
“	b	\pm 0.12	“
S/S	C at $\frac{1}{4} Q_{\text{maf}}$	\times/\div 3.19	Section 7.3.5

Table 9-4 shows the expected robustness of ESV predictions where regional regressions that have the uncertainties listed in Table 9-3 are used to estimate ESVs. For both VC and ED, the impact of larger terms and an additional term in equations (12) and (15) renders both frameworks only weakly robust. Framework 1 is also only weakly robust for SSC. However, for VC change and ED change, only the slope uncertainty contribution in the first term in the variance equations needs to be considered, and the methods are still moderately robust. Likewise the method for SSC change is still highly robust.

Hence, when using regional regressions for predicting ESVs, the methods for VC change, ED change, and SSC change could all be confidently used for setting ESV targets.

Table 9-4: Expected robustness of methods used to predict sediment ESVs for frameworks 1 and 2 for sites where sediment attributes are not measured. Robustness is estimated from the uncertainty of prediction and can be either highly robust, moderately robust, or weak. There is no method for DS and SSC in framework 2.

ESV	Framework 1	Framework 2
VC	weak	weak
VC change	moderate	moderate
ED	weak	weak
ED change	moderate	moderate
DS	weak	
SSC	weak	
SSC change	high	

9.3 Utility of frameworks for national-scale assessment of change in sediment load to meet sediment ESV targets

A key question in adopting any operational sediment management framework is the extent, at the national level, that catchment sediment loads might be required to be reduced to meet targets set around the sediment ESVs. Answering this question requires at least four components:

- a set of ESV target values, likely defined as linked thresholds and exceedance percentiles
- a methodology to translate the ESV target values to changes in sediment loads
- a GIS analysis tool that combines the above two components and delivers summary maps and tables on the required changes in sediment yields
- an analysis of the practicality and cost of implementing the indicated reductions in sediment load.

This study is the subject of the second component only, and while work to develop ESV targets is beginning as this report is being completed, there is, as yet, no ESV target data with which to “feed” into either framework 1 or framework 2. Thus in this section, we focus on summarising the utility of frameworks 1 and 2 for their intended purpose.

9.3.1 Framework 1

In summary:

- Framework 1 provides a complete mathematical framework that potentially addresses all four ESVs.
- In practice, empirical predictors have been developed in this study for the relationships between catchment sediment load, SSC, VC, and LP, but no effective relationship has been found that links DS to SSC or sediment load. While the “frozen bedload” hypothesis can

easily be included in framework 1 to relate sediment load to DS, there so far appears little basis for this in field data. Thus while it is anticipated that further research might improve the DS situation, the present state of knowledge is that DS targets cannot yet be confidently translated into a reduction in sediment load.

- By making the assumption that changes in sediment load directly relate to changes in SSC, and using assessed uncertainty in our empirical predictors, we find that framework 1 is only weakly robust in its ability to link changes in catchment sediment loads to absolute values of SSC, VC, and LP. However, framework 1 appears to be acceptably robust in linking proportional changes in sediment load to *changes* in SSC, VC, and LP. Thus the robustness of the application of the framework will be strongly influenced by what choices are made in regard to ESV target definition.
- Our research for this study has shown that the assumption that a change in sediment load causes a proportional shift in the sediment rating curve (and hence a proportional change in SSC) does not always hold at the catchment scale, because of catchment non-uniformity. However, for a national assessment it would be a reasonable approximation to assume uniform catchment responses.

9.3.2 Framework 2

In summary:

- Framework 2 provides an expedient conceptual link between optically-important constituents of the sediment load and VC and LP.
- However, at present the only way it can be linked with catchment sediment loads is via a relationship between BAC and SSC. This rather defeats its intended purpose, and renders its predictive robustness no better than that of framework 1.
- It also has no link with DS and little is known about the stability of the key linking relationship between BAC and discharge.

9.3.3 Conclusion and recommendations

Based on the above, at least for national-scale assessments, we conclude that framework 1 provides a workable approach, although it cannot be used with any confidence to assess the implications of DS targets. By comparison, framework 2 remains work-in-progress and has more limited utility. Once supplied with target values and exceedance percentiles for the ESVs, our recommendations for application of framework 1 would be to:

- Embed framework 1 in a GIS application that accesses pre-calculated results of the various parameter-prediction models developed in this study (e.g., pre-calculated into fields on the REC2 national stream network).
- Use the application to map where across the REC2 network existing estimated catchment loads cause ESV targets to be exceeded.
- Use the application to integrate regionally and nationally the reduction in sediment load required to shift the ESVs to meet their respective targets everywhere. This might best be done at the outlets of all catchments draining to estuaries or the open coast.

10 Framework for further work

10.1 Introduction

The purpose of this section is to detail a framework for further research that develops more sophisticated and accurate methods for relating ESVs to catchment loads. It includes showing how new data is needed to improve methods, test new ideas, and test assumptions. It is also required to estimate the amount of new data needed, including the number of sites, sampling methods and frequencies, and considerations for site location.

The approach adopted for this task was to run a project-team workshop³⁴ to review the results of the present study, identify shortcomings in the data used and the logical links within the working ESV-Load analytical frameworks (Section 8), and to formulate the framework for further research.

Results from the workshop are summarised in Table 10-1 and are detailed below, generally under the topics/questions of the current study.

10.2 Further research needs by topic

10.2.1 SRCs

As summarised in Section 3, we consider the Random Forest models to be reasonable predictors of the SRC parameters given the substantial natural data-scatter observed on SRCs. Several things could be improved upon, however. The first is simply to improve the geographic coverage of SRC datasets. As indicated from Figure 3-1, Northland and Marlborough are particularly data-sparse. We recommend also that new sites are established where sediment yield is (or may be in future) impacted by land use, since this will provide improvements where managing to sediment limits will be most useful. For example, while there are few sites in Fiordland's pristine terrane, it is unlikely that efforts will be directed at reducing its sediment exports (which are already very low considering its high rainfall and steep terrane (Hicks et al. 2011)).

The recommended sampling technique for new SRC sites is to use auto-samplers scheduled to provide an adequate number of samples over the full (or a targeted) range of the hydrograph (Hicks and Gomez 2016). Manual sampling during flood events, using conventional cableways and bridges³⁵ and particularly at night, has been all but rendered impractical by current health and safety policies and regulations. While powered, bank-controlled "slackline" cableways are gradually becoming more widely installed, their use, requiring good vision, is often still constrained to daylight. Auto-samples collected at a point by the bank are eminently practical, but a trade-off is that they are often not representative of the cross-section averaged SSC but typically underestimate it when there is a significant sand component to the suspended load. Hence some manual sampling to calibrate the auto-sampled SSC to the cross-section average SSC is still desirable (Hicks et al. 2013).

The second improvement, from a limits perspective, is that it may be better to focus on good fits to the low to middle flow range on the SRC (say, to flows less than a small multiple of the mean flow), which will likely be the range of the SRC that is most utilised in the sediment framework (e.g., if the SSC ESV target is set such that it is not exceeded for say 95% of the time, then the SSC exceedance will be extracted from the low-mid ranges of the flow duration and sediment rating curves). The

³⁴ Hicks, Greenwood, Clapcott, Davies-Colley, Dymond, NIWA Christchurch on 24 May, 2016.

³⁵ Such manual SS sampling was undertaken at ~ 75% of the sites used in our SRC analysis.

viability of targeting low-mid flows will only be able to be evaluated when targets have been identified for the SSC and water clarity ESVs.

A third improvement would increase the complexity of the SRC function. The present two-parameter power function (slope and intercept in log-transformed space) model simplifies the SRC shape at many sites albeit by varying degrees (e.g., the Waipaoa River shows a clearly curved SRC, Figure 4-1), so a generalised SRC model with more than two parameters may improve predictive performance³⁶.

In this regard, literature over the past decade (e.g., Basher et al. 2011; Yang et al. 2007; Warrick 2015) has confirmed that SRCs vary temporally at-a-site (within events, between events, seasonally, and inter-annually), and the focus of SRC prediction has been directed at developing models that capture the effects of antecedent hydrological conditions (particularly the effects of large storms and spates of storms associated with inter-annual climate drivers such as the El Nino – La Nina signal, e.g., Gray et al. 2015). Long-duration (e.g., 5-10 years plus) datasets with adequate sample density over the full discharge range of each year are required to develop such models; however, there are only a handful of existing datasets available in New Zealand that qualify (e.g., Motueka, Manawatu). Thus existing sediment campaigns with auto-samplers that capture regional variations in climate (e.g., Auckland, Waikato, Manawatu, Southland at least) should be continued to enable this potential improvement.

10.2.2 SSC duration curves

The present analytical framework assumes a single-value function between SSC and discharge, which enables discharge duration to be simply converted to SSC duration via the SRC. In reality, hysteresis often occurs in the SSC-discharge relationship during floods (and, indeed, is responsible for much of the scatter observed on SRC plots). This arises due to phase differences in the generation and delivery of water and sediment. Higher SSCs often occur during event rising stages (e.g., see Mararoa River example, Appendix G) due to relative exhaustion of in-channel sediment sources, particularly in small catchments (e.g., Hughes et al. 2012). However, in other catchments relatively high SSCs can linger on recessions (for example, as banks fail due to unbalanced pore-water pressure in the banks as stream water levels recede).

SRC hysteresis has the potential to change the SSC duration curve from the assumed simple transformation of the flow duration curve, but the extent and importance of this has not yet been investigated. This could be done using time-series records of SSC calibrated to continuous turbidity records which are now available from several regions over a range of river types and catchments (e.g., Motueka and Kaipara Catchments; Manawatu, Auckland, Wellington, and Southland Regions).

If hysteresis proves important, then it could potentially be captured into the analytical framework by including a dQ/dt (rate of change of discharge) term in the SRC. Alternatively, SSC duration curves could possibly be generated from more complex, time-dependent SRCs using a Monte Carlo simulation approach. Thus we recommend an investigation that would analyse existing multi-year datasets to compare the observed SSC duration curves with those predicted using simple SRCs, more complex SRCs, and also using a Monte Carlo approach (if feasible).

³⁶ We note that scatterplot smoothing methods such as LOWESS are now often used in preference to single functions for fitting SRCs on a site-by-site basis (e.g., Hicks et al. 2011, who used LOWESS to fit rating curves at 80% of 230 sites).

10.2.3 Effects of sediment load change on SRCs

The current analytical framework assumes that a change in catchment sediment load induces a simple vertical shift in the (log transformed) SRC but no change in rating slope – in which case, conveniently, the SSC at a given discharge and frequency of occurrence should change by the same proportion as the sediment load. As demonstrated in Section 4, this assumption appears reasonable in some catchments but is a simplification of the general case (e.g., Motueka example, Figure 4-1). However, there is too little information available on how SRCs in New Zealand rivers change to fully assess this assumption and to predict where it might fail at a catchment scale. Indeed, even identifying changes in sediment supply from load monitoring is problematic because of the large temporal variability in load driven by hydrological variability. Practically all of the available SRC datasets have too few data points in total, or data has been collected too sparsely over time, to reliably characterise temporal changes in catchment sediment load. What is required is more datasets and analyses that detail the temporal and (within-catchment) spatial variability of the SRC, with associated surveys of catchment sediment sources that can provide a good understanding of where the sediment load is sourced from. In particular, we need to be able to distinguish shifts in SRCs that relate to management/land use effects (which are manageable) from those that relate to climatic/flood history (which are unmanageable). We know the latter can have profound effects on SRCs but we have little information on the former.

Our working hypothesis (from Section 4) is that the extent of change expected in a SRC following a targeted reduction in sediment supply should depend on both the uniformity of the catchment in regard to dominant erosion process and where the treatment is done. Thus sites are required that capture a range in catchment uniformity, and at least two catchments (one uniform, the other non-uniform) should have at least three main-stem sites to observe downstream changes in SRC characteristics. Also, since the issue relates to the extent of changes in the SRC slope (**b**-parameter), which Section 3 has shown to be influenced primarily by catchment steepness (or unit stream power), the sampling network should ideally also cover a gradient of catchment topography types.

Thus we envisage a small network of well-monitored, representative “experimental” catchments, equipped with both auto-samplers and continuously-recording turbidity sensors, operated for a minimum of approximately five years, with attendant survey of sediment sources and changes in land use/management.

The above (providing the effects of hydrological events can be isolated) will better inform on the extent of changes in SRC parameters due to treatment of sediment sources, particularly the **b**-parameter. Should change in the SRC **b** turn out to be important, further work will also be required on developing and proving an analytical approach that ‘routes’ changes in both SRC parameters downstream following localised erosion treatment. A simple example of what is intended is illustrated in Figure 4-2. The approach should be proven with the new datasets from experimental catchments (making use of multiple sites within the same catchment).

The ultimate approach would be to use spatially distributed, time-stepping water and sediment generation and routing models, but implementation of these at a national scale would require significant investment in model development and implementation.

A routing approach, whether simple or complex, would also better deal with lake-fed rivers where the SS load in the lake outflow is greatly reduced (and sediment character markedly altered) by sediment settling, while the SSC of sediment added from relatively small tributaries may have little correlation with the main river discharge.

10.2.4 SS particle size and other physical characteristics

The physical characteristics of SS particles (particularly particle size but also particle shape and composition) exert fundamental control on the relationships between SSC and the optical ESVs, VC and LP. Thus in theory, with good knowledge of the SS constituents and their concentrations it should be possible to reliably calculate these relationships. However, with virtually no concurrent data on particle characteristics and optical ESVs available we currently rely on empiricism to define these relationships (but with very limited ability to predict spatial variation in the parameters of these empirical relationships - as found from Section 5).

Moreover, while the RF analysis from Section 6 showed that PSD is related to catchment lithology (and then land-cover), most of the variance in PSD remained unexplained, and our ability to predict PSD variables such as % mud, fine-silt, or clay remains weak - which is not surprising given that existing data is from only 59 sites (albeit widely distributed) around the country and observed temporal variability in PSD at-a-site.

Thus there is a key need for considerably more data on SS particle characteristics so that better models can be developed to predict these from catchment characteristics. This should better represent all reasonably-commonly occurring lithology types and their weathering status, and should demonstrate variability over event to annual time scales. It will also be important to measure the physical characteristics of SS delivered into channels from particularly erosion processes (e.g., eroding stream banks, gullies) so that the impact of erosion mitigation on the residual PSD can be assessed. So at least in some catchments, intensive spatial sampling of PSD during runoff events should be carried out.

Thus we recommend a widely distributed national network of sites sampled for particle characteristics analysis (we envisage ideally around one hundred sites monitored for at least two years each), with additional focus in a few key catchments where sediment sources are also sampled. In this regard, it would be sensible to include particle analysis into sampling networks operated by regional councils. As well as this SS particle characteristics network, opportunities should be taken to explore for changes in SS particle characteristics following changes in catchment sediment load (e.g., before and after forest harvesting, before and after erosion mitigation).

Multi-vertical, depth-integrated sampling would be ideal, but the health-and-safety driven practical issues of this type of sampling (see Section 10.2.1) mean that it may be pragmatic to compromise on auto-sampled point samples. PSD should be analysed using laser-based laboratory devices, ideally time-of-transit devices that measure discrete sediment grains and also image the grains for analysis of particle shape³⁷, and then allow the sample to be recovered for SSC analysis by filtering.

Most importantly, and as discussed below, monitoring of SS particle size should be undertaken in conjunction with measurements of SSC, VC, sediment composition (i.e., inorganic/organic), and coloured dissolved organic matter (CDOM).

Lastly, the modest explanation of spatial variation in particle size provided herein by existing parameters in the REC2 database suggests the need to develop or link to GIS layers of soil textural properties that better relate to SS properties.

³⁷ For example, the EyeTech-Combi particle size analyser at NIWA's Hamilton laboratory.

10.2.5 VC, turbidity, and SSC

A key finding of Section 5 was our poor ability to predict the parameters of the relationship between VC and SSC from catchment characteristics. The key intermediate information lacking concerns the physical characteristics of the SS, notably its PSD, particle shape and composition, and also the CDOM³⁸ in the water, since it is these that determine the SSC-VC relationship (e.g., Bowers and Binding 2009). Thus all of these need to be measured as an integrated package where VC is measured, and the aim should be to first calibrate relationships between these water and sediment characteristics and VC, and then relate these to catchment characteristics.

The pragmatic approach would be to add SS particle size, shape and composition (and CDOM) to the list of analyses undertaken at existing national (e.g., NRWQN) and regional (e.g., SOE) monitoring networks. We envisage a 2-5 year program.

Section 5 also showed considerable variation in the exponent of the VC-SSC relationship (i.e., the slope of the log-transformed relation), and almost equivalently, in the characteristic attenuation cross-section (c^*) among sites. The variation in parameters such as the power slope and c^* between sites indicates that the particle mixture must vary as SSC changes in a way specific to individual catchments. Better understanding of particle characteristics as they control c^* is required based on clarity data collected at high temporal resolution during runoff events. To this end, benefit will also accrue from discovering more about the relationship between VC and continuously-measurable turbidity. While turbidity is not a proper physical quantity and therefore is affected by instrument type, it nonetheless offers a proxy for monitoring VC at high temporal resolution³⁹. Moreover, while turbidity is not directly included in the sediment analytical framework, its use as a proxy for both SSC and VC (after calibration) will likely be important for implementing NOF-sediment.

Thus we recommend developing a relatively small network of sites (say 12, covering a contrasting range of lithologies, soil types, topographies, and land covers) where all relevant variables (VC, SSC, turbidity, PSD, particle shape, sediment composition, CDOM) are measured at high temporal resolution during events. The sampling approach would require auto-sampling and continuous turbidity monitoring, with the auto-samples used for the compositional analysis as well as for calibrating the continuous turbidity record to SSC and VC. Sampling would include base flow monitoring by routine monthly visits as well as event monitoring – as was pioneered by Hughes et al. (2014) in the Kaipara region. High quality data is required, implying: (i) considerable care with measuring VC to avoid the quality issues recognised in Section 5, particularly in smaller rivers and streams; and (ii) relatively large volume filtrations (up to 5 l) in clear rivers at base flow to avoid approaching detection limits with SSC measurement.

Finally, there is a need to explicitly trial the LOCS/YOCS approach (Framework 2, Section 8.4) for limiting fine sediment so as to protect optical ESVs. The approach is conceptually identical to that of developing sediment mass load estimates (Elliott et al. 2013) but instead budgets the flux of light attenuation (with units m^2/s and calculated by multiplying BAC ($1/m$) by flow (m^3/s)). Data obtained by Hughes et al. (2014) for Kaipara Harbour tributaries is being worked-up as annual light attenuation loads and yields to demonstrate the approach. However, there is a need to trial the approach elsewhere (in contrasting regions as regards optical water quality of rivers) and to further develop the approach as a potential management tool for fine sediment where its impacts are likely to be

³⁸ The CDOM can be an important control at high VC values (a higher CDOM reduces the VC for a given SSC), particularly in catchments with sources of organic matter (e.g., forests, catchments with leached soils).

³⁹ Ideally visual clarity would be monitored continuously by beam transmissometer, but these instruments are both more expensive than nephelometric turbidity sensors (by a factor of 3-5) and lack the dynamic range of turbidimeters.

mainly on VC or LP. The latter is required because how to relate the LOCS/YOCS approach to erosion mitigation and land use change remains to be determined.

10.2.6 DS

In Section 7 we found that the relationship between suspended sediment load and the various measures of deposited fine sediment was poor, despite the expectation that, other things being equal, a site with a higher mean annual sediment load should show more DS. That no good relationship was found points either to: (i) complexity in the processes causing sediment to accumulate in and over stream substrate; (ii) substantial temporal variability in DS measures (most sampled sites had few measurements over time); (iii) large uncertainty in the measurement procedures; (iv) large uncertainty in the measures of catchment properties including sediment loads; (v) dominance of sediment deposition by reach-scale processes rather than catchment sediment loads (where the extent of deposition is controlled by the local balance between sediment load and transport capacity, or (vi) a combination of some/all of the above. Our view is that (ii), (iii), and (v) share the greatest importance.

Little is known about the time-history of DS accumulations, including the role of antecedent flow conditions, and exchanges of fine sediment between the substrate and the flow in different hydraulic/geomorphic settings and substrate types. Thus there is a need for a small number of high resolution (in space and time) studies at sites collectively covering a gradient of sediment load and river morphological type. These also should have their suspended sediment load and its PSD sampled nearby (to compute the sediment load and to develop a SRC), and ideally also should be instrumented with acoustic sensors to record the flow when bedload ceases movement during floods (to better test the “frozen bedload” hypothesis discussed in Section 7.3.5). DS measurements should be made both as soon as possible after events and during base flows. We suggest the sites should include at least one with “dirty” flood recessions (such as observed in the East Cape area of the North Island) and one with clean flood recessions (such as in Westland).

There are also questions around the accuracy of the DS monitoring procedures, notably the reliability of visual estimates and the methodology of the SAM4 Quorer approach. While considerable effort went into developing these methods there has been limited effort post-development in method validation and testing for accuracy. For example, as detailed in NIWA (2013), the Quorer method involves stirring-up fine sediment inside a drum “screwed” into the substrate. Large substrates make obtaining a ‘seal’ difficult and sediment can be lost on sampling; also, different stirring strengths and depths can affect the amount of sediment re-suspended (feedback to Joanne Clapcott from regional council staff). The aerial concentration of sediment suspended (g/m^2) depends directly on the depth of bed disturbance, which, although nominally targeted to be approximately 5-10 cm, can vary (in our experience – Joanne Clapcott) anywhere from less than 1 cm to over 15 cm. Moreover, the sediment mass stirred-up into suspension also relates to the porosity of the substrate and the extent to which pore spaces are filled with sand and fine gravel too large to be homogeneously re-suspended. Thus there is the potential for large uncertainty in the results associated with taking a measurement at a single location (setting aside the question of spatial variability within a sampling reach). It may be possible that a simple measure of substrate size, such as that obtained using SAM3 (Wolman pebble count), may suffice to take substrate size into account in the Quorer method. Thus we consider it appropriate that an urgent step will be to validate and improve DS sampling methodologies – perhaps by comparison with a ‘reference’ method like the freeze-coring approach (e.g., Lambert and Walling 1988).

Collecting DS data robustly in a relatively small network of sites that have well-measured sediment loads and cover a wide gradient in catchment properties that are directly measured (rather than estimated off GIS models) should provide a better test of the DS vs sediment load relationship than the current dataset with its attendant uncertainty on both response and predictor variables (for example, the standard error on mean annual sediment yield estimators can be up to a factor of 2-3, as shown by Hicks et al. 2011).

10.2.7 Analytical frameworks

From the above, we identify three areas where the sediment mass based analytical framework (framework 1, Section 8.3) potentially requires modification. These remove dependence on the following assumptions that:

- the flow-duration curve combined with the SRC provide an adequate approximation of the SSC-duration curve
- a change in catchment sediment supply will only change the intercept (**a**-parameter) of the SRC anywhere downstream
- a change in catchment sediment supply will not change the SS particle characteristics downstream, and so will not alter the relationship between SSC and VC.

The first assumption remains untested and is a priority for further work. If need be, however, potential options to improve the existing framework include either a Monte Carlo type approach or developing SRCs that are dependent on antecedent flows and the rate of change of flow.

The second assumption fails in theory for catchments with non-uniformly distributed sediment sources or non-uniformly-treated sediment sources. However, we consider that a simple water and sediment routing model could be developed to adjust SRC parameters downstream from sub-catchments/reaches treated for erosion mitigation. This could be developed and tested using existing datasets such as those from the Motueka catchment, or by following future targets of opportunity (such as a sediment 'slug' associated with a landslide).

The third assumption is currently untested due to a paucity of information on SS physical characteristics. If it fails, then a solution would require downstream routing of individual sediment constituents, and reconstituting the ESVs at target sites from the re-assembled constituents.

The same tests/modifications are required for the alternative LOCS based framework (framework 2, Section 8.4), requiring further data to test the three assumptions listed above.

10.3 Research needs by motivation/type

The third column of Table 11-1 classifies the above research needs by type (or motivation).

Research tagged as gap-filling involves the collection of more data at more locations around New Zealand to fill gaps in geographic coverage and/or to provide a stronger basis for refinement of predictor models. Priorities include improving predictors of SRC parameters and particle size characteristics.

Research for improved understanding includes:

- integrated sampling to better understand and quantify the inter-relationships between particle characteristics, optical ESVs, and SSC

- detailed monitoring of the processes controlling fine sediment deposition.

Hypothesis testing includes collecting integrated datasets to robustly assess hypotheses such as the “frozen bedload” hypothesis.

Assumption testing includes using existing or new datasets to evaluate simplifying assumptions, including:

- that the concentration duration curve links directly to the flow duration curve through the SRC
- that changes in catchment sediment load do not change SS PSD and ESV inter-relationships.

More sophisticated model development includes:

- improvements to sub-models in the analytical frameworks to deal with (i) the general case where a change in catchment sediment load causes both SRC parameters to change downstream, and (ii) cases where the above assumptions may fail (e.g., an improved CDC predictor is required)
- developing more accurate predictors of SRCs.

10.4 Overarching needs

It became apparent at our workshop that there were overarching elements for future research that should best be dealt with through integrated studies – collecting data on all links in the analytical chain from a common network of sites. These involve:

- concurrent monitoring of optical variables and SSC in conjunction with SS physical properties (shape, grainsize, composition) at networks of sites that (i) cover gradients of controlling catchment characteristics (e.g., lithology, soils, steepness, land cover), (ii) that capture downstream variation in selected catchments
- using continuous turbidity as a proxy for *both* SSC and VC, both during runoff events and base flows
- collecting the same type of information on the characteristics of sediment generated by different erosion processes
- collecting information on DS at sites where the SS load is also monitored, including its particle size.

10.5 Value of research

The research topics/questions in Table 10-1 were rated qualitatively in regard to their scientific importance to managing sediment (low, medium, high), relative cost (low, medium, high), and their value. Value was determined as a trade-off between importance and cost, as defined in Table 10-2. For example, a topic rating “high” in importance and “low” in cost was assigned a “very high” value. As so identified in Table 10-1, the highest value obtains from addressing important questions with existing datasets (e.g., testing whether the sediment-duration-curve matches the flow-duration curve; adding complexity to sub-models in the analytical framework).

Table 10-1: Summary of further research needs to improve understanding, fill gaps in data coverage, test hypotheses, and/or develop more sophisticated models within the analytical chain linking ESVs to catchment sediment load. Importance was graded mainly on basis of current ability to predict regional variation in component relationships. Low importance reflects an issue that we consider has dubious promise. The importance of some research needs awaits the outcome of intermediate studies. VH = very high, H = high, M = medium, L = low, VL = very low, ? = can only be assessed after further research.

Topic	Issue	Research type / motivation	Aim	Number of sites	Site location	Sampling method	Sampling frequency and duration	Scientific importance	Cost	Value
SRCs	Accuracy of parameter prediction and SSC prediction	Gap-filling	Define SRCs in regions currently with sparse data	20-30	Northland, Marlborough, East Otago	Auto-sampling; manual if H&S permits	Event sampling, 5+years	M	H	L
	Accuracy of SSC and SS load prediction	More sophisticated model development	Develop more sophisticated, time-varying SRC function; Focus model-fitting on key discharge ranges	Existing	Motueka, Horizons, Auckland, Waikato datasets	Requires continuous turbidity-proxied SSC records to test	Multi-year, continuous records	M	L	H
SSC-duration curves	Does combining FDC and SRC produce reliable estimates of SSC-duration curves?	Assumption testing	Compare estimated SSC-duration curves with observed (from continuous turbidity)	Existing	Motueka, Horizons, Auckland datasets; focus on sites showing strong within-event SSC-Q hysteresis; contrast large & small catchments	Requires continuous turbidity-proxied SSC records to test	Multi-year, continuous records	H	L	VH
		Improved understanding	Characterise how SSC-Q relation varies between rising and falling stages; quantify and explain SSC recession characteristics	"	"	"	"	M	M	M

Topic	Issue	Research type / motivation	Aim	Number of sites	Site location	Sampling method	Sampling frequency and duration	Scientific importance	Cost	Value
		More sophisticated model development	Develop alternative approaches for predicting SSC-duration curves, including possible Monte Carlo approach	“	“	“	“	?	L	?
Effect of load change on SRCs	Being able to reliably predict situations where a change in sediment load will only change the SRC intercept compared to where it will change <i>both</i> SRC slope and intercept	More sophisticated model development + model validation data	Develop simple sediment and water routing procedure to predict downstream changes in SRCs; collect field data to improve understanding of temporal behaviour of SRCs, downstream change in SRCs, and to validate SRC “routing” model	Requires multiple sites within a catchment	Small number of catchments with non-uniform sediment delivery, 3+ sites per catchment, ideally capturing an erosion-mitigation area. Trial initially with existing datasets from Motueka, Manawatu, Whatawhata, Kaipara Catchments.	Auto-sampling with turbidity sensors	Event-based, 5 years	M	H	L
SS PSDs & other physical characteristics	PSD a key control of SSC-clarity relationships and deposited sediment but too little data from around NZ, particularly integrated datasets	Gap-filling, integrated data collection	Collect SS PSDs over a wide range of catchments (by lithology, land cover, steepness) in conjunction with SSC and visual water clarity. Ideally measure PSD and particle shape, but at least measure mud and organic % of SS load with every sample.	Many (~100); capture main lithological and soil classes around NZ; integrate with existing networks such as SoE, NRWQN	Widely distributed national network	Ideally manual depth integrated sampling but Health & Safety issues may constrain to auto-sampling	Events plus base flows, continued for at least 2 years	H	H	M

Topic	Issue	Research type / motivation	Aim	Number of sites	Site location	Sampling method	Sampling frequency and duration	Scientific importance	Cost	Value
	Change in PSD with sediment control measures	Improved understanding; assumption testing	Measure extent of PSD change (and resultant impact on SSC-clarity relationships) from a known change in sediment load (e.g., forest harvesting, landslide slug, riparian erosion protection)	Few	Sites of opportunity	"	"	H	M	H
	PSD characteristics of different erosion sources	Improved understanding; assumption testing	Establish if different erosion processes provide different PSDs	Many would be required to capture combinations of lithology, weathering, and erosion types	Capture erosion types over gradients in lithology and weathering	"	Event based, small number of events	H	H	M
VC, SSC, turbidity relationships	Poor understanding of controls on regional variation in relationships and ability to predict relationship parameters	Improved understanding; measurement	Link SSC-clarity relationships to sediment physical characteristics by collecting integrated datasets (PSD, shape, composition, and also CDOM where VC is sampled)	Many; combine with existing national and regional networks	Existing networks	Standard techniques	Regular monitoring capturing base flows and events; 2-5 years	H	M	H

Topic	Issue	Research type / motivation	Aim	Number of sites	Site location	Sampling method	Sampling frequency and duration	Scientific importance	Cost	Value
	Poor understanding of temporal variation in clarity-SSC at event scale – e.g., is there hysteresis that will impact on clarity-duration curves?	Improved understanding	Improve understanding of temporal variability in clarity-SSC relation as driven by particle characteristics at event scale	~ 12	Spanning range in lithology, land-cover	Auto-sampling; laser-based PSD analysis.	Events; 1-2 years	M	H	L
	Awareness of managing optical loads (Framework 2)	Demonstrate Framework 2		~ 12	At 'sediment' sites in most regions	Event (auto-samples) and base flow samples	Monthly for base flow, and over events	M	H	L
DS	Accuracy of DS measurement techniques	Improve field methodology	Validate and refine field sampling procedures; assess what is practicable in regard to resources available to regional councils for SoE monitoring; consider more sophisticated sampling as a 'reference' e.g., freeze-coring	Workshop				H	L	VH

Topic	Issue	Research type / motivation	Aim	Number of sites	Site location	Sampling method	Sampling frequency and duration	Scientific importance	Cost	Value
	Uncertain relationship with SS load	Improved understanding	Concurrent measurements of DS accumulation and SS load and characteristics in detailed studies; to capture variability and its time-scale due to large hydrological/sediment-supply events	Few	Experimental reaches with contrasting bed material size, sediment supply	Event-sampling of SSC with auto-sampler or manual sampling; apply all SAM methods, modified as needed after review workshop	Spatially detailed DS measurements immediately after events, and between events, for 5+ years	H	H	M
	Test validity of “frozen bedload” hypothesis linking re-suspendable fines to SSC at ¼ mean annual flood	Hypothesis testing; improved understanding	Validate or reject “frozen bedload” hypothesis	Few	Integrate with above	SAM4 (Quorer) measurements included with above; ideally with acoustic monitoring of bedload activity to detect discharge when bedload motion stops	“	L	M	L
Analytical framework	Assumptions relating to linking relationships	More complex model	Develop methods to better deal with predicting SSC-duration and clarity-duration curves					M	L	H
		More complex model	Develop method to “route” changes in SRC parameters downstream in catchments with non-uniform sediment supply					H	L	VH

Topic	Issue	Research type / motivation	Aim	Number of sites	Site location	Sampling method	Sampling frequency and duration	Scientific importance	Cost	Value
		More complex model	Develop method to route individual sediment constituents (should this prove an important control on SSC-clarity relationships)					?	M	?
Overarching need	A requirement for integrated datasets at-a-site that enable validation of all links in the analytical frameworks.		Improved understanding of sediment-related ESVs	Few	'sediment' sites and experimental catchments	As indicated above	As indicated above	H	H	M

Table 10-2: Value definitions for research topics in terms of relative cost and scientific importance. Value increases as importance increases and cost decreases.

		Cost		
		<i>Low</i>	<i>Medium</i>	<i>High</i>
Importance	<i>High</i>	Very high	High	Medium
	<i>Medium</i>	High	Medium	Low
	<i>Low</i>	Medium	Low	Very low

11 Conclusions

The main conclusions of this investigation are as follows.

Sediment rating curve prediction

- Sediment rating curves (SRCs) were fitted to 271 sites broadly scattered around New Zealand with a two-parameter, power-law model of the general form $C/C_{\text{mean}} = \mathbf{a}(Q/Q_{\text{mean}})^{\mathbf{b}}$. The best Random Forest (RF) regression models, relating the **a** and **b** parameters to catchment/site characteristics, were able to predict the **a**-parameter at any site to within a factorial root-mean-square error (RMSE) of $\times/\div 2.29$, which aligns with the accuracy of sediment yield estimators developed previously from a similar dataset. The RMSE on the **b**-parameter is ± 0.38 . We regard these accuracies as reasonable given the substantial natural data-scatter observed on SRCs.
- There was marginal difference in performance between RF models developed using the original suspended sediment concentration (SSC) data and values standardised by the at-a-site discharge-weighted mean concentration (which is the ratio of mean annual sediment and water discharges), thus we recommend using the predictive model based on the standardised values as a matter of convenience for the sediment framework. This model can be implemented by performing a one-off prediction of the SRC **a** and **b** parameters for every New Zealand REC2 segment within the R development environment, then accessing these results via “lookup” functions.
- The main factors determining the SRC **a**-parameter were catchment sediment yield, with lesser influences from land-cover and soil texture. The main factors determining the SRC **b**-parameter were those linked to catchment slope (e.g., unit stream power, elevation, steepness, stream order). These findings generally align with those observed in international datasets.

How do SRC parameters change when sediment load changes?

- Analysis of SRC changes in catchments experiencing events (hydrological or land management) that changed the catchment sediment load showed mixed effects on the SRC parameters. While, as expected, all experienced significant change in the **a**-parameter (by factors ranging from 1.6 to around 4) associated with the change in sediment supply, two also experienced significant changes in the **b**-parameter.
- A common factor at sites where the SRC **b**-parameter did not change was that the event’s impact was reasonably uniform over the catchment. Of the two sites where the SRC **b**-parameter did change, one had experienced changes in sediment supply over only part of its catchment, whereas the other likely experienced an increase in runoff as a result of forest harvesting. These observations, corroborated by overseas observations and a simple modelling exercise, showed that the assumption of a stable SRC **b**-parameter under changing catchment sediment loads does not hold-up in the general case where tributary SRCs and sediment load changes are not uniform within a catchment and/or the load change is accompanied by a change in runoff regime.
- In the NOF sediment context, a constant SRC **b**-parameter under changing sediment loads is a key assumption in the frameworks linking sediment loads and ESVs. The importance of violating this assumption will depend on the scale of application and the degree of change

in the SRC **b**-parameter. For national scale assessments (e.g., to map the extent of sediment load reduction required to achieve ESV targets), the assumption of uniform sediment load change across catchments (and hence constant SRC **b**-parameter) would be a reasonable approximation. However, when implementing sediment management policy at the catchment scale (e.g., when focussing erosion control at priority locations), violation of the assumption will be important where the SRC **b**-parameter change exceeds the uncertainty associated with SRC **b**-parameter prediction. For example, in the case of the Motueka at Woodstock, the change from $b = 2.03$ to $b = 1.38$ (i.e., -0.65) was larger than the RMSE error in our predictive model of **b**.

- A potential way forward with catchment-scale implementation of the analytical framework would be to add into the framework a simple water and sediment routing model that calculated downstream changes in both **a** and **b** parameters following localised changes in sediment load.

Relationships between VC, SSC, and turbidity and their prediction

- We found strong inter-relationships of VC, turbidity and SSC (as represented by TSS) in rivers, consistent with previous work. However there is appreciable variability, with about a 10-fold range in VC at a given TSS across diverse rivers. These variables are more strongly related within sites than across sites – presumably reflecting the restricted range of characteristics of fine sediment particles within particular rivers. VC and TSS are inversely related, but not linearly, such that a halving of TSS typically does not double VC but increases it only about 65%.
- Analysis of the high-quality NRWQN dataset using RF regression methods showed that the most important predictor variables were similar to those found to be important for suspended sediment particle size (including catchment elevation, sediment load, climate, geology class, and temperature). However, all RF models had only weak ability to explain the observed regional variation in parameters for the VC – TSS – turbidity relationships. Thus we conclude that catchment variables currently in the REC do not well capture the particle size, shape, and composition that theoretically together control light attenuation by sediment and thus VC. Nonetheless, while the amount of variance in the source data explained by the regressions models is small, the end-result in regard to parameter prediction accuracy still appears reasonable. For example, in the power-law relationship between VC and TSS ($VC = a TSS^b$), the standard error on the a -parameter is only $\pm 17\%$ while that on the b -parameter is ± 0.12 ; in combination, these indicate that VC can be reasonably well predicted off TSS.

SS particle size and its prediction

- Suspended sediment particle size distribution (PSD) is important because it is the main factor controlling the relationships between SSC and VC and LP. Thus understanding the factors that control PSD underpins understanding of (i) regional variation in relationships between VC, SSC, and turbidity (which can be calibrated as a proxy for both VC and SSC), and (ii) whether a change in the distribution of sediment load from a catchment (e.g., after erosion mitigation work) might also change these relationships by changing the PSD of the catchment sediment load.

- RF models were developed from sampled suspended sediment PSD data to predict both the % of SS load finer than 16 μm (i.e., fine silt and clay) and the % finer than 63 μm (i.e., the % mud). These RF models, respectively, explained 34% and 39% of the variance in the observed particle-size proportions, and predicted them to an accuracy of $\pm 16\%$ and $\pm 18\%$. The % mud is also strongly and inversely correlated ($R^2 = 0.91$) with suspended sediment median particle size, which can thus be predicted using the % mud. The main factor influencing the regional variation of % mud (and thus median size) was catchment lithology, but with some weaker control also exerted by land-cover, sediment supply, elevation and rainfall. After lithology and land-cover, the % of fine silt and clay was also influenced by temperature, elevation, and rainfall.
- While the RF analysis suggested a weak dependence of suspended sediment PSD on catchment sediment supply (which hints that if the sediment supply is changed then some change in particle size may follow), this may well be an indirect effect of catchment lithology and weathering status.

Does SS particle size distribution change when sediment load changes?

- No PSD data were available to directly assess changes in suspended sediment PSD accompanying changes in sediment load, but we used changes in specific turbidity (the ratio of turbidity to SSC) to infer changes in PSD from two catchments following extreme hydrological events, one of which (the Motueka) had a well-recorded increase in sediment load during and following the event. We conclude that changes in sediment load can cause changes in the suspended sediment PSD and so changes in the relationships between SSC and optical properties. Whether these changes are significant will depend on (i) the PSD of the sediment delivered from the affected sources compared with the catchment-average PSD, and (ii) the importance of that source to the total sediment load.
- The general utility of specific turbidity as a proxy for representative suspended sediment particle size is supported from finding across the 77 sites in the NRWQN that turbidity tended to be quasi-linearly related to SSC (with a power-law exponent averaging 1.04 ± 0.15).

Relationship of deposited sediment to sediment load and other controls

- Our analysis of deposited sediment data showed very weak dependence on catchment sediment load. While not initially expected, this is perhaps not that surprising given issues with data collection and because the bulk of catchment sediment delivery occurs at periods of high flow that effectively flush the sediment through the stream network into receiving environments. The dominant significance of low elevation and low slope in explaining variance in deposited sediment metrics in our models supports this hypothesis. A measure of antecedent flow is likely to improve our ability to link deposited sediment to catchment loads because we could then account for time since last bedload disturbance.
- Our results suggest that the supply of sediment into a stream segment from local sources and how it is 'captured' by the stream morphology may be more informative of deposited sediment than sediment load from the upstream catchment. The potential implications for management could be a requirement to focus on local habitat to minimise the chronic delivery of fine sediment that occurs during stable flows, e.g., bank erosion. However, further research is required to confirm this.

- There is a large amount of compiled data available that measures deposited fine sediment in streams. Our analyses suggest that not all of it is useful for defining the sediment load – deposited sediment relationship. In particular, estimates of fine sediment cover from the NZFFD appear particularly variable.
- The addition of environmental variables (describing elevation, slope, geology, and flow) substantially improved our ability to model deposited sediment as a function of sediment load, especially for the measures of Fines (% cover of the streambed) and SIS (re-suspended inorganic sediment using the Quorer method). However, the explanatory power of any of the models is modest and similar to that observed in a recent study exploring drivers of deposited sediment in UK streams, which (like this study) also showed that stream power was the most significant explanatory variable of deposited sediment and that the influence of sediment yield was marginal.
- There were insufficient data to robustly explore the relationship between SSC and deposited sediment and the “frozen bedload” hypothesis, however, the available data showed no relationship between the Quorer-measured aerial deposited sediment concentration and the SSC predicted at $\frac{1}{4}$ the mean annual flood discharge. Indeed, if anything the data suggests that fine sediment probably accumulates in the substrate at flows below this discharge.

Analytical frameworks linking ESVs to catchment sediment loads

- Two analytical frameworks were developed that enable ESV targets (e.g., VC threshold not exceeded more than a certain % of time) to be quantitatively related to catchment sediment load. Framework 1 services all four ESVs and links VC, LP, and DS to sediment mass load via SSC. Framework 2 only services VC and LP and links these to the load of optical cross-section (LOCS) via the BAC. In both cases, flow duration curves are combined with rating curves (sediment vs discharge for framework 1; BAC vs discharge for framework 2).

Robustness of predictions using analytical frameworks

- A worked example using data from the Wairua River showed that when using relationships based on at-a-site measurements, the robustness of predictions of VC and DS associated with a given change in sediment load was weak while the robustness of predictions of LP (represented by the euphotic depth) and SSC was moderate. The robustness improved if only the *change* in ESV was predicted from a given % change in sediment load. In the case of VC, framework 2 (working with LOCS) was more robust in predicting VC than was framework 1 (working with SSC). The main reason for this, though, was that the Wairua site had more measurements of VC than SSC, which may not be the case at other measurement sites.
- Using the national regression models developed in this study, their larger error terms (particularly in the SRC parameters) rendered both frameworks only weakly robust when predicting absolute values of the ESVs associated with given changes in sediment load. However, the predictions of *change* in ESV (from an unknown initial state) were more robust.

Utility of analytical frameworks for national-scale assessment of change in sediment load required to meet sediment ESV targets

- Framework 1, with the linking regression relationships developed in this study, provides a workable approach for making national-scale assessments of the implications to sediment loads of setting national targets for SSC, VC, and LP. However, it cannot be used with any confidence to assess the implications of DS targets because of the lack of any reliable relationships between sediment load and DS.
- The accuracy of predictions from framework 1 will be weak if targets for SSC, VC, and LP are set in terms of absolute values. This is largely due to the high uncertainty in estimating sediment rating curve parameters. Results will be more robust if targets are set in terms of *change* in ESV associated with a change in sediment load.
- Framework 2 remains work-in-progress and has more limited utility.

Framework for further research

- Our investigations have found numerous areas where further research and development is desirable to enable management of fine sediment in waters under the NPS-FM. We have collated these into a framework that distinguishes research motivation and also rates the net value of the research in regard to scientific importance (towards advancing sediment management in a NOF context) and cost.
- The motivations for the further research include data collection to fill gaps in geographic coverage to improve predictor models, data collection aimed at improving understanding of relationships and processes, assumption and hypothesis testing, and development of more sophisticated models to embed in the analytical frameworks – to improve their overall accuracy and general applicability.
- The research topics/questions of greatest value are generally those that are high in scientific importance and low in cost (which typically means use of existing datasets rather than collecting new ones). Examples include testing whether the concentration duration curve (CDC) estimated using the sediment rating and flow duration curves matches the observed CDC, and adding complexity to sub-models in the analytical framework.
- Nonetheless, we believe that in the long run the research and development priorities on sediment as an environmental stressor are best studied by collecting further field data in an integrated way, by interdisciplinary teams, at a relatively small number of dedicated ‘sediment’ sites in diverse experimental catchments. Concurrent measurements (of high quality) of SSC, VC and particle characterisation (PSD, shape and composition) would be conducted at base flow and over events (using auto-samplers) at these sites; and sampling of deposited sediment perhaps extended to experimental studies of sedimentation processes. Continuous turbidity records would be used to capture events and define non-hydrological variability and so define sediment and optical regimes and loads. Study of erosion processes and sediment (and light attenuation) sources would be ‘nested’ within these study catchments.

12 Acknowledgements

We gratefully acknowledge data provided by:

- Northland Regional Council, Auckland Regional Council, Waikato Regional Council, Bay of Plenty Regional Council, Gisborne District Council, Hawkes Bay Regional Council, Greater Wellington Regional Council, Horizons Regional Council, Tasman District Council, Nelson City Council, West Coast Regional Council, Otago Regional Council, and Southland Regional Council. Taranaki and Canterbury reported no useful data.
- Various research datasets collected/collated by NIWA, Landcare Research, and Cawthron Institute.

Martin Unwin (NIWA) kindly assisted with explaining the details of the “UnwinMfEDB” MS-Access database. Doug Booker (NIWA) provided advice on statistical analysis approaches, and made available “R” analysis code.

Les Basher (Landcare Research) provided insightful peer reviews of the preliminary and final drafts of this report.

Ton Snelder (LWP Ltd) and Carl Howarth (MfE) provided feedback on interim reports and on a preliminary draft of this report.

13 Glossary of abbreviations and terms

a	Coefficient in a power-law relation, e.g., $C = aQ^b$. When transformed to log values, the function becomes $\log(C) = \log(a) + b \log(Q)$, with $\log(a)$ the intercept of the straight-line function.
<i>a</i>	Absorption coefficient: the proportion of a collimated light beam that is absorbed per unit length of light path. An inherent optical property of water; units: 1/m.
Absorption (of light)	Conversion of light (quantum) energy to another form (ultimately heat) by molecular interaction.
Auto-sampler	An automated device that pumps water samples from a river.
b	Exponent in a power-law relation, e.g., $C = aQ^b$. When transformed to log values, the function becomes $\log(C) = \log(a) + b \log(Q)$, with b the slope of the straight-line function.
<i>b</i>	Scattering coefficient: the proportion of a collimated light beam that is scattered per unit length of light path. An inherent optical property of water; units: 1/m.
<i>BAC</i>	Beam attenuation coefficient: the proportion of a collimated light beam that is absorbed or scattered per unit length of light path. $c = a + b$. An inherent optical property of water; units: 1/m.
CDC	Concentration Duration Curve: graphical relationship between SSC and the % time that SSC is exceeded.
CDOM	Coloured Dissolved Organic Matter: humic matter in water, indexed by its absorption coefficient at standard wavelength, e.g., 440 nm.
Depth-integrating sampler	A sampling device that is traversed down to the river bed and back to the surface, collecting a water sample at a rate proportional to the ambient flow velocity.
Detection limit (DL)	The value below which a laboratory cannot confidently distinguish the analyte concentration from zero. In practice, an index of the concentration below which relative precision declines markedly. A particular problem with TSS data at base flow is that filtrations are often done on insufficient sample volume so that considerable proportions of datasets for TSS are approaching or below the DL.
Disturbance plume	A plume of turbid water produced by disturbance of fine sediment (silt and clay) deposited in the interstices of (much coarser) bed sediment in rivers. Such plumes, created by wading in channels, must be avoided for measurement of visual water clarity and water sampling for indices of SPM.
ED	Euphotic depth: depth where the light intensity is 1 per cent of that at the water surface.
ESV	Environment State Variable: a variable that captures an aspect of the state of the physical, chemical, or ecological environment.

FDC	Flow Duration Curve: graphical relationship between water discharge and the % time that discharge is exceeded.
Hysteresis	A “loop” in a relationship between two variables, e.g., when SSC is higher at a given discharge on the rising stages of a flood hydrograph compared to the falling stages.
K_d	Irradiance attenuation coefficient (in the down-welling direction as indicated by d subscript). Quantifies light penetration into waters.
LAWA	Land, Air, Water, Aotearoa: a website displaying information for more than 1100 freshwater monitoring sites throughout New Zealand.
LCDB	New Zealand’s Landcover Database v3. Classifies land cover across New Zealand in 33 different categories.
Light attenuation	Reduction in the power of a light beam (or sunlight) in water by the combined processes of absorption and scattering. Quantified by the beam attenuation coefficient (for a light beam or image-forming light) or the irradiance attenuation coefficient (for sunlight penetrating a water body).
LP	Light penetration: Sunlight (irradiance) entering water reduces with depth owing to absorption and scattering.
Lithology	Rock-type.
LOCS	Load of optical cross-section. Analogous to a sediment mass load, but measuring the ‘amount’ of light attenuation delivered (usually annually) by a river – or produced by a sediment source. First introduced by Elliott et al. (2013).
MfE	Ministry for the Environment.
MHpID	A consistent set of parameter names and descriptions derived within this project to consolidate different sample collection and processing methods.
MVDI	Multi-Vertical Depth-Integrated: An approach used to sample the discharge-weighted suspended sediment concentration over a river cross-section, using depth-integrating samplers at multiple verticals.
NRWQN	National River Water Quality Network. A monitoring network of 77 river sites run by NIWA since 1989, with an aggregate catchment about 50% of NZ’s land area (Davies-Colley et al. 2011).
NEMAR	National Environmental Monitoring And Recording. A programme sponsored by MfE to develop a national set of standard methods for collecting environmental data. Has since evolved into NEMS (National Environmental Monitoring Standards).

NSE	Nash-Sutcliffe Efficiency: a measure of the fit between observed values and model predictions, it determines the relative magnitude of the residual variance in the estimated yield compared to the measured yield variance. NSE ranges from $-\infty$ to 1, with 1 indicating a perfect match to predictions, 0 indicating that predictions are as accurate as the mean of the observed data, and negative values indicating that the observed mean is a better predictor than the model.
NZSegment	Individual river segment within REC2, with associated environmental information available. Segment boundaries occur at confluences.
Optical cross-section (of particles)	The light (beam) attenuation per unit mass concentration of particles. Units m^2/g . Has a theoretical underpinning in that the optical cross-section of perfect spheres of known composition is predictable from optical theory.
Out-of-Bag R^2 (OOB R^2)	The average proportion of the total variance explained by a Random Forest predictive model developed from n data records when the model is re-calculated n times, each time removing 1 record in turn from the derivation. Provides an estimate of the predictive power of the model for new cases.
Partial dependence plots	Show the marginal contribution of a predictor to the response (i.e., the response as a function of the predictor when the other predictors are held at their mean value) in a RF model.
Power function	A curvilinear function of the form: $Y = aX^b$, where X and Y are variables and a and b are fitting parameters.
PSD	Particle Size Distribution: the distribution of grain sizes in a sediment mixture.
REC2	River Environment Classification version 2.
RF	Random Forest. A flexible regression technique in which final predictions are based on averages across an ensemble of regression trees.
Riparian (shade)	Shade of streams and rivers by plants growing on the river bank.
RMSE	Root Mean Square Error. A measure of the precision of fit between observed values and model predictions. A lower RMSE indicates a better fit between observed and predicted values.
RSR	Ratio of the root mean square error to the standard deviation of the observed data. A dimensionless measure of the precision of fit between observed values and model predictions. A lower RSR indicates a better fit between predicted and observed values.
SAM1	Sediment Assessment Method 1: Bankside visual estimate of % sediment cover. Rapid qualitative assessment of the surface area of the streambed covered by sediment.
SAM2	Sediment Assessment Method 2: In-stream visual estimate of % sediment cover. Semi-quantitative assessment of the surface area of the streambed covered by sediment. At least 20 readings are made within a single habitat

SAM3	Sediment Assessment Method 3: Wolman pebble count. Semi-quantitative assessment of the particle size distribution, including fine sediment, on the streambed. At least 100 particle measurements are made within a single habitat.
SAM4	Sediment Assessment Method 4: Resuspendable sediment (Quorer method). Quantitative measure of total suspendable solids deposited on the streambed. Six samples are collected from a single habitat. Samples are processed in the laboratory for Total Inorganic/Organic Sediment by areal mass and/or Suspendable Benthic Solids by Volume.
SAM5	Sediment Assessment Method 5: Resuspendable sediment (Shuffle index). Rapid qualitative assessment of the amount of total suspendable solids deposited on the streambed. A score from 1-5 is assigned, where 1 is little/no sediment and 5 is excessive sediment.
Scattering (of light)	Change in direction of light photons without any change in (quantum) energy. SPM typically dominates light scattering in natural waters, although water molecules do scatter light (weakly).
Sediment load	The mass flux of sediment delivered from a catchment (typically in t/yr).
Sediment yield	The sediment load per unit catchment area (typically in t/km ² /yr).
SMA (regression)	Standard major axis regression – minimizes the variance of both the X and Y variables, in contrast to ordinary least squares regression which minimizes variance only in Y. Useful when there is no particular reason to treat either one of X and Y as the ‘independent’ variable.
SOF	Source of Flow category from REC2 (derived for REC1).
Specific turbidity	Turbidity divided by SSC.
SPM	Suspended particulate matter. Quantified (in terms of mass concentration) by TSS or (better) SSC. But there are many other characteristics of particulate matter of importance to its environmental behaviour and ecological effects, including PSD, particle shape, and chemical or mineralogical composition.
SRC	Sediment Rating Curve: graphical relationship between suspended sediment concentration and water discharge in a river.
SS	Suspended Sediment.
SSC	Suspended Sediment Concentration: mass of sediment suspended per unit volume of water (units of mg/l or g/m ³ are equivalent), measured by filtration of the <i>whole</i> a water sample, in contrast to TSS which is measured by filtration of a subsample.
SSSiteID	A unique (to this study) alphanumeric value assigned to each block of data identified as coming from a particular site by the source organisation. The code begins with the source region/organisation or database code and then with 9999 site numbers. Sites from the UnwinMfEDB begin at 5000, e.g., AKC5001 is the first site supplied by the Auckland Council within the UnwinMfEDB.

Standardised parameter	A parameter made dimensionless by dividing it by a references value or function, generally tagged with *.
Strahler stream order	The number of times a channel branches going upstream from a point minus 1.
Striping (on X-Y data plots)	Seen as 'stripes' of data, for what is really a continuous variable, plotting at integer values. Arises due to (premature) numerical rounding of data. Striping bedevils TSS data at comparatively low concentrations approaching the detection limit, when precision is low and data is rounded to perhaps only one digit to reflect that imprecision.
Suspended sediment gauging	A measurement of the discharge-weighted SSC over the full cross-section of a river.
TSS	Total suspended sediment (concentration) – measured by filtration of a <i>subsample</i> of a water sample, in contrast to SSC which is measured by filtration of the whole sample. Ideally TSS would equal SSC, but if the subsampling is not representative, typically owing to rapid settling sand, TSS may differ (and be biased).
TSS laboratory approach	A method for analysing the SSC of a water sample by filtering a small aliquot of the original field sample.
Unit stream power	The rate of dissipation of steam potential energy per unit width of channel. It equals the products of the unit weight of water, water discharge, and channel gradient, all divided by channel width.
UnwinMfEDB	Existing NIWA database containing freshwater monitoring information from regional councils and the National River Water Quality Network sites.
VC	Visual water clarity – quantified by the black disc visibility (in the horizontal direction).
Vertical	A vertical section in a river channel over which a depth-integrating sampler is traversed.
WRENZ	Water REsources of New Zealand: a GIS model previously available on NIWA's web-site.
YOCS	Yield of optical cross-section. Analogous to a sediment mass yield, but measuring the 'amount' of light attenuation delivered (usually annually) by a river – or produced by a sediment source – per unit land area. YOCS = LOCS/A where A is land area (e.g., of a catchment or a soil erosion plot).

14 References

- Allaby, M. (2004) Dictionary of Ecology. Oxford University Press.
- Ballantine, D.J., Hughes, A.O., Davies-Colley, R.J. (2014) Mutual relationships of suspended sediment, turbidity and water clarity in NZ rivers. In: Y.J. Ku, M.A. Allison, S.J. Bentley et al. (Eds). *Sediment Dynamics from the Summit to the Sea*. IAHS Press, Wallingford, UK: 265-271.
- Basher, L.R., Hicks, D.M., Clapp, B., Hewitt, T. (2011) Sediment yield response to large storm events and forest harvesting, Motueka River, New Zealand. *New Zealand Journal of Marine and Freshwater Research* 45(3): 333-356.
- Booker, D.J., Snelder, T.H. (2012) Comparing methods for estimating flow duration curves at ungauged sites. *Journal of Hydrology* 434–435: 78–94.
- Booker, D. J., Hicks, D. M. (2013) Estimating wetted width and fish habitat areas across New Zealand's rivers. Report to Department of Conservation: 33 pp.
- Bowers, D.G., Binding, C.E. (2006) The optical properties of mineral suspended particles: A review and synthesis. *Estuarine Coastal and Shelf Science* 67(1-2): 219-230.
- Breiman, L. (2001) Random Forests. *Machine Learning* 45: 5-32.
- Clapcott, J.E., Young, R.G., Harding, J.S., Matthaei, C.D., Quinn, J.M., Death, R.G. (2011) Sediment Assessment Methods: Protocols and guidelines for assessing the effects of deposited fine sediment on in-stream values. Nelson, New Zealand, Cawthron Institute.
- Clausen, B., Plew, D. (2004) How high are bed-moving flows in New Zealand rivers? *Journal of Hydrology (NZ)* 43: 19-37.
- Curran-Cournane, F., Holwerda, N., Mitchell, F. (2013) Quantifying catchment sediment yields in Auckland. Auckland Council Technical Report TR2013/042.
- Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler A., Hess K. T. (2007) Random forests for classification in ecology. *Ecology* 88(11): 2783-2792.
- Davies-Colley, R., Nagels, J.W. (2008) Predicting light penetration into river waters. *Journal of Geophysical Research* 113: G03028.
- Davies-Colley, R., Smith, D. (2001) Turbidity, suspended sediment, and water clarity: a review. *Journal of the American Water Resources Association* 37: 1085-1101.
- Davies-Colley, R. Verburg, V., Hughes, A., Storey, R. (2012) Variables for regional monitoring underpinning national reporting. Summary of recommendations. National Environmental Monitoring and Reporting (NEMaR) Variables, Step 1. NIWA Client Report (HAM2012-006) for the Ministry for the Environment. May 2012, Version 2 August 2012. 67p.
- Davies-Colley, R.J., Ballantine, D., Elliott, S., Swales, A., Hughes, A.O., Gall, M.P. (2014) Light attenuation – a more effective basis for the management of fine suspended sediment than mass concentration? *Water Science and Technology* 69(9): 1867-1874.

- Davies-Colley, R.J., Hicks, D.M., Hughes, A.O., Clapcott, J., Kelly, D., Wagenhoff, A. (2015) Fine sediment effects on freshwaters, and the relationship of environmental state to sediment load – a literature review, NIWA, NIWA Client Report HAM2015-104.
- Davies-Colley, R.J., Nagels, J.W. (2008) Predicting light penetration into river waters. *Journal of Geophysical Research* 113(G3): G03028. 10.1029/2008jg000722
- Davies-Colley, R.J., Smith, D.G. (2001) Turbidity, suspended sediment, and water clarity: a review. *Journal of the American Water Resources Association* 37(5): 1085-1101.
- Davies-Colley, R.J., Vant, W.N., Smith, D.G. (2003) Colour and clarity of natural waters. Science and management of optical water quality. Blackburn Press, Caldwell, New Jersey.
- Davies-Colley, R.M., Hicks, D.M., Hughes, A., Clapcott, J., Kelly, D., Wagenhoff, A. (2015) Fine sediment effects on freshwaters, and the relationship of environmental state to sediment load – a literature review. NIWA Client Report HAM2015-104 prepared for the Ministry for the Environment, June 2015.
- Dymond, J.R., Christian, R. (1982) Accuracy of discharge determined from a rating curve. *Hydrological Sciences Journal* 4: 493-504.
- Dymond, J.R., Herzig, A., Basher, L., Betts, H.D., Marden, M., Phillips, C., Ausseil, A-G., Clark, M., Roygard, J. (2016) Development of a New Zealand SedNet model for assessment of catchment-wide soil-conservation works. *Geomorphology* 257: 85-93.
- Elliott, A.H., Davies-Colley, R.J., Parshotam, A., Ballantine, D. (2013) Load-based approaches for modelling water clarity in streams at regional scale. *Water Science and Technology* 67(5): 1092-1096. 10.2166/wst.2013.597
- Ellis, N., Smith, S. J., Pitcher, C. R. (2012) Gradient forests: calculating importance gradients on physical predictors. *Ecology* 93(1): 156-168.
- Foster, I.D.L., Millington, R., Grew, R.G. (1993) The impact of particle size controls on stream turbidity measurement; some implications for suspended sediment yield estimation. In: *Erosion and Sediment Transport Monitoring Programmes in River Basins* (Proceedings of the Oslo Symposium, August 1992). IAHS Publ. no. 210: 51-62.
- Friedman, J. H., Meulman, J. J. (2003) Multiple additive regression trees with application in epidemiology. *Statistics in Medicine* 22(9): 1365-1381.
- Gall, M.P. (2013) Exceptional water clarity and optical purity in a sub-alpine lake. *Limnology and Oceanography* 58(2): 443-451.
- Gallegos, C.L., Davies-Colley, R.J., Gall, M.P. (2008) Optical closure in lakes with contrasting extremes of reflectance. *Limnology and Oceanography* 53(5): 2021-2034.
- Gray, B.A., Pasternack, G.B., Watson, E.B., Warrick, J.A., Goñi, M.A. (2015) Effects of antecedent hydrologic conditions, time dependence, and climate cycles on the suspended sediment load of the Salinas River, California. *Journal of Hydrology* 525: 632–649.

- Guo, Q. (2006) Correlation of Total Suspended Solids (TSS) and Suspended Sediment Concentration (SSC) test methods. Report prepared by the State University of New Jersey, Department of Civil and Environmental Engineering, for New Jersey Department of Environmental Protection, November 2006.
- Hastie, T., Tibshirani, R., Friedman, J. H. (2001) *The elements of statistical learning, data mining, inference and prediction*. New York, Springer-Verlag.
- Herbst, D.B., Roberts, S.W., Hayden, N.G. (2011) Comparison of Sediment Load Models in Predicting Sediment Deposition Patterns in Streams of the Sierra Nevada and Central Coast of California. Report to State Water Resources Control Board. Report 2 of 4.
- Hicks D.M., Shankar, U., McKerchar, A., Basher, L., Lynn, I., Page, M., Jessen, M. (2011) Suspended Sediment Yields from New Zealand Rivers. *Journal of Hydrology (New Zealand)* 50(1): 81-142.
- Hicks, D.M., Fenwick, J.K. (1993) Suspended sediment manual - field, laboratory and office procedures for collecting and processing suspended sediment data. *NZ Freshwater Miscellaneous Report* No. 91, NIWA Christchurch, June 1993, 84 p.
- Hicks, D.M., Gomez, B. (2016) Sediment Transport. In: Kondolf, G.M. and Piégay, H. (Eds). *Tools in Fluvial Geomorphology 2nd Edition*, John Wiley and Sons, Ltd: 324-356.
- Hicks, D.M., Gomez, B., Trustrum, N.A. (2000) Erosion thresholds and suspended sediment yields: Waipaoa River basin, New Zealand. *Water Resources Research* 36(4): 1129-1142.
- Hicks, D.M. (2008) Review and analysis of Amethyst River suspended sediment data. NIWA Client Report CHC2008-155 prepared for Mitton Electronet Ltd, November 2008.
- Hicks, D.M.; Basher, L.R. (2008) The signature of an extreme erosion event on suspended sediment loads: Motueka River Catchment, South Island, New Zealand. In: *Sediment Dynamics in Changing Environments* (Proceedings of a symposium held in Christchurch, New Zealand, December 2008). IAHS Publ. 325: 184-191.
- Hicks, M., Hoyle, J. (2012) Analysis of suspended sediment yields from the rivers in the Horizons sediment monitoring program. NIWA Client Report No. CHC2012-107 prepared for Horizons Regional Council. NIWA, Christchurch, December 2012.
- Hicks, M., Watson, B., Watson, J. (2013) Measurement, processing and archiving of turbidity data: Version 1.0. *National Environmental Monitoring Standards*, Ministry for the Environment, Wellington.
- Hicks, M; Quinn, J.; Trustrum, N. (2004) Sediment load and organic matter. Chapter 12 In: Harding, J.S.; Mosley, M.P.; Pearson, C.P.; Sorrell, B.K. (Eds.). *Freshwaters of New Zealand*. New Zealand Hydrological Society and New Zealand Limnological Society, Wellington, 764p.
- Holmes, R., Hayes, J., Matthaei, C., Closs, G., Williams, M., Goodwin, E. (2016) Riparian management affects instream habitat condition in a dairy stream catchment. *New Zealand Journal of Marine and Freshwater Research* in press.
- Hughes, A.O., Davies-Colley, R.J., Elliott, A.H. (2014) Measurement of light attenuation extends the application of suspended sediment monitoring in rivers. In: Y.J. Ku, M.A.

- Allison, S.J. Bentley et al. (Eds). *Sediment Dynamics from the Summit to the Sea*. IAHS Press, Wallingford, UK: 170-176.
- Hughes, A.O., Quinn, J.M. (2014) Before and After Integrated Catchment Management in a Headwater Catchment: Changes in Water Quality. *Environmental Management* 54(6): 1288-1305. 10.1007/s00267-014-0369-9.
- Hughes, A.O., Quinn, J.M., McKergow, L.A. (2012) Land use influences on suspended sediment yields and event sediment dynamics within two headwater catchments, Waikato, New Zealand. *New Zealand Journal of Marine and Freshwater Research* 46(3): 315-333.
- Kettner, A.J, Syvitski, J.P.M. (2008) HydroTrend v.3.0: A climate-driven hydrological transport model that simulates discharge and sediment load leaving a river system. *Computers & Geosciences* 34: 1170–1183.
- Kilroy, C., Biggs, B.J.F. (2002) Use of the SHMAK clarity tube for measuring water clarity: Comparison with the black disk method. *New Zealand Journal of Marine and Freshwater Research* 36(3): 519-527. 10.1080/00288330.2002.9517107
- Kirk, J.T.O. (1985) Effects of suspensoids (turbidity) on penetration of solar radiation in aquatic ecosystems. *Hydrobiologia* 125: 195-208.
- Kirk, J.T.O. (2011) *Light and Photosynthesis in Aquatic Ecosystems*. Cambridge University Press, New York: 649.
- Lambert, C.P., Walling, D.E. (1988) Measurement of channel storage of suspended sediment in a gravel-bed river. *Catena* 15(1): 65-80.
- Lane, P.W. (1955) The importance of fluvial morphology in hydraulic engineering. *American Society of Civil Engineering Proceedings* 81(745): 1-17.
- Leathwick, J., Morgan, F., Wilson, G., Rutledge, D., McLeod, M., Johnston, K. (2002) *Land Environments of New Zealand: A Technical Guide*. Landcare Research, Hamilton, 237 pp.
- MfE (2015) *Sediment attribute development scoping document*. Unpublished document, Ministry for the Environment, Wellington.
- Morehead, M.D., Syvitski, J.P., Hutton, E.W.H., Peckham, S.D. (2003) Modeling the temporal variability in the flux of sediment from ungauged river basins. *Global and Planetary Change* 39: 95– 110.
- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., Veith. T. L. (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the American Society of Agricultural and Biological Engineers* 50: 885-900.
- Naden, P.S., Murphy, J.F., Old, G.H., Newman, J., Scarlett, P., Harman, M., Duerdoth, C.P., Hawczak, A., Pretty, J.L., Arnold, A., Laize, C., Hornby, D.D., Collins, A.L., Sear, D.A., Jones, J. (2016) Understanding the controls on deposited fine sediment in the streams of agricultural catchments. *Science of the Total Environment* 547: 366-81.

- Nash, J. E., Sutcliffe, J. V. (1970) River flow forecasting through conceptual models: Part 1. A discussion of principles. *Journal of Hydrology* 10: 282-290.
- NIWA (2013) Quorer: a simple method for estimating deposited fine sediment. <https://www.niwa.co.nz/our-science/freshwater/tools/estimating-deposited-fine-sediment/quorer-field-and-laboratory-protocol>.
- Pavanelli, D., Selli, L. (2013) Effective size characteristics of suspended sediment and nutrient concentrations during flood events in the Reno River tributaries (Northern Italy). *Procedia Environmental Sciences* 19: 723-732.
- Pineiro, G., Perelman, S., Guerschman, J. P., Paruelo, J. M. (2008) How to evaluate models: Observed vs. predicted or predicted vs. observed? *Ecological Modelling* 216(3-4): 316-322.
- Pope, R.M., Fry, E.S. (1997) Absorption spectrum (380-700 nm) of pure water. II. Integrating cavity measurements. *Applied Optics* 36(33): 8710-8723.
- Quinn, J.M., Croker, G.F., Smith, B.J., Bellingham, M.A. (2009) Integrated catchment management effects on flow, habitat, instream vegetation and macroinvertebrates in Waikato, New Zealand, hill-country streams. *New Zealand Journal of Marine and Freshwater Research* 43: 775-802.
- Snelder, T. H., Biggs, B. J. F. (2002) Multi-scale river classification for water resources management. *Journal of the American Water Resources Association* 38: 1225-1240.
- Tran, V.T.T. (2014) An analysis of the suspended sediment rating curve parameters in the Upper Mississippi River Basin at the monthly and annual levels. Master's Thesis, University of Tennessee. http://trace.tennessee.edu/utk_gradthes/2854.
- Walling, D.E., Owens, P.N., Waterfall, B.D., Leeks, G.J.L., Wass, P.D. (2000) The particle size characteristics of fluvial suspended sediment in the Humber and Tweed catchments, UK. *The Science of the Total Environment* 251: 205-222.
- Walling, D.E., Woodward, J.C. (2000) Effective particle size characteristics of fluvial suspended sediment transported by lowland British rivers. In: *The Role of Erosion and Sediment Transport in Nutrient and Contaminant Transfer*, Proceedings of a symposium held at Waterloo, Canada, July 2000. IAHS Publ. no. 263: 129-139.
- Warrick, J.A. (2015) Trend analyses with river sediment rating curves. *Hydrological Processes* 29: 936-949.
- Woods, R., Hendrikx, J., Henderson, R., Tait, A. (2006) Estimating mean flow of New Zealand rivers. *Journal of Hydrology (New Zealand)* 45: 95-110.
- Yang, G., Chen, Z., Yu, F., Wang, Z., Zhao, Y., Wang, Z. (2007) Sediment rating parameters and their implications: Yangtze River, China. *Geomorphology* 85: 166-175.
- Zanevald, J.R.V., Pegau, W.S. (2003) Robust underwater visibility parameter. *Optics Express* 11: 2997-3009.

Appendix A Statement of work – Description of services

In providing the Services, the Contractor will use existing data to develop methods that will aim to link catchment sediment loads to the to the ESVs at a level of accuracy and precision that enables regional councils to set sediment related objectives, and to justifiably put in place actions and limits.

The Contractor shall provide Services to undertake the necessary research and development for the Ministry's purposes, including:

- (a) Collating all the nationally available sediment rating curves (SRC) and data (up to 30 June 2015) from which sediment rating curves can be defined held by research institutes and territorial authorities, including data previously collated as part of the Ministry-funded Auckland Council project "Integrating three regional council sediment monitoring datasets for the purposes of calibrating a sediment yield predictive model for freshwater catchments".
- (b) Collating all the nationally available data held by research institutes and territorial authorities (up to 30 June 2015) to support the development of methods to link sediment concentrations to the ESVs including: flow data and flow duration curves, measured water clarity data, measured turbidity data, measured deposited sediment data and particle size distribution (PSD) data.

The applicable data along with appropriate metadata (collectively, the Sediment Data) will be sourced by the Contractor from all readily available sources including existing datasets held by the project team organisations (NIWA, Landcare, Cawthron), territorial authorities (including data in LAWA), and Scion. The parties acknowledge there is a risk of delay when collating the Sediment Data. Accordingly, the Contractor will provide sufficient resource to complete the exercise within the requisite timeframes. It is acknowledged by the parties that supply of the data remains dependent on the source organisations, and the analysis stages will proceed with data to hand by 31 January 2016.

- (c) Carrying out an assessment of data quality of the Sediment Data and identification of a set of sites that maximises the geographic and environmental coverage of New Zealand's river catchments for which the SRC information is of sufficient quality for the analyses that follow.
- (d) Developing regional models (regionalisation(s)) of the parameters of the SRC so that these can be estimated for locations without data.
- (e) Determining how the (parameters of the) SRC change in response to changes in catchment loads.
- (f) Identifying and characterising the relationship between turbidity, water clarity, suspended sediment concentration (SSC) and light penetrations and develop regionalisation(s) to provide methods for predicting turbidity and clarity and light penetration as functions of SSC.
- (g) Using the applicable Sediment Data, examine whether PSD changes appreciably with changes in catchment load. If possible examine the extent to which PSD changes with change in load and in what circumstances. It is expected that this study would be conducted using PSD data if available or using specific turbidity as a proxy measure.
- (h) Looking for and identifying relationships between sediment loads and measures of streambed sediment deposition within the applicable Sediment Data. Based on the findings conclude if an

empirical approach has promise and, if it does, scope what new data is required to deliver functional relationships.

- (i) Where appropriate, and based on the outputs of the above studies, providing analytical frameworks for the use of these methods to determine catchment sediment load limits to achieve objectives that are enumerated in terms of the ESVs. The frameworks should include a thorough description of the “analytical chain”, including the required data, the necessary assumptions and approximations and their robustness in different circumstances and the necessary models.
- (j) Where possible, estimating and describing the sensitivity of the each step in the analytical chain and indicate the steps that most limit the accuracy of the analysis.
- (k) Providing a framework for further work to develop more sophisticated and accurate methods for relating ESVs to catchment loads. The framework should provide estimates of the amount of new data needed including the number of sites, sampling methods and frequencies and considerations for site location. The framework should describe how the data will be used to develop new methods or conduct research to test new ideas and assumptions.

These Services will be carried out in accordance with the methodology specified by the Contractor in its proposal in response to the RFP. The key outputs of the Services will be:

1. The collated and ‘cleaned’ Sediment Data used for the project (Collated Sediment Data)
2. Derived parameters required to implement predictive models (Derived Parameters)
3. A report detailing the Contractor’s methodology and findings for the project (Report)
4. A project summary workshop with the Ministry at the conclusion of the work (Workshop).

Appendix B Data-seek letter

The following letter was mailed to territorial authorities requesting data.

MFE16502

Date

Council Name and Address

Attention: Contact First & 2nd name

Dear First Contact Name

NOF-sediment Stage I: Collation of sediment-related data

Background

This follows up the letter of 11 December that you will have received from the Ministry for the Environment (MfE) regarding a request for sediment and related data held by your organisation. As explained by MfE, it has contracted NIWA to help investigate developing sediment as an NPS-FM attribute – that is, developing a framework for setting limits on river/stream sediment loads to meet thresholds for sediment-related attributes.

We are working with Landcare Research and Cawthron Institute on this project, and our first task is to compile data and metadata that will explore relationships between suspended sediment load and sediment-related attributes likely to be specified in the NOF, notably:

- water clarity
- **light penetration**
- **suspended sediment concentration (SSC, also TSS) and**
- **deposited sediment.**

Also highly relevant are:

- turbidity – which is a widely-measured, valuable, site-specific and often continuous proxy for SSC and water clarity, and
- particle-size (PS) – which is a major control on the inter-relationships between SSC and water clarity and turbidity.

Data needed

NIWA already have access to an existing compiled dataset of regional (usually monthly) state-of-environment (SoE) water quality data for rivers, including water clarity, turbidity and (sometimes) TSS up to the end of 2013. We also have data from the NRWQN, pairing TSS with water clarity and turbidity. Thus generally there is no need for you to supply data associated with either of those programmes (the exception would be supplementary data, such as data from event-sampling, which was not captured into our existing SoE database).

We do seek data from/on:

- Dedicated **sediment load/concentration monitoring** campaigns for SSC/TSS, turbidity, and water clarity. In particular with these, we are looking for:

- datasets with **concurrent measurements of combinations of SSC, TSS, turbidity, water clarity** – such as collected to ‘calibrate’ turbidity to SSC or water clarity.
- datasets relating **point to cross-sectional average SSC** – as developed to ‘calibrate’ bankside auto-sample measurements to all-of-section measurements
- ‘**rating curve**’ datasets, relating SSC (and/or turbidity and/or water clarity) to concurrent water discharge.
- **Particle characteristics of suspended sediment** that may influence the environmental behaviour of sediment and its optical properties, particularly:
 - particle size distribution (PSD),
 - particle shape
 - particle composition (especially organic content as indicated by VSS and/or POC).
- **Deposited sediment in channels:**
 - collected using any of the Sediment Assessment Methods, e.g., SAM1: visual bankside assessment of % sediment cover, SAM2: instream visual assessment of % sediment cover, SAM3: Wolman pebble count, SAM4: resuspendible sediment (Quorer method), SAM5: resuspendable sediment (shuffle index).
 - any related analyses of deposited sediment composition, including: organic content of resuspendable sediment (VSS and/or POC); sub-surface bed-material size-grading by sieving.

We do not anticipate that much data on **light penetration** exists, but if you do have some that is concurrent with water clarity, turbidity, and/or TSS, we would appreciate that too.

Metadata

A key task will be for us to assess the quality of data. Thus we request that you supply **meta-data** for each dataset, including:

- **General information:** site (name, number, reach no, coordinates), date-time, location of sample/measurement (e.g., bankside, fixed/varying location, data-collection purpose).
- **Sampling/measurement method:** e.g., deposited sediment – see SAM1 to SAM5 above.
- **Instruments used:** e.g., for turbidity - the instrument, the standard/protocol that it follows (e.g., ISO-7027, EPA-180.1, other), details of calibration (e.g., was the calibration regularly checked, updated as need be); for SSC or TSS – auto-sampled, dip-sampled, depth-integrated.
- **Laboratory procedure:** e.g., TSS or SSC procedure used for sediment concentration analysis.
- **Relevant literature:** e.g., pdf copies of reports detailing the data-collection programmes if available.

Delivery details

We would be very grateful if you could respond to this important national initiative as a priority, and provide the data and associated metadata by 31 January 2016 at the latest. As soon as possible before then would be preferable, as we have a tight timeline.

Could you please provide the data as email attachments to Kathy Walter (Kathy.Walter@niwa.co.nz, ph 03 343 7897). Ideally, the data would be in excel tables, but text files would work as well. Please communicate with Kathy if files are too large for email attachment.

Please also contact Kathy in the first instance for general queries around this request.

Thank-you in advance for assisting with this project.

Yours sincerely,

A handwritten signature in black ink, appearing to read 'Murray Hicks'.

Murray Hicks
Principal Scientist

Appendix C Database conventions

Source database

For data auditing purposes, the source database was recorded for each data record using a standard list (Table C-1). These sources were either the original NIWA database (UnwinMfEDB), the suspended sediment database (SSG), the freshwater fish database (NZFFD), the SAM database (SAM) or individuals or organisations. A second level identified the organisation that collected the data (Table C-2). An example is data extracted from the pre-existing NIWA database but collected by Auckland Council: these were coded as 'UnwinMfEDB' for the source database and 'AKC' for the organisation. Data collected by Auckland Council and provided directly by them would have both the source database and the organisation 'AKC'.

Table C-1: Source databases assigned to incoming data.

SourceDB	Description
AKC	Direct from RC, see Source table for council codes
BOP	Direct from RC, see Source table for council codes
ElectroNet	Source: Electronet Ltd
ES	Direct from RC, see Source table for council codes
GDC	Direct from RC, see Source table for council codes
GWRC	Direct from RC, see Source table for council codes
HBRC	Direct from RC, see Source table for council codes
HRC	Direct from RC, see Source table for council codes
ICM	Landcare Research's Integrated Catchment Management Research Programme
JC SAM data	Joanne Clapcott Cawthron Institute (SAM methods DB)
LAN	Direct from Landcare
MDC	Direct from RC, see Source table for council codes
NCC	Direct from RC, see Source table for council codes
NIWA	Direct from NIWA
NZFFD	New Zealand Freshwater Fish Database
SSG	NIWA suspended sediment database
TDC	Direct from RC, see Source table for council codes
UnwinMfEDB	MFE15503 database held by NIWA created by Martin Unwin
WCRC	Direct from RC, see Source table for council codes
WRC	Direct from RC, see Source table for council codes

Table C-2: Organisations assigned to incoming data.

Organisation	Description
AEL	Aquatic Ecology Limited
AKC	Auckland Council
BIO	Bioresearches Group Ltd
BML	Boffa Miskell Limited
BOP	Bay of Plenty Regional Council
CHH	Carter Holt Harvey Forests
CI	Cawthron Institute
DOC	Department of Conservation
Doyle	SSG survey, see Murray Hicks
ECAN	Environment Canterbury
ES	Environment Southland
EVK	Envirolink
FG	Fish and Game
FWS	Tom Kroos Fish & Wildlife Services
GDC	Gisborne District Council
GEN	Genesis
GOLD	Golder Associates
GWRC	Greater Wellington Regional Council
Harding	Jon Harding, UC
HBRC	Hawkes Bay Regional Council
HRC	Horizons Regional Council
KAL	Kessels and Associates
KMA	Kingett Mitchell
LAN	Landcare
Matthaei	Christoph Matthaei, UOO
MDC	Marlborough District Council
MPI	Ministry for Primary Industries
NAT	NRWQN national data
NCC	Nelson City Council
NIW	National Institute of Water and Atmospheric Research
NRC	Northland Regional Council
NZFP	New Zealand Forest Products
OPUS	OPUS consulting services
ORC	Otago Regional Council
Quinn	John Quinn, NIWA
TDC	Tasman District Council

Organisation	Description
TRC	Taranaki Regional Council
UC	University of Canterbury
UM	Massey University
UNK	Unknown
UOO	University of Otago
UOW	University of Waikato
UV	Victoria University
WCRC	West Coast Regional Council
WRC	Waikato Regional Council

Site names

A unique (to this study) SSSiteID alphanumeric value was assigned to each block of data identified as coming from a particular site by the source organisation. The code begins with the source region/organisation or database code and then with 9999 site numbers. For example, AKC0001 is the first site supplied by AKC. Sites from the UnwinMfEDB begin at 5000, e.g., AKC5001 is the first site supplied by the Auckland Council within the UnwinMfEDB.

Because SSSiteIDs were assigned while sites were incorporated into the project and because the data came from multiple sources, it is possible for the same geographic site locations to have been assigned multiple SSSiteID values. Checks for site duplicates were conducted in the sediment rating task, where multiple records from one site location could bias the results of national predictions. This was done in a multi-stage process:

- The provided site coordinates were used to calculate straight line distances between sites on the same REC2 NZSegment.
- Sites further than 100 m apart were presumed to be separate. Compiled data from sites within 100 m of each other were compared to see if data combinations looked sensible as well as checking any other independent information available about a site, e.g., site name, catchment area etc. If site names, information and provided data looked similar for sites within 100 m they were combined into one site and given a composite SSSiteID. Note that the data provided in the MS-Access database was placed there before this site compilation stage and has not been checked for multiple SSSiteIDs per site.

Parameters

Parameters were collected using different methods and described under different names by the source organisations. We developed a consistent set of parameter names and descriptions to consolidate these methods (Table C-3). These were designed to cover all methods of sediment sample collection and processing that were relevant to this project. For example, discharge may be gauged (MHPid: 201), recorded on a site visit (MHPid: 202) or the method may be unspecified (MHPid: 203; Table C-3). Likewise, turbidity may be measured in the field or laboratory on a range of different instruments, with the turbidity MHPid codes 401 to 413 designed to capture this variability (Table C-3).

Table C-3: Parameter codes assigned to different sediment collection and processing methods. A MHPid code was assigned to each data entry using the descriptions here, based on expert opinion. Unspec. = unspecified.

MHPid	Name	Sampling Method	Morphological unit	Lab methods	Instrument	Units	Comment	Parameter Type
101	SSConc	MVDI		SSC		mg/l		SSConc
102	SSConc	MVDI		TSS	APHA 2540 D	mg/l		SSConc
103	SSConc	MVDI		Unspec.		mg/l		SSConc
104	SSConc	Auto-sample/grab		SSC		mg/l		SSConc
105	SSConc	Auto-sample/grab		TSS	APHA 2540 D	mg/l		SSConc
106	SSConc	Auto-sample/grab		Unspec.		mg/l		SSConc
107	SSConc	Unspec.		SSC		mg/l		SSConc
108	SSConc	Unspec.		TSS	APHA 2540 D	mg/l		SSConc
109	SSConc	Unspec.		Unspec.		mg/l		SSConc
201	Discharge	Gauged				m3/s		Discharge
202	Discharge	Recorded				m3/s		Discharge
203	Discharge	Unspec.				m3/s		Discharge
301	WaterClarity	Black Disk				m		WaterClarity
302	WaterClarity	Sechhi Disk				m		WaterClarity
303	WaterClarity	Unspec.				m		WaterClarity
351	g340	g340		spectrophotometric absorbance				g340
352	g440	g440		spectrophotometric absorbance				g440
401	Turbidity				Unspec.-field			Turbidity
402	Turbidity				Unspec.- lab			Turbidity
403	Turbidity				Non std - field			Turbidity
404	Turbidity				Non std -lab			Turbidity

MHPid	Name	Sampling Method	Morphological unit	Lab methods	Instrument	Units	Comment	Parameter Type
405	Turbidity				ISO-field			Turbidity
406	Turbidity				ISO-lab			Turbidity
407	Turbidity				EPA-field			Turbidity
408	Turbidity				EPA-lab			Turbidity
409	Turbidity				HACH16800			Turbidity
410	Turbidity				HACH2100A			Turbidity
411	Turbidity				HACH2100N			Turbidity
412	Turbidity				HACH2100P			Turbidity
413	Turbidity				Alpha2130			Turbidity
501	%Cover	SAM1	Unspec.			%		%Cover
502	%Cover	SAMS2	Unspec.			%		%Cover
503	%Cover	Visual qualitative scale	Unspec.					%Cover
504	%Cover	SAM3:Wolman (< 2 mm)	Unspec.			%		%Cover
505	%Cover	Other1	Unspec.					%Cover
506	%Cover	SAM1	Riffle			%		%Cover
507	%Cover	SAMS2	Riffle			%		%Cover
508	%Cover	Visual qualitative	Riffle			none		%Cover
509	%Cover	SAM3:Wolman (< 2 mm)	Riffle			%		%Cover
510	%Cover	Other1	Riffle					%Cover
511	%Cover	SAM1	Pool			%		%Cover
512	%Cover	SAMS2	Pool			%		%Cover
513	%Cover	Visual qualitative (IMBED)	Pool			none		%Cover
514	%Cover	SAM3:Wolman (< 2 mm)	Pool			%		%Cover
515	%Cover	Other1	Pool					%Cover

MHPid	Name	Sampling Method	Morphological unit	Lab methods	Instrument	Units	Comment	Parameter Type
516	%Cover	SAM1	Run			%		%Cover
517	%Cover	SAMS2	Run			%		%Cover
518	%Cover	Visual qualitative	Run			none		%Cover
519	%Cover	SAM3:Wolman (< 2 mm)	Run			%		%Cover
520	%Cover	Other1	Run			%		%Cover
601	Resuspendable fines	SAM4: Quorer	Unspec.			g/m2		Resuspendable fines
602	Resuspendable fines	SAM5: Shuffle	Unspec.			g/m2		Resuspendable fines
603	Resuspendable fines	Sieve	Unspec.			g/m2		Resuspendable fines
604	Resuspendable fines	Unspec.	Unspec.			g/m2		Resuspendable fines
605	Resuspendable fines	SAM4: Quorer	Riffle			g/m2		Resuspendable fines
606	Resuspendable fines	SAM5: Shuffle	Riffle			g/m2		Resuspendable fines
607	Resuspendable fines	Sieve	Riffle			g/m2		Resuspendable fines
608	Resuspendable fines	Unspecified	Riffle			g/m2		Resuspendable fines
609	Resuspendable fines	SAM4: Quorer	Pool			g/m2		Resuspendable fines
610	Resuspendable fines	SAM5: Shuffle	Pool			g/m2		Resuspendable fines
611	Resuspendable	Sieve	Pool			g/m2		Resuspendable fines

MHPid	Name	Sampling Method	Morphological unit	Lab methods	Instrument	Units	Comment	Parameter Type
	fines							
612	Resuspendable fines	Unspec.	Pool			g/m2		Resuspendable fines
613	Resuspendable fines	SAM4: Quorer	Run			g/m2		Resuspendable fines
614	Resuspendable fines	SAM5: Shuffle	Run			g/m2		Resuspendable fines
615	Resuspendable fines	Sieve	Run			g/m2		Resuspendable fines
616	Resuspendable fines	Unspec.	Run			g/m2		Resuspendable fines
617	Resuspendable fines	SAM6: Sediment depth	Run			mm		Resuspendable fines
651	%deposited sediment cover as sand	By sieve				%		%deposited sediment cover as sand
652	%deposited sediment cover as sand	Visual est.				%		%deposited sediment cover as sand
653	%deposited sediment cover as sand	Unspec.				%		%deposited sediment cover as sand
701	D50	MVDI: Wet sieve and settling				microns	Wet sieved sand, settled mud	SSGrainsize
702	<4um	MVDI: Wet sieve and settling				%	Wet sieved sand, settled mud	SSGrainsize
703	<16um	MVDI: Wet sieve and settling				%	Wet sieved sand, settled mud	SSGrainsize

MHPid	Name	Sampling Method	Morphological unit	Lab methods	Instrument	Units	Comment	Parameter Type
704	<31um	MVDI: Wet sieve and settling				%	Wet sieved sand, settled mud	SSGrainsize
705	<63um	MVDI: Wet sieve and settling				%	Wet sieved sand, settled mud	SSGrainsize
706	<125um	MVDI: Wet sieve and settling				%	Wet sieved sand, settled mud	SSGrainsize
707	<250um	MVDI: Wet sieve and settling				%	Wet sieved sand, settled mud	SSGrainsize
708	<500um	MVDI: Wet sieve and settling				%	Wet sieved sand, settled mud	SSGrainsize
709	<1000um	MVDI: Wet sieve and settling				%	Wet sieved sand, settled mud	SSGrainsize
710	<2000um	MVDI: Wet sieve and settling				%	Wet sieved sand, settled mud	SSGrainsize
711	D50	MVDI: Settling				microns	Settling analysis for all fractions	SSGrainsize
712	<4um	MVDI: Settling				%	Settling analysis for all fractions	SSGrainsize
713	<16um	MVDI: Settling				%	Settling analysis for all fractions	SSGrainsize
714	<31um	MVDI: Settling				%	Settling analysis for all fractions	SSGrainsize
715	<63um	MVDI: Settling				%	Settling analysis for all fractions	SSGrainsize
716	<125um	MVDI: Settling				%	Settling analysis for all fractions	SSGrainsize
717	<250um	MVDI: Settling				%	Settling analysis for all fractions	SSGrainsize
718	<500um	MVDI: Settling				%	Settling analysis for all fractions	SSGrainsize

MHPid	Name	Sampling Method	Morphological unit	Lab methods	Instrument	Units	Comment	Parameter Type
719	<1000um	MVDI: Settling				%	Settling analysis for all fractions	SSGrainsize
720	<2000um	MVDI: Settling				%	Settling analysis for all fractions	SSGrainsize
721	D50	MVDI: Laser diffn				microns	Laser diffraction based settling instrument	SSGrainsize
722	<4um	MVDI: Laser diffn				%	Laser diffraction based settling instrument	SSGrainsize
723	<16um	MVDI: Laser diffn				%	Laser diffraction based settling instrument	SSGrainsize
724	<31um	MVDI: Laser diffn				%	Laser diffraction based settling instrument	SSGrainsize
725	<63um	MVDI: Laser diffn				%	Laser diffraction based settling instrument	SSGrainsize
726	<125um	MVDI: Laser diffn				%	Laser diffraction based settling instrument	SSGrainsize
727	<250um	MVDI: Laser diffn				%	Laser diffraction based settling instrument	SSGrainsize
728	<500um	MVDI: Laser diffn				%	Laser diffraction based settling instrument	SSGrainsize
729	<1000um	MVDI: Laser diffn				%	Laser diffraction based settling instrument	SSGrainsize

MHPid	Name	Sampling Method	Morphological unit	Lab methods	Instrument	Units	Comment	Parameter Type
730	<2000um	MVDI: Laser diffn				%	Laser diffraction based settling instrument	SSGrainsize
731	D50	MVDI: Laser TOT				microns	Laser time-of-transit type instrument	SSGrainsize
732	<4um	MVDI: Laser TOT				%	Laser time-of-transit type instrument	SSGrainsize
733	<16um	MVDI: Laser TOT				%	Laser time-of-transit type instrument	SSGrainsize
734	<31um	MVDI: Laser TOT				%	Laser time-of-transit type instrument	SSGrainsize
735	<63um	MVDI: Laser TOT				%	Laser time-of-transit type instrument	SSGrainsize
736	<125um	MVDI: Laser TOT				%	Laser time-of-transit type instrument	SSGrainsize
737	<250um	MVDI: Laser TOT				%	Laser time-of-transit type instrument	SSGrainsize
738	<500um	MVDI: Laser TOT				%	Laser time-of-transit type instrument	SSGrainsize
739	<1000um	MVDI: Laser TOT				%	Laser time-of-transit type instrument	SSGrainsize
740	<2000um	MVDI: Laser TOT				%	Laser time-of-transit type instrument	SSGrainsize
751	D50	Point: Wet seive and settling				microns		SSGrainsize
752	<4um	Point: Wet seive and settling				%		SSGrainsize
753	<16um	Point: Wet seive and settling				%		SSGrainsize

MHPid	Name	Sampling Method	Morphological unit	Lab methods	Instrument	Units	Comment	Parameter Type
754	<31um	Point: Wet seive and settling				%		SSGrainsize
755	<63um	Point: Wet seive and settling				%		SSGrainsize
756	<125um	Point: Wet seive and settling				%		SSGrainsize
757	<250um	Point: Wet seive and settling				%		SSGrainsize
758	<500um	Point: Wet seive and settling				%		SSGrainsize
759	<1000um	Point: Wet seive and settling				%		SSGrainsize
760	<2000um	Point: Wet seive and settling				%		SSGrainsize
761	D50	Point: Settling				microns		SSGrainsize
762	<4um	Point: Settling				%		SSGrainsize
763	<16um	Point: Settling				%		SSGrainsize
764	<31um	Point: Settling				%		SSGrainsize
765	<63um	Point: Settling				%		SSGrainsize
766	<125um	Point: Settling				%		SSGrainsize
767	<250um	Point: Settling				%		SSGrainsize
768	<500um	Point: Settling				%		SSGrainsize
769	<1000um	Point: Settling				%		SSGrainsize
770	<2000um	Point: Settling				%		SSGrainsize
771	D50	Point: Laser Diffn				microns		SSGrainsize
772	<4um	Point: Laser Diffn				%		SSGrainsize
773	<16um	Point: Laser Diffn				%		SSGrainsize

MHPid	Name	Sampling Method	Morphological unit	Lab methods	Instrument	Units	Comment	Parameter Type
774	<31um	Point: Laser Diffn				%		SSGrainsize
775	<63um	Point: Laser Diffn				%		SSGrainsize
776	<125um	Point: Laser Diffn				%		SSGrainsize
777	<250um	Point: Laser Diffn				%		SSGrainsize
778	<500um	Point: Laser Diffn				%		SSGrainsize
779	<1000um	Point: Laser Diffn				%		SSGrainsize
780	<2000um	Point: Laser Diffn				%		SSGrainsize
781	D50	Point: Laser TOT				microns		SSGrainsize
782	<4um	Point: Laser TOT				%		SSGrainsize
783	<16um	Point: Laser TOT				%		SSGrainsize
784	<31um	Point: Laser TOT				%		SSGrainsize
785	<63um	Point: Laser TOT				%		SSGrainsize
786	<125um	Point: Laser TOT				%		SSGrainsize
787	<250um	Point: Laser TOT				%		SSGrainsize
788	<500um	Point: Laser TOT				%		SSGrainsize
789	<1000um	Point: Laser TOT				%		SSGrainsize
790	<2000um	Point: Laser TOT				%		SSGrainsize
801	%Organics			VSS/TSS x 100		%	% organics by mass of total SS	SSOrganicContent
810	OrganicConc			VSS		mg/l	mass conc of particulate organics	SSOrganicContent

Source specific data information

Deposited sediment data

Data from the New Zealand Freshwater Fish Database (NZFFD) were only extracted from records that had been 'approved' in the database, and records that were recorded as in lake, wetland, pond or reservoir water body types were excluded. Note however, that some localities have 'lake' in their name. The variable 'reachtype' in REC2 can be used to exclude lakes.

Deposited sediment information was entered into the NZFFD as either %mud/silt or %sand. The %fines variable used in our analyses was calculated by summing across these two categories. The percentage of fines that was sand was retained as a separate parameter as well.

Some records within the NZFFD have reported percentage coverage of different types of morphological habitat within a fished reach. As the habitat (run, riffle, pool etc.) can influence the amount of deposited sediment we created a category for the dominant morphological habitat within a reach. If a morphological habitat category was assigned a value of >50% it was assigned as the morphological category for that site. Samples that had no morphological habitat data within the NZFFD were given a value of NA and samples where there was no dominant habitat type were left blank.

The samples were assigned MHPid values relevant to each dominant morphological habitat and associated with the 'unspecified method' (Table C-4). Morphological habitats 'cascade' and 'rapid' were incorporated with 'riffle' habitat.

Table C-4: Assignment of deposited sediment records in the New Zealand Freshwater Fish Database (NZFFD) to the new parameter categories used in this study (MHPid). Refer Table 4 for MHPid explanations.

Dominant habitat from NZFFD	No. records	MHPid	MHPParameter name	MHSampling method
Pool	1008	515	%Cover	Other1
Run	3662	520	%Cover	Other1
Riffle	1136	510	%Cover	Other1
Rapid	69	510	%Cover	Other1
Cascade	23	510	%Cover	Other1
Blank	3937	505	%Cover	Other1
NA	1243	505	%Cover	Other1

Some source organisations in the NZFFD were combined into larger groupings in the Access database. For example, some individual DOC area offices were recorded separately in the NZFFD, while we combined these to one DOC organisation code. Likewise all Fish and Game offices were combined to a single FG organisation code. If an organisation was relatively unknown (i.e., not a council, DOC, Fish and Game or large consultancy firm) and had less than 10 records it was added to the 'unknown' source organisation category (Table C-2).

Deposited sediment data from Landcare (source organisation LAN) were provided with %fines defined as %cover of sand and finer (<2mm). The values are time-space averages, i.e., averaged over a study reach and also averaged over datasets collected in 2005-2009 (5 years). All deposited sediment data from Landcare was collected from runs.

Other data

Additional information regarding data collection from other organisations can be found in the original files returned from the source organisation.

Appendix D MS-Access database tables

Main data table (DataTable)

All data used or processed for the project were stored in this table (Table 5). The DataTable is linked to the parameter description tables (MHPParameters and FullOriginalParameters) as well as to tables with site, source and catchment information (SiteInfo, SourceDB, Source, MeasureCatch and CatchmentInfo tables; Figure 2-1).

Table D-1: Column descriptions of the main data table in the MS-Access database (DataTable).

Column name	Description	Notes
mID	A unique numeric identifier per row	
SourceDB	Source database or organisation (See Table C-1)	
SSSiteID	The site id code assigned during this project	Note that codes were assigned as data arrived. Individual locations may have multiple SSSiteIDs.
Organisation	Source organisation (see Table C-2)	
Programme	The type of programme the data was collected for, if identified by source organisation	SOE: state of the environment, LU: local unspecified, RWQMP: recreational water quality monitoring programme, MOTFS: Motueka fine sediment study, SAM: SAM development, NRWQN: national river water quality monitoring, EI: Environmental impact. Others as provided by source organisations.
NZFFDcardref	Original card reference number for NZFFD records extracted for deposited sediment task	Only relevant to deposited sediment task (Section 7).
OrigSiteName	Site name as provided	
Easting	Easting in NZTM as provided	Any necessary conversions were done on the LINZ website
Northing	Northing in NZTM as provided	Any necessary conversions were done on the LINZ website
SampDate	Date sampling was conducted	Likely to be in different formats and to contain some time information
SampTime	Time sampling was conducted	Likely to be in different formats and to contain some date information
PointSample	For SS, were samples point samples or depth integrated	
MorphUnit	The type of morphological unit (run, riffle, pool etc) sampled where given by source	

Column name	Description	Notes
OrigParamID	Original name or number descriptor of parameter type	This largely corresponds to the UnwinMfEDB where there is a table of parameter ids. For other SourceDBs you may need to return to the original return files
MHPid	Unique parameter descriptor assigned in this database to different parameters (see Table C-3)	
Comment		
dval	Data value	
MfETaskNo	Task in current MfE project that the data were extracted for	Note, not all data extracted for a task were used in it.

Site information tables

Information relating to site location was stored in the SiteInfo table (Table D-2). Parameters extracted from the REC2 database for the site or its catchment were stored in the CatchmentInfo table (Table D-3), while measured information at each site was stored in the MeasureCatch table (Table D-4).

Note that the Easting and Northing coordinates for each site as included in the SiteInfo table (Table D-2) are the ones used to assign the REC2 NZSegments for that site.

Note also that Eastings and Northings are also included in the DataTable (Table D-1). These are the values as provided by the source organisations. In some cases they vary slightly within a site (e.g., a slightly different GPS location was recorded each time the site was visited).

Table D-2: Explanation of columns in SiteInfo table in MS-Access database.

Column name	Description	Notes
mID	A unique numeric identifier per row	
SSSiteID	The site id code assigned during this project	Note that codes were assigned as data arrived. Individual locations may have multiple SSSiteIDs.
SourceDB	Source database or organisation (See Table C-1)	
Organisation	Source organisation (see Table C-2)	
RiverName	Name of the river as provided by the source	
SiteName	Site name	
Easting	Easting in NZTM as provided	Any necessary conversions were done on the LINZ website
Northing	Northing in NZTM as provided	Any necessary conversions were done on the LINZ website

Column name	Description	Notes
HydroSiteNo	Hydrological gauge no, where known	Some of the OrigSiteID values may also be gauge numbers
REC2Segment	NZSegment of REC2	Assigned using a range of methods: 1) as provided, 2) matching site coordinates to nearest segment and 3) manual checking. See NZSegmentCheck column
OrigSiteID	Site identifier as provided	Some may be gauge numbers
OrigSiteName	Site name as provided	
NZSegmentCheck	Indicates if the NZSegment assignment to a site was checked manually	
Used_in_analysis	Indicates if the site was used in an analysis	Note this column may not be fully complete as individual tasks were conducted by separate researchers. Sites may have been provided but were not used in the final analysis
Combined	Indicates which sites were combined, if any. Only done for Task 4, sediment rating curves analysis.	Sites were combined if <100 m apart and site information and data matched. Only done for SRC task. No other sites were checked.

Table D-3: Extracted catchment information. Columns preceded with 'REC2' were extracted from the REC2 database if not so mentioned in Description.

Column name	Description	Units	Notes
REC2Segment	NZSegment of REC2		Segments in this table will have been checked against site locations
sedyield.t.y.	Sediment yield as extracted from WRENZ model or as observed for some segments as used in the SRC model task	t/y	For details see Hicks et al. 2011
maFlood.cumecs	From REC2	m ³ /s	
mean_flow.cumecs	From REC2	m ³ /s	
segment.slope.degrees	From REC2	degrees	
width.m..at.mean.flow	Width at mean flow as extracted from channel width model	m	For details see Booker and Hicks 2013
catch.area..km2	From REC2	km ²	
REC2_MeanFlowCumecs	Mean flow	m ³ /s	
REC2_us_slope	Upstream mean slope of the watershed	degrees	
REC2_us_hard	Upstream induration or hardness value	Ordinal scale	
REC2_us_mat	Upstream mean maximum air temperature	°C	

Column name	Description	Units	Notes
REC2_us_elev	Upstream mean elevation above sea level of the watershed	m	
REC2_downElev	Height (asl) of the downstream end of the segment	m	
REC2_us_catarea	upstream catchment area	m ²	
REC2_StreamOrder	Strahler stream order	Ordinal	
REC2_us_LCDB3_x	Upstream area of different land cover categories in the LCDB.	m ²	For details of each category see Leathwick et al. (2002)
REC2_REC1_SRC_OF_FLW	Source of Flow categories from REC1		For details see Snelder and Biggs (2002)
REC2_REC1_CLIMATE	Climate categories from REC1		For details see Snelder and Biggs (2002)
REC2_us_rain	mean annual upstream rain	mm	
REC2_us_psize	Upstream catchment average of particle size	Ordinal	
REC2_seg_elev	Segment mean elevation above sea level	m	
REC2_us_ind_forest	Upstream area with indigenous vegetation	m ²	
REC2_US_RockPhos	Average phosphorous concentration of underlying rocks	1= low to 5 = high	
REC2_USCalcium	Average calcium concentration of underlying rocks	1= low to 5 = high	
REC2_REC1_GEOLOGY	Geology categories from REC1		For details see Snelder and Biggs (2002)
Prop_us_Grassland	Composite category of LCDB classes from REC2: sum(LCDB3_40, LCDB3_41, LCDB3_44)/ catchment area	proportion	For details of each category see Leathwick, Morgan et al. (2002)
Prop_us_Scrubland	Composite category of LCDB classes from REC2: sum(LCDB3_43, LCDB3_51, LCDB3_52, LCDB3_55, LCDB3_56, LCDB3_58) / catchment area	proportion	For details of each category see Leathwick, Morgan et al. (2002)
Prop_us_Forest	Composite category of LCDB classes from REC2: sum(LCDB3_64, LCDB3_68, LCDB3_52, LCDB3_69, LCDB3_71) / catchment area	proportion	For details of each category see Leathwick, Morgan et al. (2002)
Prop_us_Agland	Composite category of LCDB classes from REC2: sum(LCDB3_30, LCDB3_33) / catchment area	proportion	For details of each category see Leathwick, Morgan et al. (2002)

Column name	Description	Units	Notes
Prop_us_Bareland	Composite category of LCDB classes from REC2: sum(LCDB3_10, LCDB3_12, LCDB3_16) / catchment area	proportion	For details of each category see Leathwick, Morgan et al. (2002)

Table D-4: Measured catchment or reach parameters as provided by the source organisation or this study. These are linked to the SiteInfo table through SSSiteID.

Column name	Description
SSSiteID	The site id code assigned during this project
Measured gradient	Channel gradient of the reach (m/m) as provided by the source
Measured width	Channel width (m) as provided by source

Original parameter descriptions table (FullOriginalParameters)

This table (Table D-5) links the original parameter descriptions and names from the source organisations to the MHPid values developed for this study. This was largely only done for the data from the UnwinMfEDB using information that was already in that database. For parameters that do not have associated OrigParamID values this information is stored in the original returned files from the source organisations.

Table D-5: FullOriginalParameters table in MS-Access database.

Column name	Description	Notes
OrigParamID	Original parameter identification number	This matches with values in the UnwinMfEDB. Many others do not have OrigParamID values and will need to be tracked through the original returned files
INTid	An intermediate identification step used in generating MHPid values	Retained for auditing purposes only
MHPid	Unique parameter descriptor assigned in this database to different parameters (see Table C-3)	
RCLabel	Parameter label as provided by the source (generally regional council)	As extracted from UnwinMfEDB
RCDescription	Parameter description as provided by the source	As extracted from UnwinMfEDB
RCComment	Comment as provided by the source	As extracted from UnwinMfEDB
RCUnits	Parameter units as provided by the source	As extracted from UnwinMfEDB
NIWAUnits	Units used in NIWA database	As extracted from UnwinMfEDB
NIWAName	Parameter name as used in NIWA database	As extracted from UnwinMfEDB

Column name	Description	Notes
NIWADescription	Description as used in NIWA database	As extracted from UnwinMfEDB
CFact	Concentration factor multiplier to convert units from RCUnits to NIWAUnits.	As extracted from UnwinMfEDB

Appendix E Random Forest regression approach

Random Forest (RF) models are an ensemble of regression trees from which a final prediction is based on the predictions averaged over all trees. They were chosen as the method to model regression coefficients for SRC and PSD predictors in this study because they have several benefits over standard linear regression techniques. Because RFs are a non-parametric method that can handle non-linear relationships explicitly, these benefits include fewer assumptions about data structure and the shape of relationships between predictors and responses than parametric methods. RFs also have inbuilt cross-validation with models tested against data held out of the set used to create the predictions (Breiman 2001, Ellis et al. 2012).

Predictive performance is measured by R^2 (Breiman 2001, Ellis et al. 2012). The measure is analogous to the coefficient of determination (often referred to as r^2) in that it has a value of 0 when the model has no predictive power and 1 when it predicts perfectly, although it applies to the predictions made for each tree that were excluded from the bootstrap samples (out-of-bag (OOB) predictions) rather than the fitted data set. A hold-one-out cross-validation (CV) procedure was performed to evaluate uncertainties for the estimated response variables (Hastie et al. 2001). Scatterplots of observed versus CV predicted values were plotted for each index. These scatterplots were overlain with a linear regression with observed values on the y-axis as recommended by Pineiro et al. (2008).

In addition to calculating R^2 , RF model performance was evaluated using three model performance metrics (Moriassi et al. 2007): Nash-Sutcliffe efficiency (NSE); percent bias (pbias); and ratio of the root mean square error to the standard deviation of observed data (RSR). NSE ranges from $-\infty$ to 1, with 1 indicating a perfect match to predictions, 0 indicating predictions are as accurate as the mean of the observed data and negative values indicating that the observed mean is a better predictor of the model. Lower RSR and RMSE values indicate a better model fit.

Each RF provides a measure of predictor importance, and is evaluated by randomly permuting each predictor in turn and predicting the response for the OOB observations. The decrease in prediction performance is the measure of importance of the original predictor. Importance represents the contribution to accuracy of independent predictions for each explanatory variable and is equivalent to the error resulting from dropping a term from a linear model (Ellis et al. 2012).

Partial dependence plots (Cutler, Edwards et al. 2007) were used to investigate the shape of the relationships between important predictor variables and the response variables. Partial dependence plots also show the marginal contribution of a predictor to the response (i.e., the response as a function of the predictor when the other predictors are held at their mean value) (Friedman and Meulman 2003). We note that these plots are not a perfect representation of the relationship between each predictor and response, particularly if there are interactions between predictors, or predictors are strongly correlated; however, they provide useful information for interpreting the model (Friedman and Meulman 2003).

Appendix F Inter-relationships between visual clarity, turbidity, and total suspended sediment

Table F-1: Key results from SMA regression analysis of inter-relationships between VC, turbidity, and TSS at NRWQN sites. Table also includes median c^* at each site. Note that the intercepts are \log_{10} values.

NRWQN site code	Flow impacted	VC-TSS R ²	VC-TSS SMA Intercept	VC-TSS Slope	VC-Turbidity R ²	VC-Turbidity SMA Intercept	VC-Turbidity SMA Slope	Turbidity-TSS R ²	Turbidity-TSS SMA Intercept	Turbidity-TSS SMA Slope	Median c^*
AK1	No	0.9476	0.6199	-0.8141	0.9686	0.8247	-0.9249	0.9181	-0.4014	1.2886	0.6998
AK2	No	0.9243	0.4183	-0.6640	0.9217	0.6459	-0.8078	0.9370	-0.3427	1.2166	0.9309
AX1	Yes	0.4035	0.8233	-0.4932	0.7810	0.5973	-0.6459	0.3667	0.4582	1.3095	0.5756
AX2	Yes	0.9090	0.9686	-0.8020	0.7748	0.3617	-0.7900	0.8829	0.7567	0.9850	0.2773
AX3	No	0.7170	0.9310	-0.7613	0.7890	0.3864	-0.7433	0.6943	0.7154	0.9764	0.2455
AX4	Yes	0.1197	0.7270	-0.6821	0.8765	0.5367	-0.8471	0.0228	0.2791	1.2419	0.5635
CH1	No	0.9796	0.7036	-0.8200	0.9780	0.5164	-0.8309	0.9702	0.2283	1.0134	0.6421
CH2	No	0.9188	0.6615	-0.7463	0.9299	0.4612	-0.7404	0.9583	0.2684	0.9921	0.5143
CH3	No	0.9717	0.6153	-0.7566	0.9448	0.4763	-0.7814	0.9267	0.1837	1.0328	0.5083
CH4	No	0.8669	0.7559	-0.7858	0.8259	0.5398	-0.7869	0.8960	0.2749	1.0015	0.3336
DN1	No	0.9043	0.6493	-0.8072	0.8582	0.5174	-0.8402	0.9117	0.1634	1.0408	0.7041
DN10	Yes	0.1599	0.6165	-0.9373	0.6573	0.5735	-0.6613	0.0976	0.0459	0.7055	1.1042
DN2	No	0.9102	0.4496	-0.6214	0.8393	0.5275	-0.8091	0.8004	-0.1253	1.3020	1.1579
DN3	No	0.9560	0.7830	-0.9106	0.9543	0.5585	-0.8520	0.9266	0.2465	0.9357	0.6323
DN4	Yes	0.9336	0.7585	-0.8920	0.9169	0.4682	-0.8300	0.9034	0.3254	0.9305	0.6559
DN5	No	0.9567	0.8740	-0.9897	0.9483	0.4859	-0.8521	0.9294	0.3922	0.8610	0.5722
DN6	No	0.8939	0.7506	-0.7728	0.7711	0.5602	-0.7149	0.6349	0.2464	0.9250	0.6755
DN7	No	0.9194	0.6757	-0.9436	0.8665	0.5823	-0.8444	0.9084	0.0989	0.8949	1.0172
DN8	No	0.9510	0.6277	-0.7673	0.9491	0.5128	-0.7968	0.8794	0.1498	1.0385	0.7476
DN9	Yes	0.9790	0.6460	-0.7943	0.8943	0.4520	-0.7977	0.9285	0.2443	1.0042	0.7840

NRWQN site code	Flow impacted	VC-TSS R ²	VC-TSS SMA Intercept	VC-TSS Slope	VC-Turbidity R ²	VC-Turbidity SMA Intercept	VC-Turbidity SMA Slope	Turbidity-TSS R ²	Turbidity-TSS SMA Intercept	Turbidity-TSS SMA Slope	Median c*
GS1	No	0.8484	0.3835	-0.6671	0.8913	0.2474	-0.6770	0.9598	0.2040	1.0149	0.4138
GS2	No	0.9474	0.6086	-0.6888	0.9564	0.5509	-0.7987	0.9659	0.0837	1.1595	0.5525
GS3	No	0.9642	0.5831	-0.7052	0.9721	0.5834	-0.8408	0.9805	-0.0003	1.1923	0.5040
GS4	No	0.9828	0.7164	-0.8415	0.9722	0.6510	-0.8878	0.9837	0.0778	1.0551	0.6029
GY1	No	0.9050	0.6348	-0.7138	0.9003	0.2641	-0.6685	0.9478	0.5192	0.9365	0.4107
GY2	No	0.9554	0.5274	-0.6918	0.9572	0.4090	-0.7894	0.9837	0.1336	1.1654	0.5354
GY3	No	0.8963	0.6266	-0.7674	0.9575	0.4023	-0.7400	0.9414	0.2922	0.9642	0.7197
GY4	No	0.7974	0.8571	-0.9585	0.9364	0.4834	-0.8860	0.6868	0.3899	0.9243	0.6704
HM1	No	0.9561	0.7010	-0.7498	0.9649	0.4124	-0.7296	0.9780	0.3848	0.9732	0.5105
HM2	No	0.9669	1.1578	-1.0681	0.8302	0.5031	-0.7604	0.7344	0.6130	0.7119	0.4213
HM3	Yes	0.8312	0.8417	-0.9698	0.7739	0.3887	-0.7470	0.7932	0.4671	0.7702	0.6732
HM4	No	0.5490	0.6862	-0.7720	0.7128	0.2690	-0.5742	0.6687	0.5404	0.7439	0.5155
HM5	No	0.5723	1.1222	-1.0071	0.8300	0.5069	-0.8450	0.5884	0.6109	0.8390	0.3429
HM6	No	0.9625	0.6915	-0.8519	0.9502	0.5306	-0.8554	0.9464	0.1889	1.0042	0.8135
HV1	No	0.9753	0.6913	-0.7647	0.9681	0.6026	-0.8024	0.9854	0.1159	1.0494	0.7303
HV2	No	0.9714	0.6295	-0.7580	0.9553	0.4743	-0.7604	0.9644	0.2048	1.0033	0.7950
HV3	No	0.9393	0.6298	-0.7803	0.9339	0.4349	-0.7609	0.9736	0.2498	0.9751	0.7320
HV4	No	0.6863	0.6921	-0.7957	0.5945	0.5712	-0.7494	0.7032	0.1557	0.9397	0.9050
HV5	No	0.9245	0.5657	-0.7020	0.6044	0.4372	-0.8179	0.6546	0.1830	1.1652	0.5653
HV6	No	0.8831	0.6711	-0.7031	0.8625	0.4778	-0.7533	0.9428	0.2749	1.0714	0.7782
NN1	No	0.6036	0.6711	-0.7031	0.9567	0.5293	-0.7867	0.7144	0.4598	1.0506	0.4384
NN2	No	0.9868	0.7170	-0.7696	0.9796	0.6967	-0.8573	0.9726	0.0324	1.1320	0.8427
NN3	No	0.9827	0.7248	-0.8380	0.9838	0.6060	-0.8620	0.9788	0.1418	1.0286	0.6319
NN4	No	0.9905	0.7447	-0.8517	0.9766	0.6238	-0.8809	0.9788	0.1420	1.0343	0.5221

NRWQN site code	Flow impacted	VC-TSS R ²	VC-TSS SMA Intercept	VC-TSS Slope	VC-Turbidity R ²	VC-Turbidity SMA Intercept	VC-Turbidity SMA Slope	Turbidity-TSS R ²	Turbidity-TSS SMA Intercept	Turbidity-TSS SMA Slope	Median c*
NN5	No	0.9836	0.8322	-0.8848	0.9685	0.6145	-0.8765	0.9739	0.2460	0.9905	0.5489
RO1	Yes	0.3525	0.7779	-0.3839	0.3290	0.5951	-0.5315	0.0048	0.4761	1.3842	0.6557
RO2	Yes	0.0556	1.2020	-0.9774	0.2692	0.4060	-0.8601	0.0214	0.9470	0.6375	0.2701
RO3	No	0.5006	1.1101	-0.9650	0.4682	0.2769	-0.5657	0.2219	0.8635	0.5862	0.3197
RO4	No	0.1534	0.8311	-0.7624	0.9418	0.3975	-0.6985	0.1348	0.5687	0.9162	0.3976
RO5	Yes	0.1580	0.6937	-0.8662	0.8630	0.4374	-0.7782	0.1510	0.2960	0.8984	0.8320
RO6	Yes	0.6687	0.8769	-0.4824	0.4124	0.4837	-0.8091	0.2299	0.8138	1.6686	0.6873
TK1	No	0.9205	0.6992	-0.8503	0.9548	0.4907	-0.7705	0.9044	0.2452	0.9062	0.8141
TK2	No	0.9394	0.6766	-0.7584	0.9591	0.4874	-0.7922	0.9369	0.2495	1.0446	0.8458
TK3	No	0.8842	0.6040	-0.8202	0.8477	0.6037	-0.8660	0.8430	0.0004	1.0557	1.0582
TK4	Yes	0.6977	0.5360	-0.7808	0.5812	0.6305	-0.8121	0.4926	-0.1210	1.0401	0.8350
TK5	No	0.8118	0.7161	-0.7742	0.9471	0.5511	-0.8036	0.7742	0.2131	1.0379	0.9873
TK6	No	0.9219	0.4856	-0.6523	0.8670	0.5764	-0.7918	0.8612	-0.1393	1.2139	0.6656
TU1	No	0.8081	0.5184	-0.5447	0.8553	0.3972	-0.6223	0.8885	0.2225	1.1425	0.6447
TU2	Yes	0.9515	0.6755	-0.5668	0.9401	0.4570	-0.6200	0.9547	0.3854	1.0937	0.8557
WA1	No	0.9832	0.4200	-0.5656	0.9696	0.3030	-0.6236	0.9671	0.2070	1.1026	0.5847
WA2	No	0.7674	0.6286	-0.5376	0.1670	0.5700	-0.7809	0.2252	0.1090	1.4528	0.9471
WA3	No	0.8810	0.5258	-0.6272	0.8020	0.4122	-0.7981	0.9052	0.1812	1.2726	0.6203
WA4	No	0.9812	0.4399	-0.6251	0.9669	0.3178	-0.6803	0.9823	0.1954	1.0883	0.4328
WA5	No	0.9812	0.6509	-0.7674	0.9013	0.5773	-0.8164	0.8977	0.0959	1.0638	0.6898
WA6	No	0.9679	0.5306	-0.6444	0.9856	0.3814	-0.6611	0.9548	0.2315	1.0260	0.5590
WA7	No	0.9592	0.4784	-0.6041	0.9730	0.4278	-0.6527	0.9489	0.0837	1.0804	0.8652
WA8	No	0.9011	0.5361	-0.6443	0.9651	0.4081	-0.6645	0.9144	0.1987	1.0313	0.6130
WA9	No	0.9016	0.7253	-0.8152	0.9279	0.3518	-0.7136	0.8892	0.4581	0.8754	0.5437

NRWQN site code	Flow impacted	VC-TSS R ²	VC-TSS SMA Intercept	VC-TSS Slope	VC-Turbidity R ²	VC-Turbidity SMA Intercept	VC-Turbidity SMA Slope	Turbidity-TSS R ²	Turbidity-TSS SMA Intercept	Turbidity-TSS SMA Slope	Median c*
WH1	No	0.7601	0.5266	-0.7039	0.9278	0.6307	-0.7486	0.7508	-0.1479	1.0634	1.3726
WH2	No	0.9482	0.5536	-0.7900	0.9336	0.6882	-0.9629	0.9690	-0.1704	1.2189	0.9490
WH3	No	0.9780	0.8733	-1.0183	0.9835	0.7291	-0.9824	0.9832	0.1416	0.9648	0.6533
WH4	No	0.8814	0.4679	-0.7193	0.8438	0.8199	-1.0089	0.8270	-0.4893	1.4027	0.9282
WN1	No	0.8873	0.5266	-0.7332	0.8948	0.4555	-0.7875	0.9598	0.0969	1.0740	1.2363
WN2	No	0.6362	0.6616	-0.5046	0.8625	0.4166	-1.0083	0.6385	0.4855	1.9981	1.4090
WN3	No	0.9561	0.6231	-0.8027	0.9584	0.5451	-0.8230	0.9846	0.0971	1.0253	0.8240
WN4	No	0.9097	0.5639	-0.7528	0.8419	0.4131	-0.7567	0.8890	0.2003	1.0052	0.7386
WN5	No	0.9473	0.6281	-0.7747	0.9768	0.5774	-0.8530	0.9551	0.0655	1.1011	1.2451

Appendix G Suspended sediment particle size variation during floods: Mararoa River example

Forty-three suspended sediment samples were collected by auto-sampler over two consecutive floods in the Mararoa River at Cliffs during May 1995. These were analysed both for SSC and PSD. The results (Figure G-1) show a progressive increase in the clay (< 4 µm) and fine silt (4-16 µm) fractions at the expense of the coarse silt fractions (16-63 µm) during both events. This most likely relates to differences in the particle size characteristics of different sediment sources in the catchment. It would have resulted in temporal variation in the relationships between SSC and water clarity and turbidity.

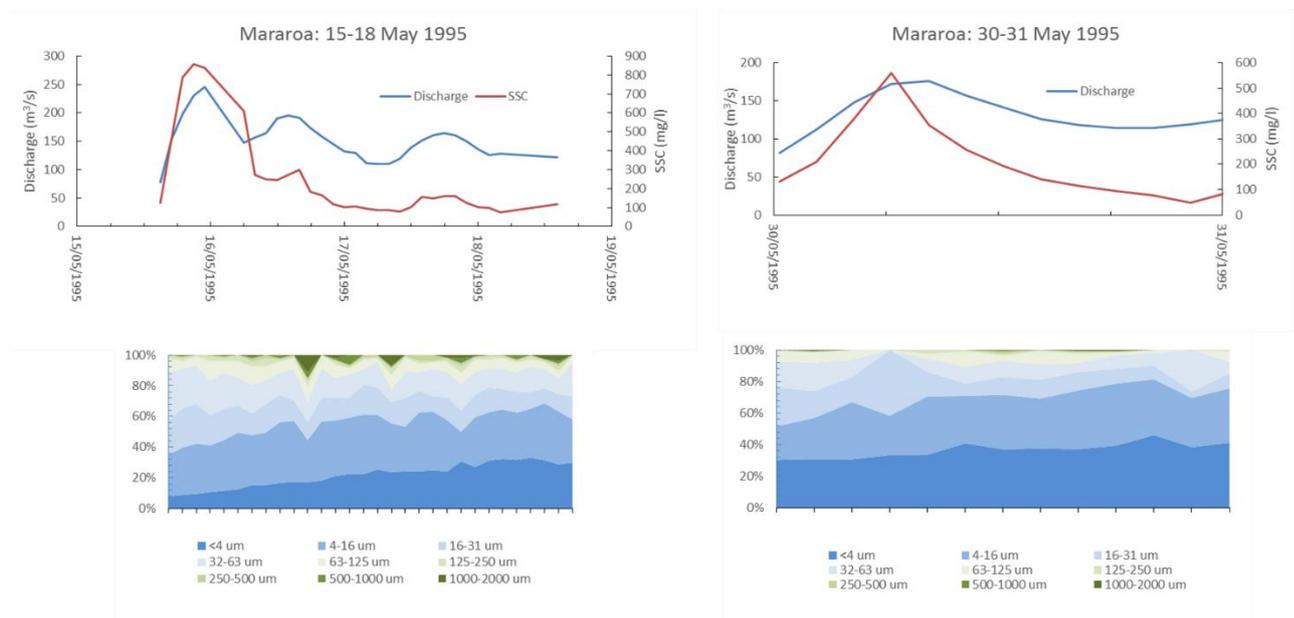


Figure G-1: Variation of SS particle size distribution through two floods in Mararoa River at Cliffs, May 1995. Upper plots show discharge and SSC records; lower plots show distribution of SS load by size fraction. Blue shadings capture the mud range (finer than 63 µm); yellow-green shadings capture the sand range. Data source: NIWA.

Appendix H Sediment Assessment Methods

Table H-1: Sediment Assessment Methods (SAMs) used for assessing the fine sediment content of stream channel beds. Summarised from Clapcott et al. (2011).

Name	Description	Variable measured	Applicability
Sediment Assessment Method 1	A rapid visual estimation from the stream bank of the proportion of the channel bed covered by fine sediment (<2 mm)	% fine sediment cover	All streams
Sediment Assessment Method 2	Semi-quantitative, in-stream visual assessment of the surface area of the streambed covered by fine sediment (< 2 mm), made by observing at least 20 locations within a single habitat	% fine sediment cover	Hard-bottomed streams
Sediment Assessment Method 3 (Wolman pebble count)	Semi-quantitative assessment of the particle size distribution, including fine sediment, on the streambed surface using a graduated template of "gravelometer". At least 100 particle measurements are made within a single habitat.	% by count of clast b-axis dimension into size fractions typically varying by a factor of 2; sediment too fine to measure (typically < 2 mm) labelled as "fines".	Hard-bottomed streams
Sediment Assessment Method 4 – (Quorer method)	Quantitative measure of total re-suspendable solids deposited on and within the streambed. A cylindrical tube is screwed into bed, the bed inside is stirred to suspend fine sediment, and the slurry is sampled and measured for suspended sediment concentration. Six samples are collected from a single habitat.	Samples are processed in the laboratory for total Inorganic/Organic sediment by area (SIS and SOS, respectively, in g/m ²) or Suspendable Benthic Solids by Volume (SBSV, g/m ³).	Hard-bottomed streams
Sediment Assessment Method 5 – (Shuffle method)	Rapid qualitative assessment of the amount of total re-suspendable solids deposited on a streambed. Made by observing turbidity created by disturbing the streambed by moving feet vigorously for five seconds.	A score from 1-5 is assigned (1 = little/no sediment; 5 = excessive sediment).	Hard-bottomed streams

Appendix I Standard error of regression line

If a normal random variable y is related to a normal random variable x through a linear regression

$$y = \alpha x + \beta \quad (\text{H-1})$$

derived from n pairs of (x, y) , then the variance of the regression line is given by

$$\text{Var}[\alpha x + \beta] = x^2 \text{Var}[\alpha] + \text{Var}[\beta] + 2x \text{Cov}[\alpha, \beta] \quad (\text{H-2})$$

Blum and Rosenblatt (1972, p442) gave expressions for $\text{Var}[\alpha]$, $\text{Var}[\beta]$, and $\text{Cov}[\alpha, \beta]$, which when substituted into (H-2), give

$$\text{Var}[\alpha x + \beta] = \frac{\sigma_y^2}{n \sigma_x^2} (x^2 + \overline{x^2} - 2x\bar{x}) \quad (\text{H-3})$$

where σ_y^2 is the variance of y , σ_x^2 is the variance of x , \bar{x} is the mean of x , and $\overline{x^2}$ is the expected value of x^2 .

Appendix J Derivation of suspended sediment percentiles from flow percentiles

The probability that flow is less than a given value x is given by

$$P(x) = \int_0^x f(Q)dQ \quad (I-1)$$

where Q is the flow ranging from zero to x .

When suspended sediment concentration C is a monotonic function of x , expressed as $C = g(x)$, then, equation (I-1) may be rewritten as

$$P(C) = \int_0^C f(g)dg \quad (I-2)$$

This shows that for a given flow percentile x with probability P of non-exceedance, the sediment concentration $C = g(x)$ has the same probability of non-exceedance and has therefore the equivalent percentile.

This result relies on being able to equate $f(Q)dQ$ with $f(g)dg$. This requires $g(x)$ to be a monotonic function. If there are errors in $g(x)$ then they need to be small and evenly distributed if $f(Q)dQ$ is to be approximately equated with $f(g)dg$. If this is not the case then a Monte Carlo simulation of concentration values from flow values could be used to estimate concentration percentiles.