AQUALINC RESEARCH LIMITED

Contaminant Load Calculator

Envirolink Project 1476- ESRC266

Prepared for Environment Southland

Report No C14098/1

June 2014

AQUALINC

AQUALINC RESEARCH LIMITED

Disclaimer:

Quality Control			
Client:	Environment Southland		
Report reference:	Title: Contaminant Load Calculator		No: C14098/1
Prepared by:	Ton Snelder, Channa Rajanayaka and Caroline Fraser		
Reviewed by:	Ton Snelder	Approved for issue by:	Ton Snelder
Date issued:	June 2014	Project No:	C14098

Document History			
Version: 1	Status: Final	Author: CR, TS, CF	Reviewer: TS
Date: 30-6-2014	Doc ID: Envirolink 1476-ESRC266_Contaminant load calculator_rpt.docx	Typist:	Approver: TS

TABLE OF CONTENTS

Page

List of Figures:

ACKNOWLEDGEMENTS

This project was funded through Envirolink Grant 1476- ESRC266.

Aqualinc Research would like to acknowledge the invaluable assistance from the following people:

- Dr Murray Hicks, NIWA
- Dr Sandy Elliot, NIWA
- Dr Jim Cooke, Streamlined Environmental Ltd
- Roger Hodson, Environment Southland
- Chris Jenkins, Environment Southland
- Andy Hicks, Hawkes' Bay Regional Council
- Bill Vant, Waikato Regional Council
- Dr Jon Roygard, Horizons Regional Council.

1 INTRODUCTION

Management of diffuse sources of contaminants (e.g., nutrients and sediment) discharging to freshwater environments in New Zealand has become an important national issue due to ongoing intensification of agriculture, in particular dairy farming, over the last 20 years (Bidwell *et al*., 2009; Davies-Colley *et al*, 2003). Government legislation in New Zealand, the National Policy Statement for Freshwater Management (NPS-FW, MfE (2011)), requires regional councils to set limits to ensure that freshwater objectives can be met. In response to the NPS-FM, regional councils are increasingly setting contaminant discharge limits, which are often expressed as annual loads. Contaminants may then be managed through allocation of a part of the annual catchment load to all land and water resource users in the upstream catchment. A fundamental component of policy development and monitoring progress towards catchment objectives is, therefore, the estimation of catchment nutrient loads.

Catchment contaminant loads are estimated based on calculations that combine discrete contaminant concentration observations with more frequent observations of flow. There are numerous calculation methods for converting water quality and flow data into contaminant load estimates and manually handling data and performing these calculations is time consuming. Regional councils need to utilise the most appropriate methods and to use tools to efficiently produce robust contaminant load estimates. Councils consider that load estimates would be more defensible if there was appropriate software that automated and standardised data processing and load calculations that was used nationally to meet NPS-FW requirements. The software tool would needs to be easy to use and scientifically robust. The software would need to accommodate datasets with spot sampling and instantaneous load estimates, as well as sites with continuous flow data combined with either spot sampling or continuous parameter data. An estimate of uncertainty in load calculation would be part of the output, as well as clear guidance on what sampling regimes should be adopted by councils to produce adequate datasets for load calculations with different levels of certainty.

The purpose of this project is to provide advice on the development of a nationally consistent approach to contaminant load calculation and software to calculate loads. This project has been initiated by Environment Southland Regional Council, is funded through Envirolink Grant 1476- ESRC266. The project objectives were:

- To identify the current methods and tools are being used for calculating contaminant loads.
- To assess the accuracy and uncertainty of contaminant load estimates produced by current methods and tools.
- To advice on the scope of study needed to enable a strategic plan for developing a contaminant load calculator that is nationally consistent.

The approach includes:

- Literature review of contaminant load calculation methods;
- Interviewing experts in the field and regional council scientists who estimate contaminate loads to identify what methods and tools are currently used and uncertainty associated with each method.

- Assessment of a most commonly used formulaic load estimation method to identify complexities associated with using a formulaic approach.

The structure of this report is:

- Chapter 2 overviews load calculation concepts and methods;
- Chapter 3 describes the methods and tools currently used nationally;
- Chapter 4 presents our analysis of current practices and tools;
- Chapter 5 includes the study recommendations.

2 LOAD CALCULATION

2.1 Concepts

For the benefit of readers who are not familiar with load calculation, this section provides a basic introduction. The concepts given hear are mostly taken from Richards (1998).

The flux (loading rate) is the instantaneous rate at which a contaminant passes a point of reference on a river (grams/second). The flow (discharge rate) is the instantaneous rate at which water is passing the reference point (cubic metres per second).

Both concentration and flow are continuous functions of time, however, they cannot be measured continuously. Therefore, the measurements we make are a sequence of discrete measurements of concentration and flow. The load can be represented as:

$$L = K \sum_{i=1}^{n} C_i Q_i \Delta t \qquad (1)$$

Where:

L	Load
K	Units conversion factor accounting for the period of record and units
C_i	i^{th} observation of concentration
Q_i	Corresponding observation of flow (i.e., i^{th} flow observation)
Δt	Interval between observations
n	Number of observations.

It is possible that observations are not equally spaced, then Equation (1) can be reformulated as:

$$L = K \sum_{i=1}^{n} C_i Q_i t_i \qquad (2)$$

Where:

t_i	Time interval represented by the i^{th} sample.

While measuring flows is not expensive, concentration measurements can range from few dollars to thousands per sample depending on the contaminant. Thus, in reality sampling is often limited to monthly or quarterly measurements. Therefore, the appropriate and pragmatic practice is to estimate "missing" concentrations to go with the flows observed at times when chemical samples were not taken. Many approaches have been developed to calculate the loads in the absence of continuous concentration measurements.

In selecting the appropriate method for a given circumstances (based on flow characteristics and contaminant), the following attributes need to be considered:

- Precision – if a method produces repeated load estimates with a small difference, the process is precise.
- Accuracy – if estimates are close to the true value, it is accurate.
- Bias - is the lack of accuracy; a measurement system which is unbiased is accurate.
- Efficiency - load estimation methods that have low bias and high precision using relatively few samples are efficient.
- Robustness - the method should be insensitive to the attributes of the data for example unusually high concentration data points.
- Objectivity – the methods should involve minimal subjectivity (e.g., professional judgment)

2.2 Load Calculation Methods

There are many methods that have been developed to calculate loads from observed flow and concentration measurements. A brief overview of eight load calculation methods is presented below.

Stratification of the data with respect to various factors such as flow magnitude or season is also used in conjunction with many of these methods (Qulibe *et al.,* 2006; Preston *et al*., 1989; Dolan *et al*., 1981; Roygard *et al*., 2012).

2.2.1 Numeric Integration

This is a simple method that uses direct numeric integration as given in Equation (2). Therefore, the load is calculated as:

$$L = K \sum_{i=1}^{n} C_i Q_i t_i \qquad (3)$$

2.2.2 Rating Curve (Regression)

Rating methods define an empirical relationship between concentration and flow. Rating curves are commonly developed using regression analysis, based on the concentration and flow data. The rating curve can be based on a univariate or multiple regression model, and concentration or flux may be used as the dependent variable. In most applications, both concentration and flow are log-transformed to create a dataset which is better suited for regression analysis. Alternatively, non-parametric regression such as LOWESS (Locally Weighted Scatterplot Smoothing; Cleveland, 1979) is used to derive a rating. The relationship between flow and concentration can be modified to account for nonlinearity, seasonal and long-term variability, censored data, biases associated with using logarithmic transformations and serial correlations in the residuals of the analysis (Cohn, 1995).

The rating curve is used to generate an estimate of concentration for every observation of flow. The load is then calculated as the product of the flow and concentration time series over the sampling period (Qulibe et al 2006).

2.2.3 Ratio Method

The ratio method is based on the assumption that the ratio of mean load to mean flow on the days that concentration was measured is representative of the average annual ratio of load to flow (Beale, 1962). Thus, the mean daily flow is multiplied by the ratio of the mean of the instantaneous monthly loads to the mean of the instantaneous monthly flows.

$$L = \mathrm{K}\frac{\left(\sum_{i=1}^{n}\frac{C_i\ Q_i}{n}\right)}{\left(\sum_{i=1}^{n}\frac{Q_i}{n}\right)}\left(\sum_{j=1}^{N}\frac{Q_j}{N}\right) \tag{4}$$

Where:

- C_i Concentration measured at some discrete interval (e.g. monthly)
- Q_j Measured flow measured at some discrete interval (e.g. 15 minutes or mean daily)
- Q_i Flow measured at the same time as the concentration sample
- n Number of concentration samples
- N Number of flow samples

When the two parameters involved are correlated, as is almost always the case with flow and load, ratio estimators are biased, and a bias correction factor is used (see Dolan *et al*., (1981) for details).

2.2.4 Averaging Methods

There are several methods that are based on multiplying the mean of the observed concentration by the mean of the observed flows (Dolan *et al.*, 1981, Defew *et al*., 2013). For example:

$$L = \mathrm{K}\left(\sum_{i=1}^{n}\frac{C_i}{n}\right)\left(\sum_{j=1}^{N}\frac{Q_j}{N}\right) \tag{5}$$

2.2.5 Integration Method

The integration method uses fixed interval, simultaneously measured instantaneous flow and contaminant concentrations data. The sampled concentrations are plotted through the time and professional hydrological judgement is used to interpolate and extrapolate between the measured concentrations (Robertson and Roerish, 1999). This method can be accurate if sufficient data is collected. However, to obtain accurate load estimates, many samples need to be collected to reflect the variability in concentration, particularly during high-flow events.

2.2.6 Flow Interval

The flow interval method developed by Yaksich and Verhoff (1983) is a semi-graphical technique. The approach begins by plotting the measured fluxes against the instantaneous flows at the time the samples were taken. The flow axis of the plot is divided into several uniform intervals covering the range of mean daily flows. The average flux for each interval is then calculated. The interval load is the product of the average flux and the number of days in the interval (and the appropriate units conversion factor).

2.2.7 Flow-Proportional Sampling

This is a largely a mechanical approach, which is completely different approach to the other load estimation methods. The method assumes that one or more samples can be obtained, which cover the entire period of interest. Each period assumed to be representing a known flow and each with a concentration which is in proportion to the load which passed the sampling point during the sample's accumulation. The load for each sample is then calculated as the flow times the concentration, if this assumption is true. A major drawbacks of this approach is precision estimates of the calculated loads cannot be obtained.

2.2.8 Worked Records

The worked records method is an expert approach and is subjective. The approach involves superimposing the sampled concentrations on a hydrograph of more regularly measured flow (e.g. daily flow). A smooth curve is drawn through these points based on the expert's experience with the relationship between concentration and flow. This interpolated curve is used to estimate a representative concentration for intervals of interest. The advantages of this method is it allows for the possible inclusion of a peak concentration greater than the largest observed concentration, and use of professional judgement.

2.3 Choice of Method

For the majority of catchments in New Zealand the best data sources from which to estimate contaminant loads are combinations of continuous (i.e. 15 minute intervals) flow records and water quality data. Water quality monitoring on a monthly basis has been common practice in New Zealand for at least two decades for the purpose of monitoring water quality state and trends. Loads calculated from monthly monitoring data are understood to be associated with large uncertainties because sampling at this low frequency fails to adequately represent the population of daily loads (e.g. Defew *et al*., 2013; Robertson and Roerish, 1999; Johnes, 2007; Philips *et al.*,1999). However, for the foreseeable future monthly samples will be the most comprehensive data from which contaminant loads can be estimated.

The most appropriate method is one that is precise and accurate (therefore efficient), and which is robust and objective. However, in reality these desirable attributes are difficult to achieve within the constraints of time and budget. In addition, it is very expensive to measure the “true” load because concentration is a continuous function of flow and therefore all methods that estimate "missing" concentrations are approximations. This means that a relatively small number of studies worldwide have attempted to evaluate the accuracy of load estimates in particular. As a result, when uncertainty is reported for a load estimate that is made from discrete concentration data, it is generally the precision of the estimate that is being reported and not its accuracy.

A number of studies have estimated the uncertainty of different load calculation methods by comparing estimated loads with the 'true load', which is generally calculated from a high temporal resolution time series (e.g. Defew *et al*., 2013; Robertson and Roerish 1999; Qulibe *et al*., 2006; Johnes, 2007; Philips *et al*., 1999; Preston *et al*., 1989). For example, Defew *et al*. (2013) used phosphorus concentration data measured at 2-hourly interval for a three month period, which they used to estimate the true load. In general, uncertainty is characterized by the average error, which comprises two components. The first component is the accuracy or bias which is the average difference between the estimated and "true" load. The second component is the precision, which is a measure of the spread or variance of the estimates.

Many studies have found that loads estimated from "infrequent" sampling are highly uncertain. For example, Defew *et al.* (2013) found phosphorus loads estimated from two weekly sampling could have errors of more than 300% compared to the true loads. In another study Johnes (2007) found loads of total phosphorus estimated from monthly data could have errors in the order of 200 to 500%. Unsurprisingly, these studies found that uncertainties decrease with increased sampling frequency and longer sampling period duration.

The uncertainty of the methods appears to depend on many factors including the method itself (Defew *et al*., 2013; Johnes, 2007; Robertson and Roerish, 1999), the frequency of sampling (Defew *et al*., 2013; Robertson and Roerish, 1999), length of the estimation period (Littlewood *et al*., 1998), the size of the watershed (Philips *et al*., 1999), the characteristics of the contaminants and the resulting the strength and form of the flow-concentration relationship (Preston *et al*., 1989; Richards and Holloway, 1987), the catchment sources of the contaminants (Johnes 2007), and the characteristics of the annual hydrograph (Johnes, 2007; Preston *et al*., 1989). The choice of method is therefore widely regarded as being context dependent (Kronvang and Bruhn, 1996, Quilbe *et al*., 2006).

It is sometimes found that methods produce load estimates with relatively high precision, but which are inaccurate. Precision, rather than accuracy may be more important in some applications such as monitoring for trends in loads or determining whether a catchment load is within a specified limit. In these applications, it is the repeatability of the estimate (i.e. its precision) that is important, rather than its absolute value. In other situations however, for example assessing the ecological effects of nutrients discharged to lakes or estuaries, an accurate estimate of the nutrient load will be desired.

In conclusion, the studies of load estimation methods indicate that loads estimated from monthly data are generally highly uncertain. Studies have found that choice of load calculation method is important and that no method is consistently superior but that a specific method will often provide the least uncertain estimate for a given site and contaminant. Thus, it is difficult, if not impossible, to recommend a particular load estimation method. Finally, load calculation methods cannot reduce imprecision or inaccuracy associated with input data. Therefore it is important to ask whether current monitoring frequency and methods will ultimately provide data that will enable loads to be calculated that are fit for purpose (i.e. of sufficient precision and accuracy).

3 METHODS AND TOOLS CURRENTLY USED NATIONALLY

There has been considerable work carried out in New Zealand that has addressed the subject of calculating contaminant and sediment loads. Dr Murray Hicks of NIWA has developed a software package, SedRate to calculate sediment loads (SedRate, 2014). SedRate capabilities have been now extended to calculating contaminant loads. SedRate uses rating curve relationships and estimates the precision of the load estimates. Horizons Regional Council developed a flow stratification method to calculate loads (Roygard *et al*., 2012). Two studies have recently reviewed the Environment Southland's suspended sediment and nutrient monitoring programmes (NIWA, 2014; Diffuse Sources and NIWA, 2012). The study by Diffuse Sources and NIWA included consideration of seasonality and methods for calculating nutrient and sediment loads. Other councils such as Waikato Regional Council, Hawke's Bay Regional Council and Canterbury Regional Council also estimate loads and are in the process of setting load limits within their regions.

We developed an overview of the methods and tools that are currently being used in New Zealand by interviewing a number of experts in the field from both research organisations and regional councils. Some of the interviews were conducted by face-to-face meetings and others were phone meetings. The following sections describe the information gathered during the interviews including the methods and tools that each of the experts use.

3.1 Experts from Research Organisations

3.1.1 Dr Murray Hicks, NIWA

Dr Murray Hicks conducts research on sediment transport and related geomorphic processes in rivers. Murray has developed a software tool, "SedRate" for estimating river contaminant loads. SedRate was originally developed for suspended sediment, however, it has been expanded to calculate loads for any contaminant (e.g. Nitrogen (N), Phosphorus (P)) (SedRate, 2014). This tool is fairly easy to use and is already used by some regional councils to estimate contaminant loads and other statistics (Bill Vant, WRC pers. comm). SedRate also provides estimates of the precision of the load estimate.

SedRate uses a rating curve approach, based on regression methods. Ratings of concentration (of contaminant) or sediment against flow can be defined using four automated regression methods: ordinary least squares (OLS), minimum variance, load weighted and LOWESS. SedRate also allows the user to subjectively fit a rating curve using mouse-clicks on the rating plot. There are two different bias correction methods available, if the user needs to deal with log-transformation bias of the rating curve.

Once the appropriate rating is established through automated methods (or user-fit) and using bias correction, it can then be applied to the site flow distribution to obtain the annual average load. SedRate uses a utility programme, GenDist to divide the flow distribution into bands (i.e., bins). A maximum number of flow bands is 200, i.e., 0.5 percent increments of the flow frequency distribution.

3.1.2 Dr Sandy Elliot, NIWA

Dr Sandy Elliot is a catchment modeller and has been undertaking research on catchment contaminant loads for many years. Sandy primarily works with data from the National Water Quality Network which have monthly concentration data and have hourly flow data. Sandy originally calculated loads using the spreadsheet software Excel. Recently Sandy has developed a procedure for calculating loads using 'R' statistical software.

Sandy generally uses a rating curve method for calculating site contaminant loads. Sandy examines input data prior to establishing the rating curve to eliminate any potential outliers. The regression is defined on log-log transformed concentration-flow data and parametric curves of linear or quadratic form are fitted to define the rating. Sandy also applies spline bias corrections to reduce bias in data.

Estimating the precision of the calculated loads was not possible under previous Excel based approach. However, the current 'R' statistical tool enables uncertainty analysis using non-parametric bootstrap methods Sandy considers that the uncertainty of load estimates for some sites can be extremely high and that accuracy with any rating curve approach can potentially lead to significant errors if samples are not collected at high flows. Sandy has found that load estimates based on sampling durations of ten years or more can substantially reduce uncertainty. He commented that if records are of reasonable length, expensive sampling techniques such as auto samplers may not be required.

Sandy considers that rivers and streams in New Zealand can be more hydrologically variable than big rivers in other parts of the world. Thus, some of the methods developed in other countries cannot be applied with the same success in New Zealand.

Sandy cautions that concentration data will generally be associated with trends. Therefore, it is important to carefully examine data to identify potential trends and de-trend data before establishing a rating curve.

Sandy has collaborated with Richard Alexander of U.S. Geological Survey (USGS) and also used LowDesk software developed by USGS to estimate loads. LowDesk enables use of many load calculation methods including rating curve method.

3.1.3 Dr Jim Cooke, Streamlined Environmental Ltd

Dr Jim Cooke has used many load calculation methods described above as part of his current work at Streamlined Environmental Ltd and previously at NIWA. He believes that accuracy of load calculations in New Zealand is primarily affected by inadequate sampling and that limitations on sampling are associated with financial constraints.

Jim considers the rating curve method to be the most widely used. He points out the method is used to extrapolate to flows for which samples were not collected, and that this can produce load estimates with large errors.

Jim recommends using the integration method (Section 2.2.5) where high accuracy is needed, particularly in small streams. Whilst this method requires more concentration data and is, therefore, more expensive than the rating curve approach, it is particularly appropriate where accurate estimates are required for short-term data.

3.2 Regional Councils

3.2.1 Roger Hodson and Chris Jenkins, Environment Southland

Both Roger Hodson (Surface Water Scientist) and Chris Jenkins (Senior Hydrologist) are involved in contaminant load estimation at Environment Southland (ES). They generally use continuous flow data (10 minute intervals), which they convert to mean daily flows, and monthly contaminant concentration or sediment samples. ES uses the rating curve method and finds that power law relationships (linear regression based on log transformed flow and concentration data) produce the best fit for their rivers. They tend to re-establish the rating curves when more data is available and often find that the previously derived rating curves were unsatisfactory (i.e. previous rating curves do not agree with the new rating curves that are based on more data). They have identified that reliable rating curves cannot be defined for some rivers due to the currently limited data. They have not estimated the uncertainty of the load estimates to date, however, are planning to assess the uncertainty when sufficient long-term data is available.

One of the main issues ES have identified is their data series are not long enough to establish an appropriate rating curve. The other issue is that the current monthly sampling regime fails to capture concentrations over the whole flow range, primarily due insufficient sampling at high flows. ES is in the process of installing auto-samplers at downstream location in major rivers. These samplers can be shifted to new locations when needed. The main aim of the auto-sampler installation is to sample concentrations over a range of flows and conditions (e.g. rising flow or changing turbidity) in addition to set time intervals. These samplers can also be triggered to sample from the office if needed. ES considers that this targeted sampling regime will enable them to collect a dataset that is more representative, primarily of high flows, and therefore to establish better rating curves for load estimation.

ES has also identified that existing concentration-flow rating curves may not remain representative over long time period due to change in land use that in turn changes concentrations. ES, therefore consider that it is important to continue with data collection and to potentially update rating curves on five-yearly basis.

ES has recently commissioned NIWA to review their SoE suspended sediment and water clarity monitoring programme (NIWA, 2014). It should be noted that the only the draft report was available at time of this study. The key findings of NIWA review relevant to this Envirolink project are:

- Most SoE sites have insufficient data for reliable suspended sediment load calculations.

- Some concentration sampling locations do not have nearby flow monitoring stations making the estimation of loads difficult and increasing the likely uncertainty.
- The recommended approach for determining the annual loads is to compile an instantaneous, all-of-record suspended sediment concentration-flow rating, and combine this with either the flow record or flow-duration table (for the year of interest) to estimate the annual load.
- Assessment of the uncertainty on rating-based estimates of annual loads should be performed with a specialised analysis package such as SedRate.
- Setting limits on sediment loads for management purposes should use an approach that allows for year-by-year hydrological variability. Options include managing a multi-year running average load or managing annual deviations from a sediment rating function.

Previously, ES had also commissioned a study to estimate the Waituna catchment loads (Diffuse Sources and NIWA, 2012). The purpose of the study was to develop a robust methodology for calculating nutrient loads for the streams discharging to the Waituna Lagoon. The study recommended using rating curve methods (regression approaches) along with SedRate software. The study recommended a bimonthly sampling strategy with additional storm sampling to improve information of concentration at high flows.

3.2.2 Andy Hicks, Hawkes' Bay Regional Council

While Hawkes' Bay Regional Council (HBRC) has been estimating contaminant loads for many years, they are in the process of establishing a more robust programme. At early stages HBRC used numeric integration method to calculate loads. However, they now use rating curve method, however, the exact details are yet to be finalised.

HBRC concentration data was mostly based on quarterly sampling, however, sampling frequency has recently been increased to generally monthly intervals. The current sampling is not targeted based on flow (e.g. high flows) or any other condition. However, they have recognised the need for additional storm sampling to improve information of concentration at high flows and are considering using auto-sampling techniques.

3.2.3 Bill Vant, Waikato Regional Council

Bill Vant has been calculating contaminant loads for Waikato Regional Council (WRC) for over 16 years. He currently uses the rating curve method and the SedRate tool.

WRC has more than 10 years of monthly data for most of the SoE sites. The current sampling approach is based on regular monthly interval sampling rather than targeted sampling (i.e, based on flow). While WRC would not rule out targeted sampling in the future using auto-sampling techniques, they are currently satisfied with the monthly sampling as long term data sample (i.e., more than 10 years) is considered to be sufficiently representative of the range of flows.

Consents for municipal sewage and industrial wastewaters in the Waikato region are subject to consent conditions requiring the monitoring of concentrations and discharge rates. Bill uses these data to calculate loads but finds the supplied data can be inconsistent in sampling frequency and formats. Calculation of loads from this data requires careful assessments and avoid bias and errors and the process can be time consuming.

Bill does not currently attempt to estimate the uncertainties associated with their load estimations. They recognise that there are errors associated with their load estimates due to many factors: including the flow data (can be up to ±8%), sampling errors, laboratory errors and errors associated with the definition of the rating curve. Due to these compounding errors, WRC considers their current resources (mainly time) are inadequate to thoroughly estimate the uncertainty of the load estimates with reasonable accuracy. They also consider that the load estimates produced under the present approach is sufficient to meet their objectives, and additional information on errors or uncertainty would not provide any knowledge that is useable in the current context.

3.2.4 Dr Jon Roygard, Horizons Regional Council

Dr Jon Roygard has developed a flow-stratified load calculation method (Roygard *et al*, 2012). Stratification includes defining 10 flow "bins" based on subdivision of the flow frequency distribution (the flow duration curve). The component of the load associated with each flow bin is estimated as the product of the mean concentration and the flow associated with the bin. These loads are then summed to find the total load. An advantage of the method is that it automatically produces the proportion of the load associated with different parts of the flow hydrograph (e.g., high flows). The flow-stratified averaging approach found to be effective in reducing bias associated with monthly sampling that does not generally consist of either very high or low flows. However, loads at low flows (i.e., below the 20%ile) were calculated by removing the loads assigned to highest two flow bins (see Roygard *et al*. (2012) for details).

4 WHAT IS NEEDED?

4.1 Regional Council View

We spoke to the regional council staff listed in the previous section about the need for nationally consistent methods and tools for calculating contaminant loads. We found some variation in the level of support for this among the council staff. In particular, HBRC and ES staff strongly supported the development of a nationally consistent tool and were willing to work with other regional councils to share knowledge and resources. Horizons Regional Council have been considering building a Hilltop compatible load calculator that incorporates their flow binning method. Horizons are willing to work with other regions and research on scoping different methods and their uncertainty. Bill Vant from Environment Waikato was willing to share his experience and assist other regions if a tool needs to be developed that is nationally consistent. However, Bill considers that load calculation methods they currently use are sufficient to meet their objectives.

4.2 Danger of Adopting a Formulaic Approach

As a result of the literature review and discussions with experts in the course of this project, we consider that there would be dangers in adopting a formulaic approach to load calculation. This is because, for the foreseeable future, monthly samples will be the most comprehensive data from which contaminant loads will be estimated for the majority of sites. The quality of these load estimates will be limited by the available data and application of any of the methods to these data involves subjective decisions. In addition, there are significant differences in the characteristics of sites and of the contaminants themselves that affect the quality of load estimates and the choices that need to be made when calculating loads. As a consequence estimating the "correct' load estimate requires expertise, professional knowledge and judgement.

To illustrate this point we show below plots of 20 years of concentration and flow data for four nutrient species at four sites with similar median flows that belong to the National Rivers Water Quality Network (NRWQN; Davies-Colley *et al*., 2011). The four nutrient species shown are Dissolved Reactive Phosphorus (DRP) and nitrate-nitrogen (NO_3N), Total Nitrogen (TN), Total Phosphorus (TP). The four sites shown in the plot are as follows:

- RO04l - Whirinaki River at Galatea.
- NN02 - Motueka River at Gorge.
- HV01 - Makaroro River at Burnt Bridge.
- DN01 - Taieri River at Tiroiti.

The plots show two types of regression relationships that have been fitted to the data (Figure 1); a linear regression, which is fitted to log transformed data so is therefore equivalent to a power law relationship, and a LOWESS non-parametric regression. The predicted values from these relationships have been plotted as lines to define a concentration-flow rating. The red rug plots on the y-axis indicate the distributions of the concentration data with each line indicating a decile (i.e. each interval contains 10% of the data).

The plots illustrate some of the difficulties and peculiarities associated with the data used for load calculations. First, the two types of regressions result in sometimes quite different rating curves and the patterns of these differences vary among the sites and nutrient species. Second, the distributions of the concentration data vary among the sites and nutrient species (as shown by the rug plots on the y-axis). Note that generally the dissolved forms of the nutrients (DRP and NO_3N) are more normally distributed than the total forms (TP and TN). The high concentration values for TN and TP are generally associated with the high flows and these variables have stronger concentration-flow relationships than DRP and NO_3N. In fact at site NNO2 there is no relationship between concentration and flow for DRP and NO_3N.

The data illustrate some of the difficulties and subjective judgements that an analyst needs to make when making load calculations, particularly those based on the rating method. For example, the plots indicate that selection of the 'best' rating curve is site and variable specific and is really a "judgement call". The data shown in these plots is for a 20 year period (1991 to 2010). For many sites, there will not be this much data available making the estimation of loads less accurate and increasing the difficulty and need to make subjective judgments.

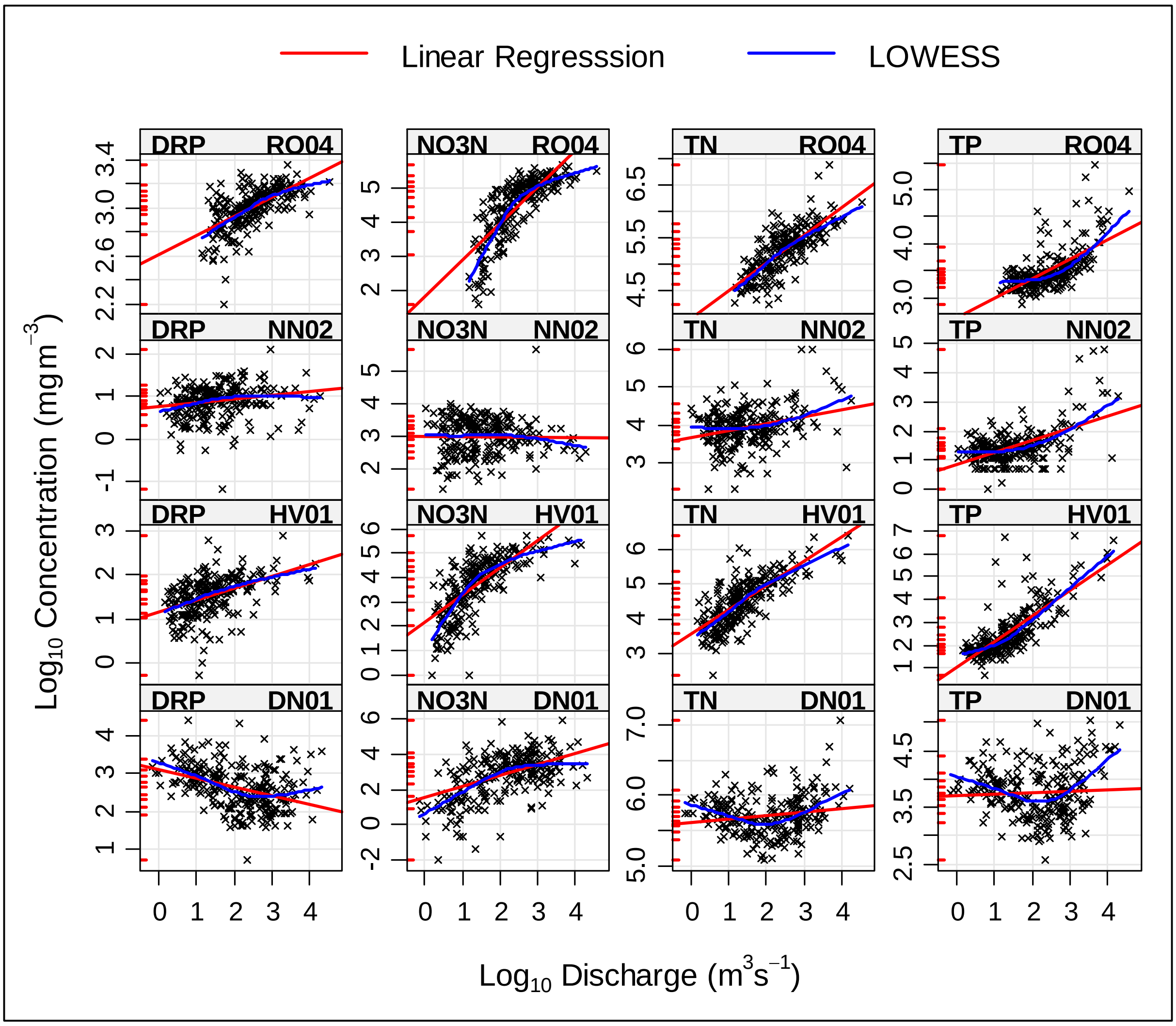

Figure 1: Examples of relationships between log₁₀ flow and concentration for four nutrient species at four sites. Two regression relationships have been fitted to the data a linear regression (power law) and a LOWESS non-parametric regression and predicted values for these have been plotted as lines. The red rug plots on the y-axis indicate the distributions of the concentration data with each line indicating a decile (i.e. each interval contains 10% of the data)

4.3 Our Conclusions

As a result of the literature review and discussions with experts in the course of this project, we hesitate to recommend a nationally consistent approach is adopted or that significant investment is made in developing new load estimation tools. Our reasons for this are:

- There are existing appropriate tools;
- The "correct' load estimate requires expertise and professional knowledge and judgement; and
- Providing simple and automated load calculation methods may lead to erroneous use of data and methods.

We consider that the existing SedRate tool is fit for purpose and supported by an organisation and individuals with considerable expertise and experience in load estimation. SedRate was developed in New Zealand and uses the load estimation method that is used by most of the regional councils and researchers we interviewed (i.e. the rating method). The two other methods that may be potentially used to calculate loads are the ratio and average methods (Section 2.2). These methods are simpler in their application than the rating methods and can be implemented with existing tools very easily.

It was beyond the scope of this project to investigate the differences between the SedRate tool and the flow stratified approach developed by Horizons (Roygard *et al*., 2012). However, the two methods are similar in that they sum, over increments of flow, the product of the flow and the conditional mean concentration. There may be modifications that can be made to SedRate that would mean it would apply the Horizons method as an option.

Another positive aspect of SedRate is that it focuses the user on the subjective aspects of load calculation. In particular, SedRate illustrates the extent to which load estimation involves extrapolation of the rating curve to flows for which concentrations have not been observed. SedRate also forces users to actively choose the type of regression model used to construct the rating. In fact, SedRate also allows the user to define the rating subjectively by inspection of a plot of concentration against flow and by using mouse clicks to define the rating.

We do however consider that there is a need for at least two areas of development around load calculations. Firstly, there is a need for more training and support for regional council staff involved in load calculation. Training should be associated with the use of existing tools that analysts work with when performing load calculations such as flow and water quality databases (Hilltop Software and Tideda), calculation tools (e.g., SedRate and Excel). Training should include aspects such as dealing with trends, censored data (i.e. detection limit issues) and outliers.

The second area of work around load calculations is more fundamental. The international literature makes it clear that loads estimated from infrequent concentration observations are likely to be highly uncertain. To our knowledge though, little work has been carried out in New Zealand on trying to understand load estimate error for nutrient contaminants in particular. Given the importance of load estimates associated with implementation of the NPS-FW, we consider that it is important that this uncertainty is understood in more detail. In particular, we consider three aspects require research:

- Quantification of the error associated with load estimates;
- Examination of the causes of this error and variability in errors among sites and contaminants; and
- Examination of the implications of the errors on subsequent management decision and actions.

5 RECOMMENDATIONS

Based on the findings of the study, we recommend that:

1. The expertise and professional knowledge and judgement required to evaluate contaminant loads is recognised and that a single load calculation method, or a simple rule based procedure for selection of methods, is not promoted.
2. There are existing appropriate tools to calculate loads and that significant investment in load calculation software is not warranted.
3. Training and support for regional council staff involved in load calculation is provided and that this training is based on the data bases and existing tools that analysts use.
4. Research on uncertainty of load calculations, its causes and implications is carried out to provide improved future guidance associated with implementation of the NPS-FW and to provide guidance for future data collection for the purpose of, particularly nutrient, load estimation.

6 REFERENCES

Bidwell, V., Lilburne, L., Thorley, M. and Scott, D. (2009). Nitrate discharge to groundwater from agricultural land use: an initial assessment for the Canterbury Plains. Environment Canterbury Technical Report, May 2009.

Davies-Colley, R J, Cameron, K, Francis, G, Bidwell, V, Ball, A, and Pang, L. (2003). Effects of rural land use on water quality, unpublished, prepared for the Ministry for the Environment. Wellington: Ministry for the Environment.

Davies-Colley, Robert J., David G. Smith, Robert C. Ward, Graham G. Bryers, Graham B. McBride, John M. Quinn, and Mike R. Scarsbrook. (2011) Twenty Years of New Zealand's National Rivers Water Quality Network: Benefits of Careful Design and Consistent Operation. 47(4): 750-771.

Defew, L. H., May, L. & Heal, K. V. (2013) Uncertainties in estimated phosphorus loads as a function of different sampling frequencies and common calculation methods. Marine and Freshwater Research 64(5), 373–386.

Diffuse Sources and NIWA (2012). Waituna catchment loads. Prepared by Diffuse Sources and NIWA for Environmental Southland, April 2012.

Dolan, D. M., Yui, A. K. & Geist, R. D. (1981) Evaluation of river load estimation methods for total phosphorus. Journal of Great Lakes Research 7(3), 207–214.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. The Annals of Statistics 7 (1): 1–26.

Johnes, P. J. (2007). Uncertainties in annual riverine phosphorus load estimation: impact of load estimation methodology, sampling frequency, baseflow index and catchment population density. Journal of Hydrology, 332(1), 241-258

Kronvang, B., & Bruhn, A. J. (1996). Choice of sampling strategy and estimation method for calculating nitrogen and phosphorus transport in small lowland streams. Hydrological Processes, 10(11), 1483-1501.

Littlewood, I.G., Watts, C.D., Custance, J.M., 1998. Systematic application of United Kingdom river flow and quality databases for estimating annual river mass loads (1975–1994). Science of the Total Environment 210–211, 21–40.

MfE (2011). *National Policy Statement for Freshwater Management 2011: Implementation Guide.* Wellington: Ministry for the Environment. ME 1077, November 2011.

NIWA (2014). Review of Environmental Southland's SoE suspended sediment monitoring programme (draft report). Prepared by NIWA for Environmental Southland, May 2014.

Ongley, E. D. (1973). Sediment discharge from Canadian basins into Lake Ontario. Canadian Journal of Earth Sciences 10, 146–156. doi:10.1139/E73-017.

Phillips, J.M., Webb, B.W., Walling, D.E., Leeks, G.J.L., 1999. Estimating the suspended sediment loads of rivers in the LOIS study area using infrequent samples. Hydrological Processes 13, 1035–1050.

Preston, S. D., Bierman, V. J., & Silliman, S. E. (1989). An evaluation of methods for the estimation of tributary mass loads. Water Resources Research, 25(6), 1379-1389.

Quilbé R, Rousseau AN, Duchemin M, Poulin A, Gangbazo G, Villeneuve J-P 2006. Selecting a calculation method to estimate sediment and nutrient loads in streams: Application to the Beau Rivage River (Québec, Canadian). Journal of Hydrology 326: 295-310.

Richards, R.P. (1998). Estimation of pollutant loads in rivers and streams: A guidance document for NPS Programs. Project report prepared under Grant X998397-01-0 U.S. Environmental Protection Agency, Region VIII, 180 pg.

Richards, R., Holloway, J., (1987). Monte Carlo studies of sampling strategies for estimating tributary loads. Water Resources Research 23 (10), 1939–1948.

Robertson, D. M., & Roerish, E. D. (1999). Influence of various water quality sampling strategies on load estimates for small streams. Water Resources Research, 35(12), 3747-3759.

Roygard, J. K. F., McArthur, K. J., & Clark, M. E. (2012). Diffuse contributions dominate over point sources of soluble nutrients in two sub-catchments of the Manawatu River, New Zealand. New Zealand Journal of Marine and Freshwater Research, 46(2), 219-241.

SedRate (2014). Overview of SEDRATE. Prepared by Dr Murray Hicks, NIWA.