Water quality trend analysis for the Land and Water New Zealand website (LAWNZ)

Advice on Trend Analysis

Date June, 2012

Dr Jon Roygard
Science Manager
Horizons Regional Council
Palmerston North

June, 2012
Report No. 2012/EXT/1245
ISBN 978-0-927189-51-4

Prepared by:
Deborah Ballantine
NIWA
PO BOX 11115
Hillcrest
Hamilton

Phone no; 078591890
Email: Deborah.Ballantine@niwa.co.nz

File: HAM2012-080

Date: June, 2012
Ref: 1049-HZLC85

CONTACT	**24hr Freephone 0508 800 800**		**help@horizons.govt.nz**		**www.horizons.govt.nz**
SERVICE CENTRES	**Kairanga** Cnr Rongotea & Kairanga-Bunnythorpe Rds Palmerston North **Marton** Cnr Hammond & Hair Sts **Taumarunui** 34 Maata Street **Woodville** Cnr Vogel (SH2) & Tay Sts	**REGIONAL HOUSES**	**Palmerston North** 11-15 Victoria Avenue **Wanganui** 181 Guyton Street	**DEPOTS**	**Levin** 11 Bruce Road **Taihape** Torere Road Ohotu
POSTAL ADDRESS	Horizons Regional Council, Private Bag 11025, Manawatu Mail Centre, Palmerston North 4442 | F 06 9522 929				

Water quality trend analysis for the Land and Water New Zealand website (LAWNZ)

Advice on trend analysis

Prepared for Horizons Regional Council

March 2012

Version

Authors/Contributors:
Deborah Ballantine

For any information regarding this report please contact:

Deborah Ballantine
Environmental Scientist
Aquatic Pollution
+64-7-859 1890
d.ballantine@niwa.co.nz

National Institute of Water & Atmospheric Research Ltd
Gate 10, Silverdale Road
Hillcrest, Hamilton 3216
PO Box 11115, Hillcrest
Hamilton 3251
New Zealand

Phone +64-7-856 7026
Fax +64-7-856 0151

NIWA Client Report No:	HAM2012-080
Report date:	March 2012
NIWA Project:	ELF12218

27 June 2012 4.55 p.m.

Contents

Tables

Figures

Reviewed by

G McBride

Approved for release by

R. Davies-Colley

Formatting checked by

AB

Executive summary

The Land and Water New Zealand (LAWNZ) website has been developed by regional councils to provide a portal for communicating environmental (and other) information from regional monitoring in a consistent manner. The initial module, which focuses on water quality, was launched at Parliament in May 2012. The website presents information on state and trends in water quality, derived from State of Environment data collected by all the regional councils and unitary authorities in New Zealand. LAWNZ is potentially the most comprehensive water quality database in the country. It is also the first time that all State of Environment water quality data has been made available in this way.

Currently there are inconsistencies between the Regional Councils' State of Environment monitoring networks which are reflected in the datasets (e.g., differences in variables measured, sampling protocols (Davies-Colley et al. 2011). These inconsistencies have proved a challenge for LAWNZ and may have impacted on the robustness of data analyses presented on the website.

The aims of this project were as follows:

- Review the current **statistical methods** underpinning LAWNZ, and establish protocols for future data analysis, including coping with censored data.
- Investigate the effect of **flow adjustment** on trend analysis results, and advise on flow adjustment in LAWNZ trend analyses.
- To provide advice on how best to amalgamate trends to facilitate clear and robust presentation.

Recommendations have been made to ensure robust present and future analysis, as follows.

- Regional Council State of Environment water quality monitoring should be harmonised as far as possible as regards sampling intervals, laboratory methods, variables measured, field protocols and quality assurance[1].
- Trends should only be calculated over the maximum time period available.
- Data should *not* be flow-adjusted before trend analysis – mainly for the pragmatic reason that flow data are currently unavailable from many SoE sites.
- Trend analysis results should be compared only when the same analysis time periods (season in TimeTrends) are used.
- Results reported as less than detection limits should be halved for LAWNZ data analysis, including trends. Where there are multiple detection limits for the same variable, and where censored data constitutes more than 40% of the data set, the higher detection limit should be applied[2].
- Plots of raw data (with trend lines superimposed) should be provided by the LAWNZ website.

[1] Note that achieving consistent water monitoring across regional councils is the main goal of the Variables workstream in MfE's NEMaR project.
[2] TimeTrends should be extended to include some options for analysis of censored data.

1 Introduction

The Land and Water New Zealand website (LAWNZ) has been developed by New Zealand's regional councils to provide a portal for communicating environmental information from regional councils with national coverage. It is envisaged that the website will eventually provide comprehensive information on water and land management in New Zealand including water quality, infrastructure, land, bio-diversity, emergency management, air and transport. The initial module focuses on water quality, and presents information from all 16 regional councils and unitary authorities.

The LAWNZ website is expected to be publicly released in early 2012. The website presents information on water quality state for approximately 900 State of Environment monitoring sites across the country and on water quality trends for over 600 of these sites. The difference between the number of sites for state and trends information in part reflects the growing nature of the size of the water quality network (a number of recently installed sites do not meet the criteria for the number of samples required for trend analysis due to the short length of record available). This is the first time that information on water quality from all regions of New Zealand has been presented on a single platform. The aggregated regional water quality dataset is the most comprehensive[3] (in terms of number of sites) currently available in New Zealand with considerable potential to inform environmental decision-making. As such, this website is a powerful communication tool for water quality information, and it is crucial that underlying data analysis methods are robust.

The presentation format and analysis methods used to date have been developed by a cross regional council group of water quality scientists commissioned by the Surface Water Information Management (SWIM) group. This project seeks to continue that collaboration and a small project 'cross-regional council' team has been formed to participate in the project involving Jon Roygard and Maree Clark from Horizons Regional Council, Bill Vant from Waikato Regional Council and Adrian Meredith of Environment Canterbury.

Currently there are inconsistencies between the Regional Councils' State of Environment monitoring networks as regards distributions of sites, site coverage and monitoring 'intensity', variables measured, sampling regimes (frequency and timing), measurement protocols, precision, and quality assurance (QA) (Davies-Colley et al. 2011). The goal of the Variables workstream of the Ministry for the Environment's National Environmental Monitoring and Reporting (NEMaR) project is to improve consistency and "dependability" of regional water monitoring as regards variables, protocols and QA. Similarly the goal of the Network's workstream of NEMaR is to improve consistency and "dependability" of regional water monitoring as regards numbers and locations of monitoring sites.

For the initial round of analysis for LAWNZ, data for trends was not flow adjusted, mainly due to availability of data, and also because of a range of opinions on the methodologies to be used. We recognise that flow adjustment is important for trend analysis as flow regime does affect concentrations of water quality variables. An important part of this project is therefore to investigate the effect of flow adjustment on trend analysis results, and on methods for flow adjustment.

[3] The NRWQN is, however, "more comprehensive" in terms of length of record, consistency, and quality assurance (Davies-Colley et al. 2011). Improving consistency and QA ("dependability") of regional water monitoring is a major goal of the Ministry for the Environment's current National Environmental Monitoring and Reporting (NEMaR) project.

The current project, focussed on the LAWNZ website, seeks to optimise water quality trend analysis and to review the current protocols for data analysis and provide recommendations for future data analysis of both state and trends at various scales (e.g., by catchment, region) so that resultant comparisons are robust. The advice will help ensure that water quality data analysis methodologies and information display used in LAWNZ are robust, transparent and defensible. We also envisage that protocols established as a result of this project, focussed on the LAWNZ website will further promote the improved consistency and dependability of regional SoE monitoring across the country.

The aims of this project were to

- review statistical methods employed in the development of LAWNZ, and to advise on any possible changes needed in advance of a data update in May 2012
- provide guidance on statistical methods that can be used to manipulate censored data robustly to permit the inclusion of censored data in on-going routine analysis
- establish protocols for future data analysis at various temporal and spatial scales so that resultant comparisons are robust
- provide recommendations around calculating trends in a robust manner accounting for the inconsistencies in regional council networks e.g., sampling frequency
- provide advice around amalgamating trends to facilitate clear presentation
- investigate the effect of flow adjustment on trend analysis results, and provide advice and make a recommendation on the need for flow adjustment in LAWNZ.

The advice received in this project will help ensure the methodologies used in the next update of the LAWNZ website are robust. This will add to the level of confidence in the information on the website for both the agencies providing the information and for those using the information. The advice also has the potential to lead to improved consistency between State of Environment monitoring networks across the country, also recommended by NEMaR (Davies-Colley et al. 2011).

2 Trend analysis in LAWNZ – current practice

The LAWNZ Working Group has developed criteria for data processing and trend analysis[4]. All trend analyses are carried out using the Seasonal Kendall test within NIWA Time Trends software. Currently there is a minimum requirement of 5 years data for trend analysis. It is anticipated that as further years data become available, current records will be updated, thus extending the records for all existing sites. Once a site has accumulated 5 years of data, it can be included in the trend analysis.

Currently, trends are amalgamated in LAWNZ using all RSKSE values (e.g., for a given catchment) regardless of significance. The average value of all the RSKSE values is calculated to obtain the average percentage change over time for that particular group. Significance is determined by examining the number of sites within the group for which there is a significant trend. If more than 50% of sites show significant trends, the overall trend is deemed to be significant.

The LAWNZ criteria state the minimum number of samples required for a site to be included in the analysis, depending on the sampling frequency as follows:

- for monthly sampling, a minimum of 45 samples is required
- for bi-monthly sampling, a minimum of 24 samples is required, and
- for quarterly sampling, a minimum of 15 samples is required.

Trends are categorised as outlined in Ballantine & Davies-Colley (2009) (following original recommendations by Scarsbrook (2006) and shown in Table 2-1. Increasing trends of visual clarity and dissolved oxygen indicate improving water quality whereas increasing trends in other variables indicate declining water quality.

Table 2-1: Trend categories used in LAWNZ.

Statistics	Visual clarity and dissolved oxygen		All other variables	
	Statistic grouping	Score	Statistic grouping	Score
$P < 0.05$ RSKSE≤-1	Significant meaningful degradation	-2	Significant meaningful improvement	2
$P < 0.05$ -1<RSKSE<0	Significant degradation	-1	Significant improvement	1
$P > 0.05$	No trend	0	No trend	0
$P < 0.05$ 0<RSKSE<1	Significant improvement	1	Significant degradation	-1
$P < 0.05$ RSKSE≥1	Significant meaningful improvement	2	Significant meaningful degradation	-2

[4] Land and Water New Zealand, Water Quality Module, Data processing and analysis. Maree Clark, Horizons Regional Council, 2010, version 20.

To describe water quality state at individual sites in LAWNZ, median values of water quality variables are used. Median values are calculated for all water quality variables at all sites. To present information on state for a particular region or land use category, the median values for any given variable at sites in that category are ranked, and quartiles are calculated. LAWNZ currently groups water quality information by topography and land use (Lowland, Upland, Lowland/Forest, Lowland/Rural, Lowland/Urban, Upland/Forest, Upland/Rural, Upland/Urban, Forest, Rural, Urban). The median value for each variable at each site is then compared with the national quartile values for the altitude or land use and categorised accordingly. Median values for individual sites are amalgamated to give overall summary statistics for land use and geographical groupings (e.g., median, upper and lower quartile).

3 Challenges for LAWNZ

LAWNZ faces a number of challenges due to inconsistencies in the regional datasets on which the website is based. In this section I aim to present the main challenges and offer recommendations for robust analysis and presentation techniques.

3.1 Variables measured

All regional council State of Environment water quality samples are analysed for a wide range of variables, which are largely similar but often not exactly the same. There are variations in how they are described, e.g., dissolved phosphorus may be known as dissolved reactive phosphorus (DRP), soluble reactive phosphorus (SRP) or filterable reactive phosphorus (FRP). Nitrate may be reported as either NO_3, NNN or NO_x-N, while ammonia may be reported as NH_4 or NH_4-N. It is important to define variables accurately, as while the same variable names may be used, they could mean slightly different things.

There may also be issues with comparing water quality data between laboratories. When data from commercial laboratories are compared, data generally falls within a similar range and comparable detection limits are used. Comparing NIWA water quality data with commercial laboratory data can reveal considerable inconsistencies. The example of NIWA and Waikato Regional Council data for Upper Waikato River sites was given. Analysis of long term data from samples collected at the same sites (on different days), yet analysed by different laboratories, revealed very different median values for water quality variables, with the NIWA medians lower than the Waikato Regional Council medians Data from the same site over a long period of time shows different results, and the major contrast in the datasets is largely attributable to differences in precision between the laboratories (e.g., detection limits are lower for the NIWA laboratory than the commercial laboratory).

3.2 Analysis methods

3.2.1 Definition of analytical methods

A variety of analytical laboratories is used to analyse State of Environment water quality samples. Multiple laboratories may use the same names for variables, but there may be differences in the actual analytical methods used. It is important to know what the actual analytical methods are, so as to be sure that the variables being compared are the same. We would recommend that the various methods used for water quality analysis be compiled. Changes in analytical methods over time should also be recorded. Analysis methods should be concordant, not necessarily identical.

3.2.2 Reporting units for *E. coli*

For *E. coli*, there can be some ambiguity about methods and units of measurements used. Thirteen councils report *E. coli* as MPN (most probable number), while two report *E. coli* numbers as CFU (colony forming units), although it is not always clear that the appellation "CFU" really does denote a traditional colony-forming test. The difference in methods and units used may not be so important, as trends in *E. coli* are rather more difficult to assess and detect. I therefore consider it satisfactory to amalgamate the datasets, regardless of the laboratory method used, because of the variability between the measurements. I recommend that the *E. coli* number be reported as n/100 ml and that the method used is clearly defined elsewhere.

3.2.3 Change in laboratory provider

Over the period of record, some regional councils will have changed laboratory provider, which may have also involved a change in some laboratory methods. This can cause artefactual trends and confound the recognition of actual time trends in water quality (e.g., Ballantine & Booker 2011). Accordingly, I recommend that the same laboratory is used to carry out water quality analyses and changes be avoided so far as possible. Any changes that are unavoidable require a period of overlap of at least a year for all sites with duplicate analysis to ensure comparability though I appreciate that this is expensive. If identical analytical tests and equipment are used at both laboratories, it may be feasible to use a selection of sites for the overlapping period. As Davies-Colley et al. (2011) point out, this expense is an impediment to change in laboratory provider.

3.3 Detection limits

Water quality data is often reported as being below a certain detection limit ('less than' or censored data). In other words, laboratory analysis has shown that the exact concentration of the variable is relatively low (not able to be measured with high certainty); it is known to be somewhere between the laboratory's reporting level and zero. As such, detection limits are defined for a variable as a function of precision.

3.3.1 Detection limits and statistical analysis

While the use of detection limits may seem sensible for laboratories, censored data makes it difficult to analyse water quality data, e.g., to compute trends or summary statistics (e.g., Porter et al. 1998), and, in general, censored data contain less information than raw data. 'Substitution' can be used to enable statistical manipulation of censored data, i.e., using half the detection limit, zero or the detection limits. Alternatively there are more complex statistical tools available with which censored data can be manipulated for use in trend analysis and summary statistics (see Ballantine & Booker 2011 for a worked example).

3.3.2 Detection limits in LAWNZ water quality data analysis

In LAWNZ, the current practice is to halve the detection limit. I recommend that this should be continued until other suitable options are available, for example, options for analysing censored data could be incorporated into TimeTrends. Because detection limits for any particular variable will vary between the councils, this process should be done on an individual variable and individual council basis.

Often, over the monitoring period, the detection limit for a variable will have changed due to laboratory or method change. Where this has occurred, and where more than 40% of the data for a particular variable is censored, I recommend that the higher of the two detection limits be used and applied not only to the censored data, but to the other values that are below the higher detection limit. Where there are multiple detection limits, and where censored data constitutes less than 40% of the data set for a particular variable, I recommend that censored values are treated individually and that they are halved. Other measurements should not be amended.

I also recommend that the different precision levels (and detection limits defined in terms of these precisions) for variables used by regional councils are compiled for reference. Changes in detection limits over time should also be noted, especially where the changes have been associated with a change in laboratory.

Occasionally a laboratory may provide results reported to a higher detection for a sample where there has been an error related to dilution. I recommend that such anomalous results are omitted from data analysis.

3.3.3 Detection limits and suitability for analysis

Where results for two variables have to be added together to give a value for another related variable (e.g., the addition of nitrate and nitrite to give total organic nitrogen, or total kjeldhal nitrogen and total organic nitrogen to give total nitrogen), and the results for one of the variables is censored, I recommend that the censored value be replaced with half the detection limit to enable addition.

Where there is more than 70% censored data for a variable at a site, I recommend the variable is not included in the analysis.

3.4 Frequency of monitoring

Throughout New Zealand, protocols for State of Environment water quality sampling vary widely. In particular the frequency of monitoring differs between the different regional councils. Furthermore, some councils have relatively short water quality records, while other regional councils have data sets collected using consistent protocols that span long time periods, which provide a robust basis for trend analysis.

3.4.1 Sampling interval

Sampling intervals (that is, frequency of sampling) varies between the regional councils. Some regional councils undertake water quality sampling on a monthly basis, while others, due to budgetary or other constraints, collect samples bimonthly or quarterly. This difference in temporal resolution of data makes comparison of results difficult as it is not appropriate to compare trends calculated for different sampling frequencies.

Because of the different sampling frequencies, trend analysis in LAWNZ is currently carried out at a range of temporal scales (monthly, bi-monthly, quarterly). It has been previously observed that, using bi-monthly or quarterly data, trends are either not observed, or are weaker than they would be with monthly data (Stansfield 2001). This is because smaller data sets have larger standard errors. In the case of quarterly data, the standard error may occasionally be large enough to discount a trend that becomes evident when using monthly data. Stansfield (2001) detected fewer trends and found that, at the same site, trends detected using quarterly data were different when compared with the trends detected with monthly data. Trend analysis is best carried out using monthly data, and monthly data is available for a wide range of sampling sites throughout the country.

Where trends at a particular site are calculated using monthly data, I recommend this is highlighted on the webpage for that site. Equally, where quarterly data has been used for trend analysis, I recommend this is clearly stated. In any documentation giving advice for users of LAWNZ, I recommend that the difference in time periods for trend calculation is highlighted, so that users are aware that comparing trends calculated from data of sampling frequency is not appropriate.

To illustrate the influence of temporal resolution on trend detection, an example is provided based on data from three NRWQN sites (Table 3-1). Meaningful increasing trends were observed for nitrate-nitrogen (NO_x-N) at all three sites over an 11 year time period (DN5, 7, and 8). There was one significant trend when trends were calculated using bi-monthly data over the same time period, while reducing the data frequency further to quarterly meant that no significant trends were observed in the data. The Sen slope, RSKSE values and median concentrations also changed with the sampling interval, although these changes are expected to be the result of stochastic 'noise' rather than bias.

Table 3-1: Comparison of trends calculated for NO_x-N at 3 NRWQN sites for monthly, bi-monthly and quarterly data (2000-2011).

Site	Sampling frequency	Median concentration	P value	Sen slope	RSKSE
DN5	Monthly	1064.10	**0.010**	19.90	1.87
	Bi-monthly	1053.47	0.078	14.78	1.40
	Quarterly	1036.89	0.069	17.21	1.66
DN7	Monthly	541.50	**0.001**	20.24	3.74
	Bi-monthly	530.73	0.008	17.25	3.25
	Quarterly	546.84	0.107	15.08	2.76
DN8	Monthly	987.21	**0.038**	19.59	1.98
	Bi-monthly	942.46	0.208	16.00	1.70
	Quarterly	827.17	0.069	25.47	3.08

3.4.2 Seasonality and the Seasonal Kendall test

Seasonality is a source of variability in water quality data that might prevent the detection of trends. Concentrations of some water quality variables vary throughout the year. Using monthly data and the Seasonal Kendall test, the effect of seasonal variation is reduced by comparing only water quality values that are collected during the same season or month of each year. Using monthly data, this means that all January data is compared, all February etc. The test is performed for individual 'seasons' of the year, and then the individual tests are combined into one overall test to ascertain if there is a monotonic trend (a change in a consistent direction over time) (Hirsch et al. 1982). Where sampling is monthly, the number of seasons used in the Seasonal Kendall test should be set to 12, i.e., the number of annual samples. Monthly data can be analysed using fewer seasons within TimeTrends however when this is done, the median of the three values making up each quarter will be used in the calculations and so information is lost. Where sampling is bi-monthly or quarterly, the number of seasons used in the Seasonal Kendall test should be set to 4.

To compensate for variations in data density resulting from changes in sampling frequency, the number of annual seasons should be selected to reflect the years with the smallest sampling frequency. Accounting for seasonality is a major reason why monthly monitoring is strongly recommended by NEMaR (e.g., Davies-Colley et al. 2011).

3.5 Time period for trend analysis

Regional council data sets are of highly variable lengths, ranging from over twenty years to just a few years. Data stored on the LAWNZ website is for 2004 to 2011 and will be added to as new data becomes available, however not all of the information for each site is complete for the full time period. It is difficult to compare trends that are calculated over different time periods, for example, comparing trends that have been calculated with a 5 year data set should not be compared with trends calculated over an eight year period. To avoid users comparing trends calculated over a shorter time period with those calculated over the maximum time period, I therefore recommend that trend analysis should only be carried out and displayed on the website for sites with data for the maximum time period. As shown below, trends at the same site calculated for different time periods may differ (Table 3-2).

I recommend that:

- Trends are calculated only when data for the maximum time period held in LAWNZ is available.
- Only trends calculated for the same time period are compared between sites.

3.5.1 Length of record for trend analysis – long and short term trends

For trend analysis, it is necessary to have long and continuous data records. I, and other authors (Burt et al. 2008, Howden et al. 2011), have found that trends on data from periods greater than ten years are more robust than short term trends.

Table 3-2: Comparison of NO_x-N trends at 4 NRWQN sites calculated over 5, 10 and 22 years.

Site	Time period	Median	*P* value	Sen slope	RSKSE
DN5	5 years	1098	0.83	5.77	0.53
DN5	10 years	1064	0.01	19.90	1.87
DN5	22 years	960	0.01	22.62	2.36
DN6	5 years	258	0.36	6.18	2.39
DN6	10 years	255	0.15	2.24	0.88
DN6	22 years	230	0.01	3.97	1.73
DN7	5 years	614	0.44	14.29	2.33
DN7	10 years	541	0.01	20.24	3.74
DN7	22 years	418	0.01	18.14	4.34
DN8	5 years	1057	0.29	-39.57	-3.74
DN8	10 years	987	0.04	19.59	1.98
DN8	22 years	758	0.01	27.62	3.64

Table 3-2 shows trends in NO_x-N for 4 NRWQN sites, calculated over 5, 10 and 22 years. Trends were not significant over the 5 year period; most were significant over the 10 year time period and all were significant for the 22 year period. At these four sites, NO_x-N concentrations have consistently increased over the period of record (Figure 3-1). Having trends calculated at longer and shorter time scales gives information on short term variability within the longer time period.

Short term trends can give very different information from the longer trends as shown in Table 3-3 and Figure 3-2. This example shows that DRP concentrations have been increasing significantly over the 22 year time period, however over the 5 and 10 year periods, they have been decreasing, though only significantly for the 10 year period. Without seeing the time series for the full time period, it would be easy to assume that water quality has improved significantly at this site, when, in truth, concentrations have merely fallen back to values characteristic of when monitoring commenced. Short term trends therefore need to be interpreted with care and should be set in the context of the longer term data and accompanied by data plots.

Table 3-3: Trends in DRP for NRWQN site DN5 over 5, 10 and 22 years.

Period	Median	P value	Sen Slope	RSKSE
5 years	13.51	0.72367	-0.15	-1.13
10 years	17.26	0	-0.79	-4.57
22 years	14.54	0.00011	0.24	1.62

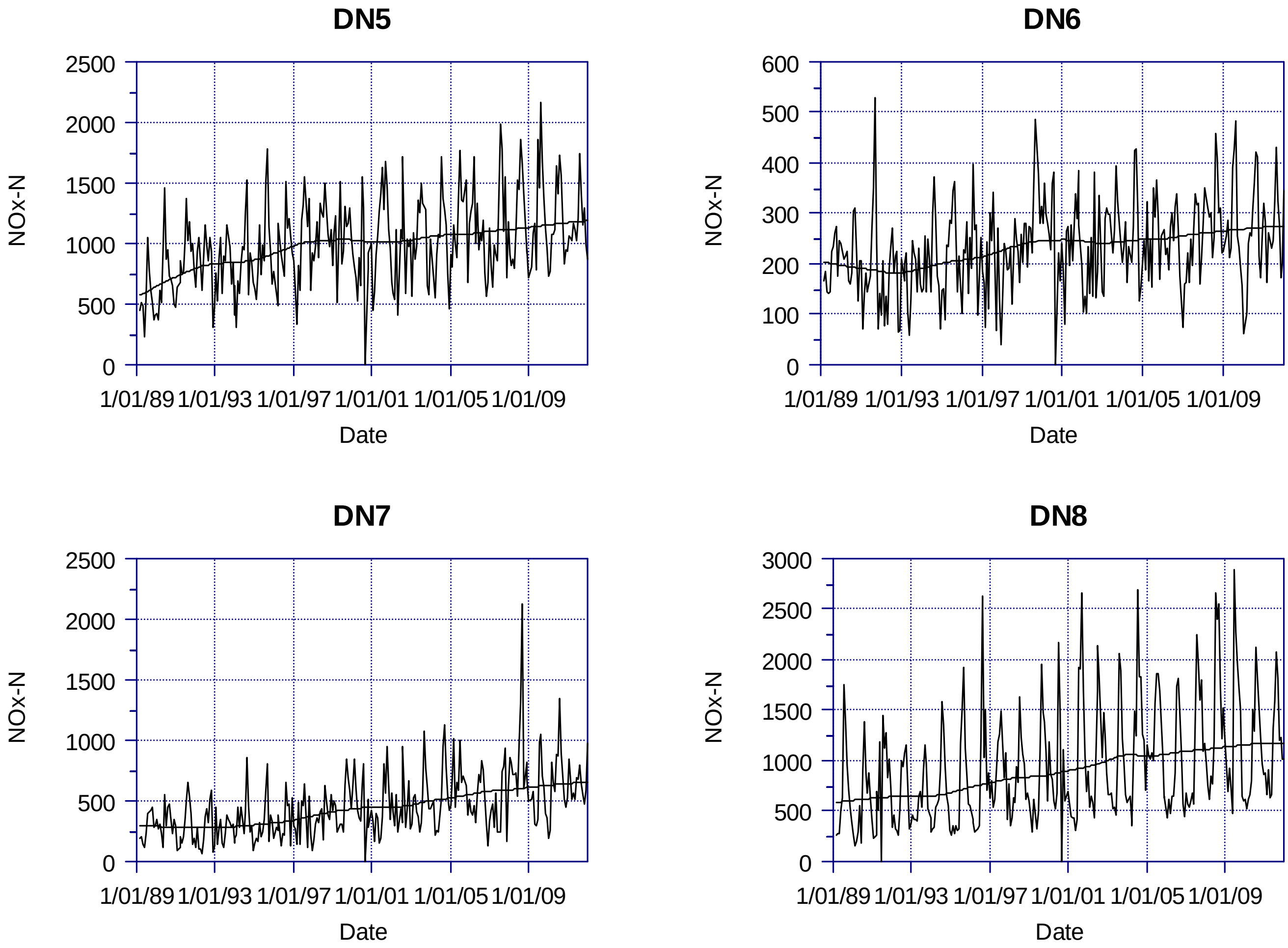

Figure 3-1: Time series of NO_x-N concentrations at 4 NRWQN sites for 1989-2011.

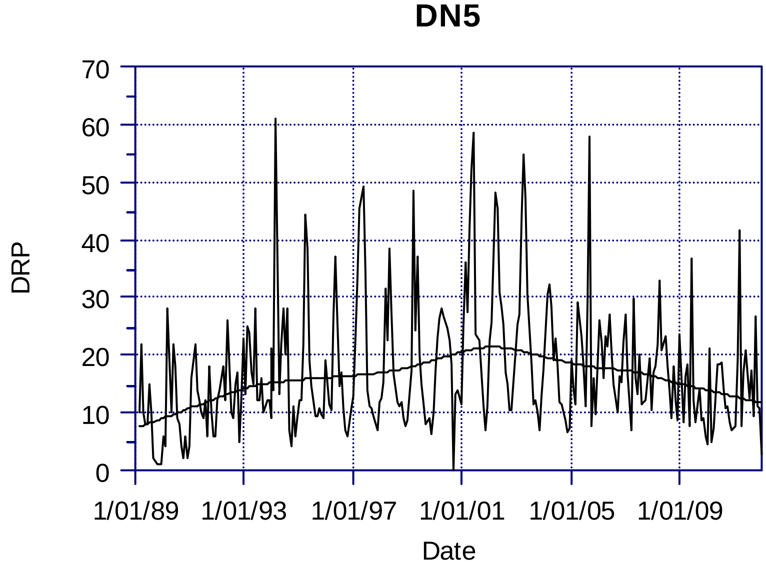

Figure 3-2: DRP concentrations at NRWQN site DN5 from 1989-2011.

3.6 Flow adjustment – to flow adjust or not

3.6.1 Background

Most water quality variables are correlated with flow, either positively or negatively. Smith et al. (1996) discuss typical relationships of variables measured in the NRWQN with flow. The nature of the relationship between streamflow and concentration varies by constituent and individual river basin. In catchments dominated by diffuse source pollution, pollutants (like total phosphorus or *E. coli*) are generally positively correlated with flow, and concentrations increase in high flows because of wash-in during rainstorms or mobilisation from in-channel stores. The loading of pollutants is relatively constant in point source dominated catchments, so an increase in flow will generally result in reduced concentrations due to dilution. Some solutes (major ions and consequently conductivity) tend to dilute with increasing flows. For analysis of trends, concentrations are often adjusted to remove the effects of streamflow.

3.6.2 Is flow adjustment appropriate for LAWNZ?

When analysing water quality data for trends over time, there is the question of whether or not to flow adjust data to take account of relationships with flow. The user can define whether or not to flow-adjust data within TimeTrends.

The decision will depend on the question being asked of the data. If the main question has to do with the actual concentrations in the water body, then data should not be flow adjusted before trend analysis. Non-flow adjusted trends are used extensively to quantify changes in water quality conditions affecting a particular environmental resource and are appropriate for determining changes in water quality at the monitoring site and examining the overall response of the ecosystem to changing water quality.

If, in contrast, the concern is the amount of pollution coming from a catchment, data should be flow adjusted as this allows investigation of whether associated loads are increasing or decreasing. An important point is that the flow-adjusted trend does not necessarily represent all the water-quality changes that result from human influence and management actions; it only describes those separate from flow.

The LAWNZ website is not concerned with fluxes of pollutants, but rather what is actually happening in the river at a particular point. It is therefore satisfactory to use non flow adjusted trends on the website.

3.6.3 Flow adjustment method

In flow adjustment, a line or curve relationship is developed between the variable and covariate (flow) and then each variable is adjusted according to the value of the covariate (flow) and the selected relationship. Flow adjustment can be done within TimeTrends using various methods. Data should be plotted before the adjustment method is chosen, and the most plausible method (e.g., LOWESS, log-log) should be chosen. LOWESS (Locally Weighted Scatterplot Smoothing) has been found to be a suitable method for flow adjustment and has been used in various trend assessments (e.g., Ballantine & Davies-Colley 2009). It was also considered by Smith et al. (1996) to be a good compromise between under- and over- smoothing for the Network data when monthly data was available for a period of 5 years. I would therefore recommend that LOWESS is used for flow adjustment as long as there is sufficient data.

3.6.4 Comparison of flow adjusted and non-flow adjusted data

Flow adjusted and non-flow adjusted trends have been compared for NRWQN sites for a range of variables for data spanning 22 years (Table 3-4). Comparison of flow adjusted and non-flow adjusted Sen slope values shows that, regardless of significance, they are closely related and r^2 values are high. Comparing significance between flow adjusted and non-flow adjusted trends implies that marginally more significant trends are detected with flow adjusted data than non-flow adjusted data – because flow adjustment reduces data variability.

It is worth remembering that, if there is a strong trend in flow over the period of investigation, this may influence non flow adjusted trends in variables with strong relationships with flow. A cautionary note should be added to the website to indicate that, if there was a flow trend over the period, trends should be interpreted carefully as the trend may have been caused by flow trends.

Table 3-4: Comparison of flow adjusted and non-flow adjusted for 77 NRWQN sites for 1989-2011.

Variable	R^2	% Non-flow adjusted/flow adjusted
Nitrate-nitrogen (NO_x-N)	0.9697	90.4 (47/52)
Total nitrogen (TN)	0.9653	83.7 (41/49)
Dissolved reactive phosphorus (DRP)	0.9847	80 (32/40)
Total phosphorus (TP)	0.9471	78.8 (26/33)
Percent dissolved oxygen (% DO)	0.9465	89.7 (35/39)
Clarity	0.782	86.7 (39/45)

3.6.5 Challenges for flow adjustment of LAWNZ data

Actual flow measurements (or reliable flow estimates) are not available for most water quality sampling sites in New Zealand. For example, Waikato Regional Council has actual flow measurements for 25% of water quality sites, and estimated flow for the remaining 75%. I recommend that councils should flow-stamp all sampling visits so flow adjusted trends can be calculated and compared with non-flow adjusted trends. By "flow stamping" is meant a reasonably precise (+/- 15%) estimate of flow at the time of sampling from a nearby rated site on the same main-stem river (not necessarily a continuous level recorder, although a flow rating and a staff gauge will be needed).

3.7 Amalgamating trends and estimating overall trends in LAWNZ

Currently the overall trends are estimated by calculating the average RSKSE value for all the sites within a grouping. The overall trend is deemed to be significant if there are significant trends at more than 50% of sites within the grouping. I recommend this procedure is changed so that the median RSKSE value is used as the measure of central tendency. The median is a non-parametric measure and, unlike the mean, will not be skewed by outliers. Also, with overall significance, I recommend this be determined through use of a binomial test if the data is suitable (i.e., if there are no zero trends). The current method does not distinguish sufficiently between positive and negative significant trends at individual sites.

3.7.1 Estimating the overall national trend

To **estimate** the overall national trend for any given water quality variable, all of the slopes (RSKSE values) for the variable should be included in the assessment, regardless of significance and regardless of sampling frequency. The median of all the slopes will give the overall direction of the trend at the national level. No statistical test should be done on this data set as the trends will have been calculated using data with different sampling frequency (i.e., monthly, bi-monthly, quarterly) and with different seasonal resolution (12 seasons or 4 seasons) in the Seasonal Kendall test. The overall trend for a region, land use or New Zealand is the median RSKSE value for the chosen grouping.

3.7.2 Significance of amalgamated trends

Only trends that are calculated from data with the same sampling frequency should be amalgamated to determine whether or not the overall trend for a grouping is statistically significant. This means that the significance of the overall national trend should not be calculated because not all regions in New Zealand do monthly sampling. Significance of the overall trend direction can be calculated for individual regions or environmental groupings as long as all sites have been sampled at the same frequency (i.e., monthly, quarterly) and the same number of seasons (i.e., twelve or four) has been used in the Seasonal Kendall test. The binomial test can be applied to test for significance using the RSKSE values for each site, but because this is a two way test, it can only be applied if there are no zero slopes. It may be appropriate to develop a multinomial test of the statistical significance of this national level result when there are zero slopes, as zero slopes are possible with the non-parametric (rank-based) methods being used. For the time being, I recommend that statistical significance is determined only when there are no zero RSKSE values.

For LAWNZ, I recommend:

- that the overall direction of the national trend is estimated using the median value of all calculated trend slopes (RSKSE values) regardless of sampling frequency and statistical significance
- that statistical significance for groups of sites should only be calculated when trends have been calculated on data with the same temporal resolution. This can be done with a binomial test using all the RSKSE values for a given grouping, regardless of significance. Where there are zero RSKSE values, the statistical significance should not be tested. This means that the statistical significance of the overall national trend cannot currently be robustly estimated because the underlying data has not been collected at the same frequency.

3.7.3 Time series of data

The LAWNZ website at present provides numerical summary statistics on water quality variables at sampling sites. While this is useful information, I recommend that time series graphs of variables should be added to the webpages. Time series graphs would provide an indication of how water quality variables have changed over time and, along with summary statistics, would give users very comprehensive information on water quality. Because some variables vary appreciably from point-to-point, it would be useful to routinely smooth the data so the underlying trend direction is visible (e.g., LOWESS). Time series graphs of TN and NO_x-N concentrations at 1 NRWQN site have been provided as examples (Figure 3-3, Figure 3-4.). Data has been smoothed using a LOWESS smoother.

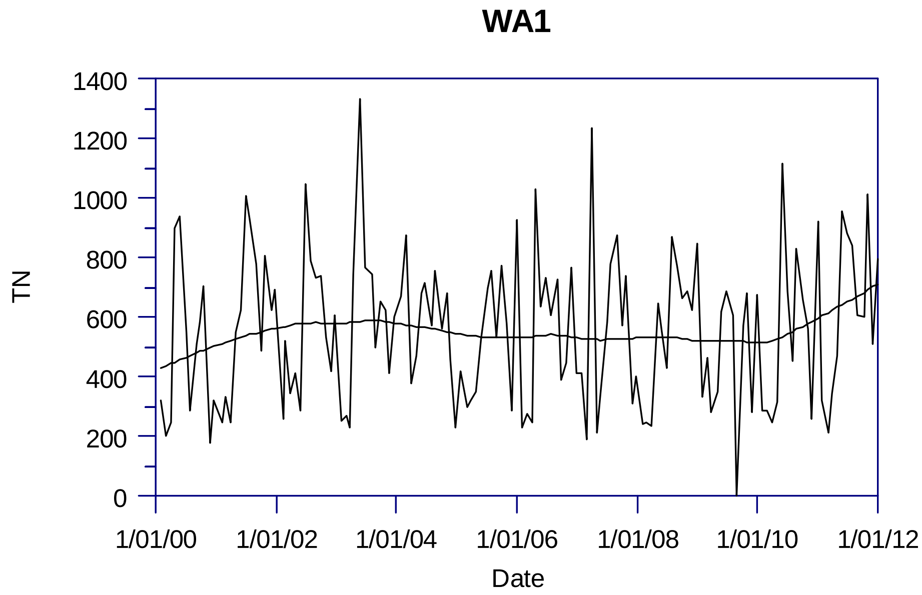

Figure 3-3: TN concentrations at NRWQN site WA1 for 2000-2012.

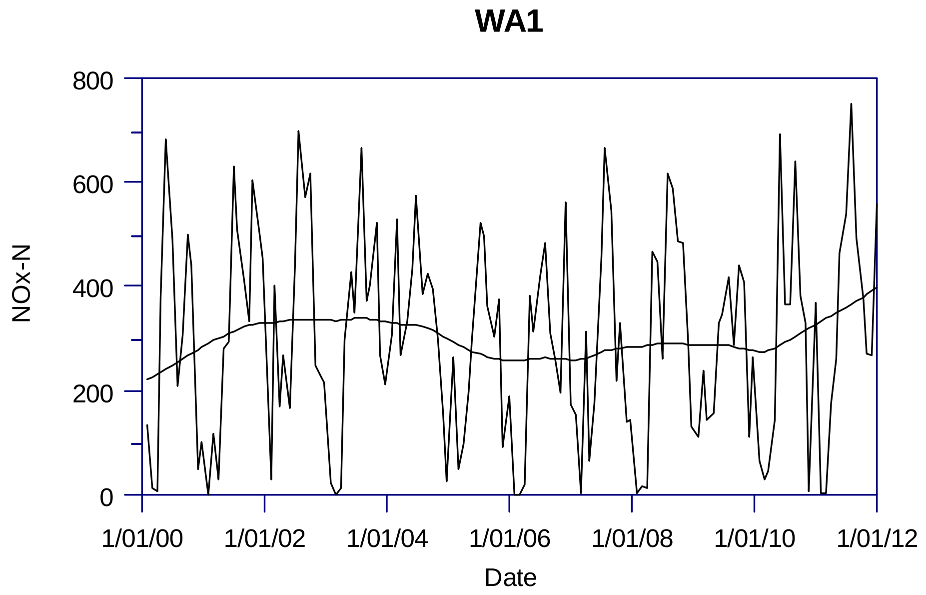

Figure 3-4: NO_x-N concentrations at NRWQN site WA1 for 2000-2012.

3.7.4 Harmonised sampling frequency

A monthly sampling frequency in both lakes and rivers is strongly recommended for the LAWNZ dataset, particularly to support robust water quality state, trend analysis and direction into the future.

3.7.5 Analysis information

The LAWNZ website is constructed so it is suitable for a range of users, from the general public to Chief Executives of Regional Councils. It is envisaged that it will become an increasingly useful resource for the scientific community. With scientific applications in mind it would be beneficial if comprehensive information on data analysis and compilation was available to fully meet the needs of scientific and technical users. Data analysis information should be readily available for website users.

4 Recommendations for LAWNZ

4.1 Core variables

I recommend:

- ‘Core’ variables are identified that all councils will commit to consistently and indefinitely measuring. Care needs be taken to ensure variables are correctly and consistently defined. A common nomenclature for water quality variables needs to be adopted by all regional councils for ease of understanding and to ensure the same variables are being compared.
- Quality assurance (QA) is given much higher priority in regional monitoring, both of the actual field monitoring activity and processing of data.

4.2 Analysis methods

For LAWNZ, I recommend that:

- The various analytical methods used by each of the laboratories for sample analysis are documented.
- Changes in methods over time should be documented as they may be useful in explaining unexpected results from state and trend analysis.
- *E. coli* numbers are reported as n/100 ml and the method used defined.
- As far as possible, changes in laboratory provider are avoided by regional councils. There should be a period of overlap (1 year suggested) with duplicate analysis at all sites in the event of a laboratory change.
- Precision be reported with each batch of sample analysis.

4.3 Detection limits

I recommend that:

- Until better data analysis methods are available (e.g., analysis options in TimeTrends), data reported as “less than the detection limit” should be halved to facilitate analysis in LAWNZ.
- Where there is more than one detection limit for any variable within the data for an individual council, and where more than 40% of the data is censored, halving should be applied as the half of the higher detection limit. Where more than 70% of the data for a variable is censored, I recommend the variable is omitted from the analysis.
- Anomalous data (often censored at higher values than the detection limit) that occur due to issues around dilution during laboratory analysis should not be included in the analysis.
- Detection limits and precisions for water quality variables by different councils should be compiled. Changes in detection limits and measurement precision over the monitoring period should be documented.

4.4 Frequency of monitoring

For the LAWNZ website, I recommend the following:

- Trends should be calculated only when the data for the maximum period held in LAWNZ is available. Shorter term trends can be misleading and should be set in the context of longer term trends calculated for the same site. When calculating trends using monthly data, the Seasonal Kendall test should be calculated using 12 'seasons' to account for seasonal variability. Where sampling frequency is less than monthly, quarterly trends should be calculated (using four seasons within TimeTrends).
- To avoid users comparing trends calculated for different (shorter) time periods at different sites, only trends calculated over the maximum time period should be displayed on LAWNZ. It is important to only compare trends calculated over the same time period.
- For comparison of trends between sites and regions, the website should have a filtering facility so that only monthly trends can be compared between sites for which monthly trends have been calculated, and quarterly trends between sites with only quarterly trends calculated.
- The percentage of sites for which trends are calculated using monthly, bi-monthly and quarterly data be added to the website.
- Information on water quality variables monitored by Regional Councils and sampling frequency should be collated to ensure the data set is used to its maximum potential.

4.5 Time period for trend analysis

For LAWNZ, I recommend that:

- Trends should only be calculated when data for the maximum time period is available. Short term trends can be misleading so it is best to calculate trends for the longest period of data possible. Equally trends can only be robustly compared when calculated over the same time period, with data collected at the same sampling frequency and analysed with the same number of seasons in the Seasonal Kendall test.

4.6 Flow adjustment

For LAWNZ, I recommend that:

- Data for trend analysis is not flow adjusted (for now, only because water quality samples are not all flow stamped).
- A cautionary note is added to the website to remind users that a trend in flow may have caused a trend in a related water quality variable.
- Councils provide "flow-stamped" data to LAWNZ where possible, and the trend for flow be given as well as the trend for water quality.

- In future, if flow-stamping becomes standard practice, as recommended by NEMaR reports, trends could be routinely given for flow-adjusted as well as non-flow adjusted data. LOWESS should be used for flow adjustment.

4.7 Amalgamating Trends

- The median should be used as a measure of central tendency.
- All of the slopes (RSKSE values) for a variable should be included in the assessment regardless of significance and sampling frequency. The median of all the slopes will give the overall direction of the trend at the national level. No statistical test should be done on this data set because of the variation in sampling frequency.
- Trends calculated from data from the maximum time period with the same sampling frequency can be amalgamated to determine whether or not the overall trend for a grouping is statistically significant. The significance of the overall national trend should not be calculated because not all regions in New Zealand do monthly sampling. The binomial test can be applied to test for significance using the RSKSE values for each site, but because this is a two way test, it can only be applied if there are no zero slopes.
- Where there are zero slopes, a statistical test should not be applied. A multinomial test should be developed to accommodate zero slopes.

4.8 Time series of data

I recommend that:

- Time series data is presented for each site.
- A smoothed trend line is also presented routinely.

4.9 Harmonising sampling frequency

I recommend that:

- Monthly sampling should be carried out at State of Environment sites.

4.10 Analysis information

The LAWNZ website is constructed so it is suitable for a range of users, from the general public to Chief Executives of Regional Councils. It is envisaged that it will become an increasingly useful resource for the scientific community. With scientific applications in mind it would be beneficial if comprehensive information on data analysis and compilation was available to fully meet the needs of scientific and technical users.

I recommend that:

- Detailed information on how the data is analysed is compiled in a technical document that is readily accessible (downloadable .pdf) from the website.

This report has acknowledged the need for harmonisation of water quality sampling protocols, and data analysis procedures. In that respect, the future of the LAWNZ website is strongly tied to the recommendations of the current NEMaR project which seeks to improve consistency and “dependability” of regional SoE data.

5 Conclusions

The LAWNZ website is a useful and unique portal for presenting environmental information from New Zealand's regional and unitary councils. It is the first attempt at amalgamating all Regional and Unitary Council State of Environment water quality data. Throughout this process, various challenges have arisen. To ensure current and future analysis and presentation of this complex data set is robust, I have made the following recommendations:

Variables measured:

- 'Core' variables be identified that all councils will commit to consistently and indefinitely measuring.

Analysis methods:

- The various analytical methods used by each of the laboratories for sample analysis are documented, along with any changes over time. Care should be taken with variable definitions so as to avoid ambiguity and confusion.

Detection limits

- Data reported as "less than the detection limit" should be halved to facilitate analysis in LAWNZ. Where there are two detection limits for any variable within the data for an individual council, halving should be applied as the half of the higher detection limit. Changes in detection limits over time should be documented.

Frequency of sampling

- Into the future, water quality sampling should be harmonised and done monthly by all councils.

Trend analysis

- Trends be calculated only when data for the maximum period held in LAWNZ is available. Only trends calculated for the same time period can be compared between sites.

- Where there is monthly data available, trends should be calculated using twelve seasons within the Seasonal Kendall test so as to make use of all the available data. Where data resolution is less than monthly, quarterly trends (four seasons within TimeTrends) should be calculated. The period of time over which trends are calculated should be mentioned on individual webpages within LAWNZ, e.g., monthly or quarterly.

- Data for trend analysis is not flow adjusted for the time being. Into the future, Regional Councils should work towards flow stamping all water quality sampling. LOWESS should be used for flow adjusting data.

Amalgamation of trends

- To estimate an overall trend, all of the RSKSE values for a variable should be ranked, regardless of significance and sampling frequency. The median of all the slopes will give the overall direction of the trend at the national level. No statistical test should be done on this data set.

- To determine the significance of an overall trend, RSKSE values from each site, calculated from data from the maximum time period with the same sampling frequency, can be amalgamated. The binomial test can be applied to test for significance. This is a two way test, and can only be applied if there are no zero slopes. Where there are zero slopes, the binomial test should not be used. A multinomial test should be developed to accommodate zero slopes. The significance of the overall national trend should not be calculated because not all regions in New Zealand do monthly sampling.

Standardisation of data is important to allow clear reporting, such as by the LAWNZ website, however scientific innovation will be on-going and may lead to new approaches being introduced, either for sample collection and analysis or data processing. This inevitably introduces tension between consistency and innovation, and so there is a need to recognise that some change in procedures, and also perhaps variables measured, is probably inevitable over time. However, when changes are being considered, there should be a careful consideration of the likely consequences, including for trend analysis.

6 Acknowledgements

Funding for this study was received from MSI's Envirolink Fund. I thank the working group from the regional councils who provided a sounding board for ideas and advice. The draft was reviewed by Graham McBride and Rob Davies-Colley.

7 References

Ballantine, D.J.; Booker, D. (2011). Greater Wellington State of Environment river water quality data assessment and trend analysis (2003-2010*). NIWA Client report HAM2011-126* for Greater Wellington Regional Council. 65 p.

Ballantine, D.J.; Davies-Colley, R.J. (2009). Water quality trends at National River Water Quality Network sites for 1989-2007. Ministry for the Environment Wellington. 43 p., http://www.mfe.govt.nz/publications/water/water-quality-trends-1989-2007/index.html

Burt, T.P.; Howden, N.J.K.; Worrall, F.; Whelan, M.J. (2008). Importance of long-term monitoring for detecting environmental change: lessons from a lowland river in south east England. *Biogeosciences 5(6)*: 1529–1535.

Davies-Colley, R.; Larned, S.; Unwin, M.; Verburg, P.; Hughes, A.; McBride, G.; Ballantine, D.; Hudson, N.; Daughney, C.; Hamill, K. (2011). Dependable monitoring of freshwaters for national-scale environmental reporting. *Client report HAM2011-055* prepared on behalf of the Ministry for the Environment. 142 p.

Davies-Colley, R.; Verburg, P.; Hughes, A.; Storey, R. (2012). Variables for regional water monitoring underpinning national reporting. *NIWA client report HAM2012-006* for Ministry for the Environment. .

Hirsch, R.M.; Slack, J.R.; Smith, R.A. (1982). Techniques of trend analysis for monthly water quality data. *Water Resour. Res. 18*: 107–121. <http://dx.doi.org/10.1029/91wr00259>

Howden, N.J.K.; Burt, T.P.; Worrall, F.; Whelan, M.J. (2011). Monitoring fluvial water chemistry for trend detection: hydrological variability masks trends in datasets covering fewer than 12 years. *Journal of Environmental Monitoring 13(3)*: 514–521. <http://dx.doi.org/10.1039/c0em00722f>

Hudson, N.; Ballantine, D.; Gibbs, M.; de Winton, M.; Storey, R.; Verburg, P.; Hamill, K. (2011). Investigation of single indicators for water quality assessment and reporting. *NIWA client report HAM2011-066* prepared on behalf of the Ministry for the Environment. 170 p.

McBride, G.B. (2005). Using statistical methods for water quality management. Wiley, New Jersey.

Porter, P.S.; Ward, R.C.; Bell, H.E. (1998). The detection limit. *Environmental Science and Technology 22*: 856–861.

Scarsbrook, M.R. (2006). State and trends in the National Water Quality Network (1989-2005). *NIWA Client Report HAM2006-131* to MfE.

Smith, D.G.; McBride, G.B.; Bryers, G.G.; Wisse, J.; Mink, D.F.J. (1996). Trends in New Zealand's river water quality network. *New Zealand Journal of Marine and Freshwater Research 30*: 485–500.

Stansfield, B. (2001). Effects of sampling frequency and laboratory detection limits on the determination of time series water quality trends. *New Zealand Journal of Marine and Freshwater Research 35*: 1071–1075.

Appendix A Key features regarding archiving and interpreting water quality and aquatic biology data

In 1989 a product called AQUAL (Aquatic Quality Unified Archival Library, McBride & Shankar 1989) was developed, intended for use by DSIR, Regional Councils, and consultants. It was never taken up (most likely because of high implementation costs), but some Councils (especially Waikato regional Council) implemented parts of it. Nevertheless its design had considered many aspects of archive requirements, and these may be worth re-stating given their relevance to LAWNZ and other environmental information initiatives. . One particular feature that emerged is the necessity to minimise the possibility of archiving ambiguous data, which seems to be especially easy to do with water quality and ecological data (if haphazard data entry is allowed) because appropriate data interpretation needs to take account of a lot of meta information. A (not exhaustive) list of features and issues that particularly need to be addressed follows.

1. Sample location: Water depth? Position across a river's width? Where in the lake/estuary?
2. Sampling method: Size of filter? Grab or composite?
3. Who did the sampling and when?
4. What organisation gathered this datum?
5. Field or laboratory measurement?
6. Measurement/laboratory method?
7. Reporting units (e.g., is nitrate reported as N or as NO3? Is ammonium reported as N or NH4? What does "Ammonia" mean?)
8. Who is able to enter data? Not just anybody?
9. Who checks the data for consistency and failures of various "laugh tests"? For example, transcription errors can result in pH = 71, but it should have been 7.1 (any laugh test should be able to screen that out). DRP should be less than TP.
10. What is mandatory input? That is, if we don't require input of key meta-data we could end up in a sea of ambiguous data. But if we go overboard on mandatory data we may discourage data entry—no simple solution to this one!
11. Maybe flag all data on entry as "provisional", with that flag being changed to "reliable" (or some such word) once it has been checked.
12. Entry of Regional Council and (particularly) consultant's/client's data may often be lacking desired meta-data.
13. Entry of historical data to be allowed? They may often be lacking desired meta-data. But NIWA has good documentation of the NRWQN data and metadata going back 20 years.

14. Drifting nicknames, i.e., (i) using the same nickname over the years even though the laboratory method has changed; (ii) multiple organisations using the same nickname but using different laboratory methods. (The AQUAL manual has an extensive list of suggested attribute codes and nicknames, in Appendix 5.)