BEFORE THE PROPOSED NATURAL RESOURCES PLAN HEARINGS PANEL

IN THE MATTER	of the Resource Management Act 1991
AND	
IN THE MATTER AND	Water quality
IN THE MATTER	of Right of Reply evidence to matters raised during Hearing Stream 4

STATEMENT OF RIGHT OF REPLY EVIDENCE OF ANTONIUS HUGH SNELDER ON BEHALF OF WELLINGTON REGIONAL COUNCIL

TECHNICAL – REGIONAL WATER QUALITY TRENDS

4 May 2018

TABLE OF CONTENTS

1. INTRODUCTION

1.1 My name is Antonius (Ton) Hugh Snelder. I am a director of LWP Ltd and consultant/researcher in the field of water resources managment.

1.2 I was engaged by Wellington Regional Council to provide evidence relating to water quality trends.

1.3 My qualifications and expertise are described in my primary evidence.

1.4 I wrote the technical reports Snelder 2017a and 2017b that concern analyses of trends in river water quality in the Wellington region. In my primary evidence I described evidence of water quality improvement across the region over the past decade that was based on these reports.

1.5 I have also been party to a joint statement of evidence with Dr Adam Canning and Ms Kate McArthur that described caucusing by the three of us concerning their rebuttal of my primary evidence.

1.6 I have also provided supplementary evidence to the panel that provided evidence for water quality improvement at the level of the individual Whaitua (i.e. sub-regions defined for management purposes) that make up the Wellington region.

2. CODE OF CONDUCT

2.1 I confirm that I have read the Code of Conduct for Expert Witnesses contained in the Environment Court Practice Note and that I agree to comply with the code. My evidence in this statement is within my area of expertise. I have not omitted to consider material facts known to me that might alter or detract from the opinions which I express.

3. SCOPE

3.1 This right of reply evidence relates to matters raised by submitters and the Panel since my primary evidence was released and covers the following:

(a) Clarification of changes to statistical analyses presented in my primary evidence but not included in my technical reports.

(b) Further analysis of the representativeness of GWRC's river water quality monitoring network to address comments made by Doctor Canning when he gave his primary evidence to the panel.

(c) Quantification of statistical uncertainty associated with regional water quality improvement to address comments made by Doctor Canning when he gave his primary evidence to the panel.

4. METHODOLOGY

4.1 All my evidence, including that presented in this Right of Reply, is based on a regional scale assessment of trends made by Snelder (2017a). Snelder (2017a) aggregated the results of a separate study of trends performed by Snelder (2017b) on river water quality monitoring data collected at 61 state of environment (SoE) sites in the Wellington region over the last decade (see Table 1 of my primary evidence). The regional scale assessment (Snelder 2017a) included trends in up to 18 water quality variables at each site, and for two time-periods of five and ten years; both ending at the end of 2016.

5. CLARIFICATION OF CHANGES TO STATISTICAL ANALYSES PRESENTED IN MY PRIMARY EVIDENCE

5.1 In my primary evidence I presented results of binomial tests that considered whether the numbers of improving trends that were observed across the region over the 5 and 10 year time-periods were greater than if increasing and decreasing trends were equally likely. I interpreted a statistically significant test as strong evidence of regional improvement in the water quality variables under consideration.

5.2 In his primary evidence (paragraph 59), Dr Canning pointed out that when multiple hypotheses are tested, the probability of getting a "false positive" or "false negative" increases. In expert conferencing we agreed that this issue would be addressed by correcting the binomial test statistic p-values for false discovery. I corrected all binomial test p-values for false discovery and presented these to the panel in the expert conferencing joint witness statement on the topic of water quality trends (McArthur *et al.,* 2018).

5.3 The result of correcting the tests for false discovery was that all designations of “regional trends” remained the same, except nitrate-nitrite-nitrogen (NNN) and total nitrogen over the five-year period changed from “improving” to “not significant.” No regional trends were designated as “degrading”.

5.4 Dr Canning raised this issue of false discovery when presenting his evidence. It is important to be clear that the term “false discovery” is a statistical term and that Dr Canning’s concern had been addressed in our joint witness statement.

5.5 In the joint witness statement, the conclusion was made that the additional analysis increased the robustness of the findings. The experts agreed that there is no evidence of *region-wide* degradation over the ten-year or five-year time periods.

5.6 It is also important to note that after correcting for false discovery, the significant regional trends that I noted in my primary evidence, except for nitrate-nitrite-nitrogen (NNN) and total nitrogen over the five-year period, remained significant. Therefore, the basis for my overall conclusion; that there has been a dominance of improving river water quality across the region over the past decade, was largely unaffected by correction for false discovery.

5.7 In his evidence (paragraph 58), Dr Canning also raised the issue of “pseudo replication” of monitoring sites because the monitoring network has catchments with multiple sites. Dr Canning suggested that sites located in the same catchment are influenced by the same conditions and a component of the water at downstream sites is measured at the upstream sites and that this raises the possibility that my assessment was unduly influenced by pseudo replication.

5.8 As part of expert conferencing, I conducted a sensitivity analysis to assess the degree to which my original results may have been influenced by pseudo replication. The sensitivity analysis indicated that controlling for pseudo replication does not change my original conclusions (Snelder, 2017a).

5.9 The sensitivity analysis was detailed in the joint witness statement along with our conclusion that there is no evidence of *region-wide* degradation over the ten-year or five-year time periods (McArthur et al., 2018).

6. REPRESENTATIVENESS OF GWRC'S RIVER WATER QUALITY MONITORING NETWORK

6.1 In our joint witness statement (paragraph 12), we agreed that the sites in GWRC's monitoring network may not be adequately representative of the Wellington Region. Because my analysis was based on these sites, there is the possibility that it's results provide a misleading representation of regionwide trends. No further analysis of this issue was made as part of expert conferencing and Dr Canning raised this issue when he presented his primary evidence.

6.2 In this Right of Reply, I present results of additional analysis of the representativeness of GWRC's river water quality monitoring network. The approach I took to this has been taken in other studies (Larned *et al.*, 2016; Larned and Unwin, 2012). The character of rivers in the Wellington region was described using six environmental variables that are strongly associated with variation in water quality (Table 1). These environmental variables are available for each segment of a digital river network that represents all of New Zealand. The rivers of the Wellington region are represented by 18,000 segments (Snelder 2017b). Each site in the monitoring network was associated with the same six environmental variables corresponding to the segment on which the site was located. The representativeness of the network was evaluated by comparing how closely the distribution of environmental variables describing the sites in the monitoring network is matched by the distribution of the same variables across all segments in the region. A close match between the two distributions for a specific variable indicates that the site network is representative of the regional variation for the environmental variable under consideration (i.e. that the site network is not biased to rivers with particular environmental characteristics).

Table 1. Environmental variables used to describe the character of rivers of the Wellington region. Each monitoring site was associated with the same variables based on values pertaining to the river network segment on which it was located.

Environmental variable name	Characteristic	Relevance
segAveElev	Elevation (m ASL)	High elevation sites generally have higher water quality than lower sites.
usArea	Log (base 10) transformed catchment area (km^2)	Depending on catchment land cover, water quality varies according to size of upstream catchment.
usCatElev	Mean elevation of upstream catchment (m ASL)	High elevation catchments generally have better water quality than low elevation catchments.
usUrban	Proportion of upstream catchment occupied by urban land cover types (proportion)	Catchments with a high proportion of urban land generally have poor water quality.
usPastoral	Proportion of upstream catchment occupied by pastoral land cover types (proportion)	Catchments with a high proportion of agricultural land generally have poor water quality.
usIFS	Proportion of upstream catchment occupied by indigenous forest and scrub land cover types (proportion)	Catchments with a high proportion of indigenous forest and scrub generally have good water quality.

6.3 The comparison of the distribution of the six environmental variables for sites in the monitoring network and the Wellington river network itself are shown in Figure 1. Each panel represents an environmental variable. On each panel, the blue bars represent the sites and the red bars indicate the entire regional river network. Adjacent blue and red bars with the same values (indicated on the vertical axis) indicate that the monitoring network has sites in close to the same proportion as the entire network for the indicated values of that environmental variable. For example, approximately 40% of both network segments and monitoring sites are associated with low values (0 – 0.1) of usPastoral (Figure 1). Furthermore, when all pairs of blue and red bars have similar values (indicated on the vertical axis); this indicates that the monitoring network is representative of the regional variation of that environmental variable.

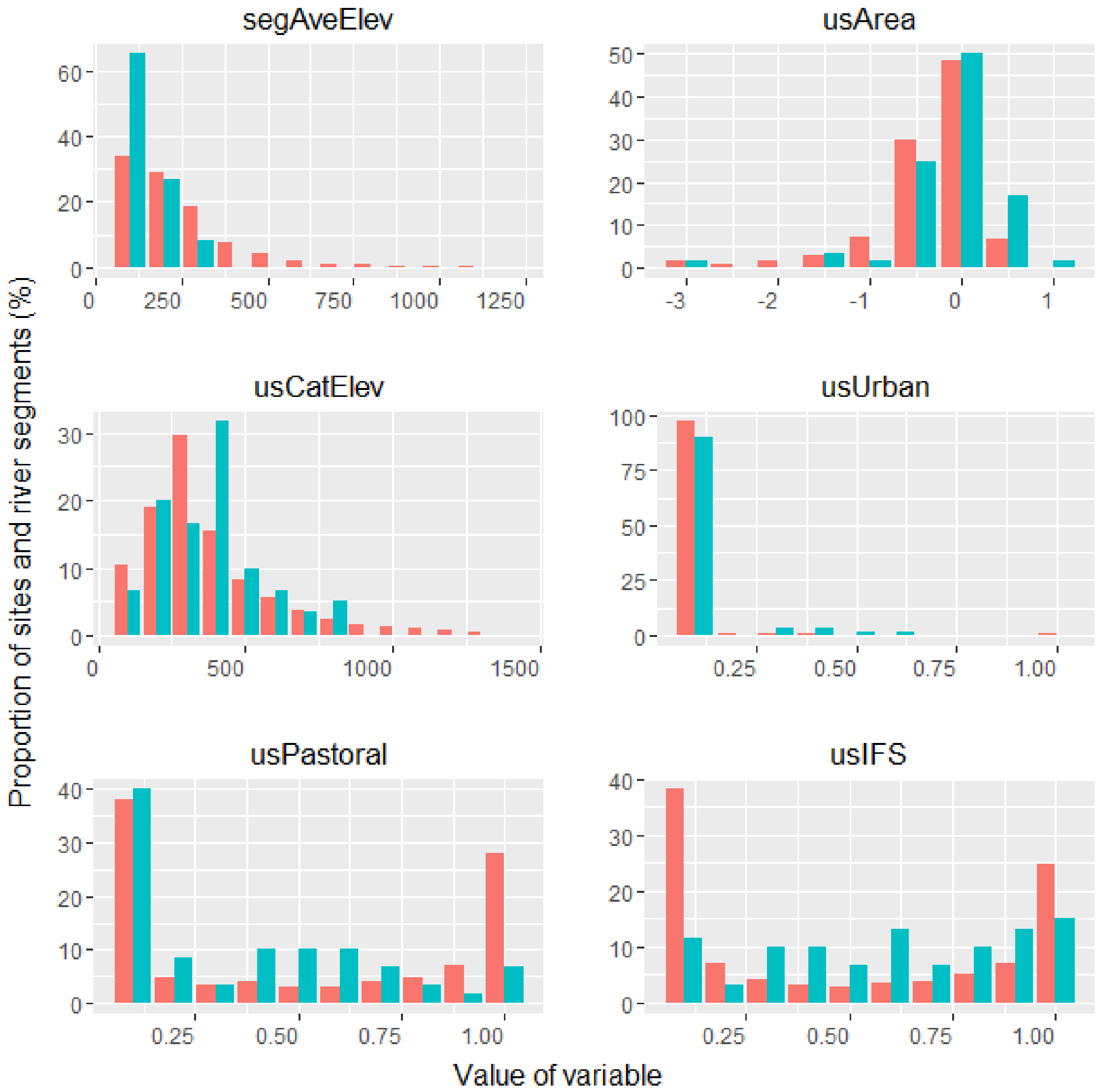

Figure 1. Comparison of the distribution of the six environmental variables for segments of the river network in the Wellington (red bars) and the monitoring site network (blue bars).

6.4 Most paired bars (red and blue) shown in Figure 1 have reasonably closely matched values on the y-axes. This indicates a reasonable distribution of monitoring sites over the range of each of the environmental variables represented by the entire river network. There are some exceptions to this, which indicate a degree of bias in the monitoring sites' representation of regional variation for some environmental variables. For example, Figure 1 indicates that there is a lower proportion of monitoring sites that are associated with very high values of usPastoral (blue bars) than network segments (red bars). Conversely, there are a lower proportion of monitoring sites that are associated with very low values of usIFS (blue bars) than network segments (red bars). There is also over-representaion (i.e., bias) of monitoring sites with low elevation locations (segElev) and with low to mid-elevation catchments (usCatElev). The monitoring sites also over-represent catchments with very high urban land cover (usUrban).

6.5 In his evidence (Paragraph 58), Dr Canning expressed the concern that many monitoring sites are located in headwaters and therefore have excellent water quality. My analysis (Figure 1) shows that this is not the case and in fact the monitoring network is biased to lower elevation locations, and therefore, sites that are more likely to be impacted.

6.6 It is not possible to define a monitoring network that is perfectly representative of the region's rivers for several reasons including the practicalities of sample collection and the need for monitoring to provide for a variety of types of information. In my opinion, and based on the above analysis, I consider the existing network is reasonably representative of the range of river environments in the Wellington region. Therefore, I conclude that my analyses provide a reasonably representative picture of the water quality changes that have occurred in the Wellington region over the past decade.

7. ASSESSMENT OF UNCERTAINTY OF PROPORTIONS OF IMPROVING TRENDS ACROSS A REGION

7.1 In Snelder (2017a), and in my statement of evidence, I evaluated 'overall water quality trends' across the Wellington region by aggregating individual site trends assessed for 18 water quality variables across up to 61 sites. The evaluation was based on the proportion of sites with improving trends and indicated that there has been a dominance of improving trends (i.e., >50% of sites) at the regional level over the past decade. In Snelder (2017a) and my statement of primary evidence I concluded that these results are strong evidence of overall water quality improvement at the regional level over the past decade.

7.2 In the assessments, site trends were aggregated and summarised using various statistics (e.g., the proportion of sites with improving trends and regional trends). These statistics were derived using all trends, irrespective of the confidence in the trend direction assessed for each site and variable combination. In Snelder (2017a) and my statement of primary evidence, I considered that this was justifiable because over many sites, incorrect classifications of direction will cancel out (i.e., as many sites will be misclassified as increasing as sites misclassified as decreasing). However, the inclusion of ‘uncertain’ site trends (i.e., those with low confidence in the evaluated trend direction) in assessments in this manner implies an uncertainty in the derived statistics. For example, if the proportion of improving trends is the statistic being estimated, the estimate has an uncertainty that is associated with the use of the uncertain site trends.

7.3 In his evidence, Dr Canning said that the inclusion of ‘uncertain’ site trends in the evaluation meant that there was low confidence in my conclusions. It is noted that Dr Canning’s opinion was not based on an analysis of uncertainty.

7.4 I have carried out additional analysis to respond to Dr Canning’s comment that there can be only low confidence in my conclusions. The analysis evaluated the statistical accuracy of the regional assessment of the proportion of improving trends for each water quality variable. The evaluation is based on a Monte Carlo analysis, which is a commonly used method for estimating accuracy of models and statistical assessments. The statistical accuracy of the regional assessment of the proportion of improving trends provides an objective measure of the degree of confidence in the statistics on which I base my conclusions. A detailed explanation of the Monte Carlo analysis is appended to this evidence.

7.5 The Monte Carlo analysis indicated that 62% of all combinations of sites and variables had improving trends over the 10-year periods. This is consistent with my original assessment that 63% of site-variable trends indicated improvement (Snelder, 2017a). Therefore, my new analysis does not change the original assessment, which was the basis for the conclusion I reached in my statement of primary evidence; that there is strong evidence of overall water quality improvement at the regional level over the past decade.

7.6 The Monte Carlo analysis indicated that the 95% confidence intervals for the proportion of all trends that were improving is the range 52% to 71%. This indicates that there is high confidence that the majority of site and variable trends are improving over the 10-year period.

7.7 The best estimate of the proportion of improving sites for each variable is provided by the third column of Table 2 and the 95% confidence interval is shown in the fourth column (both assessed from the Monte Carlo analysis). The 95% confidence intervals range from 12% to 24%, depending on the water quality variable (Table 2). For example, 89% of sites are estimated to have improving clarity and the 95% confidence interval ranged from 83% to 95% (Table 2). This means we can be 95% confident that between 83% and 95% of sites have improving clarity trends.

7.8 The estimated proportion of sites with improving trends was less than 50% for only three variables (Fils-Mean, MCI, QMCI; Table 2). However, it is notable that for these three variables, the upper confidence intervals were greater than 50% (Table 2). For example, MCI is estimated to be improving at 48% of sites and the 95% confidence interval ranged from 38% to 58% (Table 2). This means that it is possible, that the majority of sites have improving trends for these three variables. It is also noted that the estimated proportion of improving sites was greater than 50% but the lower 95% confidence interval was below 50% for seven variables; TOC, %EPT_Taxa, Fils-Max, DRP, Mats-Mean, E. coli, %EPT (Table 2). This means we cannot exclude the possibility, at the 95% level of confidence, that the majority of sites have degrading trends for these seven water quality variables.

Table 2. Results of Monte Carlo analysis of the estimated proportion of sites with improving trends and the 95% confidence interval of the estimate. The variables are ordered from top to bottom by the proportion of improving sites.

Water quality variable	Number of sites	Proportion of improving sites (%)	95% confidence interval (%)
Clar	52	89	83 - 95
TN	40	84	76 - 92
TP	45	77	69 - 85
NNN	55	71	63 - 79
NO3-N	50	70	62 - 78
Turb	56	64	54 - 74
Chla	47	64	52 - 76
Mats-Max	45	62	50 - 74

TOC	51	59	49 - 69
%EPT_Taxa	54	59	49 - 69
Fils-Max	45	58	46 - 70
DRP	35	55	45 - 65
Mats-Mean	45	55	43 - 67
E. coli	55	52	44 - 60
%EPT	54	52	42 - 62
Fils-Mean	45	48	36 - 60
MCI	54	48	38 - 58
QMCI	54	44	34 - 54

7.9 The evaluation shows that there is a high level of confidence that a majority of sites have improving trends over the past decade for most variables. The conclusion that can be drawn from these results are consistent with those presented in Snelder (2017a) and in my statement of primary evidence; that there is strong evidence of overall water quality improvement at the regional level over the past decade.

7.10 It is noted that 'overall water quality' here is based on aggregating across up to 61 sites and 18 water quality variables. As such, the results are a general picture and all sites and water quality variables have been given equal weight. Particular sites may be more 'significant' (in the sense of being 'important' rather than statistically significant) than others and some water quality measures may also be considered more important than others. This analysis was based on aggregating data across variables and sites without making any judgement about relative significance or importance of sites and variables.

8. CONCLUSIONS

8.1 In my Right of Reply I have clarified that additional analyses were conducted to assess the degree to which my original results may have been influenced by two statistical issues; false discovery and pseudo replication. These analyses indicate that neither issue changes my original conclusions that there is strong evidence of regional improvement in the water quality variables under consideration. It is noted that in the joint witness statement, the conclusion was made that the additional analysis increased the robustness of my original findings. As experts, we agreed that there is no evidence of region-wide degradation over the ten-year or five-year time periods.

8.2 I have also presented additional analyses that show that GWRC's monitoring network is reasonably representative of variation in the character of rivers in the Wellington Region. Because my analyses of water quality trends are based on the sites in this monitoring network, I conclude that my analyses provide a representative picture of the water quality changes that have occurred in the Wellington region over the past decade.

8.3 In his evidence, Dr Canning said that the inclusion of 'uncertain' site trends in the evaluation meant that there was low confidence in the conclusions I presented in my statement of primary evidence. I have carried out additional analysis to respond to Dr Canning's comment and have evaluated the statistical accuracy of my assessment of the proportion of improving trends for each water quality variable. The evaluation shows that there is a high level of confidence that a majority of sites have improving trends over the past decade for most variables. This is consistent with the results and conclusions in Snelder (2017a) and in my statement of primary evidence; that there is strong evidence of overall water quality improvement at the regional level over the past decade.

References

MacArthur, K, Snelder, T., Canning, A., Expert conferencing joint witness statement to the hearing panel. Topic: water quality trends. 19 February 2018

Snelder, T., 2017a. Analysis of regional-scale river water quality trends in the Wellington region; Period 2007 to 2016. LWP Client Report: 2017-07, Christchurch.

Snelder, T., 2017b. Analysis of Water Quality Trends for Rivers and Lakes of the Wellington Region. LWP Client Report: 2017-01, Christchurch.

Larned, S., T. Snelder, and M. Unwin, 2016. Water Quality in New Zealand Rivers; Modelled Water Quality State. NIWA CLIENT REPORT, NIWA, Christchurch, New Zealand.

Larned, S.T. and M. Unwin, 2012. Representativeness and Statistical Power of the New Zealand River Monitoring Network. NIWA Client Report, NIWA, Christchurch, New Zealand.

Appendix 1: Details of the Monte Carlo analysis

Introduction

1. Snelder (2017a) assessed ‘overall water quality trends’ across the Wellington region by aggregating individual trends assessed for 18 water quality variables across up to 61 sites. The individual site trends were analysed by Snelder (2017b) using the method of Larned *et al.* (2016). The analysis of individual trends provided two pieces of information: the estimated direction (which was classified as improving or degrading); and the probability that the trend’s true direction was decreasing (the ‘probability’). This probability reflects the degree of confidence in the classification of the trend direction given the data. Traditionally, when considering a specific trend, a high degree of confidence (i.e. 95%; or correspondingly a low misclassification error risk (5%)) is required to declare that the direction can be inferred.

2. Snelder (2017a) aggregated the individual site trends by variable to assess overall regional trends in two ways: the proportion of sites with improving trends at various levels of confidence, and an assessment of ‘regional-trends’ based on the binomial test. In both assessments, site trends were aggregated, irrespective of the confidence in the classification of trend direction. This is a common approach (e.g., Scarsbrook et al., 2003). The assumption is justifiable because over many sites, incorrect classifications of direction will cancel each other out (i.e., as many sites will be misclassified as increasing as sites misclassified as decreasing). However, the inclusion of ‘uncertain’ site trends (i.e., those with a misclassification error risk of greater than 5%) in assessments in this manner implies an uncertainty in the derived statistic. For example, if the proportion of improving trends is the statistic being derived, the estimated proportion has an uncertainty that is associated with the use of the uncertain site trends.

3. The analysis presented here quantifies the proportion of improving trends for each water quality variable based on aggregating all site trends. The uncertainty of this assessment is quantified by performing a Monte Carlo analysis that is informed by the directions and probabilities of the individual site trends.

Data

4. The present analysis was performed using results of analyses of site trends for 18 water quality variables for the 10-year period ending at the end of 2016. Trends were used only if the input data complied with the inclusion rules used by Snelder (2017a). The inclusion rules restricted site and variable combinations to those for which 90% of the years had had observations for at least 90% of the sampling occasions. The evaluated trends were not flow adjusted (i.e., they were raw trends) because this allowed the largest number of sites to be included in the analysis and because Snelder (2017a) showed conclusions about regional trends were similar for flow adjusted sites.

5. Snelder’s (2017a) results indicate that the proportion of sites across the Wellington Region with improving trends for the 10 year period finishing at the end of 2016 exceeded 50% for all variables except Mean-Fils, MCI and QMCI (Table 3). This evaluation was based on counting trends based on their face values (i.e., ignoring their uncertainty) and does not include an assessment of the uncertainty of the proportion of improving sites. Hereafter, the Snelder (2017a) assessment is referred to as the ‘count-based evaluation’ of the proportion of improving trends.

Table 3. Results from count-based assessment of raw ten-year time-period site trends reported by Snelder (2017a). The number of sites with decreasing and increasing ten-year trends and the proportion of improving trends (%) evaluated from the face value of the trend direction by water quality variable. The last two columns show the cumulative proportion of sites (%) with improving raw trends with at least the level of confidence indicated. These values were assessed from the probabilities that the true trend represents improvement that were obtained from the individual trend analyses.

Variable	No. sites	No. decreasing	No. increasing	Proportion improving	Likely	Likely as not
Clar	52	5	47	90	88	90
Turb	56	35	21	62	55	66
DRP	35	20	14	57	49	66
TP	45	36	10	80	71	87
NO_3-N	50	35	14	70	70	72
NNN	55	40	14	73	71	75
TN	40	36	5	90	82	92
TOC	51	33	19	65	45	75
E. coli	55	29	26	53	47	58
Chla	47	32	15	68	55	68
Fils-Max	45	28	17	62	40	76
Fils-Mean	45	19	26	42	29	53
Mats-Max	45	27	17	60	44	80
Mats-Mean	45	24	21	53	31	73
%EPT	54	24	30	56	41	57
%EPT_Taxa	54	20	33	61	50	63
MCI	54	26	27	50	35	52
QMCI	54	33	20	37	22	39

Method

6. For any given site and variable combination (S_i,V_j), the original trend analysis provided the probability the true trend (i.e., the trend in the population from which the samples were drawn) was decreasing ($p_{i,j}$). The probability the true trend was decreasing was converted to the probability the trend was improving ($p_{i,j}'$) by considering whether decreasing represents improvement or degradation, recognising that this varies between commonly used indicators of water quality. This result was used to parameterise a binomial distribution representing the trend direction such that success (trend direction was

improving) has probability $p_{i,j}'$ and failure (trend direction was degrading) has probability $1 - p_{i,j}'$. This description of the trend for each site and variable was used in a 'probabilistic evaluation' of the proportion of improving trends based on a Monte Carlo simulation. The Monte Carlo simulation proceeded in the following steps:

1. For a given site (i) and variable (j) (S_iV_j), a classification of trend ($T_{i,j}$) is randomly sampled from the binomial distribution with a probability of success ($p_{i,j}'$).
2. Step 1 is repeated for all 1:n sites, to produce a binary distribution of trends across all sites ($T_{1:n,j}$), which is one "realisation" of the Monte Carlo analysis.
3. The proportion of improving trends (PIT_j) is calculated as the number of improving trends for the realisation divided by the total number of sites: $PIT_j = \Sigma(T_{1:n,j} = \text{"Improving"})/n$.
4. Steps 1:3 are repeated for 1000 realisations.
5. The estimated proportion of improving trends for variable (j) is calculated as the mean PIT_j across all 1000 realisations.
6. The standard error of the proportion of improving trends is quantified by the standard deviation of PIT_j across all 1000 realisations. The uncertainty of the probabilistic evaluation is expressed by the 95% confidence interval, which is evaluated from the standard error.
7. Steps 1:6 are repeated for all 1:m variables.

Results

7. The standard deviation of the estimated proportion of sites with improving trends varied from 3-6% across the individual variables (Table 4). As expected, the smaller standard deviations (e.g., 3-4%) were associated with the variables for which there were more sites and for which a large proportion of trends had high probabilities (of trends representing improvement, e.g., Clar, Turb, TP, NO3-N, NNN, TN; see Table 3). The larger standard deviations (e.g., 5-6%) were associated with the variables for which there were fewer sites and for which a large proportion of the trends had assessed probability close to 50% or "as likely as not" (e.g., Chla, Fils-Max, Fils-Mean and Mats-Max; see Table 3).

8. The Monte Carlo analysis produced similar estimates of the proportions of improving trends to the count-based evaluation of the proportion of improving trends (Figure 2). The 95% confidence interval derived from the Monte Carlo analysis always included the one to one line of the plot comparing the two sets of evaluations. This demonstrates that the count-based evaluation of the proportion of improving trends is consistent with the probabilistic evaluation.

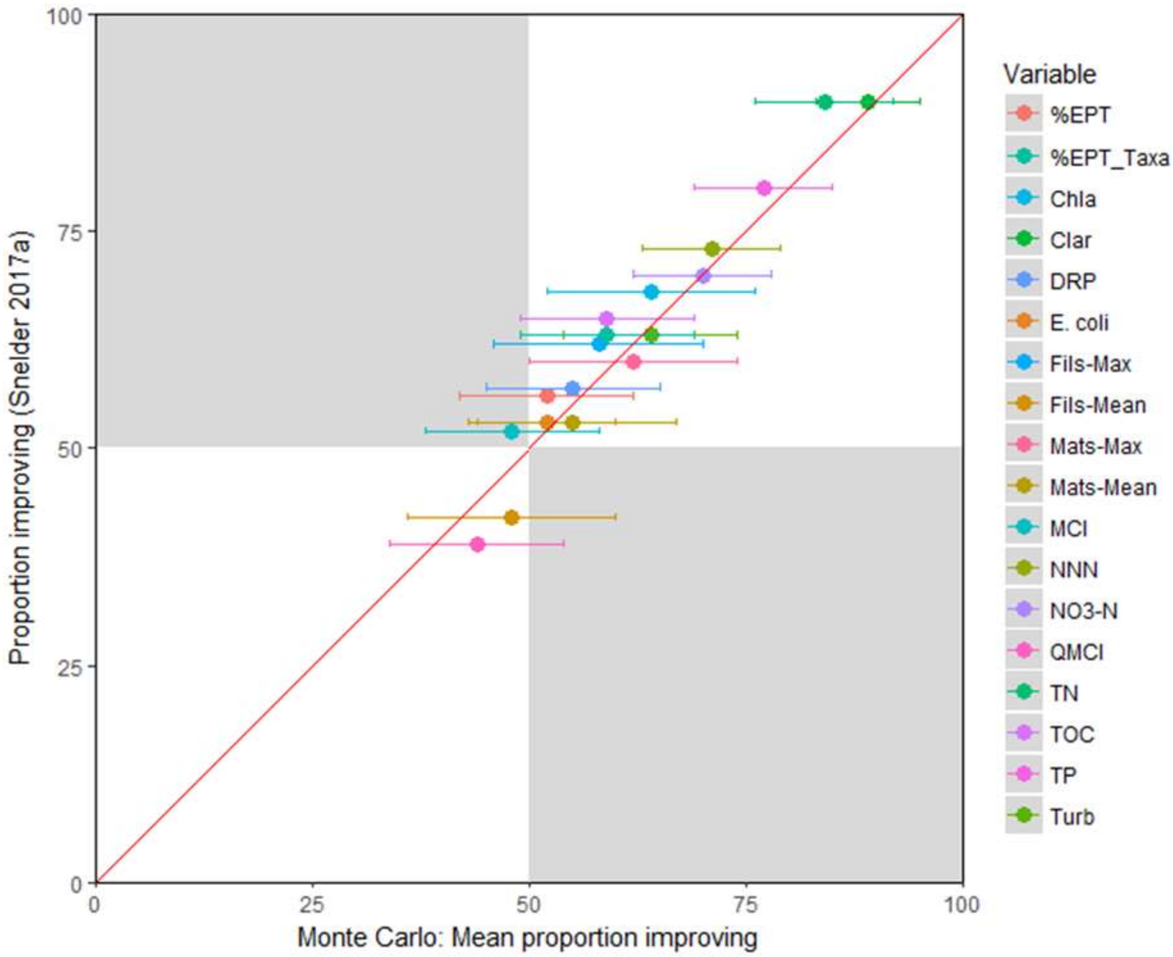

Figure 2: Comparisons of proportions of sites improving between the Monte Carlo average and the count-based evaluation. Solid dots are the evaluated proportions of improving sites using both methods. Error bars indicate the 95% confidence interval assessed using the Monte Carlo simulations. Grey areas indicate areas of disagreement in the majority trend direction between the two evaluations.

9. The results of the simulation indicate that on average over all 18 variables, 62% of site trends are improving. This is consistent with the count-based evaluation reported by Snelder (2107a), for which 63% of sites were improving (Table 3). The two evaluations resulted in disagreement in the majority trend direction for only one variable; MCI (Figure 2). That is, the count-based assessment concluded that 50% of sites had improving MCI trends (Table 3), whereas the probabilistic assessment concluded 48% (a minority) of sites had improving trends (Table 4). However, the uncertainty bounds for the probabilistic

evaluation included the count-based evaluation. This indicates that the count based evaluation is consistent with the probabilistic evaluation.

10. The analysis indicates that the 95% confidence intervals are small relative to the estimated proportion of sites with improving trends. For example, 89% of sites are estimated to have improving Clarity and the 95% confidence interval ranged from 83% to 95% (Table 4). Nitrate (NO_3-N) was evaluated to be improving at 70% of sites and the 95% confidence interval ranged from 62% to 78%. It is notable that all variables have upper confidence intervals for the proportion of sites with improving trends that are greater than 50%. For example, MCI is improving at 48% of sites and the 95% confidence interval ranged from 38% to 58% (Table 4). For eight variables, the lower confidence interval is >50%, which indicates that, for these variables, that there is very high confidence that the majority (>50%) of sites are improving.

Table 4. Results of Monte Carlo simulation. The variables are ordered from top to bottom by the proportion of improving sites.

Variable	No. sites	Proportion of improving sites (%)	Standard deviation of proportion Improving (%)	95% confidence interval (%)
Clar	52	89	3	83 - 95
TN	40	84	4	76 - 92
TP	45	77	4	69 - 85
NNN	55	71	4	63 - 79
NO3-N	50	70	4	62 - 78
Turb	56	64	5	54 - 74
Chla	47	64	6	52 - 76
Mats-Max	45	62	6	50 - 74
TOC	51	59	5	49 - 69
%EPT_Taxa	54	59	5	49 - 69
Fils-Max	45	58	6	46 - 70
DRP	35	55	5	45 - 65
Mats-Mean	45	55	6	43 - 67
E. coli	55	52	4	44 - 60
%EPT	54	52	5	42 - 62
Fils-Mean	45	48	6	36 - 60
MCI	54	48	5	38 - 58
QMCI	54	44	5	34 - 54