Comparison of MCI – nutrient relationship analyses of Canning and Snelder

23 June 2019

Ton Snelder, Adam Canning

Introduction

Canning and Snelder made separate analyses of the relationship of DIN and DRP with MCI at the national scale. Both workers regressed MCI observed at regional council monitoring sites against DIN and DRP concentrations measured over the 5-year period ending 2017 (dataset differences in Table 1). In both analyses, the MCI data were obtained from observations at regional council monitoring sites. In Snelder's dataset, only 426 sites at which both water quality (i.e., DIN and DRP) and invertebrate assemblages had been measured over the 5-year period ending 2017 were used. Canning used a dataset of 1852 sites that included many sites where there was no water quality monitoring, thus modelled nutrient concentrations were used. Both workers used the regression line to calculate the median concentrations of DIN and DRP at which MCI had values of 80, 100, 120.

Table 1. *A comparison of datasets used by Snelder and Canning in their initial investigations*

Component	Snelder dataset	Canning dataset
MCI scoring	Environment Aotearoa dataset – calculated by individual regional councils	Cawthron dataset – calculation nationally standardised
MCI central tendency	Five-year median	Five-year average
Number of sites	426	1852
Nutrient data	Observed median of monthly samples for 5 years ending 2017	Modelled average annual median (for 5 years ending 2017) (Whitehead, 2018).

Differences in the datasets used by Canning and Snelder (Table 1) led to differing results that the STAG has asked to be explained. Snelder used MCI scores that were supplied by each regional council and used the median of the observed monthly nitrate nitrogen and DRP to represent DIN and DRP. Snelder accessed the MCI data from the national dataset that was assembled for Environment Aotearoa 2019 (EA 2019) as described by Larned *et al.,* (2018). Snelder originally reported an analysis to STAG at their meeting on 16 April 2019. Subsequent to that date, a change was made to the calculation of site median MCI score in the EA 2019 dataset and the publicly available dataset was updated accordingly. The change involved calculating the median MCI score of samples that fell within a single year before calculating the median over the 5-year period. This affected some sites and councils that sample invertebrates on more than one occasion annually and means that all council invertebrate datasets represent a single annual value. The new MCI score data has been used in this analysis and therefore results vary slightly compared to those originally reported to STAG.

Canning used MCI scores that had been calculated by Cawthron from the regional council's raw taxonomic data (Clapcott et al., 2017). These MCI scores were, therefore, consistently calculated compared to the MCI scores used by Snelder, whose calculation were subject to a degree of regional variation. Canning used predicted values of DIN and DRP instead of the observed values. The predicted values were made using random forest models for a dataset that included a larger set of sites than are represented by the MCI data as described by Whitehead (2018).

For the analyses presented here, Cawthron's recalculated MCI scores were merged with the original EA2019 dataset. There were 375 sites of the original 426 sites in the EA2019 data for which there were MCI score that were calculated by Cawthron. This reduction in sites with MCI scores is presumably due to difficulties in obtaining the raw taxonomic data for some sites or differences in site name that could not be reconciled. The merged dataset where there are sites in common is referred to hereafter as the "reduced dataset". A comparison between the MCI scores in the EA 2019 data and the Cawthron recalculated MCI scores is shown in Figure 1 and indicates that there are some large differences at some sites but that overall agreement is reasonable.

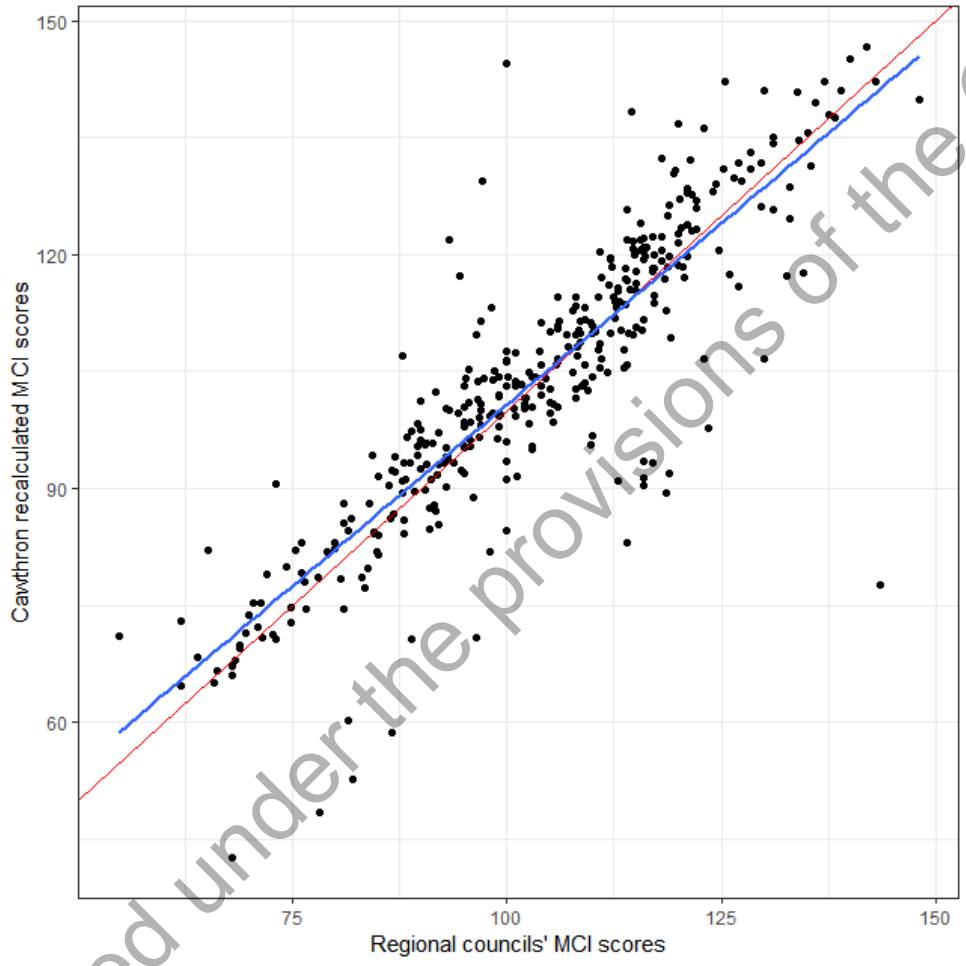

Figure 1. Comparison between the regional councils' MCI scores (in the EA 2019 dataset) and the Cawthron recalculated MCI scores. The red line is one to one and the blue line is a linear regression ($r^2 = 0.76$).

The modelled DIN and DRP used by Canning are compared with the observed on Figure 2 and Figure 4. The same results in log-log space are shown for DIN on Figure 3, which indicates that the model residuals are log normally distributed. This means the absolute errors of the two models are larger for higher values.

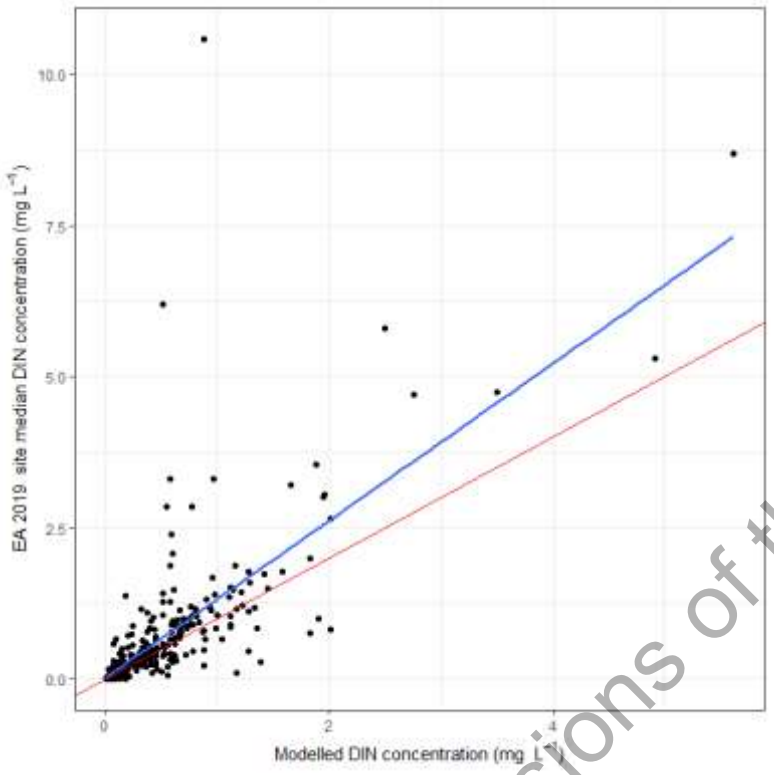

Figure 2. Relationship between the site median DIN concentrations in the EA 2019 data and the modelled DIN in the Canning dataset. The red line is one to one and the blue line is a linear regression ($r^2 = 0.57$). Note in the plot below the same data is shown in log(10) space.

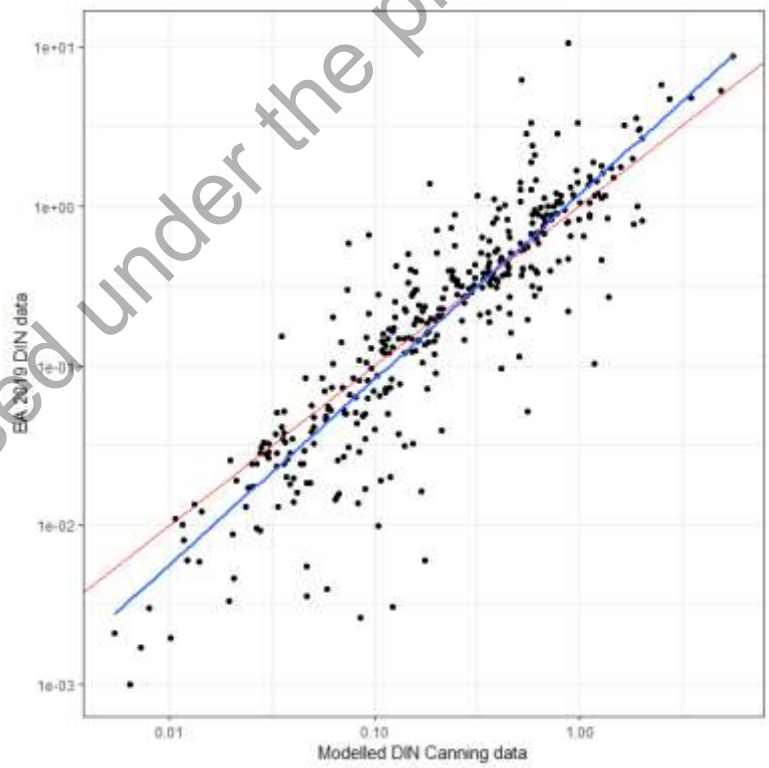

Figure 3. Relationship between the site median DIN concentrations in the EA 2019 data and the modelled DIN in the Canning dataset shown in log (base 10) space. The red line is one to one and the blue line is a linear regression ($r^2 = 0.79$).

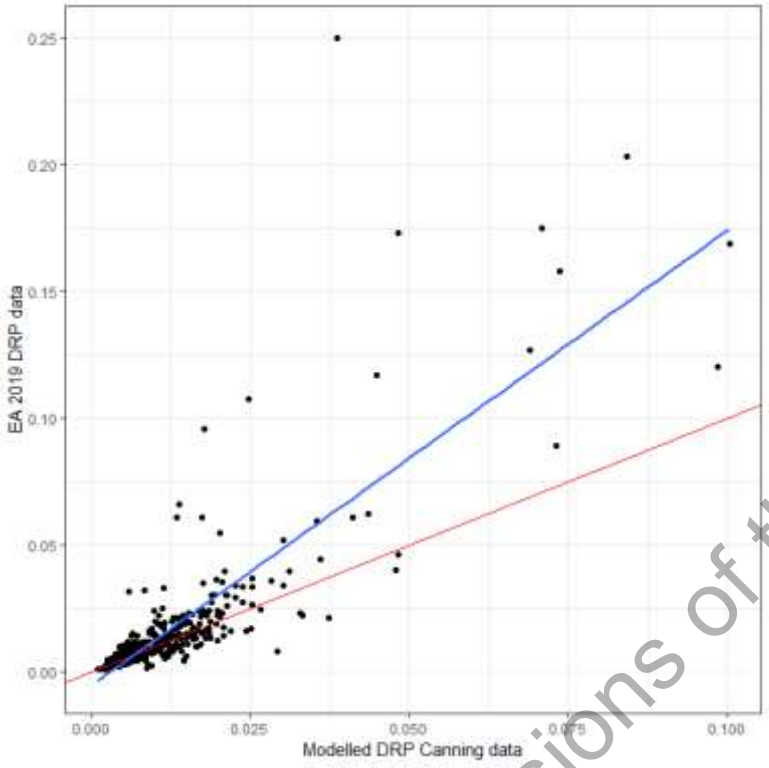

Figure 4. Relationship between the site median DRP concentrations in the EA 2019 data and the modelled DRP in the Canning dataset. The red line is one to one and the blue line is a linear regression ($r^2 = 0.67$). Note that the regression r^2 in log-log space was 0.79).

Results of analyses

Observed concentrations, Regional Councils' MCI scores, Full EA 2019 dataset, post-updates to calculation of MCI median

The relationships between the Regional Councils' MCI scores and site median values of DIN and DRP are shown on Figure 5. Statistics describing these relationships and median concentrations of DIN and DRP at which MCI had values of 90, 100, 120 are shown in Table 2. These are similar to those reported to STAG at their meeting on 16 April 2019 but differ slightly due to the updating of site median MCI scores in the EA 2019 dataset. Note that in this analysis, the median DIN concentration is associated with an MCI score of 90 in this analysis is 4.5 mg L^{-1} whereas the value reported to STAG at their meeting on 16 April 2019 was 5.5 mg L^{-1}.

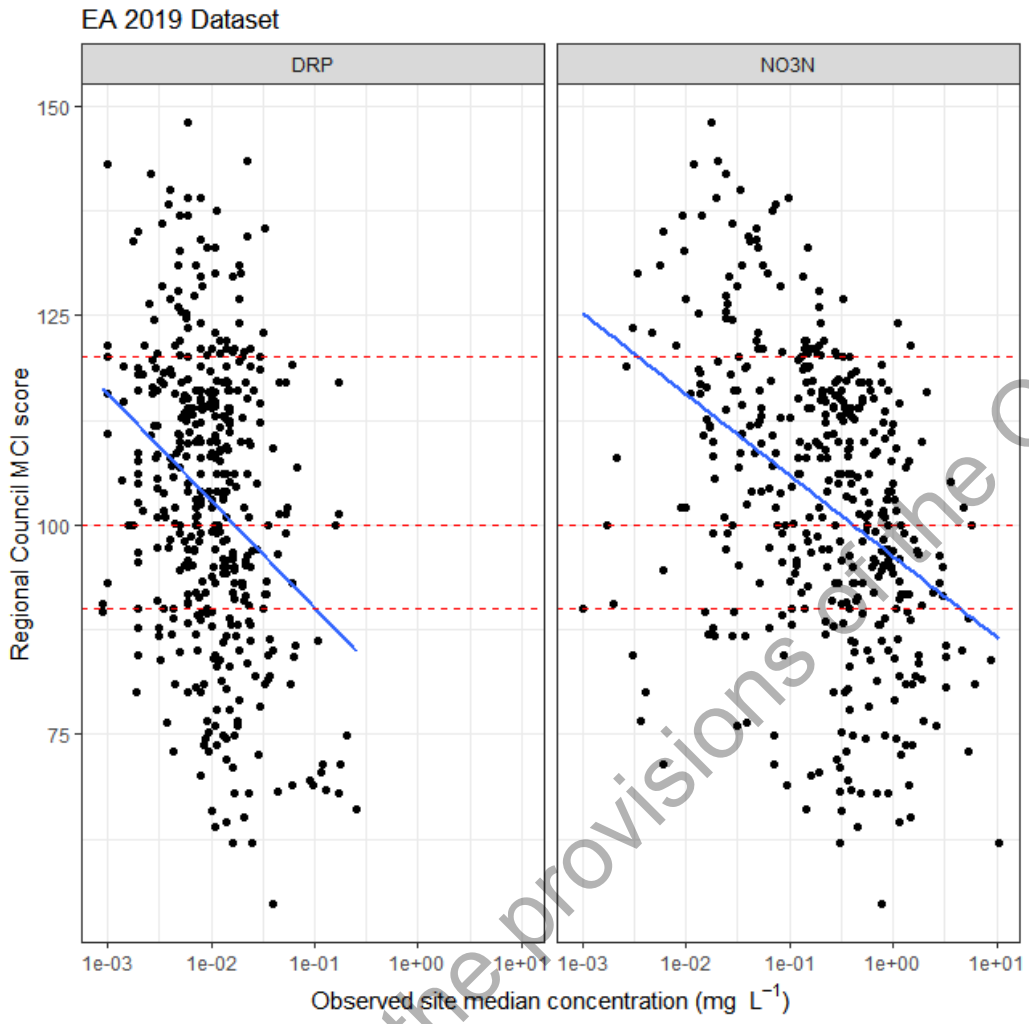

Figure 5. Relationships between the Regional Councils' MCI scores and the observed site median DIN and DRP for the full EA 2019 dataset (with observed nutrients and MCI scores) of 426 sites. The blue lines are linear regressions (see Table 2 for details).

Table 2. Results of linear regression of EA 2019 dataset median MCI scores against the observed site median nutrient concentrations at the national scale. The concentrations of DIN and DRP that were associated with MCI scores of 90, 100 and 120 are shown.

Nutrient	No. samples	Coefficient	r^2	p-value	Concentration at MCI values		
					90	100	120
DIN	426	-9.67	16	<0.0001	4.456	0.412	0.004
DRP	426	-12.85	10	<0.0001	0.101	0.017	0.000

Observed concentrations, Regional Councils' MCI scores, reduced dataset

The relationships of observed site median DIN and DRP concentrations with Regional Councils' MCI scores for the reduced dataset of 375 sites are shown on Figure 6. Statistics describing these relationships and median concentrations of DIN and DRP at which MCI had values of 90, 100, 120 are shown in Table 3. These are similar to those shown in Table 2 and reported to STAG at their meeting on 16 April 2019. Note that the median DIN concentration associated with an MCI score of 90 in this analysis is 5.1 mg L^{-1} whereas the value reported to STAG at their meeting on 16 April 2019 was 5.5 mg L^{-1}.

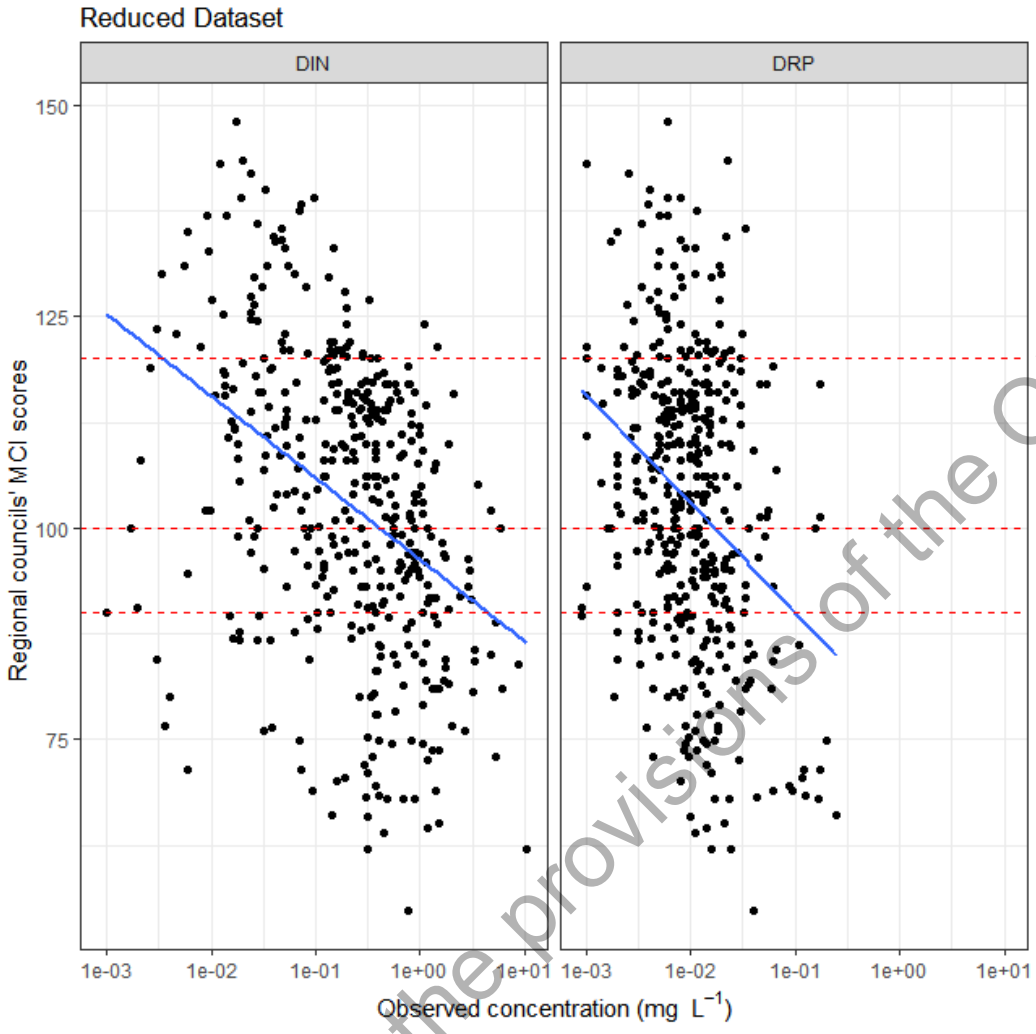

Figure 6. Relationships between the Regional Councils' MCI scores and the observed site median DIN and DRP for the reduced dataset of 375 sites. The blue lines are linear regressions.

Table 3. Results of linear regression of Regional Councils' MCI scores against the observed site median nutrient concentrations at the national scale for the reduced set of 375 common sites used in the analyses of Canning and Snelder. The concentrations of DIN and DRP that were associated with MCI scores of 90, 100 and 120 are shown.

Nutrient	No. samples	Coefficient	r^2	p-value	Concentration at MCI values		
					90	100	120
DIN	375	-9.2	15	<0.0001	5.143	0.421	0.003
DRP	375	-14.26	11	<0.0001	0.079	0.016	0.001

Observed concentrations, Cawthron MCI scores, reduced dataset

The relationships of the observed site median DIN and DRP concentrations with the recalculated Cawthron MCI scores for the reduced dataset of 375 sites are shown on Figure 7. Statistics describing these relationships and median concentrations of DIN and DRP at which MCI had values of 90, 100, 120 are shown in Table 4. These are similar to those shown in Table 2 and Table 3 and those reported to STAG at their meeting on 16 April 2019.

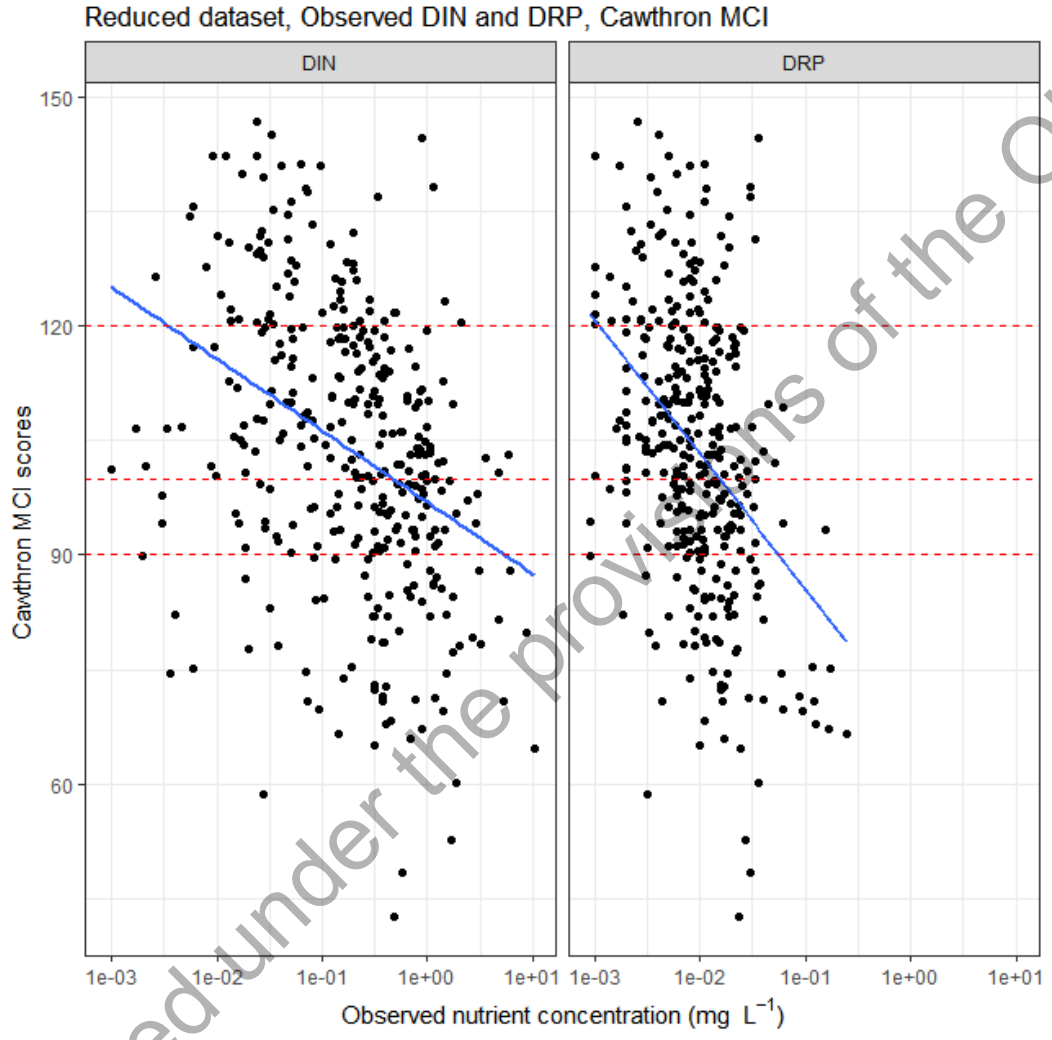

Figure 7. Relationships between the recalculated MCI scores in the observed site median DIN and DRP concentrations for the reduced dataset of 375 sites. The blue lines are linear regressions.

Table 4. Results of linear regression of recalculated Cawthron MCI scores against the observed site median nutrient concentrations at the national scale for the reduced set of 375 common sites used in the analyses of Canning and Snelder. The concentrations of DIN and DRP that were associated with MCI scores of 90, 100 and 120 are shown.

Nutrient	No. samples	Coefficient	r^2	p-value	Concentration at MCI values		
					90	100	120
DIN	375	-9.38	14	<0.0001	5.485	0.471	0.003
DRP	375	-17.55	15	<0.0001	0.056	0.015	0.001

Modelled concentrations, Cawthron MCI scores, full dataset

The relationships of the modelled DIN and DRP concentrations with the recalculated Cawthron MCI scores for the full dataset of 1852 sites are shown on Figure 8. Statistics describing these relationships and median concentrations of DIN and DRP at which MCI had values of 90, 100, 120 are shown in 5. Note that the median DIN concentration associated with an MCI score of 90 in this analysis is 1.0 mg L^{-1}.

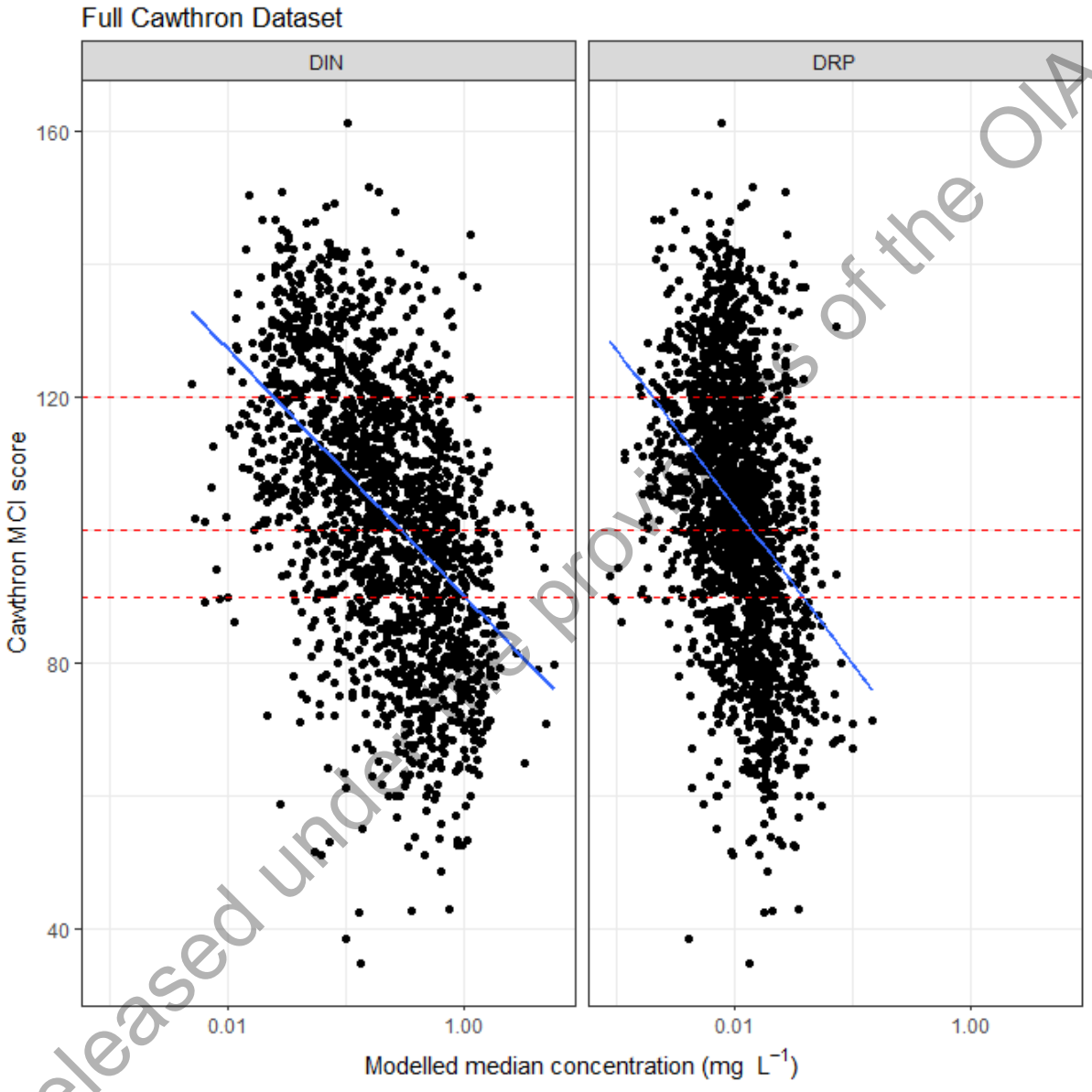

Figure 8. Relationships between the recalculated MCI scores in the modelled DIN and DRP concentrations for the full Cawthron dataset of 1852 sites. The blue lines are linear regressions.

Table 5. Results of linear regression of recalculated Cawthron MCI scores against the modelled nutrient concentrations for the full Cawthron dataset of 1852 sites. The concentrations of DIN and DRP that were associated with MCI scores of 90, 100 and 120 are shown.

Nutrient	No. samples	Coefficient	r^2	p-value	Concentration at MCI values		
					90	100	120
DIN	1852	-18.64	23	<0.0001	1.002	0.292	0.025
DRP	1852	-23.72	13	<0.0001	0.038	0.014	0.002

Modelled concentrations, Regional Councils' MCI scores, reduced dataset

The relationships of the modelled DIN and DRP concentrations with Regional Councils' MCI scores for the reduced dataset of 375 sites are shown on Figure 9. Statistics describing these relationships and median concentrations of DIN and DRP at which MCI had values of 80, 100, 120 are shown in Table 6. Note that the median DIN concentration associated with an MCI score of 90 in this analysis is 1.6 mg L^{-1}.

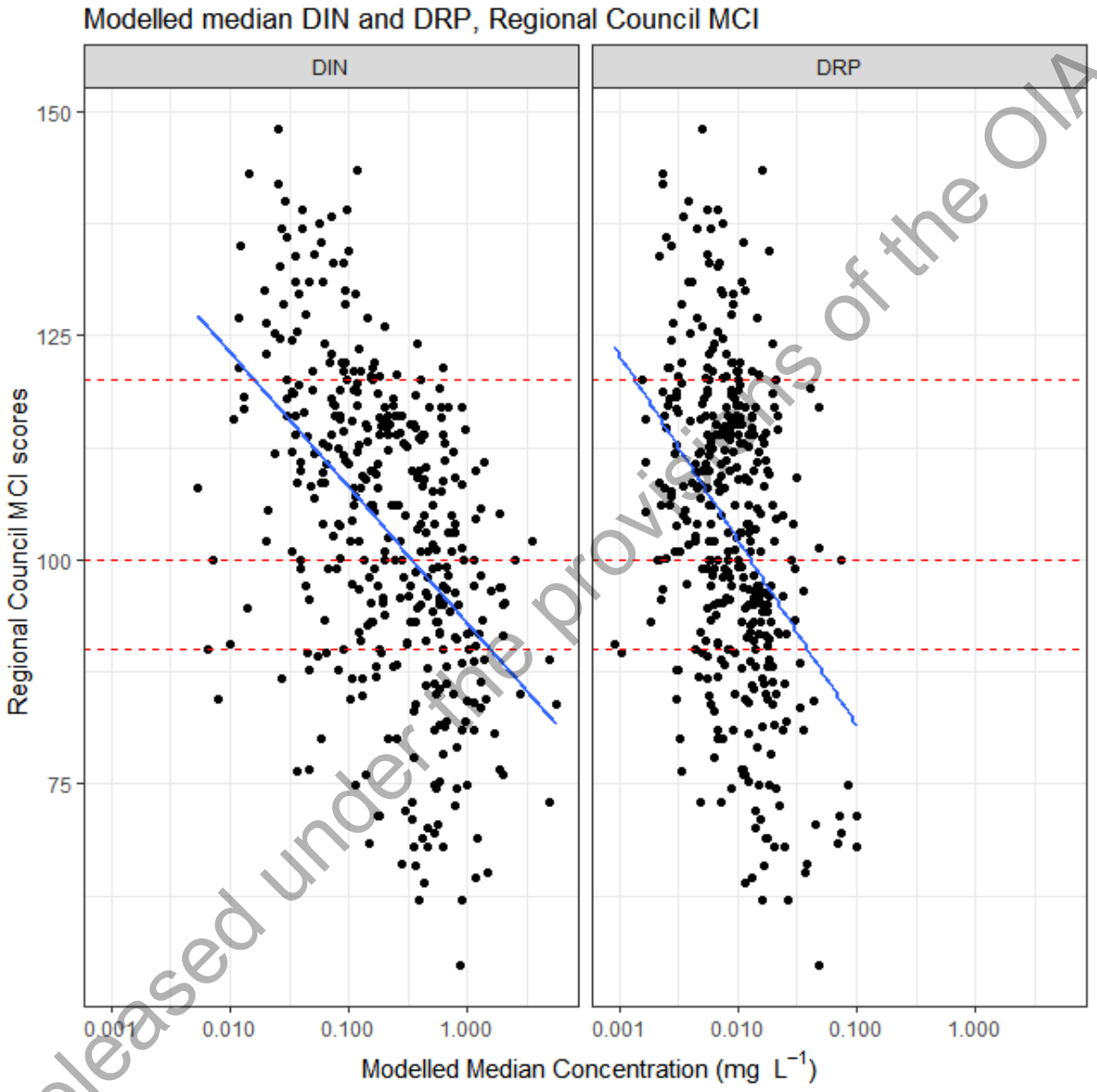

Figure 9. Relationships between the recalculated MCI scores in the modelled DIN and DRP concentrations for the reduced dataset of 375 sites. The blue lines are linear regressions.

Table 6. Results of linear regression of Regional Councils' MCI scores against the modelled nutrient concentrations for the reduced set of 375 common sites used in the analyses of Canning and Snelder. The concentrations of DIN and DRP that were associated with MCI scores of 90, 100 and 120 are shown.

Nutrient	No. samples	Coefficient	r^2	p-value	Concentration at MCI values		
					90	100	120
DIN	375	-14.67	23	<0.0001	1.622	0.338	0.015
DRP	375	-21.71	17	<0.0001	0.035	0.012	0.001

Modelled concentrations, Cawthron MCI scores, reduced dataset

The relationships of the modelled DIN and DRP concentrations with the recalculated Cawthron MCI scores for the reduced dataset of 375 sites are shown on Figure 10. Statistics describing these relationships and median concentrations of DIN and DRP at which MCI had values of 80, 100, 120 are shown in Table 7. Note that the median DIN concentration associated with an MCI score of 90 in this analysis is 1.6 mg L^{-1}.

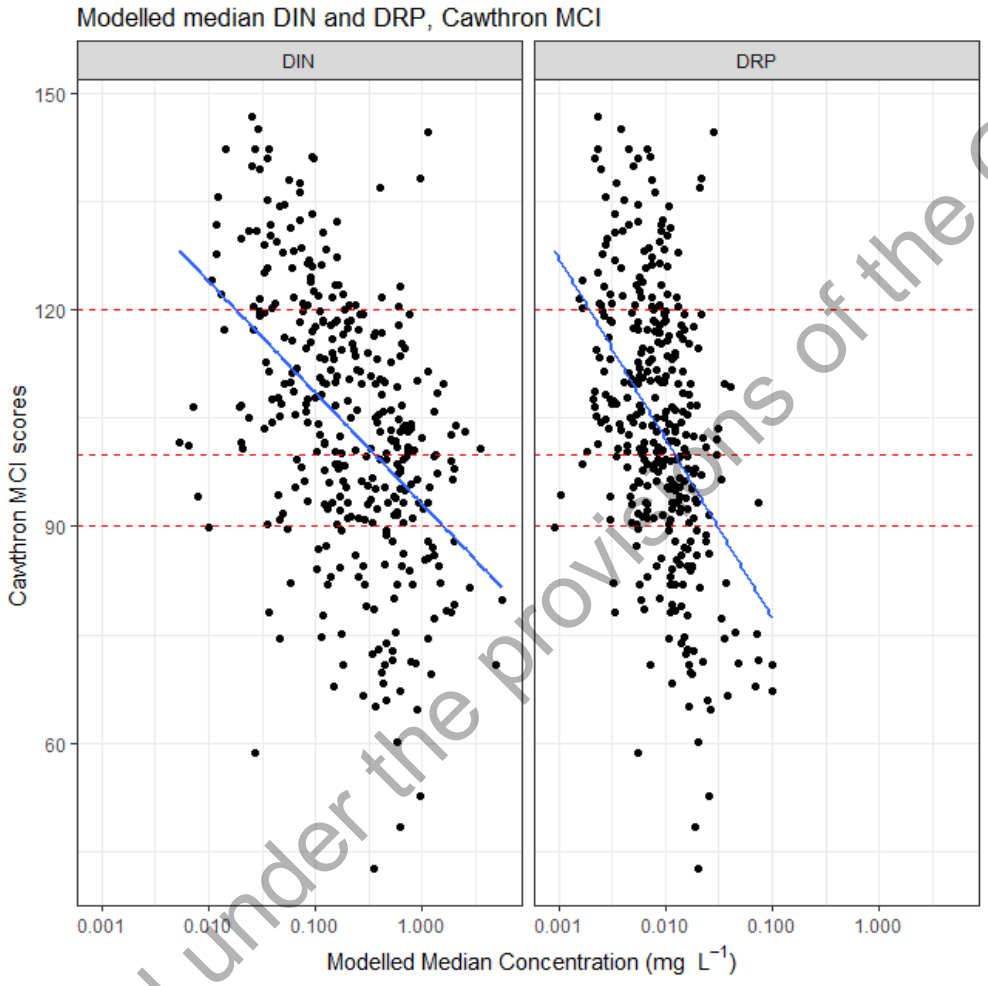

Figure 10. Relationships between the recalculated MCI scores in the modelled DIN and DRP concentrations for the reduced dataset of 375 sites. The blue lines are linear regressions.

Table 7. Results of linear regression of the recalculated Cawthron MCI scores against the modelled site median nutrient concentrations at the national scale for the reduced set of 375 common sites used in the analyses of Canning and Snelder. The concentrations of DIN and DRP that were associated with MCI scores of 90, 100 and 120 are shown.

Nutrient	No. samples	Coefficient	r^2	p-value	Concentration at MCI values		
					90	100	120
DIN	375	-15.45	22	<0.0001	1.576	0.355	0.018
DRP	375	-24.97	19	<0.0001	0.031	0.012	0.002

Summary and conclusions

These results indicate the main reason for the differences between the analyses performed by Canning and Snelder are associated with the choice of nutrient concentration data. Snelder used the

observed values of DIN and DRP at the sampling sites. Canning used modelled concentrations of DIN and DRP. As shown in Figure 3, the residuals of the DIN and DRP models are log normally distributed. This means the absolute errors of the two models are large for higher values. This detail is influential in this analysis because the high values of DIN and DRP anchor the end of the regression model and thus determine the values of DIN and DRP that are associated with an MCI score of 90.

Table 8. Results summary table

Figure	Water quality dataset	MCI dataset	Median N concentration associated with MCI score of 90 (mg L^{-1})
N/A (in previous report to STAG)	Observed median DIN and DRP	Regional Councils' MCI calculation, full EA dataset, pre-updates. Reported to STAG 16 April 2019.	5.5
5	Observed median DIN and DRP	Regional Councils' MCI calculation, full EA 2019 dataset, post-updates to calculation of MCI median	4.5
6	Observed median DIN and DRP	Regional Councils' MCI calculation, reduced dataset	5.1
7	Observed median DIN and DRP	Cawthron – consistent MCI calculation (reduced dataset)	5.5
8	Modelled median DIN and DRP	Cawthron – consistent MCI calculation (full dataset)	1.0
9	Modelled median DIN and DRP	EA2019 – Regional Councils' MCI calculation	1.6
10	Modelled median DIN and DRP	Cawthron – consistent MCI calculation (reduced dataset)	1.6

References

Clapcott J, Wagenhoff A, Neale M, Storey R, Smith B, Death R, Harding J, Matthaei C, Quinn J, Collier K, Atalah J, Goodwin E, Rabel H, Mackman J, Young R 2017. Macroinvertebrate metrics for the National Policy Statement for Freshwater Management. Prepared for the Ministry for the Environment. Cawthron Report No. 3073. 139 p. plus appendices.

Larned, S., A. Whitehead, C. Fraser, T. Snelder, and J. Yang, 2018. Water Quality State and Trends in New Zealand Rivers. Analyses of National-Scale Data Ending in 2017. NIWA, NIWA, Christchurch, New Zealand.

Whitehead, A., 2018. Spatial Modelling of River Water-Quality State. Incorporating Monitoring Data from 2013 to 2017. NIWA Client Report, NIWA, Christchurch, New Zealand.