Test of an EU toolkit to develop nutrient standards in lakes

Research note from
DCE - Danish Centre for Environment and Energy

Date: 27th August 2018

Thomas Alexander Davidson & Martin Søndergaard

Department of Bioscience

Claimant:
The Danish Environmental Protection Agency
No of pages: 17

Referee:
Torben L. Lauridsen
Quality assurance, DCE:
Poul Nordemann Jensen

AARHUS
UNIVERSITY
DCE – DANISH CENTRE FOR ENVIRONMENT AND ENERGY

Tel.: +45 8715 0000
E-mail: dce@au.dk
http://dce.au.dk

Contents

1 Introduction

Together with an ECOSTAT working group, the EU commission by the Joint Research Center (JRC) has developed a toolkit that can be used as a best practice guide (BPG) to set nutrient (nitrogen and phosphorus) standards in lakes. Concentrations of nitrogen and phosphorus should be used in the implementation of the European Water Framework Directive as supporting elements in the overall ecological classification of lakes. Denmark was asked to comment on the best practice guidance and to run a test of the guidance using data from the Danish lakes.

The aim of this project is to:

- Give comments to "draft Best Practice Guide": https://circabc.europa.eu/w/browse/6a4a9f9f-f5d6-4b7c-803e-ea5242ce9a85
- Run a test of the toolkit in order to develop nutrient standards in lakes (toolkit for boundary setting, Excel File, R Scripts & example data files) are available at: https://circabc.europa.eu/w/browse/6a4a9f9f-f5d6-4b7c-803e-ea5242ce9a85
- Participate in an international workshop held in Berlin on the use of the EU-toolkit.

2 Data and methods

2.1 Data

The Danish Environmental Protection Agency supplied data used for the test of the toolkit. The data set has previously been used in an analysis by Larsen et al. (2017) of Danish lakes included in the current water management plans ("Vandområdeplaner 2016-2021").

For all the lakes included in the management plans, this data set includes data on lake type, nutrient concentrations (summer means) and the ecological quality ratio (EQR) of the three biological quality elements used (submerged vegetation, phytoplankton and fish). Data is not available from all elements in all lakes.

2.2 Methods

In testing the tool kit, we used the analyses suggested in the BPG. This comprised different regression models as described in the following section.

3 Results

3.1 Comments to the "draft Best Practice Guide"

The draft of "Best practice guide on establishing nutrient concentrations to support good ecological status" by Phillips et al. (draft), circulated to the ECOSTAT and nutrient experts in August 2017, is a very thorough and well-presented document on the background, rationale and procedures that may be used for establishing nutrient boundaries for good ecological status. The document is well written and introduces the complexity of the nutrient - ecological status relationship. Overall, the tool kit provides a very good and useful guide on how nutrient standards in lakes can be developed based on existing data.

The comments provided by other member states highlight some of the issues connected with its application. The most important of these is the problem of setting nutrient boundaries in the presence of multiple stressors.

3.2 Test of the tool kit

We applied the toolkit from the BPG to the data analyses using R, an open source programming language and software environment for statistical computing and graphics, as suggested in the guidance. The analyses were conducted on data from shallow high alkalinity lakes (lake type 9) for the ecological assessments based on vegetation, phytoplankton, fish and then chlorophyll-a as an additional element.

Regression based analysis

The first suggested option for determining nutrient concentrations to support good ecological status in the BPG consists of three regression analyses. The first two use ordinary least squares (OLS) regression, which assumes no variability in the predictor. Of those two the first uses the EQR as the predictor (x) and nutrients as the response (y), the second uses the nutrient concentration as the predictor (x) and the EQR as the response (y). The third type of regression is termed type II regression or ranged major axis (RMA) regression and assumes an equal amount of error in both predictor and response. The three different regression approaches give different results and the BPG document suggests how these model outputs can be synthesised into setting boundary values.

Each method can be used to determine the mean value of total phosphorus (TP) and total nitrogen (TN) at the boundary of ecological status. The quantiles of the residuals of the model are then calculated to determine upper and lower quantiles of TP and TN for the specific EQR boundary. The regression-based methods rely on a number of assumptions, chiefly that there is a linear relationship between EQRs and nutrient concentrations. In addition, the BPG states that the R^2 values of the model should be greater than 0.36. These assumptions are largely met by the Danish lake data, although the vegetation model has an R^2 of 0.33. The toolkit in the BPG provides the codes needed to examine whether the data meet the assumptions required for the regression analyses and to identify outliers that may be excluded from the analyses. For the vegetation data, a stricter cut off limit for exclusion of outlier samples would increase the R^2, without altering results, in terms of the identified values at the boundary of each ecological status and so the results presented here can be considered robust.

The results of the three different types of regression applied to vegetation, phytoplankton and fish derived from EQRs can be seen in figures 1-10 (appendix) and the summary results can be found in table 1 for the good to moderate boundary and table 2 for the high to good ecological status boundary. Here we focussed on TP in the report as in fresh waters TP is thought to be more important for determining EQR. However, the methods can equally be applied to determine TN values associated with changes in EQR.

Table 1. Summary results of regression analyses for nutrient concentrations at the good to moderate boundary. **The results in bold have an r^2 higher than 0.36.**

			Nitrogen µg/l							
		Veg.	**Phyto**	**Fish**	**Veg.**	**Phyto**	**Fish**	**Veg.**	**Phyto**	**Fish**
	Model		**mean**			**lower quartile**			**upper quartile**	
EQR predicting N	OLS model 1	763	**617**	501	368	**444**	245	1364	**852**	966
	RMA model	983	**792**	898	748	**626**	690	1276	**1027**	1131
N predicting EQR	OLS model 2	1017	**891**	966	791	**721**	747	1323	**1114**	1173
			Phosphorus µg/l							
		Veg.	**Phyto**	**Fish**	**Veg.**	**Phyto**	**Fish**	**Veg.**	**Phyto**	**Fish**
	Model		**mean**			**lower quartile**			**upper quartile**	
EQR predicting P	OLS model 1	50	**33**	**30**	23	**17**	**16**	109	**56**	**59**
	RMA model	65	**46**	**45**	41	**30**	**29**	99	**74**	**73**
P predicting EQR	OLS model 2	72	**62**	**49**	47	**40**	**29**	117	**90**	**78**

Table 2. Summary results of the regression analyses for nutrient concentrations at the high to good ecological status boundary. **The results in bold have an r^2 higher than 0.36.**

			Nitrogen µg/l							
		Veg	**Phyto**	**Fish**	**Veg**	**Phyto**	**Fish**	**Veg**	**Phyto**	**Fish**
	Model		**mean**			**lower quartile**			**upper quartile**	
EQR predicting N	OLS model 1	370	**301**	270	178	**217**	132	661	**416**	520
	RMA model	807	**526**	755	615	**416**	580	1048	**683**	952
N predicting EQR	OLS model 2	896	**686**	859	697	**555**	665	1167	**857**	1044
			Phosphorus µg/l							
		Veg	**Phyto**	**Fish**	**Veg**	**Phyto**	**Fish**	**Veg**	**Phyto**	**Fish**
	Model		**mean**			**lower quartile**			**upper quartile**	
EQR predicting P	OLS model 1	17	**10**	**13**	8	**5**	**7**	37	**16**	**27**
	RMA model	38	**19**	**28**	24	**13**	**18**	57	**31**	**45**
P predicting EQR	OLS model 2	51	**38**	**33**	33	24	20	82	**55**	**53**

In table 3, we have given an example of how to synthesise the three regression models. Generally, the RMA regression type has the lowest error (the least difference between lower and upper quartile) in predicting nutrient boundaries, suggesting that, where the assumption of a linear relationship between EQR and nutrients is met, and the R^2 is sufficiently high, that this is the most reliable method to apply for the Danish lake data (see decision tree in appendix 2). Following the methods set out in the BPG, table 3 shows a "predicted" value which is the mean RMA model value, the "range" is the mean values OLS mod1 and OLS mod2 methods and the "possible range" is the lower and upper quartile of the RMA regression. In the case of the high/good boundary in table 3, the lower limit of the possible range and the range are the same. Normally, the possible range is larger than the range; however, in this case for vegetation, the mean OLS model 1 value (the lower end of the range) was lower than the lower quartile of the RMA model. As the range cannot be larger than the possible range, 17 µg TP/l was used for both.

Table 3. Examples of how to synthesise the three regression method results to set boundary values (good-moderate and high-good ecological state) for lake type 9. The table shows a summary of predicted TP boundary values (µg/l) and the range and the possible range using vegetation. $R^2 = 0.33$, $p<0.001$.

	Most likely boundary				
	Predicted	**Range**		**Possible range**	
Good/Mod	65	50	72	41	99
High/Good	38	17	51	17	57

Categorical and logistic regression methods

Where a linear relationship between the EQR and nutrient levels is not present or the R^2 value is not high enough, then regression-based analyses are not suitable (see decision tree in appendix 2). In addition, if there are multiple stressors shaping the community, the regression techniques may not be appropriate. In this case, the BPG suggests the use of categorical methods, which do not assume a gradual gradient response, but instead place the lakes into categories or groups. We applied categorical methods of averaged adjacent quartiles and also the 75th quartile of the EQR group, the latter is the method previously used to suggest nutrient boundaries in Danish lakes (Søndergaard & Lauridsen, 2015). In table 4, the average mean of adjacent quartiles is the mean of the average of two TP values, 1) average of the 75th quartile of TP values for one category (e.g. good) and 2) the lower 25th quartile of TP values for the adjacent category (e.g. moderate). The Average median quartile is similar in that it is the average of the median values of the 75th quartile of the good category and the median of the 25th quartile of the moderate category. The 75th quartile class is simply the TP associated with the 75th quantile of a particular class; the BPG states that this method provides the least stringent estimate for the good/moderate boundary.

Furthermore, we applied logistic regression and mismatch analysis to determine the concentrations of TP and TN (again only TP shown here) likely to support good ecological status based on vegetation-based EQR in lake type 9. The results of the averaged adjacent quartile analysis can be seen in table 4 below and can be compared with the results of the regression analyses in tables 1 and 2.

Table 4. Categorical method results to set boundary values for TP (µg/l) for lake type 9 based on vegetation-based EQR and phosphorus data. The average mean adjacent quartile gives the average of two values, 1) mean nutrient values of the 75th quartile of one class (e.g. good) and 2) mean values of TP for the lower 25th quartile of the adjacent class above (e.g. moderate). The average median quartile is the average of the median values of the 75th quartile of TP in good status and the median of the 25th quartile of TP in moderate status, and the 75th quartile class is the simply the TP values at the 75th quartile of the lakes classified as good – in the case of the good/moderate row.

	Average mean of adjacent quartiles	**Average median of adjacent quartiles**	**75th quartile class**	**Mismatch analysis**	**Logistic regression**
Good/moderate	88	76	115	108	82
High/good	48	50	56	27	19

Comparison of regression and categorical methods

The guideline document suggests that the type II regression model (the ranged major axis (RMA) method) should be used to predict the most likely nutrient concentrations that occur at the EQR boundaries. The two OLS regression models provide alternative upper and lower predictions and the true value will lie within this range.

The regression methods produced lower values of TP for both the good/moderate boundary and for the high/good boundary compared with the percentile-based categorical methods, although most values fell within the broader possible range (table 3). The mismatch and logistic regression analysis produced higher values of TP compared with the other methods at the good to moderate boundary but much lower values at the high to good boundary (table 4).

All methods suffer from the fact that it is a combination of TP and TN that shapes ecological status. In treating each nutrient separately, the definition of the upper limits of the good to moderate boundary may be overestimated. This is because lakes maintaining good status at higher P are likely to be N limited and thus have lower N. It might be useful to apply methods that can take both nutrients and their potential interactions into account to define the nutrient concentrations most likely to support good ecological status. Such methods might include boosted regression trees or Random forests to model the potential interactions between N and P in shaping lake EQR.

Chlorophyll-a analysis

The WFD stipulates which biological group should be used to define EQR; however, chlorophyll-a has also been routinely used to define EQR in Denmark and therefore may be useful to define nutrient boundaries for the EQR as data exist for many more sites. Here we applied the same regression and categorical methods to chlorophyll-a data (Table 5) that can be compared with the regression analysis-based results from phytoplankton data (Table 6) and the categorical methods (Table 7).

As stated above the predicted value (see Tale 5 & 6) is the mean of the RMA regression and the "range" values are the mean values of the two OLS regression techniques. The "possible range" is provided by the lower and upper quartiles of the residuals from the type II RMA regression. The interquartile range of the residuals includes 50% of the water bodies in the modelled data set, so 75% of water bodies would be expected to have values less than the upper range value, while 75% would have values greater than the lower "range" value. Comparing the values of TP associated with the change in EQR as determined by chlorophyll-a and phytoplankton shows that the values for the Good/moderate boundary are similar between the two. In contrast phytoplankton give much lower values of TP for the High/Good boundary.

Table 5. Boundary levels of TP (µg/l) based on the three regression techniques for using chlorophyll-a for lake type 9. R^2=0.455, $p<0.001$.

	Most likely boundary				
	Predicted	**Range**		**Possible range**	
Good/Mod	52	43	65	33	76
High/Good	41	32	56	27	61

Table 6. Boundary levels of TP (µg/l) based on the three regression techniques for using phytoplankton for lake type 9. R^2=0.401, $p<0.001$.

	Most likely boundary				
	Predicted	**Range**		**Possible Range**	
Good/Mod	46	33	62	30	74
High/Good	19	10	38	10	38

Table 7. Categorical method results to set boundary values for TP (µg/l) for lake type 9 based on chlorophyll-a EQR.

	Average quartiles	Average median	75th quartile class	Logistic regression
Good/Mod	72	70	86	60
High/Good	42	41	42	29

3.3 Conclusions from test of the tool-kit

The toolkit presented here and the Best Practice Guide (BPG) have been applied to data from Danish lakes of typology 9, shallow high alkalinity lakes. We used data on the three quality elements: vegetation, phytoplankton and fish plus the concentration of chlorophyll-a that previously was the only element used to define EQR in Danish lakes.

For the vegetation-based analysis, which had the most data of the three quality elements, the assumptions of the regression methods were not quite fulfilled, although the R^2 value was very close to the acceptable value (0.36) in the BPG. As the R^2 values were so close to acceptable the boundaries of the EQR groups shown in table 3 are reasonable: however, if the R^2 is deemed too low, then the results from the categorical methods shown in table 4 could be used. The categorical method has already been applied earlier to chlorophyll-a data and produced near identical results (Søndergaard & Lauridsen, 2015).

We would not recommend use of the mismatch analysis for setting the boundaries at the good to moderate boundary as it overestimates the TP values supporting a good EQR. The reason is probably that setting the TP boundary is vulnerable to the influence of TN.

Although chlorophyll-a is not a quality element, we applied the methods in the toolkit to chlorophyll-a also, because of the much larger data set and because chlorophyll-a as has been used frequently to set the EQR in Danish lakes. The summary result for the regression analysis can be seen in table 5 and the categorical methods in table 7. The values for the phytoplankton EQR-based regression analyses are very similar to those of the chlorophyll-a analysis, if a little lower (table 6).

To conclude:

- If the BPG methods are followed and the decision tree employed (see appendix 2), then the regression analyses values for phytoplankton (and chlorophyll-a, if treated as a quality element) should be used to define the good to moderate boundary. A decision on which of the values "predicted", "range" and "possible range" to be used would need to be made, but the highest value of total phosphorus concentration as supporting element is the upper end of the possible range, which is 76 µg l^{-1} for Chlorophyll-a and 74 µg l^{-1} for phytoplankton.

- It can be considered important not to choose too restrictive a value for a supporting element and the 75^{th} quartile class has been used previously (Søndergaard & Lauridsen, 2015). This defined the TP value for the moderate to good boundary in lake type 9 at 85 µg l^{-1}, which is identical to the 75^{th} quartile in the data here (86 µg l^{-1}). The regression analyses for both chlorophyll-a and phytoplankton set the boundary for good to moderate status lower, as did the other categorical methods, the adjacent quartiles and the logistic regression (tables 5, 6 & 7). Accordingly, the BPG also states that the 75th quartile class provides the least stringent estimate as for the good/moderate boundary.

- The BPG tool-kit approach recommends the use of liner regression methods. This means that values from table 5 or 6 are appropriate for TP boundary settings between good and moderate if using chlorophyll-a and phytoplankton. However, given that the nutrient boundary values are a supporting element using a less restrictive nutrient concentration, such as selecting the 75th quartile value for chlorophyll-a, may also be considered.

3.4 Notes on Ecostat meeting Berlin 26-27 September 2017

The aim of the Ecostat meeting in Berlin was to find out how different member states could use the proposed tool-kit as guide to develop nutrient standards for lakes. Thomas Davidson from AU participated in the meeting.

We contributed with our experience and assessment of the 'Best Practice Guide on establishing nutrient concentrations to support good ecological status' and the statistical tool-kit designed alongside. For lakes, we found that there was a relatively strong relationship between BQE and nutrients and so the simpler linear methods were largely applicable. We also tested the categorical methods that can be applied when the linear relationships were absent and these methods also performed relatively well. The discussion of multiple drivers was interesting and some of the alternative methods may be useful to integrate both phosphorus and nitrogen into the same model. Furthermore, they may be useful in the analysis of stream data.

4 References

Larsen S.E., Søndergaard M. & Johansson L.S. 2017. Metode til vurdering af risikoen for skift i tilstandsklasser i søer ved uændrede belastningsforhold. Notat fra DCE, 29 sider.

s. 3: Phillips et al. (draft) Best Practice for establishing nutrient concentrations to support good ecological status – available here : https://circabc.europa.eu/w/browse/6a4a9f9f-f5d6-4b7c-803e-ea5242ce9a85

Søndergaard, M. & Lauridsen, T.L. 2015. Anvendelsen af kvalitetselementer i ikke-interkalibrerede danske søtyper. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi, 48 s. - Videnskabelig rapport fra DCE - Nationalt Center for Miljø og Energi nr. 139. dce2.au.dk/pub/SR139.pdf

Appendix 1

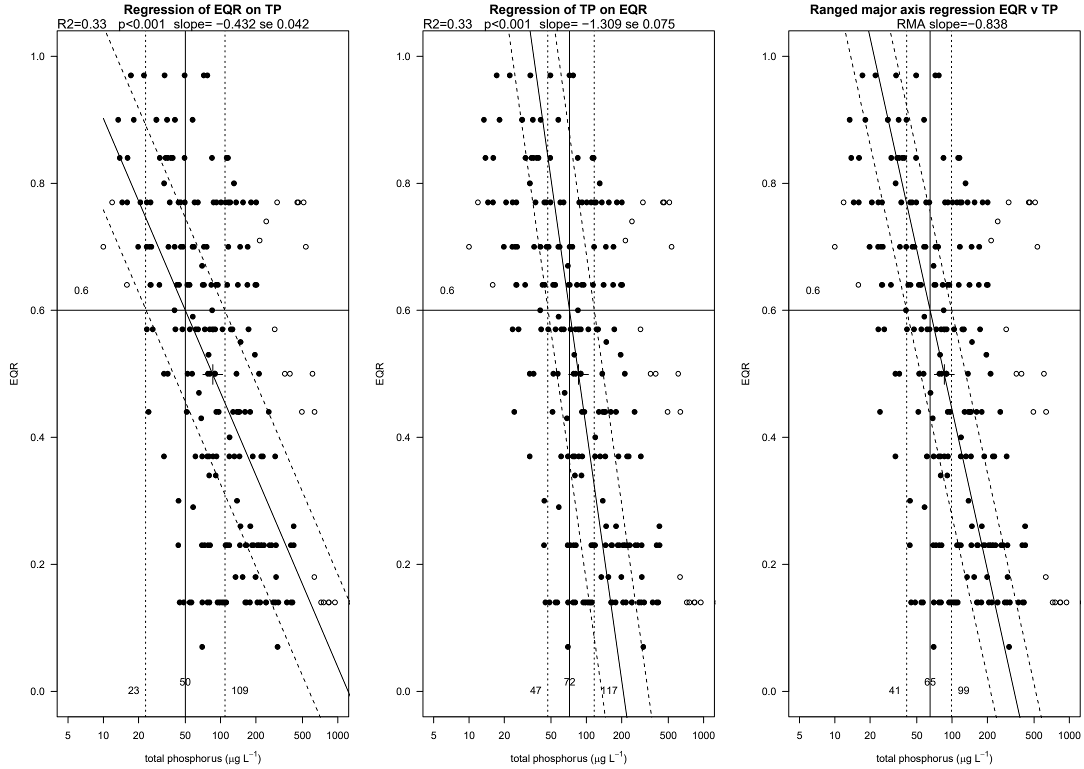

Figure 1. Vegetation analysis – regression plots for the three types of regression – EQR predicting TP, TP predicting EQR and ranged major axis regression of TP to EQR. The mean, upper quartile and lower quartile values of TP are given in table 1.

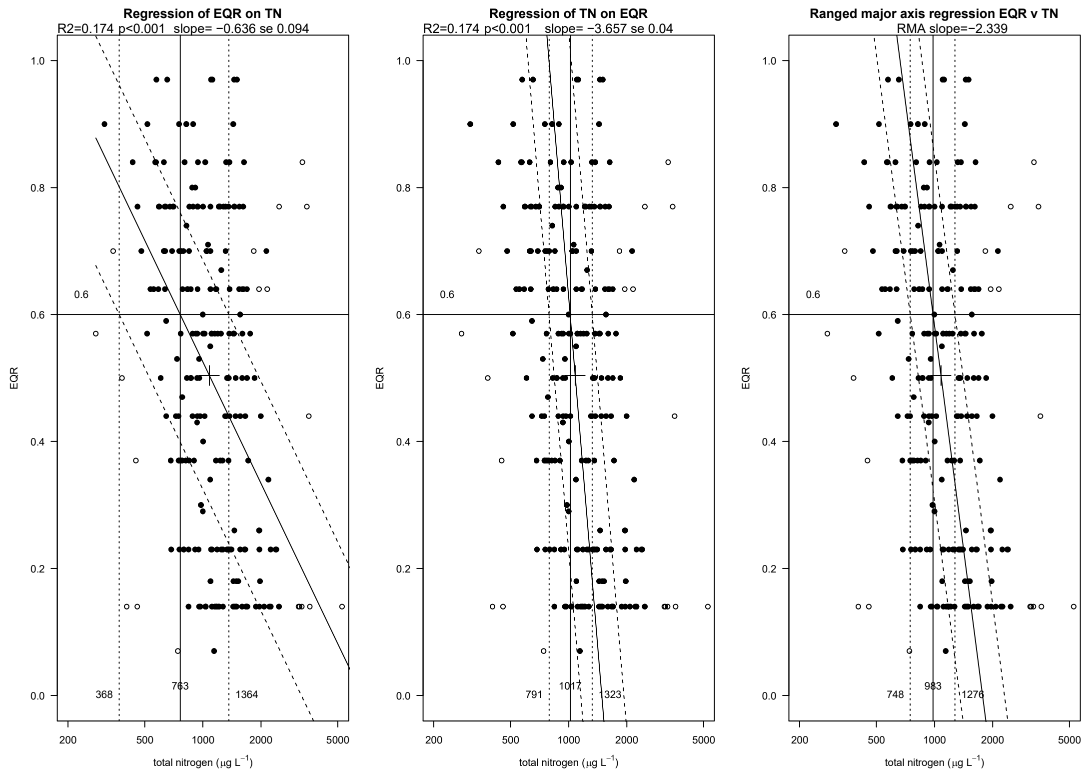

Figure 2. Vegetation analysis – regression plots for the three types of regression – EQR predicting TN, TN predicting EQR and ranged major axis regression of TN to EQR. The mean, upper quartile and lower quartile values of TP are given in table 1.

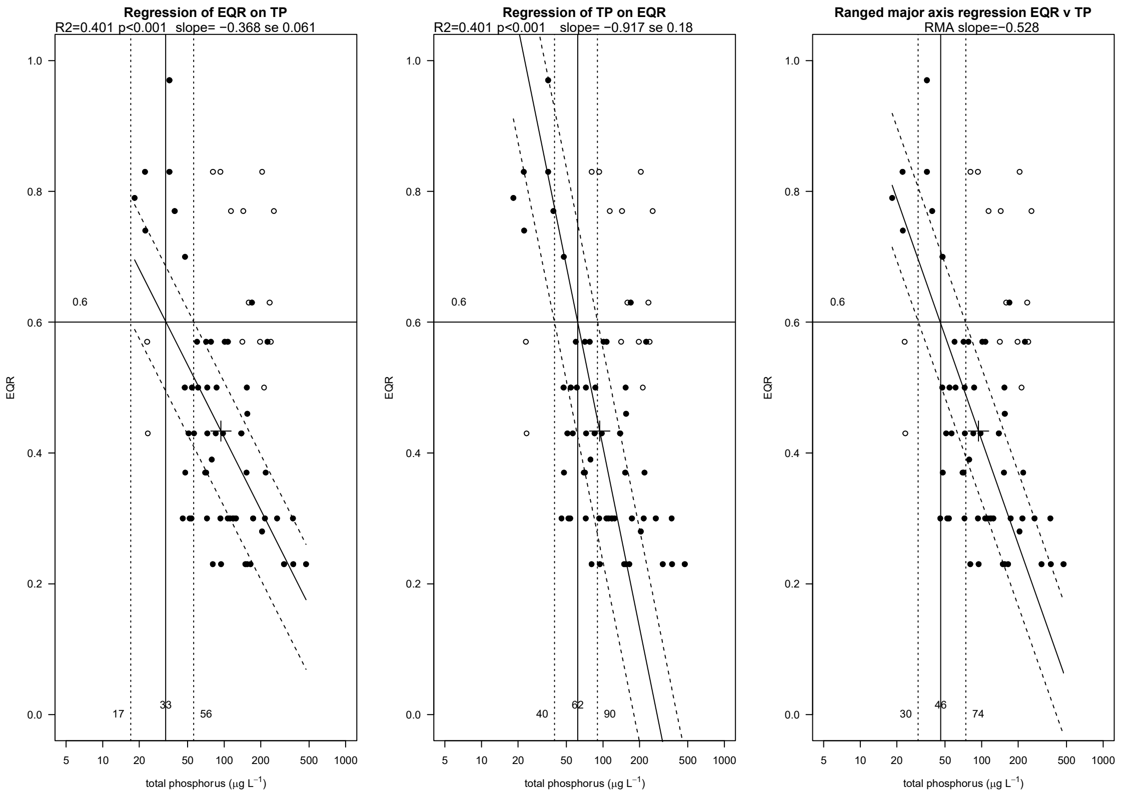

Figure 3. Phytoplankton analysis – regression plots for the three types of regression – EQR predicting TP, TP predicting EQR and ranged major axis regression of TP to EQR. The mean, upper quartile and lower quartile values of TP are given in table 1.

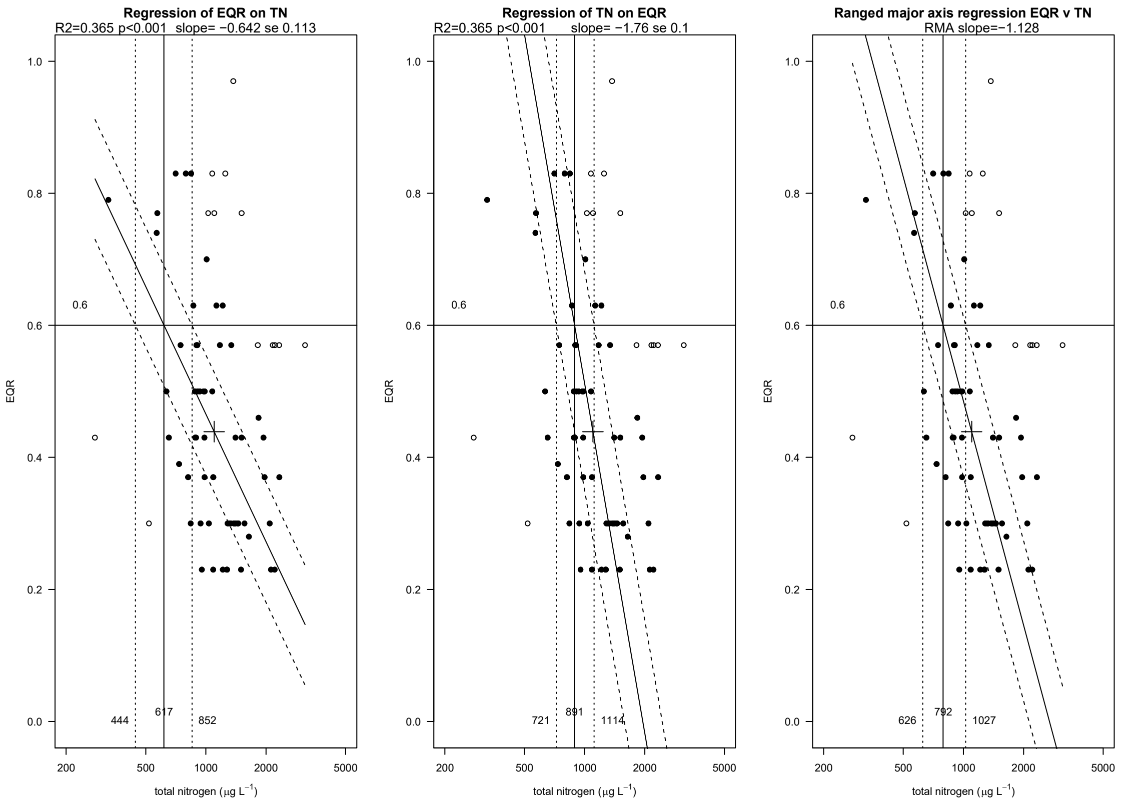

Figure 4. Phytoplankton analysis – regression plots for the three types of regression – EQR predicting TN, TN predicting EQR and ranged major axis regression of TN to EQR. The mean, upper quartile and lower quartile values of TP are given in table 1.

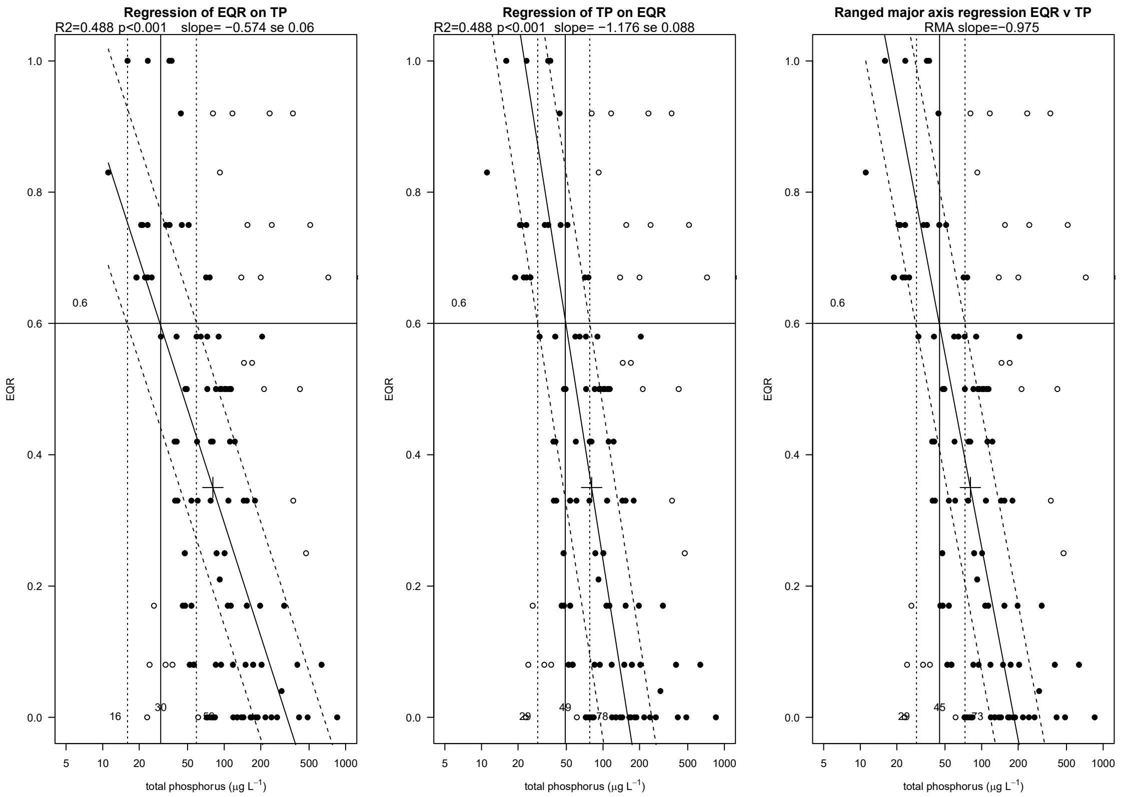

Figure 5. Fish analysis – regression plots for the three types of regression – EQR predicting TP, TP predicting EQR and ranged major axis regression of TP to EQR. The mean, upper quartile and lower quartile values of TP are given in table 1.

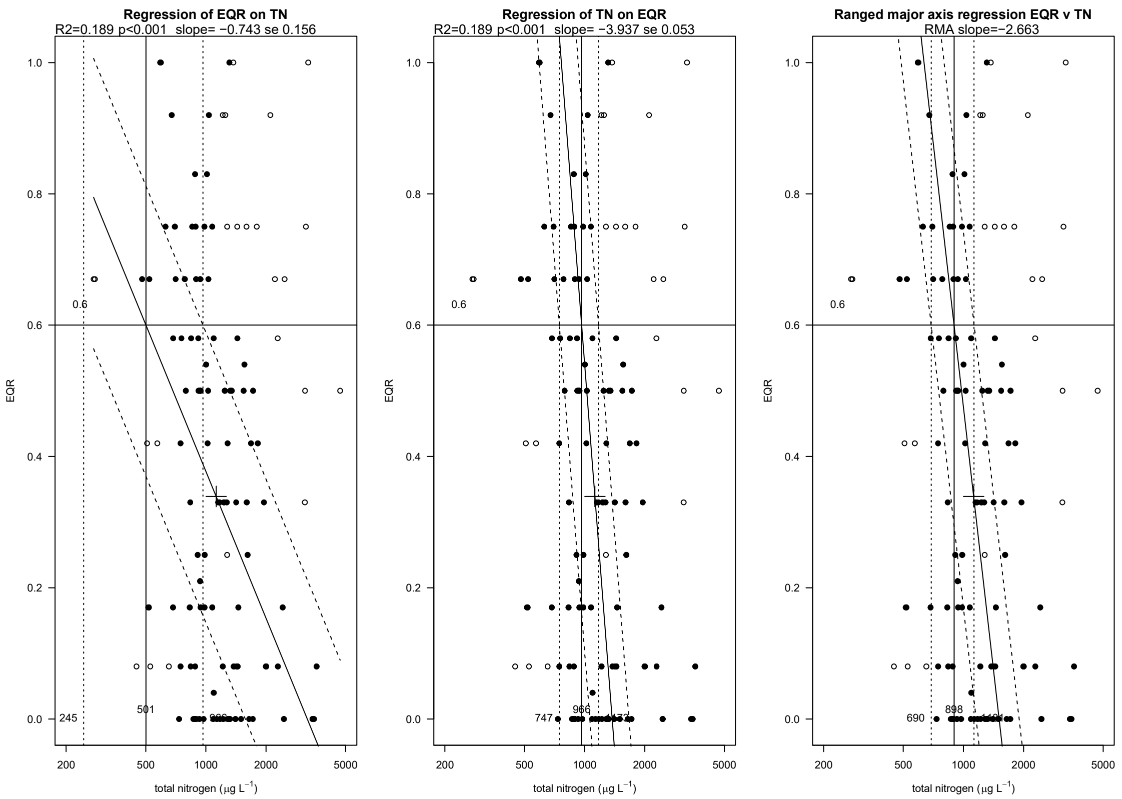

Figure 6. Fish analysis – regression plots for the 3 types of regression – EQR predicting TN, TN predicting EQR and ranged major axis regression of TN to EQR. The mean, upper quartile and lower quartile values of TP are given in table 1.

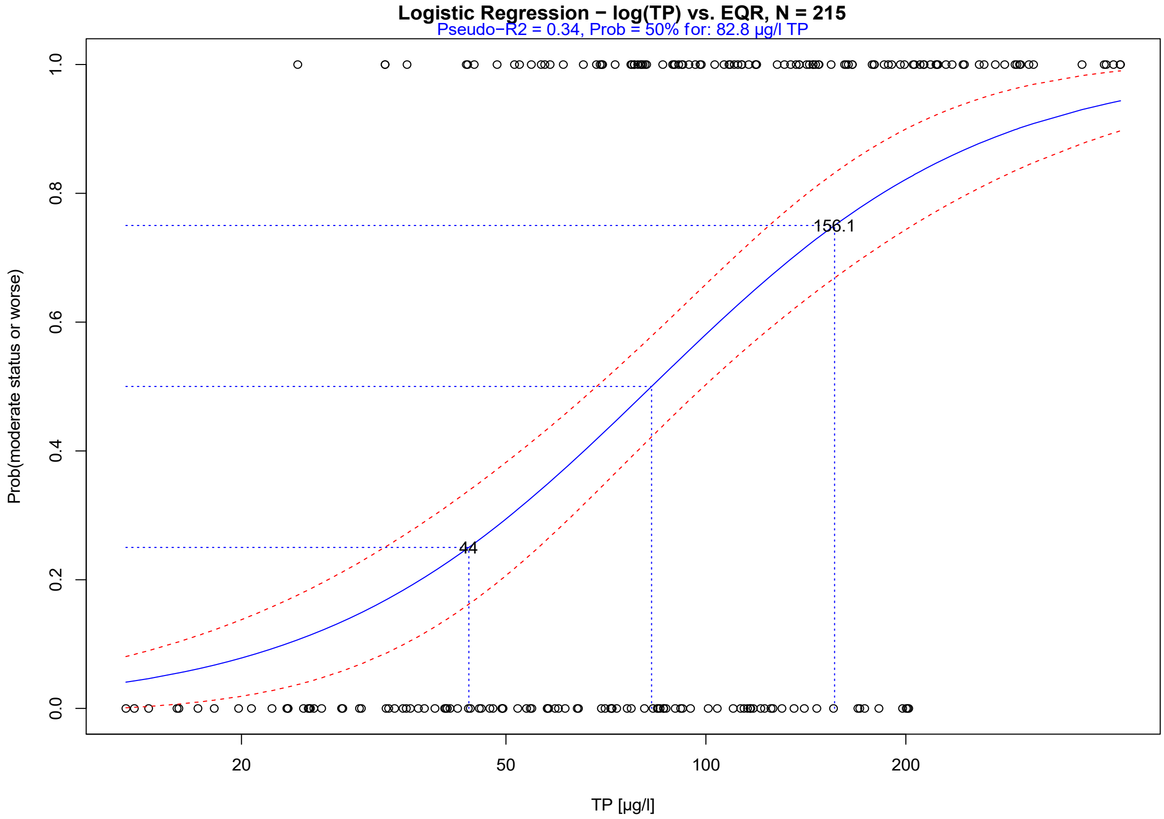

Figure 7. Vegetation analysis – logistic regression to calculate the likelihood of moderate or worse status based on TP (µg/l) concentration.

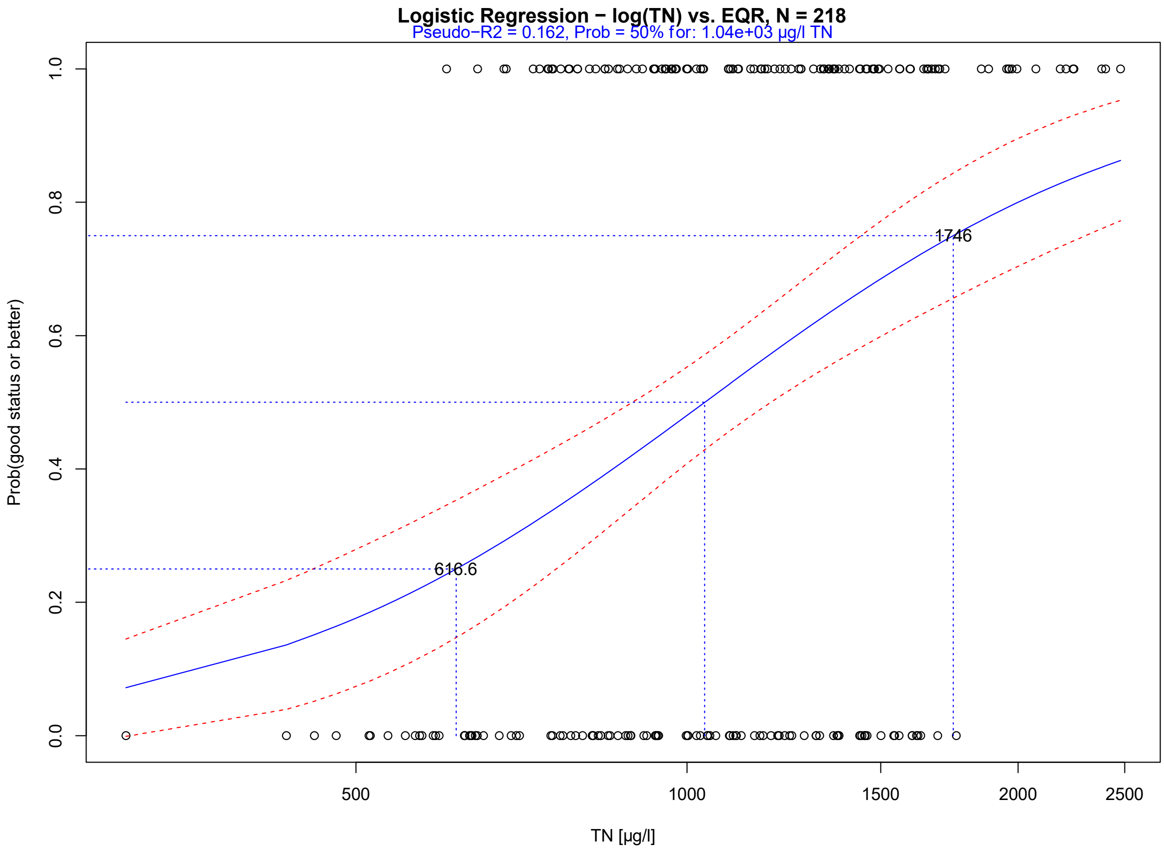

Figure 8. Vegetation analysis – logistic regression to calculate the likelihood of moderate or worse status based on TN (mg/l) concentration.

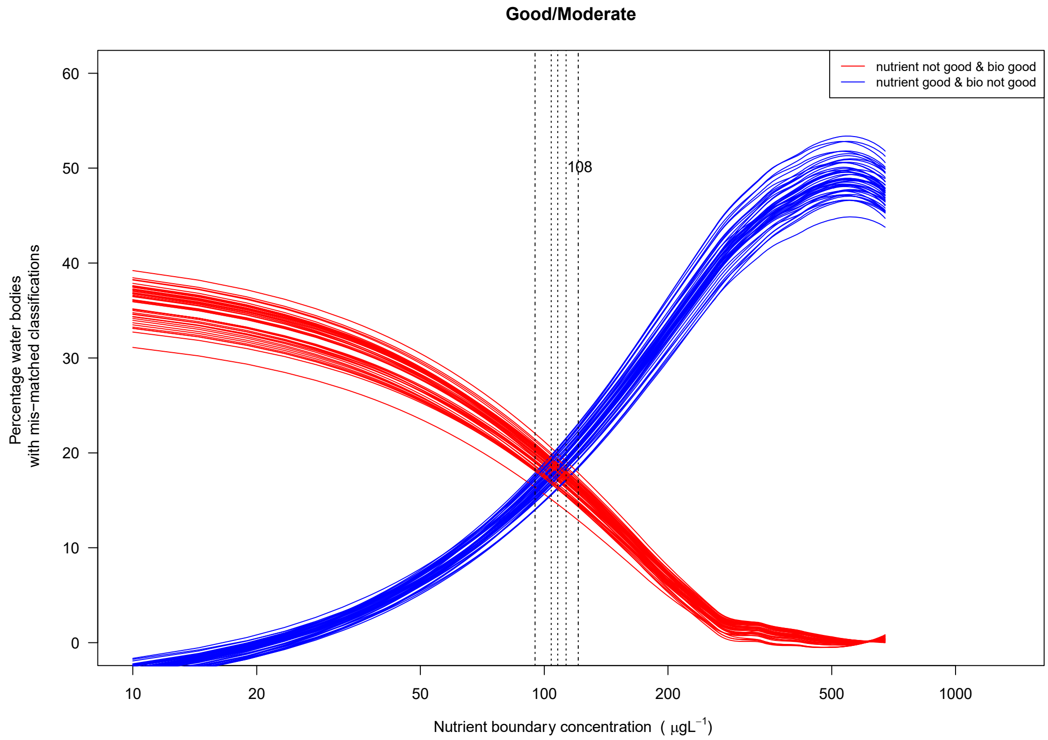

Figure 9. Vegetation analysis – mismatch analysis for the good to moderate status with TP (µg/l) concentration. The many lines results from the fact the analysis is iterative and repeatedly randomly selects a portion of the data and to make a model which it then test on the remaining data and then calculates the percentage of misclassifications.

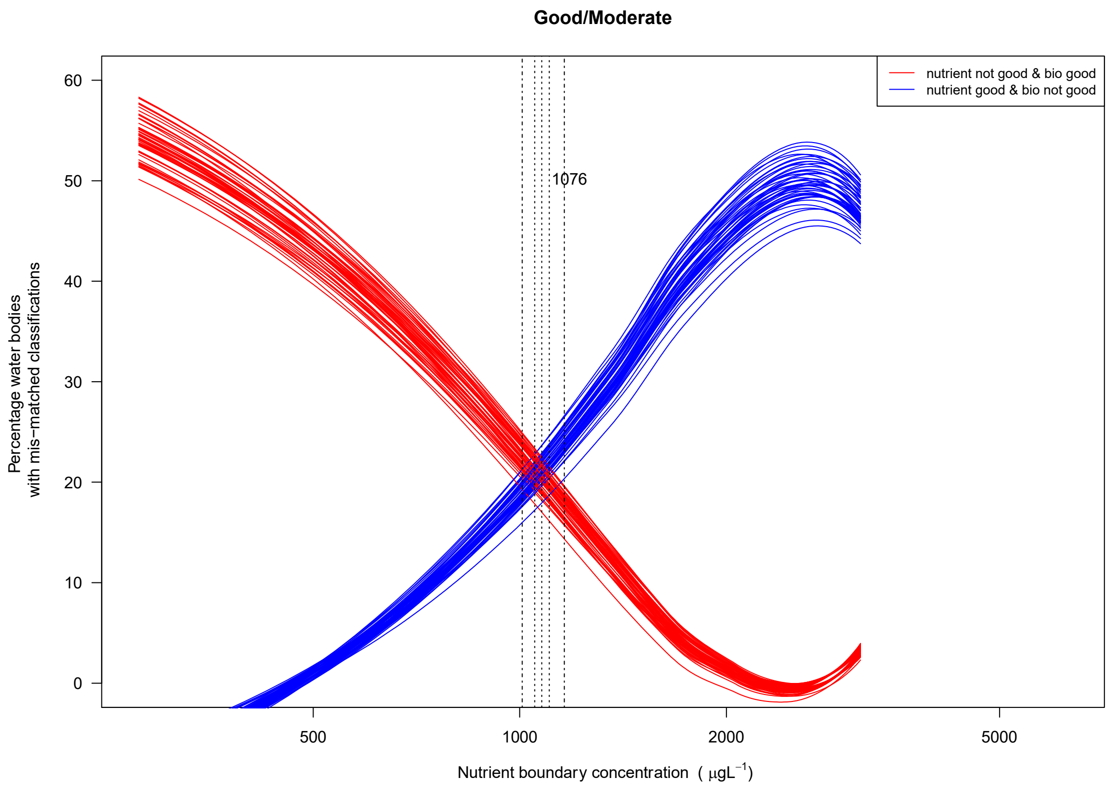

Figure 10. Vegetation analysis – mismatch analysis for the good to moderate status with TN (mg/l) concentration. The many lines results from the fact the analysis is iterative and repeatedly randomly selects a portion of the data and to make a model which it then test on the remaining data and then calculates the percentage of misclassifications.

Appendix 2 Decision tree for choice of approaches – for an explanation see the best practice guide document

Biological EQR data available

N

Use categorical methods
Go to Section C

Only categorical methods can be used.

Y

Relationship is linear for at least H, G and M and has a high r^2

Y

Use regression methods but compare with categorical methods. Appendix 1 A3 & A4.3

The key point here is that the relationship is "anchored" at both ends of the status gradient, allowing linearity of the response between nutrients and biology to be assessed.

N

Relationship is linear and has a high r^2 but does not extend beyond the G/M or H/G boundary

Y

Use regression methods to predict boundary values. Appendix 1 A3 & A4.3

In this situation, categorical methods may be impossible or unreliable as data points in high or moderate class are too few

The greater the difference between the mean EQR and the boundary, the greater the uncertainty of the prediction

N

Relationship is linear and has a low r^2 but does not extend beyond the G/M or H/G boundary

Y

Consider combining data from similar type or neighbouring country. Appendix 1 A5.1

N

Models with low r^2 should not be extrapolated. The best option is to extend the data set. Minimisation of mismatch may be possible if data for at least two classes around the boundary of interest are available

Data available for four classes H, G, M and P but linear relationship has low r2

Y

Go to Section B

Choice of approaches B

Data available for four classes H, G, M and P but linear relationship has low r^2

N

If the data set does not span four classes it may be possible to use a categorical approach. This, however, necessitates good definitions of H, G and M status

Y

Confounding variable identified; values available in data set

Y

Use of regression with categorical variable (Appendix 1 A5.1.1) or use of multiple regression approach (Appendix 1 A5.2.1.3)

N

An example of this type of situation may be where there is good evidence of N and P co-limitation; threshold values may be interdependent and multiple regression allows values of one variable to be predicted for the full range of values of the second

Relationship is "wedge shaped" (i.e. asymmetric with respect to line of best fit)

Y

Use quantile regression (Appendix 1 A5.2.1.1)

N

The point here is to identify an upper or lower surface to the scatter and fit a line to it. A lower surface would produce a precautionary boundary value if other factors prevent the expression of nutrient pressure. An upper surface a least precautionary boundary where other stressors cause a lower biological status despite low nutrients

Relationship is "cloud-shaped" (i.e. approximately symmetrical about "line of best fit")

Y

Data has high uncertainty. Categorical method minimisation of mismatch likely to be the best approach, but try regression and compare outcomes. Consider combining normalised EQRs for different BQEs or if the use of a metric rather than EQR improves r^2

N

Too few data. Use other approaches (see 6.0)

Choice of approaches C
(Categorical Methods)

Data available for three classes H, G and M & box plots for adjacent classes do not substantially overlap

Y

Compare the following three methods

a) average of upper and lower quartiles of adjacent classes
b) average of median of adjacent classes
c) minimisation of mismatch

Currently, the minimisation of the mismatch method appears to be the least sensitive to outliers and non-linear relationships and is likely to be the most robust method

N

Data available for either classes H & G or G & M & box plots for adjacent classes do not substantially overlap

Y

As above, but only for the H/G or G/M boundary

N

Box plots substantially overlap or too few data in any class:
categorical analysis likely to be unreliable; use other methods (see 6.0)