COMMON IMPLEMENTATION STRATEGY FOR THE WATER FRAMEWORK DIRECTIVE AND THE FLOODS DIRECTIVE

Best practice for establishing nutrient concentrations to support good ecological status

Disclaimer

This technical document has been developed through a collaborative framework (the Common Implementation Strategy) involving the Member States, European Free Trade Association countries and other stakeholders, including the European Commission. The document reflects the informal consensus position on best practice acknowledged by the EU Water Directors. However, the document does not necessarily represent the position of any of the partners.

To the extent that the European Commission's services provided input to this technical document, such input does not necessarily reflect the views of the European Commission.

The technical document is intended to facilitate the implementation of Directive 2000/60/EC and is not legally binding. Any authoritative reading of the law should only be derived from Directive 2000/60/EC itself and other applicable legal texts or principles. Only the Court of Justice of the European Union is competent to authoritatively interpret EU legislation.

Foreword

In 2000 the EU Member States, Norway and the European Commission jointly developed a Common Implementation Strategy (CIS) for implementing Directive 2000/60/EC, the Water Framework Directive (WFD), to ensure consistent implementation. The focus is on developing a common understanding of the technical and scientific implications of the WFD. One of the objectives is the development of non-legally binding and practical guidance documents on various technical issues in the Directive. These are targeted at experts who are directly or indirectly implementing the WFD in river basins. The structure, presentation and terminology are therefore adapted to their needs and formal, legalistic language is avoided wherever possible.

In 2009 the CIS *Guidance document on eutrophication assessment in the context of European water policies — Guidance document No 23* was published, providing guidance for evaluating the impacts of nutrient enrichment, a major cause of failure to achieve good status under the WFD. However, an apparently wide range of nutrient boundary values to support good ecological status had been established by the Member States. The Water Directors requested that the CIS Working Group on Ecological Status investigate this issue, and the subsequent work has been led by the United Kingdom (freshwaters), Germany (saline waters) and the Joint Research Centre. The aim of the work was to establish the reasons for differences between Member States in the development and application of nutrient boundaries, leading to the production of this document on best practice. This work is an addition to, and not a replacement for, the earlier guidance on eutrophication assessment.

In developing this document a number of tasks were undertaken. The range of nitrogen and phosphorus boundary values in use by Member States, and the methods used to derive those values, has been reported for both fresh and saline waters. Further work was undertaken to investigate the relationships between nutrient pressure and biological response in the different surface-water categories. This work was then used to inform the development of this document and the associated statistical toolkit. During the project a series of workshops were held involving nutrient experts nominated by Member States. These experts contributed to the development and testing of the technical document and toolkit, and provided details of alternative methods of boundary-setting in use in some Member States.

The purpose of this report is to provide technical guidance to enable Member States to establish new, or review existing, boundaries for phosphorus and nitrogen to support good ecological status. This should facilitate the establishment of comparable and consistent boundaries across all Member States. However, it is recognised that alternative methods of arriving at boundary values may be valid, and use of this document and the associated statistical toolkit is ultimately a decision for the Member State.

The responses of biological elements to nutrient availability are complex, and vary between water categories. This document is therefore not a substitute for the application of ecological knowledge and understanding at a local level. Furthermore, responses to nutrients may be confounded by the impact of other pressures acting on a water body, and our understanding of how to account for multiple stressors is still developing. The document does not specifically address how the nutrient boundaries are used to derive an overall classification or to drive action to control nutrients, both of which may be relevant to the level at which the boundaries are set.

Contents

The statistical toolkit that accompanies this document can be found at:

https://circabc.europa.eu/w/browse/822756c8-3f08-4bf8-b5af-94ca378fe28b

The 'Shiny' version of the toolkit is available at:

http://phytoplanktonfg.okologia.mta.hu:3838/Tkit_nutrient/

Four appendices to this technical document are also available here

Summary

1. Elevated concentrations of nutrients are a major factor contributing to the failure of many water bodies to achieve good ecological status, and Member States need to determine levels appropriate to their own territories.

2. This document describes statistical methods for determining appropriate concentrations to support ecological status. These statistical methods should be set in a broader framework that also encompasses chemical, ecological and regulatory aspects relevant to the type of water body under consideration.

3. Three approaches to setting boundary concentrations are included. These are:

    - regression analysis, using a continuous relationship between an ecological quality ratio and nutrient concentration;

    - categorical analysis, using the distribution of nutrient concentration within biological classes;

    - minimisation of the mismatch of classifications for biology and nutrients.

4. The choice of method depends upon a number of factors, including the length of the gradient that available data sets cover and the statistical strength of the relationship between the explanatory and response variables. In some cases Member States may be better able to achieve the statistical prerequisites for methods by working with neighbours that share similar water-body types.

5. Microsoft Excel and a 'toolkit' written using the "R" statistical software package"R" are provided to make calculation of boundary concentrations more straightforward.

6. Options for situations where none of these methods are appropriate are also described.

7. Finally, some practical issues associated with the use of these boundary concentrations for regulation are discussed.

List of abbreviations and acronyms

BQE	biological quality element
Chl*a*	chlorophyll *a*
CI	continental influences
CIS	common implementation strategy
CTRW	coastal and transitional waters
CW	coastal waters
DIN	dissolved inorganic nitrogen
DIP	dissolved inorganic phosphorus
EQR	ecological quality ratio
EU	European Union
GES	good ecological status
GIG	Geographical Intercalibration Group
G/M	good status/moderate status boundary
Helcom	Helsinki Commission (Baltic marine environment protection commission)
H/G	high status/good status boundary
IC	intercalibration
LOESS	see glossary
M/P	moderate status/poor status boundary
MS	Microsoft
MSFD	marine strategy framework directive
NEA	north-east Atlantic
nEQR	normalised EQR
OLS	ordinary least squares (regression method, see glossary)
OSPAR	Oslo-Paris Convention (Convention for the Protection of the Marine Environment of the North East Atlantic)
P/B	poor status/bad status boundary

TON	total oxidised nitrogen (= nitrate-N + nitrite-N)
RF	random forest
TN	total nitrogen
TP	total phosphorus
TRP	total reactive phosphorus
TRW	transitional waters
WFD	Water Framework Directive

Glossary

Bayesian networks	A type of statistical model which uses the probability of various possible outcomes associated with stages in a process to predict the likelihood of a particular event occurring. Also known as 'Bayesian belief networks'.
Binomial logistic regression/model	A type of regression model which can be applied to categorical, rather than continuous, data.
Cyanobacteria	A group of bacteria capable of oxygen-evolving photosynthesis. They can be abundant in both marine and freshwater systems, either attached to surfaces or free-living. Formerly known as 'blue-green algae'.
Ecological quality ratio	A means of expressing the ecological condition of a biological quality element in which the observed state is presented as a ratio of the state that would be expected in the natural or near-natural state.
Eutrophication	The enrichment of water by nutrients, especially compounds of nitrogen and/or phosphorus, causing an accelerated growth of algae and higher forms of plant life to produce an undesirable disturbance to the balance of organisms present in the water and to the quality of the water concerned.
Hypolimnion	The lower part of the water column, beneath the thermocline and, therefore, cooler than water in the upper layers ('epilimnion'). In European lakes differentiation into epi- and hypolimnia is a seasonal phenomenon.
Intercalibration	The formal process that ensures comparability between the classification results of the Water Framework Directive assessment methods developed by the Member States for the biological quality elements. The intercalibration exercise establishes values for the boundary between the classes of high and good status, and for the boundary between good and moderate status, which are consistent with the normative definitions of those class boundaries given in Annex V of the Water Framework Directive.
LOESS regression	A polynomial regression technique for fitting a smooth curve between two variables. It is a development from locally weighted scatter-plot smoother (Lowess), which used a weighted linear least squares method to fit curves. LOESS (which is not strictly an acronym) uses a weighted quadratic least squares method to fit the curve.

Macroalgae	Multicellular algae (or aggregations of unicellular algae) living attached to substrata and visible with the naked eye.
Macrophyte	Larger freshwater plants which are easily seen with the naked eye, or which usually form colonies, including all aquatic vascular plants, bryophytes, stoneworts (*Characeae*) and macro-algal growths.
Phytobenthos	All phototrophic algae and **cyanobacteria** living on or in close contact with surfaces in aquatic environments.
Phytoplankton	The community of free-living, suspended, mainly photosynthetic organisms in aquatic systems comprising **cyanobacteria** and microscopic algae.
Ordinary least squares regression	The most familiar type of linear regression model in which a linear function (the 'line of best fit') is computed which predicts the parameters of an (unknown) 'dependent' variable from values of a (known) 'independent' variable. The independent variable is assumed to be free of error.
Quantile regression	An extension of linear regression which, rather than predicting the mean value of the dependent variable for any value of the independent variable, predicts the median or other quantiles of the dependent variable. It is particularly useful for predicting the highest (upper-quantile) or lowest (lower-quantile) values likely to be encountered at any level of the dependent variable.
Random forest	A statistical technique which builds 'decision trees', splitting data sets into subsets that simultaneously maximise the similarity within the set and the dissimilarity with records outside the set. In order to overcome statistical uncertainty, this process is repeated many times until the most likely breaks in the data set become apparent.
Ranged major axis regression	A form of **Type II regression** similar to major axis (orthogonal) regression that minimises the sum of the squared perpendicular deviations of the regression line but uses a transformation of the x and y axes to standardise their ranges, which overcomes some of the defects of major axis regression.
Secchi depth	The depth at which a Secchi disc— a circular disc divided into two black and two white quadrants — ceases to be visible from the surface: a measure of water transparency.

Transitional water	Bodies of surface water in the vicinity of river mouths and partly saline in character as a result of their proximity to coastal waters, but which are substantially influenced by freshwater flows.
Type 1 error	Erroneous rejection of the null hypothesis (i.e. reaching the conclusion that a difference exists when, in fact, it does not).
Type 2 error	Erroneous retention of the null hypothesis (i.e. reaching the conclusion that there is no difference when, in fact, a difference exists).
Type I regression	A form of linear regression model which assumes that the independent variable is measured without error. See '**Ordinary least squares regression**'.
Type II regression	A form of linear regression model which does not assume that the independent variable is measured without error. See '**Ranged major axis regression**'.

1. Introduction

1.1. Purpose of the document

The purpose of this document is to help Member States achieve good ecological status (GES) in surface waters. It complements the Common Implementation Strategy (CIS) *Guidance document on eutrophication assessment in the context of European water policies* (European Commission, 2009) by providing advice on how to link nutrient concentrations in surface waters to specific policy objectives. Elevated nutrient concentrations are a major factor contributing to the failure of many water bodies to achieve GES; however, the links between these nutrients and ecosystem functioning are complex. This creates uncertainty in relationships between biology and nutrients and, in turn, creates difficulties in setting realistic targets for nutrient concentrations that would enable GES to be achieved.

The Water Framework Directive (WFD) specifies that to be at GES 'nutrient concentrations [should] not exceed the levels established so as to ensure the functioning of the ecosystem and the achievement of the values specified for the biological quality elements (WFD Annex V, Section 1.2). This places the onus on Member States to determine levels appropriate to their own territories. However, a review of these values revealed that a wide range of concentrations are currently being used (Phillips and Pitt, 2015; Dworak et al., 2016). Some of this variation will reflect the substantial differences in background concentrations and the sensitivities of water bodies to nutrient enrichment that exist within and between Member States. However, it is also possible that the variety of methods used to set boundaries and the inherent uncertainties, coupled with different regulatory regimes, have combined to create nutrient targets that are likely to make it more difficult for Member States to achieve GES in water bodies.

One major achievement of WFD implementation has been the establishment of a common view of ecological status through the intercalibration (IC) exercise (Birk et al., 2013; Poikane et al., 2015). This has ensured that the concepts of ecological status are transferable between groups of organisms (fish, invertebrates, macrophytes, algae, etc.) and between the Member States of the European Union (EU). This in turn provides a robust view of GES (and other status classes) that can be used as the starting point for the development of nutrient boundaries through the development of statistical models relating nutrient concentrations to ecological status assessed using biology.

The Marine Strategy Framework directive (MSFD) also set criteria to deal with eutrophication (descriptor 5) in order to achieve 'Good Environmental Status' in marine waters. The MSFD requires that 'human-induced eutrophication is minimised, especially adverse effects thereof, such as losses in biodiversity, ecosystem degradation, harmful algal blooms and oxygen deficiency in bottom waters'. Many of the indicators adopted by this directive are already used for assessing specific biological quality elements in the WFD (e.g. phytoplankton or aquatic flora) or as supporting indicators in the WFD, for example: nutrients or chlorophyll *a* in the water column; dissolved oxygen in the bottom of the water column; opportunistic macroalgae in benthic habitats; or photic limit (transparency) of the water column (see European Commission, 2017 for the full set of D5 indicators).

Commission Decision 2017/848/EU (European Commission, 2017) stipulates that the assessment of eutrophication in marine waters needs to take account of the assessment of coastal and transitional waters under the WFD. This implies that in coastal waters, the criteria for MSFD assessments should be selected in accordance with WFD best practices. Marine experts have considered that where the WFD assessed status is clearly related to nutrient pressure, then the WFD assessment could be adopted as MSFD descriptor 5 assessments where the jurisdictions of the two directives overlap (Palialexis et al., 2016). However concerns were raised in relation to situations in which pressures other than nutrients are observed or when GES assessments would not include all relevant aspects of eutrophication (e.g. high nutrient concentration without direct or indirect effects resulting in transboundary nutrient transport and eutrophication elsewhere; Palialexis et al., 2016). Finally, the different assessment time period/cycle and the level of ambition between the two directives have also been identified as possible obstacles for directly transposing the WFD status assessment to the MSFD. A recent review of the MSFD's implementation by Member States identified the need to set quantifiable boundaries for nutrient concentrations in the water column, where possible (Hoepffner and Palialexis, 2015). Work presented in this document may overcome some of these concerns and contribute to better integration between the two directives, particularly where their jurisdictions overlap.

Establishing the nutrient concentration that would support GES is the responsibility of individual Member States (although international collaboration will be necessary in the case of water bodies that span territorial boundaries). However, sufficient experience has now accumulated that it is possible to provide some general guidance on the steps required to test existing views of supporting nutrient concentrations and to develop new ones. Key lessons learned are that water bodies respond differently to nutrient enrichment depending on category, type and geographical location, and that the influence of confounding factors on the underlying nutrient–biology relationship can also vary considerably. This means that there is no single prescription for developing relationships with nutrients. Instead, the document provides a range of options to cover different eventualities. It is also recognised that situations will exist where, for various reasons, relationships cannot be developed within a single Member State using the methods presented here. In some cases, collaboration between neighbouring Member States will be necessary; in others, novel approaches may be required. There are also situations in which multiple pressures influence biological status, which are common in rivers and transitional waters. There is currently insufficient understanding of this issue to provide firm guidance, but methods that can be used to quantify relationships are demonstrated and their interpretations discussed. One consequence of multiple pressures is that it is currently impossible to separate clearly the influence of nutrients from other pressures, and in this situation Member States will need to consider carefully how the relationship between nutrient concentration and ecological status is translated to management thresholds that can aid the delivery of WFD objectives.

How Member States implement the concept of nutrient concentrations which support GES is beyond the scope of this document. Rather it aims to achieve a clearer understanding of the problem and of methods that can quantify uncertainty, and how these can be used to support the aims of the WFD. For this reason this document refers to values that can be associated with a specific likelihood of biological quality elements achieving GES as 'boundary' concentrations, rather than 'targets'. The word 'standard' is used only where a value already has a specific regulatory purpose.

1.2. Development of a toolkit

This document is designed to assist Member States in determining the concentrations of phosphorus and nitrogen that are likely to support GES. It can be used to check existing boundary values or to develop new ones.

The document is supported by a toolkit, which provides the statistical models, in the form of both a Microsoft (MS) Excel ([1]) workbook and a series of scripts which can be run using R, an open-source language widely used for statistical analysis and graphical presentation (R Development Core Team, 2016). The toolkit provides the full R code, together with a series of examples which can be used to explore the methods. A web-based 'Shiny' application([2]) has also been developed, providing an interactive interface to the R scripts, accessible to those unfamiliar with the R environment. Details of how to obtain these can be found on page 4 of this document.

The toolkit has been subjected to extensive testing by Member State experts from across all water categories (lakes, rivers, transitional and coastal waters).

The key principles that need to be considered are highlighted in Chapter 2 of this document, while Chapter 3 provides a brief overview of the process and Chapter 4 introduces the analyses that are included in the toolkit.

1.3. Limitations

Management of eutrophication is not straightforward. There are many pathways of nutrients from land to water and the biological response can take many forms, not all of which will be captured by the methods currently used for ecological assessment. There may also be a time lag before the effect of an increase (or decrease) in nutrients is manifest in the biota. Establishment of a nutrient boundary inevitably requires working at a regional or national level, and nutrient management within individual catchments may require a more nuanced approach. Furthermore, the data sets used to develop relationships between biology and nutrients are mostly developed from spatial surveys which assume biology and nutrients to be in equilibrium, and make a 'space-for-time substitution' (i.e. that variation observed between sites will translate into actual variation over time, if nutrient concentrations change). The rate at which a response to change is seen may well vary between groups of organisms. While this is a benefit when using multiple indicators it is yet another complication when establishing nutrient boundaries and comparing the values established by different Member States for the same type of water body. Thus, for a variety of reasons the relationship between nutrients and biological status is uncertain. Recognition of this is crucial as it influences the selection of the data used, the modelling approaches adopted and the interpretation of the results.

Chapter 2 outlines the key principles underlying the procedures described in this manual. In essence, setting nutrient boundaries requires that a causal relationship between the nutrient ('pressure') and a biological variable ('response') can be expressed in a statistically meaningful manner. In practice, there is a high level of variability in both 'pressure' and 'response' that needs to be accounted for during any calculations. Moreover, nutrient pressures rarely exist in isolation, thus users also have to

---

([1]) Requires Excel 2007 or later

([2]) The Shiny application was a parallel development and may require further work as it had not been as widely tested by users at the time of publication.

be aware of the consequences of interactions with other stressors influencing the outcome, and that an apparently significant relationship may not necessarily indicate underlying causality. In particular, multiple stressors are emerging as an important issue for rivers and transitional waters, and currently there is no simple statistical solution that can overcome the issue. **The tools described here are designed to help ecologists who already have a good understanding of processes at play in their water bodies, and are not a substitute for that ecological understanding.** In particular we advocate validation of boundary values using independent evidence (see Chapter 6).

Several statistical methods are available and we provide (Chapter 3) a guide to selecting the most appropriate (a 'road map'). However, in most cases more effective boundaries will be obtained when several alternative models are compared. This may include different statistical models generated from a single biological quality element (BQE), but should also consider how other BQEs respond. Combining predicted values using a mean or minimum could be appropriate, particularly when statistical relationships are weak.

Problems encountered while establishing transferable good status concepts during IC means that, in some cases, nutrient concentrations were used, directly or indirectly, as part of the process (e.g. Birk and Hering, 2008). This could introduce circularity into the process of establishing nutrient boundaries. Under such circumstances there is perhaps little to be gained by making statistical relationships between biological status and nutrient concentration, and it would be more important that thresholds be set using other approaches (see Chapter 5) or validated by independent means (see Chapter 6).

2. Key principles

2.1. Ecological

The origins of this work can be traced back to attempts to classify the trophic state of lakes based on their nutrient concentrations (Vollenweider, 1968). Importantly, even at this early stage Vollenweider recognised that the type of lake was important, suggesting 'permissible' (= oligotrophic below this level) and 'critical' (= eutrophic above this level) loadings of nitrogen and phosphorus based on the mean depth of the lake. This work was developed by scientists (including Vollenweider) working under the auspices of the Organisation for Economic Cooperation and Development (Anon, 1982) both to incorporate additional variables into the classification scheme and to develop an 'open boundary' system that reflected the uncertainties involved in the classification process (see Harper, 1992). Many of the elements of the current work were already in place at this early stage, including international cooperation, the use of regression analysis to establish relationships and recognition of the need to incorporate statistical uncertainty and for boundaries to be specific to particular types of lakes.

The underlying assumption, derived initially from studies on phytoplankton in deep lakes, was that a causal relationship between nutrients and biology could be encapsulated in a straightforward linear regression model. This, however, proved to be optimistic, particularly in shallow lakes, where macrophytes make a substantial contribution to primary production and habitat structure. The result is complex food webs, the interactions between which confound the straightforward cause–effect relationship between nutrients and phytoplankton that was central to the Organisation for Economic Cooperation and Development's approach (Anon, 1982). The so-called alternative stable states hypothesis (e.g. Dent et al., 2002) emphasises the role of factors other than nutrients in determining the ecological condition of shallow lakes. While nutrient control is still a prerequisite for lake recovery, a bivariate regression will, inevitably, account for less of the total variation in a biological variable than would be the case for deep lakes.

A further complication in many ecosystems is that nutrients may not be the only pressure influencing the biota (Nõges et al., 2016; Hering et al., 2014). A significant linear regression may indicate a biological response to a single causal agent, but there may also be other stressors that correlate with the variable of interest, the effect of which on the biological response can be difficult to disentangle. There are two issues here: first, a pressure–response relationship may appear to be statistically stronger than it really is, as the variable of interest is one of a number of variables contributing to a compound pressure gradient and, second, the response of that variable may be modified by the presence of other stressors, which may act antagonistically, synergistically or hierarchically (Vinebrooke et al., 2004; Wagenhoff et al., 2011; Piggott et al., 2015; Gunderson et al., 2016). Page et al. (2012) and Harris and Heathwaite (2012) emphasise the complexity of interactions between nutrients, biota and other stressors that may arise and the problems that this creates when attempting to derive meaningful boundary values. Feld et al. (2016) provide a 'cookbook' aimed at understanding the hierarchy of stressors and their interactions within data sets, and this may be a useful preliminary step before the methods described here are used. However, the reality is that in many situations the relationships between nutrients and biological status is uncertain, and thus potentially a relatively wide range of concentrations could be used as boundaries.

Two conclusions can be drawn from this brief overview.

1. The effective use of methods in this manual will depend upon a good understanding of local ecology and, in particular, the relative importance of nutrients compared to other pressures.

2. There is a limit to the sensitivity of the methods described in this manual. They are necessary for strategic planning, for prioritising water bodies and for exploring the costs and benefits of nutrient control options. **Users need to remember that nutrient boundaries are intended to help Member States achieve GES, bearing in mind that classification should be based on a combination of biological and physicochemical quality elements (WFD Annex V, Section 1.4.2).**

2.2. Regulatory

Annex V of the WFD states that, for good ecological status, 'nutrient concentrations [should not] exceed the levels established so as to ensure the functioning of the ecosystem and the achievement of the values specified […] for the biological quality element'. Decisions about how to set boundaries will also be influenced by how these are used, and this also differs between Member States. These issues are addressed in the CIS eutrophication guidance document (paragraphs 139 and 187) and also in the CIS classification guidance document (Chapter 4).

Using the analogy of a car dashboard, the BQEs are equivalent to the speedometer, giving the driver an indication of their performance in relation to ecological status boundaries (equivalent to the legal 'speed limit'), while the supporting elements represent the other dials and warning lights on the dashboard that allow the driver to diagnose possible reasons for the biological 'engine' not running as smoothly as desired. A low ecological quality ratio (EQR) for a nutrient-sensitive BQE may indicate a problem with nutrients, but there may be other explanations (macrophytes, for example, are sensitive to both nutrients and hydromorphology). The combination of a low EQR **and** exceedance of a nutrient boundary, therefore, is a stronger indication that nutrients may be responsible for the failure to achieve GES than a low macrophyte EQR alone. This allows broad-scale overviews of problems and the likely costs for dealing with these to be established at a regional or national level. What the exceedance of supporting element boundaries does not do is provide an unambiguous indication that the ecological status of any particular water body is compromised solely by one pressure and that others may not also play a role.

The methods described in this manual are designed to minimise the mismatch between classifications based on biology and nutrients. The key word is 'minimise', as this mismatch is unlikely to be eliminated entirely and, therefore, the regulators have to incorporate residual uncertainty into their decision-making. The situation is most easily envisaged when best-fit regression lines are used (Figure 2-1). The supporting element boundary for good status may be set at the point where the biological boundary intersects the chemistry (Figure 2-1a) or at a position above or below this point (the upper or lower 95 % confidence limit, for example). The use of the upper line gives a low probability of restoring water bodies back to good status, but minimises the risk of a water body being wrongly downgraded (i.e. chemical boundary is exceeded, while biology is still at good status (Figure 2-1b). The lower line is more precautionary, giving a high probability of restoring water bodies back to good status, but will result in more water bodies being wrongly downgraded (Figure 2-1c). There are, in other words, trade-offs between the 'false positives' and 'false negatives' ('type 1' and 'type 2' errors) that a particular boundary will produce.

The scale of this problem will increase as the predictive power of the regression equation decreases. In rivers, transitional waters and coastal waters many more pressures other than nutrients influence biological status than in lakes, particularly when BQEs that do not react directly to nutrients (e.g. invertebrates) are considered. In these cases, relationships between nutrient concentration and biological status have a very high level of uncertainty. A scatter plot may show a 'wedge'-type relationship to which an upper-quantile line can be fitted which provides an estimate of the highest level of nutrient that is theoretically consistent with good status (Figure 2-2a). However, if other pressures (e.g. barriers increasing water-retention time in rivers, removal of shade in riparian areas or warming by climate change) enhance the sensitivity to nutrients, such an upper-quantile line will not be consistent with good status. An inverted wedge (Figure 2-2b) can also occur where other factors mitigate the effect of nutrient enrichment. In a lake this might be grazing by zooplankton; in rivers and estuaries it might be shade or flow reducing primary production, or the toxic effects of herbicides or metals. In some cases, particularly in rivers and estuaries, both effects may occur, resulting in a poor relationship with a single pressure.

(a) (b) (c)

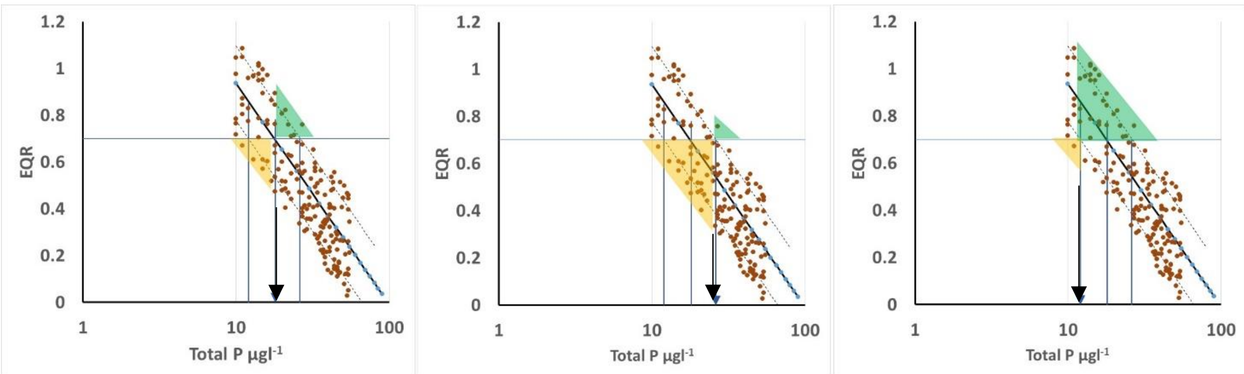


Figure 2-1: Hypothetical relationship between total phosphorus and biological EQR, showing regression line with confidence intervals (dotted lines). Horizontal line shows the biological good/moderate boundary (0.7 in this example); vertical lines show intersection with regression line ± confidence intervals marking potential good/moderate boundary values for total phosphorus using (a) intersection with best-fit line, (b) upper confidence line, (c) lower confidence line. Triangles mark areas where classification mismatches occur, green (biology good but phosphorus moderate) and yellow (biology moderate or worse but phosphorus good) using three different approaches to interpretation.

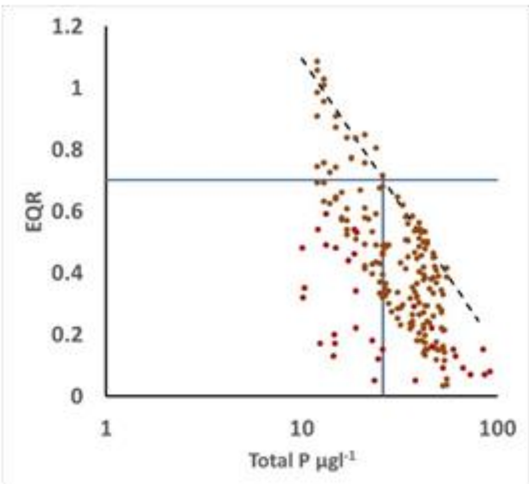
Box 1: Points to consider when developing nutrient thresholds: ecology

- Do you have an intercalibrated metric for a biological quality element that is known to respond to nutrients?
    - Intercalibration ensures that ecological status concepts are consistent amongst Member States.
        - If nutrients were used directly or indirectly to establish good status concepts then thresholds must be validated by independent means (see Chapter 6).
    - Methods may still be suitable in situations where intercalibrated metrics are not available; however, extra care will be needed.
    - Is there independent evidence to demonstrate a causal relationship, rather than just a statistical association?
    - If there is more than one BQE with a causal relationship to nutrients they should be tested separately. It is also possible that a stronger relationship will be obtained if BQEs are combined rather than by using them separately. An appropriate combination rule (typically the average or minimum of the constituents) will need to be devised.
- What other pressures may confound understanding of the pressure–response relationship?
    - Choose the biological metric that is most sensitive to nutrients (or, conversely, least sensitive to other pressures), for example phytobenthos rather than macrophytes in some rivers.
    - Be aware that when multiple pressures are present simple statistical methods are likely to underestimate boundary values, as the other pressure may reduce the observed EQR more than it would as a result of the nutrient pressure. However, the alternatives (use of upper quantiles, either regression-based or categorical) may overestimate the boundary value as it is impossible to estimate the uncertainty of the true relationship. The use of these methods will require value judgements.
    - Is it valid to assume a 'dose–response' relationship or are there other (e.g. top-down) factors (e.g. grazing, shade or discharge) that may affect the expression of eutrophication in a water body?
    - The use of these methods will require knowledge of the nature of the interactions between the different pressures and their relative importance (guidance and tools from the EU-funded MARS project are available at http://fis.freshwatertools.eu/index.php/guidance.html).
- Value judgements should also be included in setting the thresholds, including considerations that using upper-quantile thresholds as management targets will result in more nutrients moving downstream and increase the risk of not achieving good status in downstream water bodies.

The important point is that, given the uncertainty of even national type-specific relationships, there will always be a range of potential boundary values from which a Member State can choose. The choice will also reflect how the boundary value is used within the Member State, and therefore it is important to recognise this as a further factor influencing variation between national standards. Broadly, the following two strategies may be adopted.

- Action (e.g. programme of measures) is triggered as soon as the nutrient boundary is exceeded. Under such circumstances a higher boundary value may be appropriate, to minimise the instances in which elevated nutrient concentrations trigger measures despite biology being at good status (i.e. Figure 2-1b). Additional considerations that should be made before measures are implemented include current proximity to the good/moderate (G/M) boundary for sensitive BQEs, likely trends in nutrient concentrations if no action is taken and whether there are other factors that might reduce the sensitivity to nutrients (e.g. shade, high flow, grazing, toxic substances).
- An exceedance of the nutrient boundary is one of a number of strands of evidence that are considered before a programme of measures is triggered. Under such circumstances a more precautionary (lower-concentration) boundary value may be selected (e.g. Figure 2-1c); however, the regulator would then check that that a nutrient-sensitive BQE was also failing in the water body under consideration prior to taking action, or that there was evidence that it might do so in the future (e.g. if there was a trend of increasing nutrient concentrations, or the likelihood of an increased sensitivity to nutrients due to climate change, removal of shade, lower flow or reduced grazing pressures).

(a)



(b)

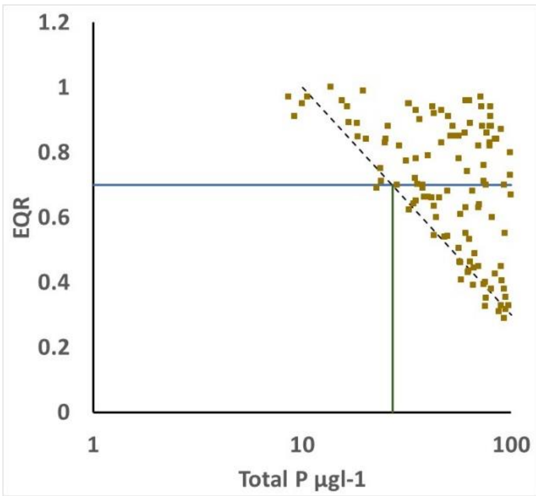


Figure 2-2: Hypothetical relationship between total phosphorus and biological EQR where multiple pressures occur. Figure 2-2a shows the regression of an upper quantile (e.g. 95th percentile). Figure 2-2b shows the regression of a lower quantile (e.g. 5th percentile). Horizontal lines show the biological good/moderate boundary, vertical lines show intersection with line marking potential good/moderate boundary values for total phosphorus.

This, in turn, raises questions about the role of supporting element boundaries. It is clear that the relationships are not always strong enough to indicate convincing cause–effect relationships between nutrients and BQEs. Indeed, the scale of uncertainty in the relationships is a timely reminder that we are attempting to detect the effect of a single stressor within a multi-stressor environment. There is, nonetheless, a need for regulators to unpick the knot of ecological interactions in order to identify those stressors most likely to be responsible for BQE failures.

Box 2: Points to consider when developing nutrient thresholds: regulation
❖ How are nutrient boundaries used in enforcement and regulation? Is the purpose to ... ➢ maximise the chance that biology achieves good status? ➢ minimise the chance that nutrient status is worse than biological status? ➢ minimise the mismatch between nutrient and biological status classes?

2.3. Chemical

General considerations

The previous section discussed ‘pressure’ in general terms. In practice, pressure, for the purpose of this manual, can be defined more specifically as an excess supply of the limiting nutrient. This, in turn, presumes that the limiting nutrient is known and has been measured (and summarised) in an appropriate way (see below). In both cases, users need to consider the quality of the insights and the representativeness of measurements.

In practical terms, any consistent approach to sampling and analysis of nutrients provides a basis for comparisons within a region, especially if this approach can be justified via a good understanding of ecological dynamics within the region and if users are aware of the limitations. If options are available (e.g. use of annual versus summer average concentration), the test should not be ‘which explanatory variable gives the highest r^2’ but, rather, ‘which variable is most likely to provide a measure of the likely benefits of regulation, considering the problem in its entirety (including nutrient pathways within the catchment and risk of secondary effects)?’ It is, however, important to recognise that differences in approaches to sampling, analysis and averaging of data may complicate comparisons between Member States.

A note on units: in this document we use examples from both fresh and saline waters and we follow the most frequently used conventions with respect to units. Thus for freshwater we use mass units (e.g. $\mu g\ L^{-1}$) while for saline waters we mostly use molar units (e.g. µM or $\mu mol\ L^{-1}$). This may cause confusion, but molar units can be converted to mass units by multiplication of the molar mass (N = 14, P = 32). Thus, a P concentration of 32 $\mu g\ L^{-1}$ is equivalent to 1.0 µM, and for an N concentration of 1.4 $mg\ L^{-1}$ is equivalent to 1.0 mµM.

Freshwaters

The prevailing assumption that phosphorus is the limiting nutrient in freshwaters has been challenged in recent years (Maberley et al., 2002, 2003; Moss et al., 2013) with the consensus now being that both phosphorus and nitrogen are capable of contributing to eutrophication and, therefore, that both need to be managed. Some (e.g. Schindler, 2012; Schindler et al., 2009) have questioned the evidence base used to support the role of nitrogen and argue that ecosystem-scale verification is required before proceeding with expensive polices to reduce nitrogen.

Total phosphorus (TP) is the most widely-used phosphorus parameter for lakes, with a roughly even split in Member States between those who use annual and growing-season means (Phillips and Pitt, 2015). Complications arise because it is not always easy to differentiate the 'pressure' (i.e. the nutrients available to the aquatic biota) from the 'response' of the biota (Schneider et al., 2016). The reality is nuanced, with the biota, in turn, influencing the concentrations of nutrients by direct uptake. A large part of the TP in a lake, for example, may be bound into phytoplankton cells, while aquatic macrophytes are also capable of taking up nutrients from the sediment, and photosynthesis by charophytes can lead to the co-precipitation of phosphorus with calcite (Schneider et al., 2016). In theory, winter concentrations of soluble nutrients should be a good measure of the load available to the biota in lakes, as they are in coastal waters (see Section 2.3.3). However, this parameter is rarely used for regulation in freshwaters.

There is less consistency in rivers, with countries using either total or 'soluble' phosphorus, in a few cases 'total reactive phosphorus' (i.e. phosphorus in unfiltered water that reacts with molybdate). Annual means are preferred but a few Member States use growing-season means (Phillips and Pitt, 2015). There is usually a close relationship between annual and growing-season means for reactive phosphorus (Figure 2-3a), probably because phosphorus uptake is often balanced by sediment release and, particularly where there are point sources, the effect of a continued point source supply along with lower dilution in summer. However, while it is sometimes necessary to exclude data from extreme events (e.g. floods), it should be recognised that annual means are rarely based on sampling regimes with a frequency of greater than once a month while much of the nutrient load, particularly in catchments influenced by diffuse sources, may be delivered in episodic pulses likely to be missed by routine sampling (Ockenden et al., 2016). There is, in addition, evidence that straightforward measurements of 'dissolved' ('soluble', 'filterable' or 'ortho-') phosphorus may underestimate the supply that is available to the biota (Whitton and Neal, 2010). The most appropriate determinand may depend on the type of stream or river under consideration.

The situation for nitrogen is similar to that for phosphorus, with total nitrogen (TN) preferred in lakes and, again, about half of the countries using annual rather than growing-season means (Phillips and Pitt, 2015). For rivers, there is an even split between those using 'dissolved' fractions of inorganic nitrogen and those using TN with, in this case, most preferring to use annual rather than growing-season means**.** In UK rivers, growing-season means of total oxidised nitrogen (TON — mainly nitrate-N) tend to be lower than those for winter means due to the combined effects of biological uptake and reduced supply (Figure 2-3b).

Although measures of central tendency (e.g. mean/median) are recommended, there may also be situations where the 90th percentile of nutrient measurements is appropriate (in Austria, for example, much of the dissolved phosphorus is removed and the 90th percentile was considered to give a better indication of the size of the nutrient pool before algal uptake (Karin Deutsch, pers. comm.)).

(a) (b)

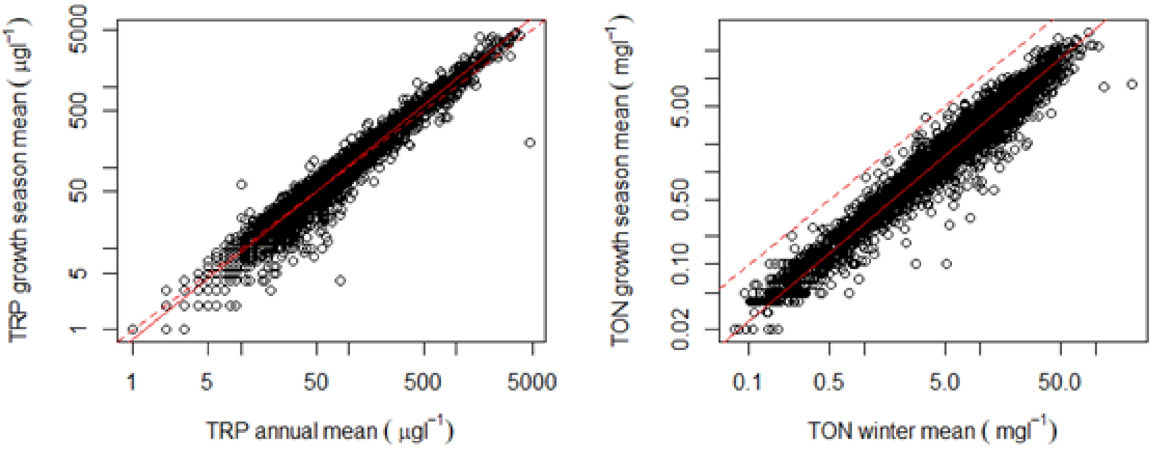


Figure 2-3: Relationship between (a) annual and summer mean values for total reactive phosphorus (TRP) and (b) winter and summer mean values for total oxidised nitrogen ([3]) (TON) from UK rivers. Dotted red line shows 1:1 relationship, solid red line shows type II RMA regression fit.

Transitional, coastal and marine waters

Nitrogen or phosphorus limitation?

There are many differences in the nutrient parameters assessed, the assessment period (summer, winter or all year round, i.e. annual), and in the statistic used (mean, median or 90th percentile) when Member States assess the condition of saline waters. These differences can be observed between Member States, within the four marine ecoregions defined by the MSFD and even within Member States between transitional, coastal and marine waters (Dworak et al., 2016).

Nitrogen, rather than phosphorus, is considered to be the most likely limiting nutrient in many temperate coastal waters (Tsirtsis, 1995). However, a number of exceptions occur (Table 2-1) and, in some situations, other nutrients play a role. Limitation due to silica, for example, can lead to a shift from diatom to flagellate algae which are frequently poor food for grazers, leading, potentially, to rapid increases in algal biomass and associated secondary effects (Officer and Ryther, 1980). Nutrient limitation is also influenced anthropogenically and it can change due to an unbalanced reduction of nitrogen and phosphorus, as documented for the North Sea (Burson et al., 2016).

When looking for correlations between the biology and nutrient concentrations in order to derive nutrient thresholds it is important to know the limiting nutrient, since a strong correlation can only be expected to exist between the limiting nutrient and the biology. However, the patterns of nutrient limitation reviewed in Table 2-1 suggest that determination of the limiting nutrient can be challenging, since limitation can vary on small spatial scales and between seasons. Hence boundary

([3]) TRP is the amount of phosphorus that will react with molybdate in an unfiltered sample of water. In most cases it is very similar to soluble reactive phosphorus but may also include loosely bound forms of P on fine particulate material.
TON = NO_3-N + NO_2-N; in practice, NO_3-N is the dominant form).

values should be established for, and monitoring programmes should consider, both nitrogen and phosphorus.

Table 2-1: Patterns of nutrient limitation in European marine and coastal waters.

Region	Comments
Adriatic Sea	Mostly phosphorus-limited (Rinaldi, 2014). Along the coast of the northern and central Adriatic Sea, 90 % of the overall chlorophyll *a* variability is explained by TP (Giovanardi et al., 2016). The high N:P ratios in the Adriatic sea (>50) demonstrate that nitrogen does not limit algal growth. The Po River has a major effect on the whole Adriatic basin, determining patterns of both spatial and temporal variation.
Atlantic Ocean	Nitrogen is generally assumed to be the nutrient that limits primary productivity in most oceans (Carstensen et al., 2011; Tyrrell 1999); however recent studies show that phosphorus may limit primary production in some areas (Karl et al., 1998; Karl 1999, 2000), particularly in the Loire (Guillaud et al., 2008) and the Gironde plumes (Labry et al., 2002). Furthermore phosphorus limitation is not always due to low phosphorus concentrations but sometimes to high N:P ratios.
Baltic Sea	Generally, nitrogen limits phytoplankton growth in the open and coastal waters of the Baltic Sea. However, depending on the area and season, phosphorus, and in the case of diatoms, silica, can also be limiting nutrients (Helcom, 2009). Bioassay results show primary production to be mostly phosphorus-limited in the Bothnian Bay (Andersson et al., 1996; Tamminen and Andersen, 2007) and mostly nitrogen-limited in the Kattegat (Granéli et al., 1990); nutrient limitation patterns switch during seasons (Tamminen and Andersen, 2007), depending on proximity to freshwater sources (Pitkänen and Tamminen, 1995), and during cyanobacteria blooms (Lignell et al., 2003; Nausch et al., 2004). The relationship between loads and nutrient concentrations is not as simple as for enclosed systems. Loss mechanisms (sedimentation, denitrification) and retention time play key roles but can obscure cause–effect relationships. There is, nonetheless, generally an excess of nitrogen relative to phosphorus, supporting the current findings that the Bothnian Bay is generally phosphorus-limited. In the Bothian Sea, Helcom (2009) showed that the ratio of dissolved inorganic nitrogen (DIN) to dissolved inorganic phosphorus (DIP) in winter has

Region	Comments
	decreased over the past 10 years, suggesting a switch towards more co-limitation in the coastal zone and stronger nitrogen-limitation in the open sea. Meanwhile, in the Baltic proper, the winter DIN:DIP ratio has been decreasing since 1993, indicating increased nitrogen-limitation of the spring phytoplankton bloom.
Black Sea	The influence of rivers decreases to the south along the coast and offshore for most of the year due to photosynthetic consumption of dissolved inorganic nutrients and sedimentation within the north-western and western shelves. The river supply gives rise to a high nitrogen:phosphorus ratio within the north-western shelf that makes phosphorus the primary limiting nutrient along the coastal zone. Weakly nitrogen- or phosphorus-limited systems are associated with the outer shelf, but the interior basin and major part of the sea is strongly nitrogen limited (Black Sea Commission, 2008). The seasonal alternation does not show a regular pattern, but varies between years.
Mediterranean	Phosphorus is often the limiting nutrient (Margalef, 1963; Berland et al., 1980; Lazzari et al., 2016; Thingstad et al., 2005), although it is closely followed by nitrogen in this limiting role (Estrada, 1996). The dissolved nitrogen to phosphorus ratio in the Mediterranean has been reported to be about 21 to 23 in the western part (Bethoux et al., 1992), and even higher in the eastern basin (Krom et al., 1991), which is quite different from the ratio of 15 found in the global ocean (Tyrrell, 1999). Experimental studies in the eastern Mediterranean have always shown phosphorus limitation, both at coastal stations (Bonin et al., 1989) and in deep waters (Berland et al., 1980, Zohary and Robarts, 1998). In addition, experiments performed in the north-west Mediterranean Sea coastal waters showed phosphorus as the limiting factor for phytoplankton growth (Jacques et al., 1973, Fiala et al., 1976). This has been corroborated by the model-based reconstruction of inorganic phosphate and nitrate distributions presented by Lazzari et al. (2016). The model demonstrated that when nutrient limitation occurs, in the vast majority of cases, phosphorus is the limiting nutrient, with the notable exception of the Alboran Sea, which is mainly nitrogen limited, and the southwest basin, in which both

Region	Comments
	nitrogen and phosphorus can limit plankton growth. Ramirez et al. (2005) showed nitrogen-limitation in the upper layers (top 20 m) of the north-west Alboran Sea during the winter, summer and autumn while Dafner et al. (2003) suggested phosphorus limitation in the Strait of Gibraltar area, but phosphorus limitation proposed by these authors in the upper layers of this area is not due to very low phosphorus concentrations but rather to a very high nitrogen:phosphorus ratio to the east of Gibraltar, caused by the upwelling of deep Mediterranean waters with high nitrogen and phosphorus levels relative to silica.

Total or inorganic dissolved fraction? Assessment period? Statistics?

During phytoplankton blooms, dissolved inorganic nutrients in surface layers may be almost completely consumed, leading to nutrient limitation at periods of peak biological activity. This results in large seasonal variability of nutrient concentrations and, for this reason, DIN and DIP are usually measured and assessed during winter, when biological activity is lowest.

Winter DIN is generally regarded to be the single most important contributor to anthropogenically-induced change in phytoplankton communities (Devlin and Bonne, 2016). Winter measurements indicate the size of the nutrient pool available for the spring phytoplankton bloom (Andersen and Conley, 2006). Metrics, therefore, must attempt to match the spring-summer-autumn biological data with concentrations of DIN from the previous winter. Winter DIN is used because biological activity (e.g. phytoplankton growth) is low during winter. Inorganic winter nutrients are, as a result, used very widely (e.g. OSPAR, 2009; European Environment Agency, 2012). There are, however, exceptions: monitoring winter nutrient concentrations is not a good practice for the western coastal areas of the Black Sea, as nutrient concentrations here peak in April-May at the time of highest Danube discharge (Black Sea Commission, 2008).

TN and TP, which include all forms of nitrogen and phosphorus compounds, are also important parameters that should be assessed in addition to the dissolved nutrients, as is already common practice for example by Helcom and in Swedish, Finnish and Estonian coastal waters (Helcom, 2009). The value of adding total nutrients alongside inorganic nutrients as core indicators strengthens the link from nutrient concentrations in the sea to nutrient enrichment. In particular these parameters allow consideration of climate change in the eutrophication assessment since higher temperatures will lead to year-round phytoplankton proliferation and/or possible changes in zooplankton communities. In the Baltic Sea, for example, a decrease in winter DIN concentrations has been identified in the Bornholm Basin since the 1990s, but TN concentrations have remained high. A possible reason for this observation could be that in winter more nutrients are bound in the phytoplankton due to the higher water temperatures. In such a situation, assessing only dissolved inorganic concentrations gives the wrong impression that nutrient concentrations seem to be declining, while, in fact, they are stable or increasing as can be seen when also assessing total concentrations (Helcom, 2017).

In addition, there are other considerations that are not directly linked to setting nutrient thresholds but that are nevertheless important. Total nutrients are essential for determining nutrient budgets (an estimation of how much nutrient enters and leaves an area). Such budgets have particular importance in coastal and marine waters that are influenced by transboundary nutrient transport and receive nutrient inputs from other countries. This is a common situation in the North Sea. Furthermore, total nutrients are also essential parameters for establishing nutrient reduction targets. This means that monitoring and assessing both total and dissolved nutrients is necessary if a good understanding of the trend in nutrient concentrations in the marine environment is to be obtained.

To enable a consistent management approach, it is important that the nutrient parameters that are monitored and assessed are consistent between transitional, coastal and marine waters, at least within a region or subregion.

A final consideration is the choice of statistical measures used to aggregate nutrient samples from a chosen assessment period to determine indicator concentrations. Most of the Member States and Regional Seas Conventions use mean concentrations and, in order to aid comparisons across Europe, we recommend this. However, there might be cases where using the median is more robust, since it is less influenced by outliers. The choice of the appropriate statistics depends very much upon sampling size and monitoring quality.

Box 3: Points to consider when developing nutrient thresholds: chemistry

- How strong is the evidence that the nutrient of interest limits primary production in the water bodies under consideration?
- Are nutrients analysed using reputable laboratories with thorough quality assurance procedures?
- Is the detection limit for the method appropriate for the water body type? (There should not be a high proportion of values below the detection limit.)
- How well does the choice of variable (i.e. total versus soluble/filtered), and the summary metric used (spot value, average, summer average etc.) capture the range of variability likely to be encountered within a water body?

2.4. Statistical

Practical issues for setting nutrient thresholds

Before developing nutrient boundaries, a data set needs to be established with, at the very least, nutrient concentrations and classifications, but ideally metric values, of nutrient-sensitive BQEs. It is worth plotting some preliminary graphs to visualise the relationship between nutrients and biology. Where long runs of historical data are available, it may be possible to detect change points that can be used to establish nutrient thresholds (see Section 5.1); however, in most cases, it will be more appropriate to use spatial data sets.

The key criterion in establishing robust nutrient thresholds is that the data span a long gradient, ideally at least four status classes, in order to ensure that the response of the biological variable is understood across the status spectrum. This is important as relationships are not necessarily linear throughout their range, and the validity of any fitted lines can then be checked by eye. As the focus of this work is on establishing thresholds for high and good status, biology must span at least high, good and moderate status. This may be achieved within a single national type, but there may be situations where national types can be merged (i.e. using 'broad types') or where collaboration with neighbouring countries may be the most productive option. It is also important that there is not a high proportion of 'less than' values in the data set as these constitute 'censored' data which incorrectly 'anchor' regression relationships at a single lower point on the pressure gradient, exerting undue influence on the modelled gradient. Where this is the case specialist advice should be obtained, as there are methods for fitting models to censored data which are beyond the scope of this technical document and toolkit.

The best method for determining nutrient thresholds is to use regression models fitted to data that span the pressure gradient. An alternative is to use categorical methods. These may be less sensitive to the requirements of the regression model (e.g. linearity); however, they are sensitive to class width and assume a robust assessment of biological status. They also have only limited ability to illustrate how the biological status changes across the pressure gradient. They are however intuitively simple to understand and may be the best approach where relationships are weak or the gradient is short, provided there are significant differences between the distribution of nutrient concentrations in adjacent classes.

In most cases a better categorical approach is to fit a binomial logistic model to data that are classified into two groups (e.g. good or better and moderate or worse, in the case of the G/M boundary). This type of statistical model also allows uncertainty to be assessed, although it is important to note that this is still a categorical method and is thus dependent on a robust delineation of biological status.

The simple regression-based methods assume a linear response between the variables ([4]) (Table 2-2). This can often be achieved by log transformation of nutrient concentration data; however, even after this, visual inspection may reveal nonlinearity, often with sigmoid responses (i.e. with regions at the extremes of the distribution, where there is little response of the biology to changed concentrations of nutrients). In the statistical toolkit, segmented regression methods are provided to test for linearity within the data set; users need to be sure that the thresholds of interest are captured within the linear portion of the graph.

In many data sets, particularly from rivers or transitional and coastal waters, scatter plots often suggest a wedge-shaped relationship, where the range of EQRs increases at one end of the pressure gradient (see Figure 2-2). This is most likely a result of other pressures causing increased or decreased sensitivity of BQEs to nutrients. In this situation it is inappropriate to fit a regression model, as the requirement of normally distributed residuals would clearly not be met. A potential solution could be to remove sites thought to be influenced by other pressures, or at least to stratify the data into similar levels of other pressures. Where this is not possible, fitting a regression line to a quantile of the data is likely to be a more appropriate method. However, the choice of quantile

([4]) Nonlinear models — including generalised additive models — could also be used, but are beyond the scope of this document, although the principles outlined here would be the same.

needs to be considered carefully as this will influence the relative precaution of predicted nutrient thresholds. The percentile used will be a trade-off between the relative influence of other pressures and the uncertainty (error) in the true biological response to nutrients. A full consideration of fitting quantile models is beyond the scope of the current toolkit, but given the emerging importance of multiple pressures we have included an example R script (see Appendix 1, Section A4.2.2).

Truncating EQRs at 1.0 (which is standard practice in many countries) will influence the statistical properties of the relationship and introduce curvature in the response. When working with data from a single type, the denominator for EQR calculations should be constant (except in a few cases where site-specific reference values are used) and, therefore, it may be preferable to work with raw metric values rather than EQRs. This could cause complications if several countries are collaborating to produce a common nutrient boundary and clearly requires a common view of the metric boundary value. Where EQRs are used, the effect of 'normalisation' of the EQR scale on the properties of the relationship (to ensure that class boundaries are equidistant) also needs to be considered. Was normalisation applied across the whole EQR gradient or only at the upper end to constrain values greater than 1? In the latter case, intercalibrated boundaries would still apply and non-normalised data should be used to develop relationships. However, in some situations (e.g. when several measures are combined to form a multimetric index) normalisation will apply across the gradient and, in such cases, normalised metrics will be necessary in order to derive nutrient thresholds. Combining data sets from several countries to extend pressure gradients will also require data to be normalised. Various methods can be used, but the piecewise linear transformations developed for IC are perhaps the easiest to use (see MS Excel TKit_Normalise.xlsx SectionA2).

In general, large data sets are more likely to yield robust predictions than small ones. However, this may not be possible if there are few water bodies of a particular type within a territory. One way around this problem, widely used in lakes, is to use data from several years, treating each as a separate data point in a relationship. However, each of the 'lake years' will not be strictly independent from other data points from the same lake and users need to be aware of the dangers of 'pseudoreplication' (Hurlbert, 1984). It is not possible to provide a simple ratio of 'lakes' to 'lake years' that is acceptable, however it is important that the data are not dominated by a minority of sites with multiple years, or these will over-influence the relationships. A dialogue between an experienced ecologist and a trained statistician may be essential. Collaboration with nearby countries with similar types of water body is a further option.

Box 4: Characteristics of an ideal data set

- Data span a long gradient, ideally high to moderate/poor, if shorter data are then evenly balanced around the boundary of interest.
- Nutrient concentrations are a summary metric, preferably one that measures central tendency, e.g. a mean or median value, although other metrics (e.g. 90th percentile) may be appropriate in some situations.
- Data using replicate years for the same water body should be avoided. If multiple years are used, there are a similar number of years from each water body.
- Summary metrics should exclude extreme events, e.g. floods.
- Biological and nutrient data for a water body are from the same (or a similar) year.
- Biological data do not use truncated EQRs (i.e. values > 1.0 are allowed).
- There are only a small proportion of 'less than' values for nutrient concentrations contributing to the summary metric.
- The data set consists of sufficient records. It is difficult to be specific, but ideally ≥ 50.

Type I or type II linear regression?

Regression models allow the relationship between nutrients and biological status to be established. However, one of the issues with the use of regression is that ordinary least squares (OLS) regression minimises the variation in the dependent variable and thus assumes no uncertainty in the predictor variable. This is often the case for experimental studies, but unlikely to be so when using data from monitoring programmes such as are used for the WFD. Thus, when using OLS regression to quantify the relationship between nutrient concentration and biological status we have to make a choice concerning whether biological status (EQR) or nutrient concentration are considered to be the dependent variable. The choice of the dependent variable is important — where both variables contain error, an OLS regression will underestimate the true slope of the relationship (Legendre, 2013) and thus influence the nutrient concentration we determine for the biological boundary.

As the purpose of the model is to predict the nutrient concentration that occurs at a given ecological status, for example the G/M boundary, it might be logical to make the dependent (y) variable nutrient concentration, with biological status as the independent (x) variable. However, when considering the relationship between nutrients and biological status we generally assume that the nutrient concentration 'causes' the ecological status, which is why we seek to establish the nutrient concentrations that will support good status. Thus, it is also logical to make the dependent variable biological status, predicted from nutrient status, with boundary values subsequently determined by rearranging the regression equation. The fact that nutrient concentrations are also influenced by the biology through uptake should not be completely ignored, especially when dissolved inorganic nutrients are used in the regression.

Table 2-2: Advantages and disadvantages of statistical approaches used in this manual.

Method	Advantages/disadvantages
Linear regression	• Less reliant on class width than categorical methods (see below). • Requires linearity, at least in the region around which thresholds are being inferred. • Least sensitive to position of data cloud relative to boundary of interest.
Quantile regression	• Allows lines to be fitted that define the edges of the data cloud, which can be used to allow for the influence of other pressures or environmental factors. • No objective way to determine quantile used as there is currently insufficient understanding of pressure interactions relating to nutrients (guidance is currently being drafted and will be made available at: http://fis.freshwatertools.eu/index.php/guidance.html). • Requires a value judgement as, if an upper quantile is used, the approach risks setting too high a nutrient boundary value by overestimating the influence of other pressures. • Least sensitive to position of data cloud relative to boundary of interest.
Categorical methods 1: using boxplots	• Less dependent upon linearity. • Requires a significant difference between nutrient concentrations in adjacent classes. • Establishes thresholds based on quantiles, so needs ample data points spread throughout the classes around the boundary of interest to ensure robust estimates of parameters. Width of class can also influence position of quantiles. • Sensitive to position of data cloud, relative to the boundary of interest.
Categorical methods 2: binomial logistic regression	• Allows estimates of boundary values for different probability of class. • Potentially appropriate for multiple pressures, by use of higher probability of class. • Uncertainty assessment is possible. • Less sensitive to position of data cloud relative to boundary of interest.
Categorical methods 3: decision trees	• Simple to interpret. • Less dependent upon linearity and outliers. • Appropriate for multiple pressures • Allows importance of other pressures to be assessed
Categorical methods 4: Mis-match approach	• Simple to understand. • Excel tool is unable to estimate uncertainty of thresholds, however R script using bootstrapping overcomes this. • More sensitive to position of data cloud relative to boundary of interest than logistic regression, but less sensitive than boxplot methods.

The choice of regression approach depends on the degree of asymmetry in the relative uncertainty of the dependent and predictor variable (McArdle, 2003; Smith, 2009). It is clear that estimates of both the biological EQR and nutrient concentration will contain errors due to sampling, however this is not the only source of uncertainty we need to consider. In addition, the uncertainty in the relationship between nutrients and biology — sometimes called 'equation error' — also needs to be taken into account (McArdle, 2003). As other environmental factors also influence the biology, the relationship between nutrients and biology is likely to be asymmetric in relation to uncertainty, increasing the error of the EQR. It is reasonable to conclude, therefore, that the total uncertainty in the biological EQR is often greater than that of nutrients. However, the issue is whether it is 'much greater', as required for the use of OLS regression. Where r^2 values are high (>0.6) there is little practical difference in the nutrient boundaries resulting from a regression of EQR on nutrient or nutrient on EQR, but for less certain relationships the differences are more substantial ([5]).

The alternative is to use a type II regression (Sokal and Rohlf, 1995), where the fitting procedure minimises the variation of both dependent and independent variables. The disadvantages of a type II regression are that it is less appropriate where the purpose of the model is to make predictions (Legendre and Legendre, 2012); it is more difficult to interpret uncertainty (Smith, 2009); it is less readily available in statistical software and it can only be used with a single predictor variable. It is also important to only apply type II regression to relationships with a correlation of at least 0.6 (r^2 = 0.36) (Jolicoeur, 1990 cited in Smith, 2009) as some methods will generate a line with a slope significantly different to zero from random data. It should be noted however, that if the boundary EQR being predicted is close to the mean EQR of the data, the choice of regression method will have little effect as both type I (i.e. OLS regression) and type II fitted lines will pass through the mean of x and y.

Quantile regression

Linear regression is unlikely to yield meaningful results in the situations illustrated in Figure 2-2, which will occur when nutrient-biology interactions are confounded by other stressors. In such cases variance around the mean of the response variable is itself a function of the explanatory variable, leading to a wedge-shaped distribution. Under these circumstances, quantile regression may be more appropriate. This is a variant of conventional least squares regression analysis. Whereas least squares regression aims to predict the mean of the response variable for a given value of the predictor variable, quantile regression aims to predict different aspects of the statistical dispersion of points. In particular, there are many situations in ecology where it is useful to understand rates of change in the response variable along the upper or lower boundary of the conditional distribution of values (Cade and Noon, 2003).

Quantile regression is beyond the scope of this toolkit. It can be implemented through packages such as 'quantreg' (Koenker, 2016) within R and we have provided some example scripts in Appendix 1 Section A4.2.2 that could be adapted for other uses. The values produced by an upper quantile of a relationship between EQR and nutrients will be inherently less precautionary than those produced by the least squares approaches detailed above, so they need to be interpreted with caution. In effect, an upper quantile defines the maximum value of a response variable for a given value of the

(5) The slope of the OLS of a regression of x on y is given by r^2/slope of y on x, so the difference in slopes when expressed relative to the same axis is proportional to the reciprocal of r^2.

explanatory variable and is useful where one or more additional pressures drive the response variable and override the influence of nutrients. Such pressure(s), by lowering the value of the response variable (EQR, in this case) that is observed for a given value of the explanatory variable (e.g. nutrient concentration) (Figure 2-2a), could be wrongly interpreted as indicating a response to nutrients. However, there are also situations that may reduce the BQE sensitivity to nutrients, preventing its full expression, e.g. shade, toxic substances, flushing or grazing. In this case, the use of a lower quantile, which will produce a precautionary boundary value, may be appropriate.

While quantile regression may describe the relationship between nutrients and biology, its use in setting thresholds needs to be considered with care. A wedge-shaped distribution might, for example, indicate that nutrients are not the primary factor influencing the biota for sites included in the data set. This, in turn, might provoke investigations into the role of other stressors and better regulation of these might need to take priority over nutrient control. The upper quantile will, nonetheless, provide a value that can serve as an interim target and, more importantly, to prevent deterioration of sites. The confidence with which the slope and intercept can be estimated will decrease towards the extreme of the distribution, due to a likely variation of the 'conditional density of the response' (Koenker, 2011). The selection of the quantile to use for boundary setting is essentially a value judgement, partially conditioned by data distribution, but it should be based on knowledge of the importance of nutrients versus other pressures and their interactions affecting the nutrient sensitivity of the BQEs. In general, where the nutrients are considered to be the most important pressure and thus control biological status, a lower value for the upper quantile is needed. It is difficult to recommend a threshold, but we suggest that values of the 25th and 75th percentiles are most likely to be appropriate for data with inverted wedge- or wedge-shaped scatter plots, respectively. Where an upper-quantile approach is used, leading to less precautionary thresholds, the validation step (Chapter 6) is particularly important.

Categorical approaches

The previous section uses EQRs to derive nutrient thresholds. In practice, ecological status assessment collapses the EQR, a continuous variable, into five categorical ecological status classes and it is also possible to derive nutrient thresholds directly from these. The phosphorus or nitrogen concentrations associated with a particular ecological status class (e.g. GES) could be expressed as a distribution from which an upper quantile might be chosen to indicate a phosphorus or nitrogen concentration above which GES was very unlikely to be achieved, or a lower quantile below which GES was very likely to be achieved, if nutrients are the main drivers of status. However, the variation inherent in biology-nutrient relationships means that there will be many instances where lower concentrations of nutrients are not associated with GES. The risk of misclassification could, therefore, be reduced by also considering the distribution of nutrient concentrations in the adjacent class (moderate, in this case), where a lower quantile could be adopted to indicate the nutrient concentration below which moderate status was unlikely (and GES was likely to be achieved).

Mismatch in classifications could be reduced further by using the average between the lower quantile of nutrient concentrations associated with moderate status and the upper quantile associated with good status. One advantage of these approaches is that linearity is not as important (Table 2-2), although it is important to check that there are significant differences between adjacent classes. Because the method depends on the range of values within each class, different outcomes may be obtained when discrete classes are used, compared to the situation when classes are aggregated into ‘good or better’ and ‘moderate or worse’. They are also sensitive to the position of the data cloud relative to the boundary of interest (See Appendix 3). Also, if the moderate class is wide, there is a possibility that the biological response to increased nutrient concentrations will reach an asymptote, which may have the effect of increasing the lower quantiles of the moderate status class and thereby influencing the position of the boundary.

Categorical methods, in other words, may be a valid option in a few situations, particularly if high, good and moderate status classes are well defined but there are few heavily impacted sites with which to ‘anchor’ a regression model. However, the precision of estimates will not be any greater when the relationship is very noisy than would be the case if a regression was used. The categorical approach is, in effect, the same as making a regression of nutrients on biology, because it assumes that all the uncertainty is in nutrients and that biology is the (error-free) ‘predictor’. Problems will also arise if there are few water bodies in each category or if there are missing categories.

A better categorical approach is the use of binomial logistic regression, which provides a method for fitting a logistic model to categorical data using a binary response, for example biology moderate or worse = 1 or biology good or better = 0. This approach has the advantage of being less sensitive to the position of the data cloud, and the quality of the statistical model can be tested using a variety of methods. The resulting model can also be used to determine boundary values at different levels of probability of being ‘moderate or worse’ (Figure 2-4). This approach has been subject to limited testing, but results obtained using simulated data and the example data sets provided with the toolkit suggest it is likely to be the best alternative to linear regression models, provided that the data cloud is not wedge shaped (i.e. where multiple pressures are influencing biological status).

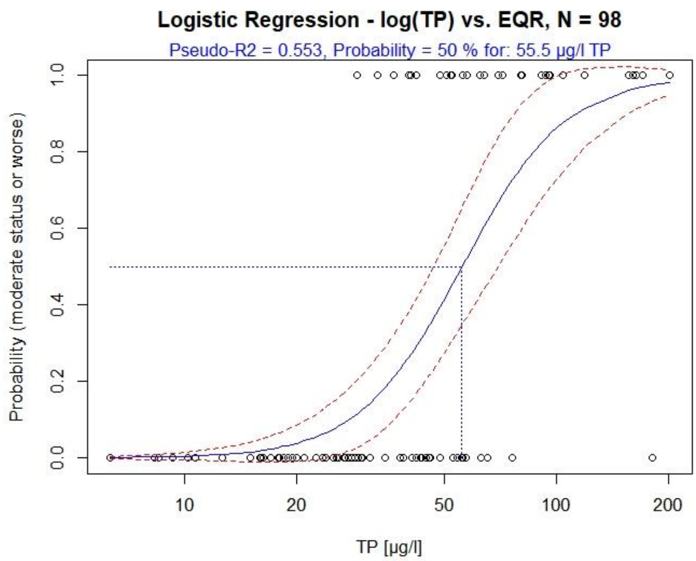


Figure 2-4 Binomial logistic regression of total phosphorus on probability of biology being moderate (or worse) status, using a simulated data set. Boundary value of total phosphorus shown for probability = 0.5.

Decision tree methods such as classification and regression trees also can be used as alternatives to logistic regression (Figure 2-5). These enhance the predictive models with high accuracy, stability and ease of interpretation. They work by iteratively splitting the data into distinct subsets, with the splits chosen in such a way that entropy in the resulting subsets is minimised. This allows the structure of a data set to be explored by highlighting the most important variables. Decision tree output is very easy to understand even for people with non-statistical backgrounds. Over-fitting is a practical difficulty when dealing with decision tree models, but this can be managed by using random forests. Random forest methods 'grow' many classification trees using a random bootstrap sample from the original data. They use multiple models in order to give a better performance than can be obtained when just using a single tree model in multiple stressor environments. We have developed sample code for this method in the toolkit using R (See Appendix 1 Section A5.4).

A final approach is simply to set a nutrient boundary that minimises the mismatch between ecological status and the supporting element. This is a pragmatic approach and testing using synthetic data (See Appendix 3) has shown that it is less sensitive to data uncertainty than the simple categorical approaches, although more sensitive than logistic regression. The method has been developed within the R version of the toolkit to include bootstrap sampling and a LOESS curve fit to provide an assessment of uncertainty (Appendix 1 Section A5.2).

In all situations except those with highly significant regressions, we recommend that all methods be used and compared. If they provide very different results, then the user needs to think carefully about why, and then consider which method is most appropriate. In many cases the different methods provide similar boundary values, which increase confidence in the results.

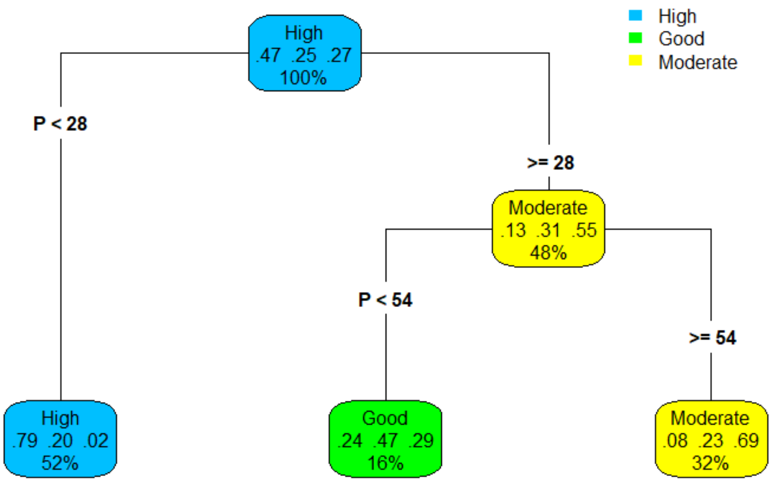


Figure 2-5 Classification decision tree of total phosphorus on biological classes (high, good, moderate). Boundary value of total phosphorus for high status < 28; boundary value of total phosphorus for good status < 54. Each node shows the predicted class, the predicted probability of each class and the percentage of observations in the node (high, good, moderate).

Box 5: Points to consider when developing nutrient thresholds: statistical

- Are data spanning at least four status classes available?
  - The length of the gradient is important to ensure that linear portions can be identified and that regressions are securely 'anchored' at both extremes.
  - It may be possible to produce relationships with shorter gradients, so long as these straddle the biological boundaries of greatest interest (i.e. high/good (H/G) and G/M).
  - Non-regression methods may also be suitable when there are adequate data and a significant relationship, but the full gradient is not available.
- Is there a linear relationship across the range of interest (i.e. where H/G and G/M status classes are located)?
- How much uncertainty exists in the independent (explanatory) variable? Is type I or type II regression most appropriate? (NB: this decision is particularly important when r^2 is low).
- Has the maximum EQR value been rounded down to 1.0? Are uncapped EQRs (i.e. extending to > 1) available? If the denominator for the EQR calculation is constant within a type, raw metrics rather than EQRs could be used.
- Avoid using several summary metric values from the same site but different years, as there is a risk of 'pseudoreplication'. If site-years are used ensure that the number of years is similar at all sites.
- Is there evidence of wedge-shaped scatter, where the variability of EQR increases at one end of the nutrient gradient? This is typically an increase of EQR variability as nutrient concentrations decrease, suggesting multiple pressures.
- What is the predictive power of the relationship (see Road Map in Section 3)? Is a regression method appropriate or would a categorical method be more productive?
- Is it practicable to merge national types, or to combine data with that from neighbouring countries to produce a larger and more robust data set?

3. Overview of process ('road map')

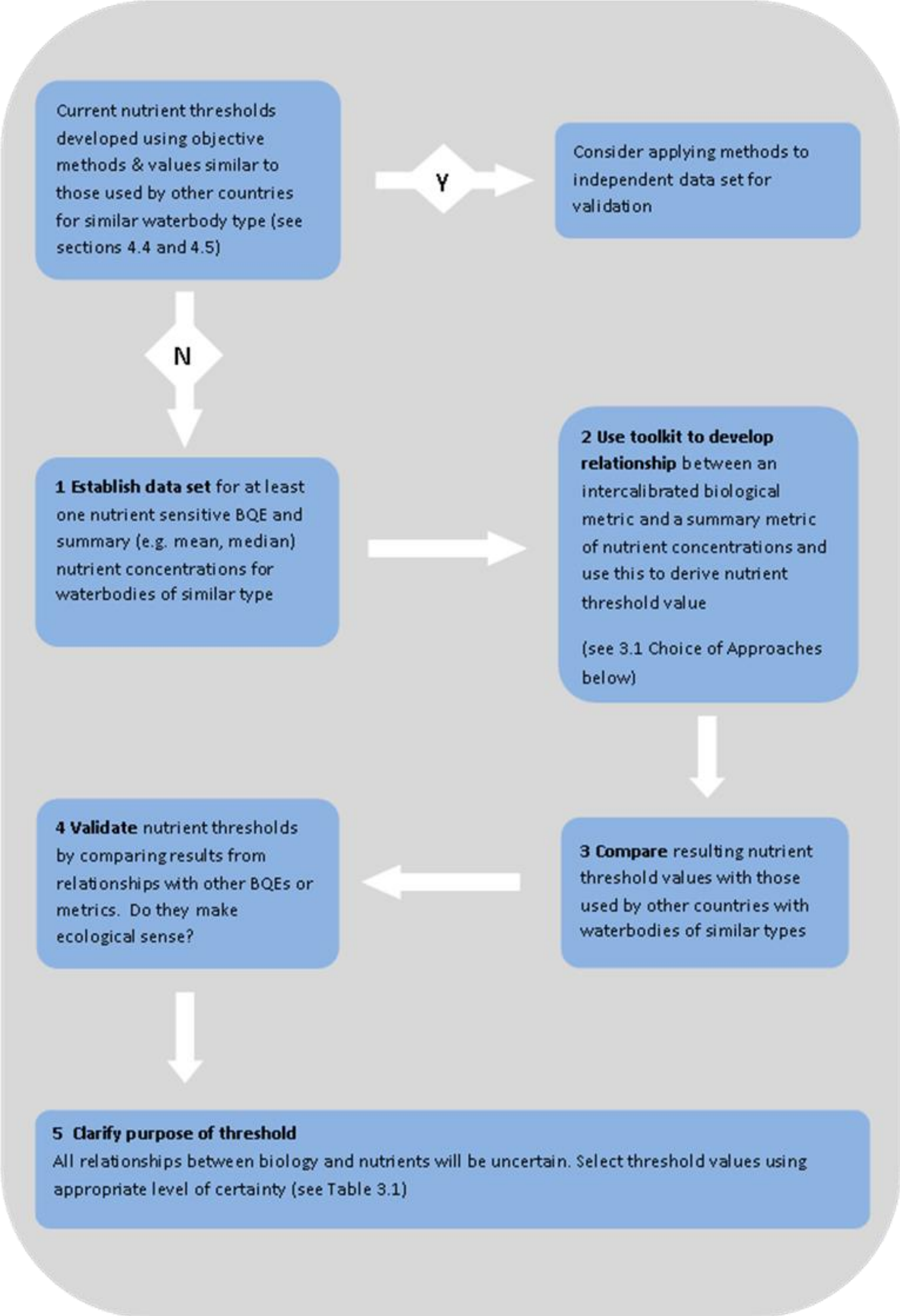


Table 3-1: Approaches to establishing nutrient thresholds to meet specific objectives (see Chapter 2) (more robust methods indicated in bold).

Purpose	Approach
Minimise mismatch between biological and nutrient element	**Best fit line**, **minimisation of mismatch** or average of adjacent classes, **binomial logistic regression with error P= 0.5** (6)
Minimise probability of incorrectly assuming a nutrient effect on BQE when there is none (type 1I error, i.e. when biological element is predicted to be below good status when it is, in fact, in good status)	**Upper quantile of regression residuals**, upper quantile of nutrient concentration in class or **upper probability value of binomial logistic regression (P > 0.5)**
Maximise probability that biology is good or high (minimise type 2 errors), by setting more stringent nutrient boundaries.	**Lower quantile of regression residuals**, lower quantile of nutrient concentration in class **or lower probability value of binomial logistic regression (P < 0.5)**

In the following section, we provide a flow chart to guide the selection of methods. **This is best used to illustrate the decisions needed when selecting methods for specific data sets rather than as a prescriptive flow diagram**.

The ideal situation is where there is likely to be a strong effect of the nutrient pressure, with only minor influence of other pressures or confounding environmental variables. Experience to date suggests this is most likely to be the situation in lakes and coastal waters. In this situation, we recommend regression methods as they provide a general statistical model of the pressure-impact response.

3.1 deals with methods where biological data in the form of an EQR, or a continuous metric, are available.

3.2 deals with situations where the relationship between nutrient and biology is weak, which is likely where multiple pressures occur or other environmental gradients that are not removed via a typology occur. This is most likely to occur in rivers and transitional waters. The toolkit provides approaches for these situations, but cannot offer simple solutions.

3.3 contains categorical methods. These methods place more reliance on an appropriate class width and do not allow for situations where nutrient concentrations do not span the entire status class. They have the advantage over linear regression of not assuming linearity, but should be accompanied by significance testing of differences of nutrient concentrations between the classes.

(6) Note that to minimise mis-classification rates the distribution of the data needs to be considered when selecting the p value to use. This is currently beyond the scope of the toolkit but can be achieved using R package modEva.

Several of the methods should be tested (on several BQEs, if possible) and the resulting boundary values compared. If they are very different, use the flow chart to help determine which are likely to be the most robust and consider how the values might be validated.

The methods have been tested and compared using synthetic data sets (see Appendix 3 for details). These data sets were designed to have linear relationships and normally distributed errors, so represent an ideal situation; however they illustrate the following points:

(1) The categorical methods were sensitive to the distribution of the data cloud used. Where the mean of the data lies at a higher quality than the boundary of interest, the categorical methods tend to underestimate boundary concentrations, and where it is at a lower quality they overestimate concentrations.

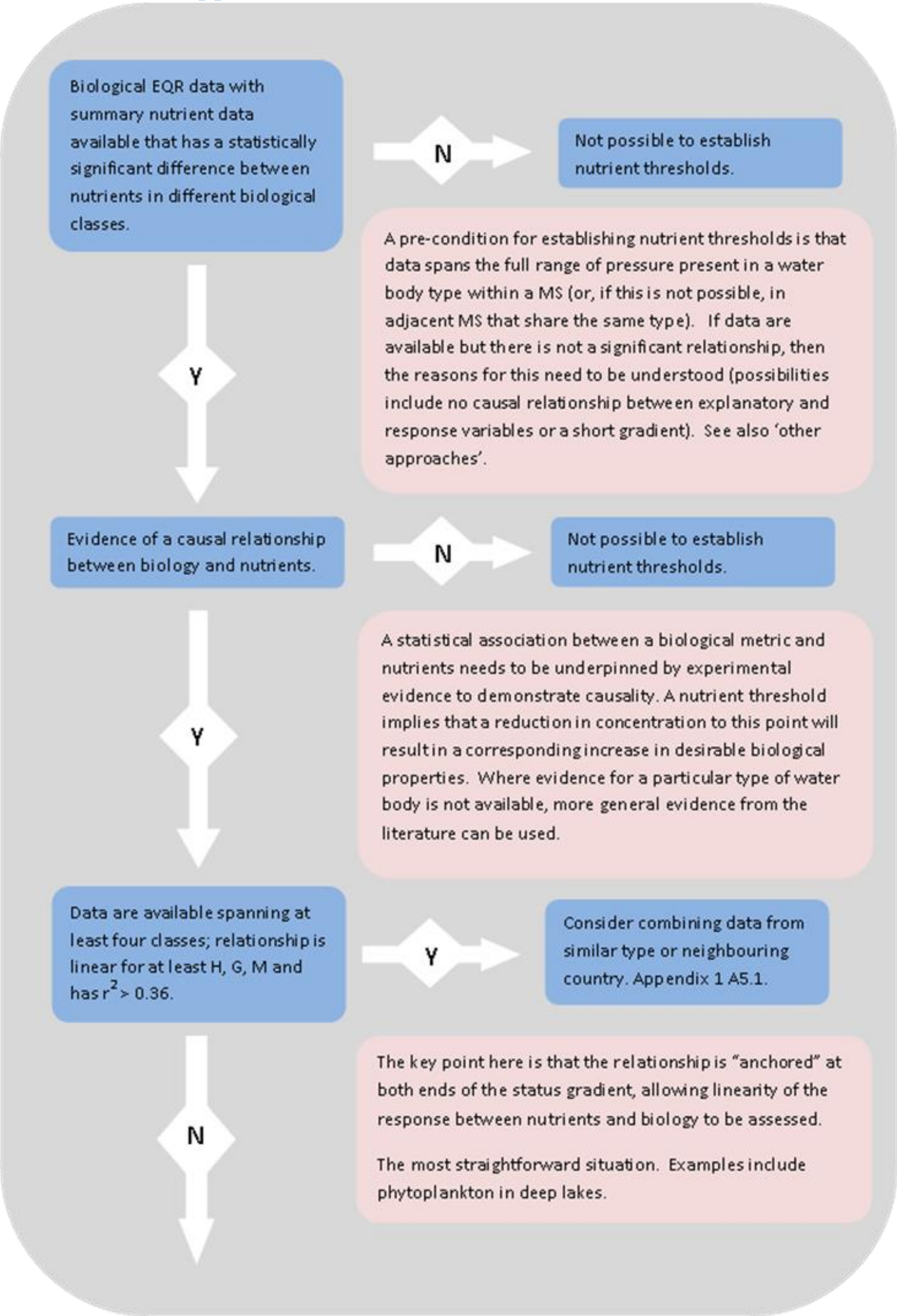
(2) Linear regression and the binomial logistic regression were largely uninfluenced by this and thus are probably the most reliable methods.

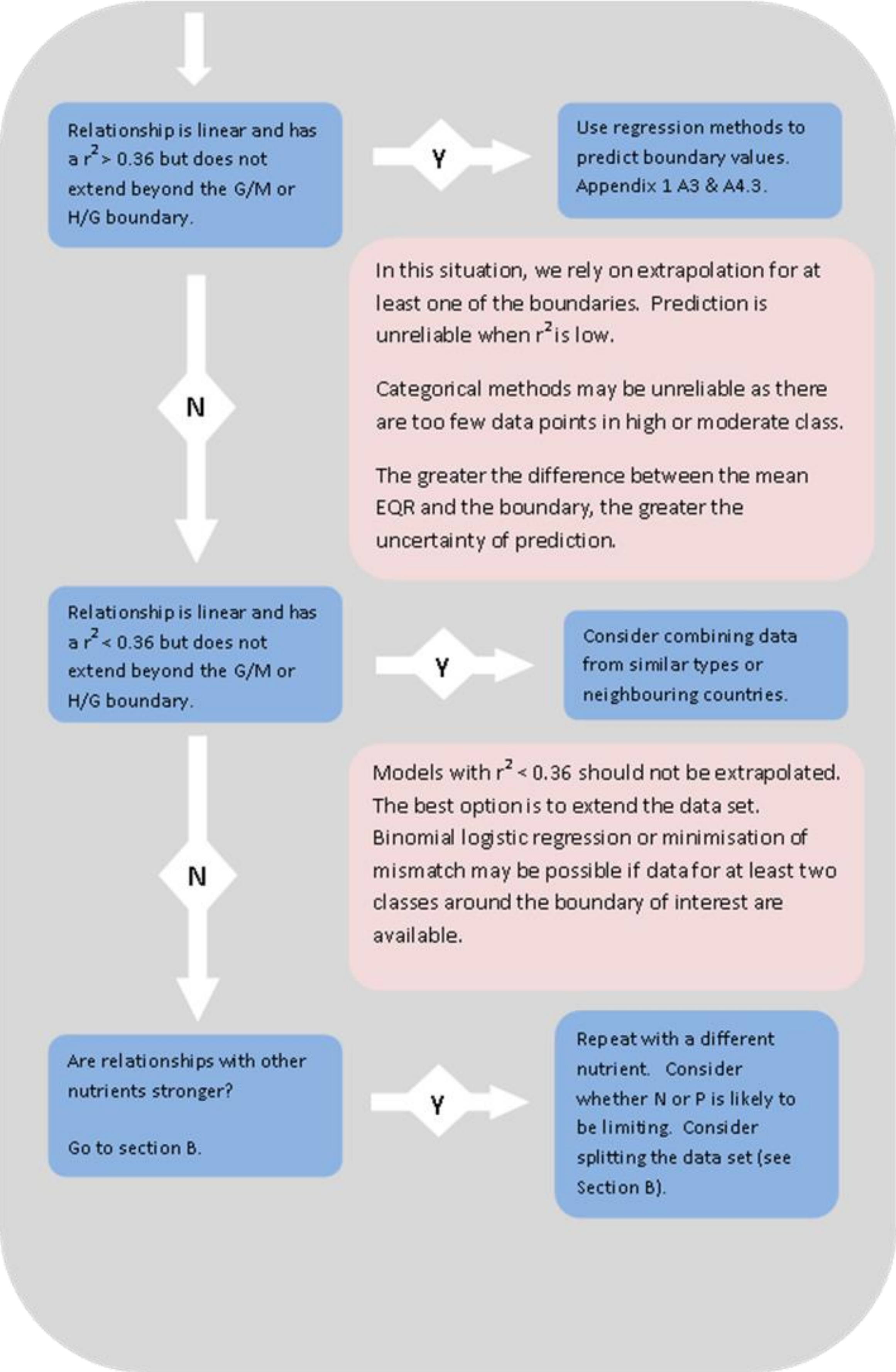
(3) The minimisation-of-mismatch method was also little affected by data distributions compared to the other categorical methods, and is preferred to the simpler categorical methods.

(4) For wedge-shaped data clouds, where multiple pressures could influence the observed biological EQR values, linear regression and binary logistic regression (using a P = 0.5) consistently underestimate boundary values and are not an appropriate method. The categorical methods performed better, but only when the data cloud was evenly distributed around the boundary of interest.

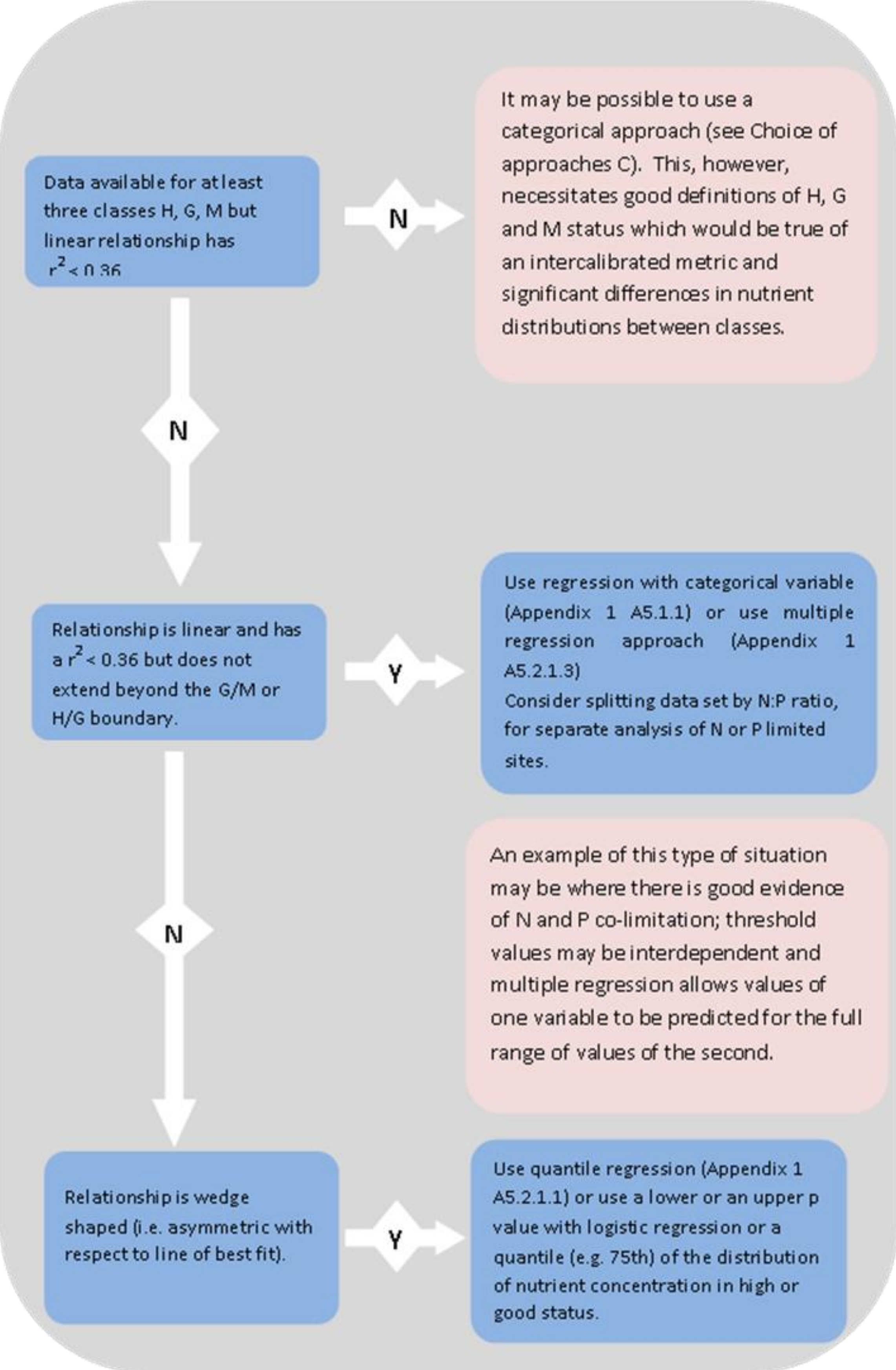
(5) For wedge-shaped data clouds quantile regression was the most reliable approach; however, it is difficult to identify an appropriate quantile to model. There is a risk of selecting too high a quantile, which does not allow for uncertainty in a true relationship between nutrient and biology.

3.1. Choice of approaches A




Relationship is linear and has a $r^2 > 0.36$ 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.
N
In this situation, we rely on extrapolation for at least one of the boundaries. Prediction is unreliable when r^2 is low.
Categorical methods may be unreliable as there are too few data points in high or moderate class.
The greater the difference between the mean EQR and the boundary, the greater the uncertainty of prediction.
Relationship is linear and has a $r^2 < 0.36$ but does not extend beyond the G/M or H/G boundary.
Y
Consider combining data from similar types or neighbouring countries.
N
Models with $r^2 < 0.36$ should not be extrapolated. The best option is to extend the data set. Binomial logistic regression or minimisation of mismatch may be possible if data for at least two classes around the boundary of interest are available.
Are relationships with other nutrients stronger?
Go to section B.
Y
Repeat with a different nutrient. Consider whether N or P is likely to be limiting. Consider splitting the data set (see Section B).

3.2. Choice of approaches B, weak linear relationship, $r^2 < 0.36$)



N

The objective is to fit a line to the upper or lower surface to the scatter. A lower surface would produce a precautionary boundary value, if other factors prevent the expression of sensitivity to nutrient pressure. An upper surface would produce a less precautionary boundary where other stressors cause a lower biological status despite low nutrients. **In this case it is very important to validate the boundary (see chapter 6).**

Relationship is linear and has a $r^2 < 0.36$ but does not extend beyond the G/M or H/G boundary.

Y

Data has high uncertainty. Binomial logistic regression may be the best method. Compare with other categorical methods and minimisation of mismatch, but try regression and compare outcomes.

Consider if a combination of normalised EQRs for different BQEs or the use of an absolute value scale of a metric rather than EQR improves r^2.

N

Too few data. Use other approaches (see section 5).

3.3. Choice of approaches C (categorical methods)

Data available for three classes H, G, M and box plots for adjacent classes do not substantially overlap. Wilcoxon Rank Sum Test confirms significant difference between adjacent classes.

Y

Compare following five methods:

a) binomial logistic regression;
b) minimisation of mis-match;
c) average of median of adjacent classes;
d) average of upper and lower quartiles of adjacent classes;
e) 75th percentile of high (for H/G) or 75th percentile of good (for G/M).

Binomial logistic regression is likely to be the most reliable method for categorical data.

Minimisation of mismatch is simple to understand, but testing with synthetic data shows it to be sensitive to the degree of scatter of the data.

Box plots are a pictorial representation of the distribution of values within a subset of comparable water bodies, and the quartiles shown in the box plot will depend upon the range of all values in the class. The subset, in turn, is assumed to be representative of the population. If the subset of measurements from a particular status class does not span the whole nutrient range of that class, then threshold values of nutrients derived using the methods in the toolkit may be misleading. Analysis of synthetic data suggests these methods are the most sensitive to scatter in the data and the spread of the data cloud relative to the boundary.

N

Data available for three classes H, G, M and box plots for adjacent classes do not substantially overlap. Wilcoxon Rank Sum Test confirms significant difference between adjacent classes.

Y

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

4. Descriptions of procedures

4.1. Options for analysis

Three approaches to analysis are presented:

- regression analysis, using a continuous relationship between EQR and nutrient concentration;
- categorical analysis, using the distribution of nutrient concentration within biological classes;
- minimisation of mismatch of classifications for biology and nutrients.

The first two methods are similar to each other in that they both depend on having data that span an adequate range of quality, including the moderate class. The categorical analysis may appear to be less sensitive to nonlinearity in the data; however, estimation of quartiles depends upon having values that span the full range of the status class in question. For example, if moderate status is only represented by samples at the lower end of nutrient concentrations for the class (i.e. closer to good status), then the quartiles (and, therefore, the boundary) will be lower.

The method that is least influenced by nonlinearity is the minimisation of mismatch of class. In the MS Excel tool this method, like the categorical methods, cannot provide an uncertainty estimate of the class boundary, which is a significant disadvantage. However, an R script has been provided that uses a bootstrap approach to provide an estimate of uncertainty (see Appendix 1 Section 5.2).

The road map in Section 3 provides guidance on selecting methods, however all methods should be tested and results compared, wherever possible.

The guidance document also provides a summary of ranges of nutrient concentrations supporting good status derived from analysis of Geographical Intercalibration Group (GIG) and other data sets (see Sections 4.4 and 4.5). These can be compared with results obtained by following the procedures in this document and provide an initial validation of values. Similarly, values can be compared with those used by other Member States for similar water body types. Finally, it is recommended that a further validation be carried out to ensure that the predicted nutrient concentrations that support good status meet broader ecological expectations (Chapter 6).

4.2. Overview of the stepwise procedure

This section provides an overview of the procedure. See Appendix 1 for a step-by-step guide.

Step 1: assemble a data set

The data set needs to consist of water body level biological quality element (BQE) EQR values, their WFD class and matching summary nutrient concentration. Any BQE sensitive to nutrients may be used and we suggest that results obtained from different BQEs be compared. As ecological status is decided on a ‘one out all out’ (worst BQE) basis, selecting the BQE most sensitive to pressure will ensure that nutrient concentration standards support overall status. It is also possible that combining nutrient-sensitive BQEs will yield a stronger relationship with nutrients than any BQE in isolation. Bear in mind, too, that the BQE that has the strongest statistical relationship to nutrients may not necessarily yield the most precautionary boundaries.

One of the key issues in developing relationships between BQEs and nutrient concentrations is minimising variability. **It is strongly recommended that summary nutrient concentrations such as mean or median be used rather than spot samples, and that nutrient and biological samples be taken from similarly representative locations and cover similar time periods.** Nutrient samples should be representative of the summary period to minimise the influence of seasonal variation.

To facilitate analysis data should be entered into the toolkit template *DataTemplate.csv*

Step 1 (see Appendix 1 toolkit details A1)

- Assemble a data set using DataTemplate.csv
- Use summary data not spot samples
- Ensure nutrient and biological data are spatially and temporally matched

Step 2: inspect the data

The first and most important step of the analysis is to check the data (e.g. using basic chart functions in MS Excel) and determine if they are adequate for the purpose. An ideal data set would have the following features:

- Data span a range of biological quality from high to poor, as it is important to avoid extrapolation of pressure response relationships;
- EQRs would have a clear response to nutrient concentration and be linear over most of the pressure gradient.

Few data sets will meet these requirements perfectly. Experience suggests that relationships between phytoplankton and nutrients in lakes are the easiest to deal with as phytoplankton respond directly to nutrients and there are usually few other significant pressures, particularly in larger oligotrophic to mesotrophic lakes. Relationships for rivers, estuaries and coastal waters are significantly more uncertain as the BQEs are likely to be experiencing multiple stressors and subject to a higher number of other factors that will influence response.

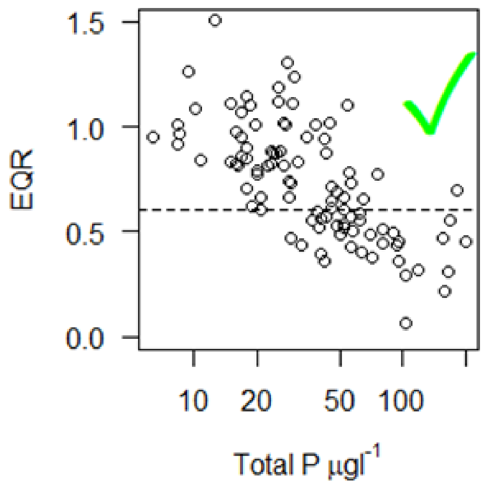
Step 2a: check the data using a scatter plot

Use either Excel tool (Appendix A2.3) or R script *TKit_check_data.R,* (Appendix A3.2.2).

These data meet the requirement:

- a reasonable relationship between EQR and P;
- extend from high to poor status.

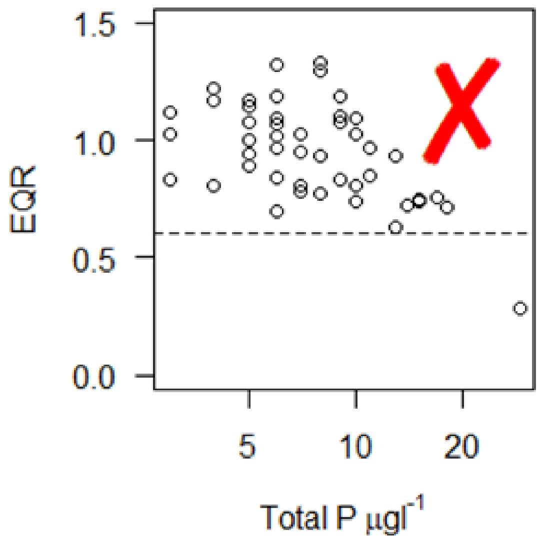
Next check for outliers and then assess linearity.



These data do not meet the requirement:

- a poor relationship between EQR and P;
- only a single point worse than good status.

More data are needed. Try adding data from another MS or from other water body types that are similar to the actual water body type covered by the original data (Appendix A4.1).



The data can also be plotted as a box plot, which is useful for visualising the categorical methods

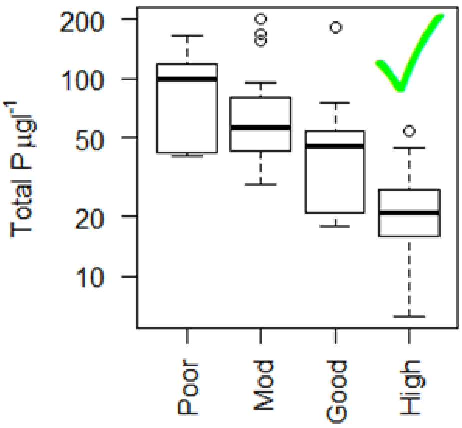
Step 2b: check the data using a box plot

Use R script *TKit_check_data.R,* (Appendix A3.2).

These data meet the requirement:

- 4 biological classes poor to high;
- minimal overlap between high and good, and good and moderate.

Note the outliers: it may be necessary to remove these (e.g. if they represent extreme hydrological events).



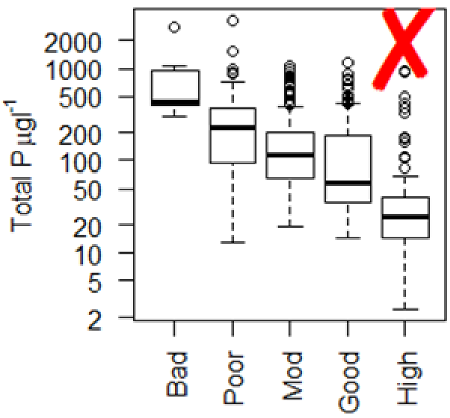
These data do not meet the requirement:

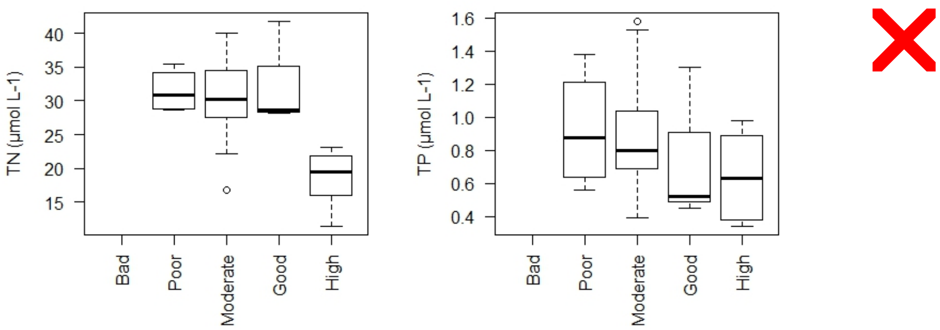
- 5 biological classes bad to high;

but

- large overlap between good and moderate.

Note more outliers. These data require further investigation, see Appendix A4.2.





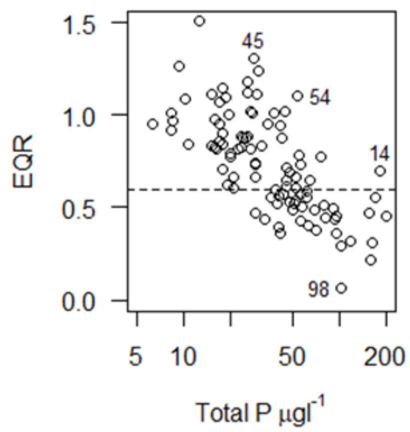
These data do not meet the requirements either:

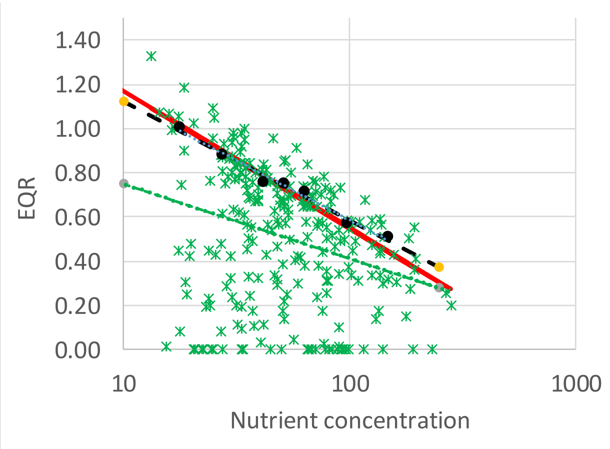
- TN shows a flat response between good and poor status;
- TP data shows a weak gradient and with high overlap between good and moderate classes.

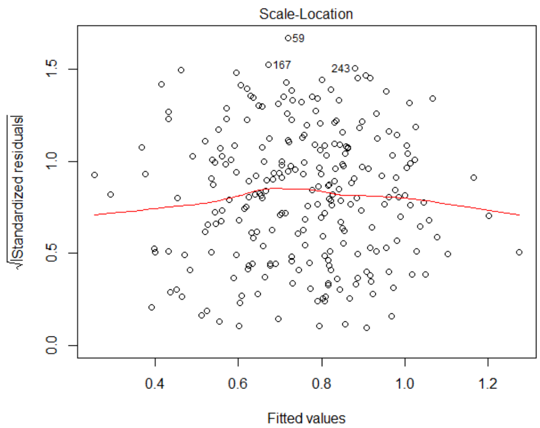
Step 2c: scatter plot suggests a wedge shape — check for outliers

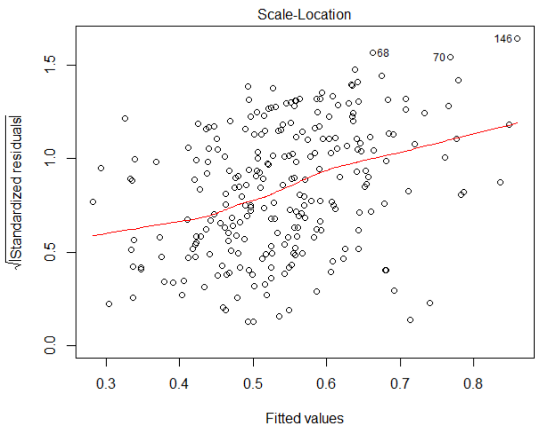
- Use scatter plot to inspect the data,
- Do not identify outlier points simply to achieve a better fit; however, some may be associated with atypical conditions, justifying removal (e.g. data collected during extreme events).
- See Appendix 1 Sections A2.3 and A4.2.1.
- The use of an upper-quantile regression (black dotted line) may reflect the true relationship between nutrient and EQR (red line). (Appendix 1 Section 4.2.2).

1 A simple linear OLS regression (green dotted line) may underestimate threshold boundary values. Also note that the residuals will not be normally distributed, compare the 'Scale-location' plots below. (Left hand using data with no indication of a wedge, right hand using wedge-shaped data. Note the standardised residuals increase with fitted EQR values).









It is important to check that the response of EQR to nutrient is linear before fitting a linear regression model. This can only be done visually using the Excel tool (see Appendix 1 Section A2.4), but using the R script a generalised additive model is fitted which gives clear indications of nonlinearity (see Appendix 1 Section A4.3.1.2).

Step 2e: is the response linear?

- Fit a generalised additive model to scatter plot.

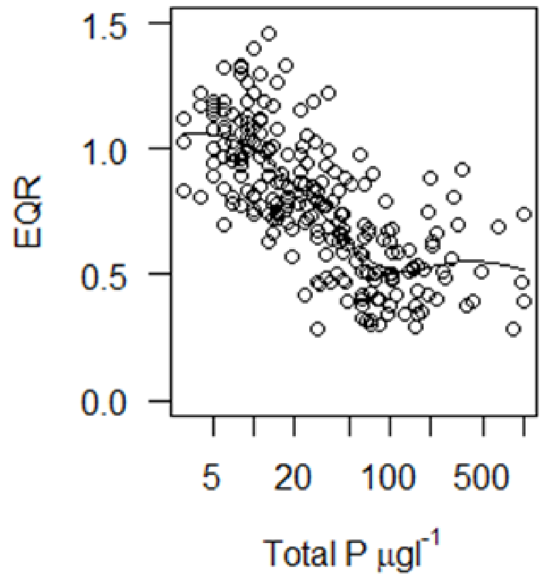
(Use R script *TKit_check_data.R* Appendix A3.2.2).

- Or assess scatter by eye

(use Excel tool).

- Determine the linear range to be used for regression modelling.

In this example the response is linear within the TP range (10-100 µg l^{-1}), with a flat response beyond this range.



Note that in the above relationship EQRs were not truncated at 1.0. EQRs >1.0 occur if reference metric values are set using the mean metric value in reference sites, as some of the values would have be higher than the mean. Different approaches are used to meet the WFD requirement that EQRs should be within the range of 0-1, for example rescaling or normalisation, but if truncation is used (e.g. an EQR of 1.3 is set to 1.0) the value is said to be 'censored'. If many values in the data set are censored in this way they will influence the regression slope and the uncertainty, tending to cause a flattening of the relationship. Such data should be treated with caution.

The R script *03_TKit_check_linearitydata.R* will also fit a series of stepwise linear models which can be used to assess changes in slope of the relationship. These may be useful for the identification of a linear range. It is also suggested that as the slope of the relationship is critical for determining boundary values, the effect of using different values be assessed. The Excel tool provides a quick way of achieving this. Also, note that the further the predicted boundary value is from the mean value of the data used, the more uncertain the value will be.

Step 3: fit linear regression models

The toolkit allows three regression models to be fitted to the linear portion of the data:

- an Ordinary Least Squares (OLS) regression of EQR v nutrient concentration; assumes all uncertainty is in measurement of the EQR (underestimate of slope);
- an OLS regression of nutrient v EQR; assumes all uncertainty lies in measurement of nutrient concentrations (overestimate of slope);
- a type II regression; assumes equal uncertainty in measurement of both EQR and nutrient (slope between the two OLS regressions).

The boundary values predicted by the regression models depend on the slope of the relationship and the difference in the slopes produced by these relationships depends on the r^2: the lower the value the greater the difference. There is no clear statistically valid cut-off for what represents a low

r^2 but we suggest that if the model r^2 is > 0.36, then the relationship is good enough to make predictions.

It is not possible to test the significance of the slope of a geometric average regression, as used in the Excel toolkit. For this reason, it is important that there is a significant correlation and as r > 0.6 is recommended by Smith (2009), we propose that an r^2 of > 0.36 is appropriate. This can be relaxed if type II ranged major axis regression is used (see Appendix 1 for details).

The true slope of the relationship lies between the lines predicted by the models. We suggest that the type II regression model 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 type value will lie within this range. Remember that the higher the r^2 value, the smaller the differences between the model slopes.

The output also uses the interquartile range of the residuals of the model to provide an indication of the range of nutrient concentrations that any particular water body of the type modelled might have. 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 lower than the upper range value, while 75 % would have values greater than the lower range value.

Step 3: fit linear models

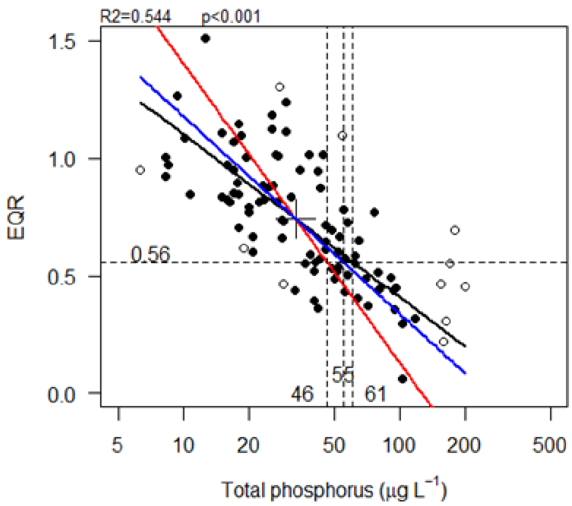
Use either Excel tool (Appendix 1, A2.5) or R script *04_TKit_fit_lin_model1.R* (Appendix 1, A3.4).

- R^2 of model is >0.36.

Model is adequate to predict nutrient concentration at EQR boundaries (see Appendix 1 for details)

Model slope determines the predicted TP concentration; true slope lies within the range of the two OLS models (black and red lines).

- TP range at G/M boundary: 46-61 µg l^{-1}
- Most likely value given by type II regression (blue line): 55 µg l^{-1}.
- This value lies close to the mean of the TP data, and the different model slopes have a relatively small effect, as all lines intersect at the means $\bar{x}$ and $\bar{y}$.



Black and red lines show slopes of 2 OLS models.

Blue line shows slope of type II regression model using Ranged Major Axis regression.

Mean of EQR and TP marked by cross.

Step 4: calculate categorical methods and step 4a: compare results

Although regression methods provide the most robust method of estimating the nutrient concentration that occurs at the EQR boundaries and should be used wherever possible, it is useful to compare these with values provided by categorical methods.
Five categorical estimates are provided.

Step 4: categorical methods

Use either Excel Tool (Appendix 1, section A2.6 & A2.7), or R scripts (section A5).

R scripts *09a_TKit_P_Categorical.R, 09b_TKit_N_Categorical.R, 10a_ Tkit_mismatch3_GM.R, 10c_TKit_mismatch3_HG.R.*

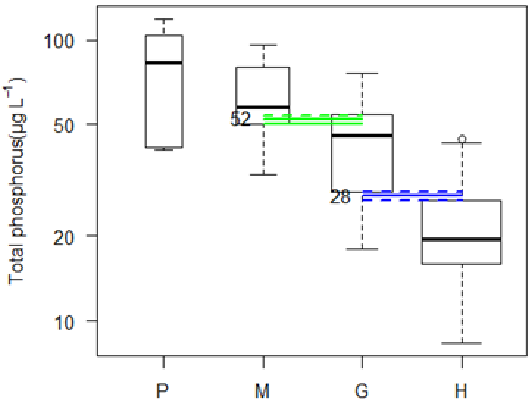
- Average of upper and lower quartiles of nutrient concentration of water bodies in adjacent biological classes; i.e. for G/M the average of the 25th percentile of the nutrient concentration of water bodies in moderate and the 75th percentile of the nutrient concentration of water bodies in good.

- Average of the median nutrient concentration (typically on a log_{10}-transformed scale) of water bodies of adjacent biological classes.
- Upper 75th percentile of nutrient concentration at good biological status.
- The nutrient concentration that gives the lowest mismatch between classifications based on biology and on nutrient concentration.
- The use of binomial logistic regression (only available in R).

Step 4a: average adjacent quartiles

Calculate the median and interquartile range of nutrient concentration in each biological class.

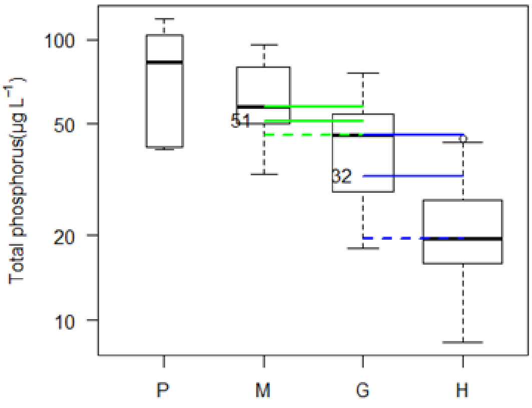
Average the upper 75th quantile of good and the lower 25th quantile of moderate.



Step 4b: average adjacent medians

Calculate the median and interquartile range of nutrient concentration in each biological class.

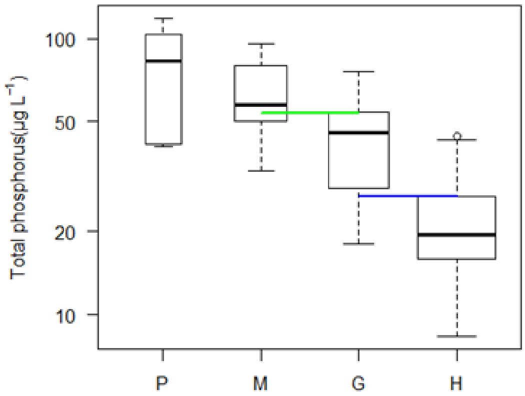
Average the median of good with median of moderate.



Step 4c: upper 75th quantile of better class

This method provides the least stringent estimate as for the G/M boundary it takes the concentration below which 75 % of water bodies are in good status.

It is similar to using the intersection of the upper error bar in regression or the use of quantile regression (see 4.3).



Note

(a) There should be a significant difference (Wilcoxon test) between the distributions in adjacent classes.

(b) These box plot methods are sensitive to the overall distribution of nutrients; they are most reliable when the mean of the data cloud EQR is close to the boundary of interest.

Step 4d: minimise mismatch

Make a binary classification of biology and nutrients (‘good or better’ and ‘moderate or worse’).

Use series of nutrient concentrations to define nutrient class.

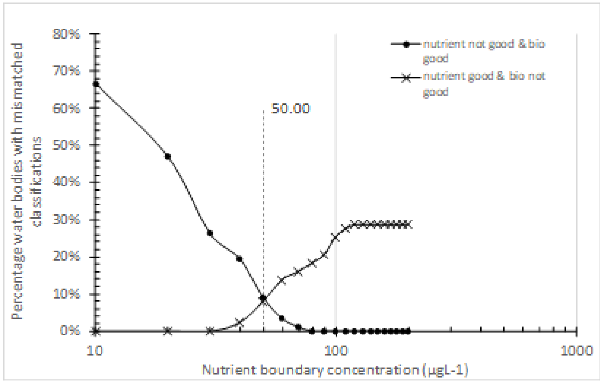
Plot rate of mismatches.

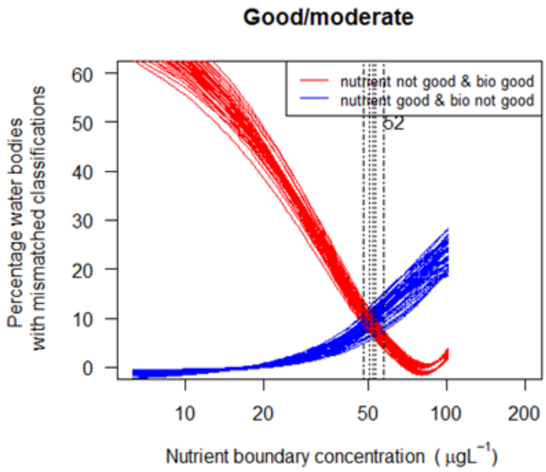
Point of intersection identifies nutrient boundary concentration for minimum mismatch (upper graph on the right).

(Appendix A2.6 & A5.2).

This method is the least sensitive to non-linearity and outliers and is recommended where there is a significant correlation but high scatter.

The R version of the tool allows uncertainty to be estimated (middle graph on the right).





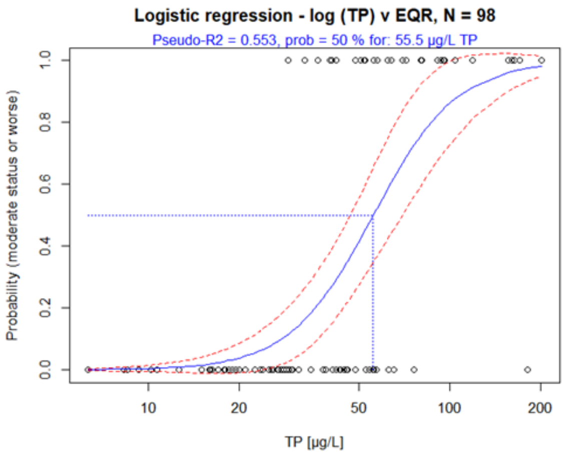
Step 4e: fit binomial logistic regression

(R Script *11a_LogisticRegGood.R* and *11b_LogisticRegHigh.R*, see Appendix A 5.3)

Check pseudo r^2.

Select probability (of being moderate or worse) and read off the nutrient concentration at this point. (e.g. at a probability of 0.5 TP concentration would be 56 µg l^{-1}). Red dotted line shows uncertainty band ± SE.

For well balanced data sets where the scatter of points is not wedge shaped p = 0.5 is appropriate.



Step 4f: compare regression and categorical results

The MS Excel tool can be used for this, although this will not show the logistic regression results.

- Predicted boundary concentrations are very similar.
- Any approach could be used.

In **Example 1** (boxes below), the relationship between phytoplankton EQR and TP in shallow high alkalinity lakes (broad type 3) is used to predict boundary value using regression models. For the G/M boundary the best fit predicted value is 55 µg L^{-1} and is most likely to fall within the range 46-61 µg L^{-1} but 25 % of water bodies in the type will have values < 82 µg L^{-1} and 25 % > 34 µg L^{-1}.

The categorical approach provided very similar results, with the minimisation-of-mismatch method suggesting a value of 50 µg L^{-1}. Note that the logistic regression provides a slightly higher value of 56 µg L^{-1} (see above as this is not shown in the Excel summary table).

This is a relatively small range, and for this lake type the TP concentration that is most likely to occur at the G/M boundary determined by phytoplankton falls within the range of 50-56 µg L^{-1}. The average of these values is 53 µg L^{-1}, which could be conveniently rounded down to 50 µg L^{-1}.

Example 1:

Phytoplankton	**BT3 (L-CB1)**	**Total P µgl-1**		R^2	0.544	<0.001
Summary of predicted boundary values						
		Most likely boundary			Possible range	
	Predicted	Range				
Good/mod	55	46	61		34	82
High/good	28	27	30		18	42

Categorical methods	Ave adj class quartiles	Ave adj class median	75th quartile class	minimise mis-match
Good/Mod	52	51	54	50
High/Good	28	33	27	30

Step 5: compare with existing nutrient boundaries

Having computed one or more potential boundary values using the methods described above the next step is to compare these values with any current regulatory standards for similar water bodies or values obtained from analysis of GIG data sets (see Sections 4.4 and 4.5). In the example above, how do the values of 50 µg L^{-1} compare with values predicted from GIG data sets and those used by most other countries for this lake type?

A TP concentration of 50 µg L^{-1} is slightly higher than the most likely range of values for broad type 3 (Table 4-1), although it is well within the range of possible values. It is above the median but within the interquartile range of values used by other Member States (Table 4-4). Although the lower of the 'most likely' values of 46 µg L^{-1} might be more typical of conditions at the G/M boundary, 50 % of lakes of this type are likely to have values between 34 and 82 µg L^{-1}.

See Chapter 7 for further discussion on the issues to consider regarding the use of boundary values for management given the relatively high uncertainty of the relationships between nutrients and biological status.

Step 5: compare boundary values

- Compare with values of tables for nutrient boundaries in section 4.4 and section 4.5 respectively for freshwater and transitional and coastal waters systems.
- The predicted G/M boundary values are within the broad type range and the interquartile range of national boundary values for the type.

Step 6: validate

It is important to consider additional validation approaches (see Chapter 6). For example, in the shallow calcareous lakes (broad type 3) used in the example above it might be appropriate to compare boundary values with changes in key components of the macrophyte flora. Free et al. (2017), for example, used changes in the relative frequency of charophytes to check the boundary for marl lakes.

Step 6: validate boundary values

Consider what independent data could be used to demonstrate that the predicted concentrations meet ecological expectations and are likely to protect key taxa associated with good status conditions.

4.3. How to treat data where relationships are poor

When relationships are poor, further steps may be needed to investigate the reasons for this. It might be worth, for example, returning to the preliminary investigation of the data to confirm that there are no glaring errors such as extreme unexpected nutrient concentrations in a sample. Are sample data rather than summary metrics being used for nutrient concentrations? Was there adequate temporal frequency of sampling for nutrient data?

4.3.1. Wedge-shaped relationships — quantile regression

In some cases, the scatter plot may be wedge-shaped, i.e. asymmetrical with respect to a line of best fit. For example, other factors, such as shade or limitation by a different nutrient, may be preventing the expression of nutrient impacts (an 'inverted wedge') allowing good or better status to occur at high nutrient concentrations. Alternatively, other pressures may be depressing biological status so that moderate or worse biological status occurs at low nutrient concentrations (a 'wedge'). It is also possible that other pressures can act synergistically with nutrients causing enhanced nutrient sensitivity and lower than expected biological status at low nutrient concentrations, e.g.

deterioration of riparian areas causing bright light and warmer water in a river, or barriers causing increased water retention, giving more time for algal biomass to increase. In any of these cases, it may be more appropriate to fit lines to percentiles of the data, thus defining the upper and lower surfaces of the scatter plot, using techniques such as quantile regression.

The choice of which surface to fit (upper or lower), and the most appropriate percentile to use (the proportion of points that remain beyond the line), depend on the shape of the scatter (see below) and the assumptions made about the uncertainty of the true relationship with nutrients and the effectiveness of the other pressure. For an inverted wedge, where other factors prevent the expression of nutrients, it is more appropriate to fit the lower quantile, producing precautionary boundary values. Where other pressures are directly depressing biological status, it may be necessary to fit an upper quantile to avoid unnecessarily penalising nutrients. However, if these pressures are acting synergistically to enhance the impact of nutrients, a lower percentile may be the best quantile to use. Such an approach would also be needed to counteract the negative impacts of warming (climate change).

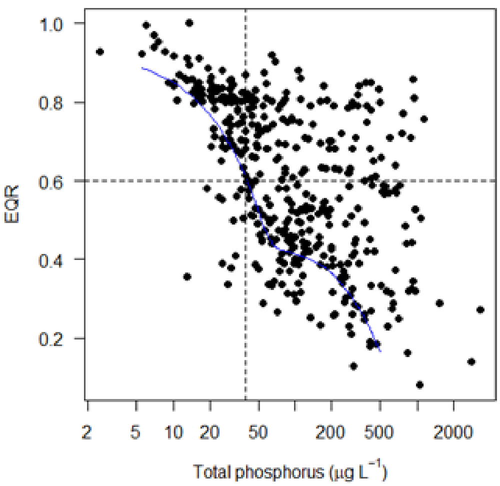
Given the difficulties of identifying and quantifying the impact of multiple pressures the choice of percentiles will always be difficult, and the decisions will be influenced by value judgements and the way the boundary values will be used for management. The issues presented by quantile regression are similar to the choice of the 'best-fit' regression line, or an upper/lower error bar used in linear regression, but as the slopes of the upper-/lower-quantile lines are different, the choice of upper or lower line and the actual percentile used have greater impact on the predicted boundary value.

Are data distributed in a wedge shape?

An **inverted wedge**

A line fitting the lower surface of these data can be fitted using quantile regression.

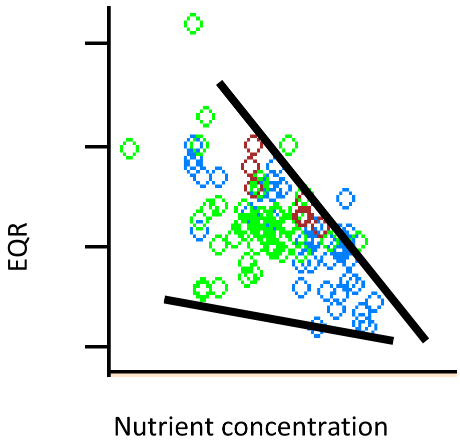
Water bodies that occur above this line may be subject to factors that prevent the expression of the phosphorus response, for example limitation by nitrogen (see Appendix 1 Section A4.2.2 for further details).



A **wedge**

An upper surface may be appropriate when other pressures cause biological quality to be low.

Quantile regression can be used to fit an upper-quantile line which could be used to determine boundary values for nutrients above which there would be an effect of nutrients.



The R package 'quantreg' (Koenker, 2016) can be used to fit linear or additive models to quantiles, and example scripts showing its use are provided in Appendix 1 Section A4.2.

Categorical approaches using the distribution of nutrient concentrations by class could also be used. Where multiple pressures are suspected the upper 75th quantile of class approach might be appropriate, however, as for quantile regression there is a risk of deriving a non-precautionary boundary value. The minimisation-of-mismatch approach is a further possibility as simulations have suggested this is least sensitive to the uncertainty of the true relationship with nutrients. Another approach would be to use logistic binary regression, but selecting a different cut-off probability. However, although there are potential approaches ([7]) it is not possible to recommend a method that would help select the most appropriate probability value to use.

Dealing with wedge-shaped data remains a challenge as there are currently no simple solutions (see Section 7.4 for further examples).

4.3.2. Interactions between explanatory variables

It may be helpful to investigate if there are interactions between the explanatory variables. These can be investigated by comparing relationships for different categories of a second explanatory variable as in Figure 4-1. Interaction co-plots are a useful tool to explore complex relationships within the data and are available in the R scripts provided in the toolkit (see Appendix 1 Section A3.2.3). The example presented here shows evidence of a poor relationship between EQR and nitrogen in a phosphorus limited system (in Figure 4-1a most data points have N:P (molar ratio) > 20 (> 9 by mass), showing that an effect of TN on EQR is apparent only at higher concentrations of the limiting nutrient, i.e. TP (top right panel in the Figure 4-1a). In those conditions however, the wedge-shape distribution of the data also seems to indicate that a quantile approach could be

([7]) Links between mismatch and the logistic binary model can be made by calculating a 'confusion matrix' which shows the proportion of sites that are classified the same or differently by biology and nutrients. There are several indices that summarise such a matrix (% positives correctly identified, % negatives correctly identified), see Fielding and Bell (1997) and these could potentially be used when multiple pressures are suspected to identify the most appropriate p value to use in the logistic model. However, further work would be needed to develop this approach.

appropriate for deriving nutrient boundaries for TN in this case. Figure 4-1b shows more clearly that where phosphorus is the limiting nutrient (i.e. N:P ratio > 20; top panel), an increase in phosphorus concentration leads to a decrease in the EQR. However, as in the previous example, the wedge-shape distribution of the data limits the approaches that can be used to derive meaningful nutrient boundaries. This example also shows some differences in nutrient patterns between regions within the water type. The region of Abruzzo (Figure 4-1b bottom panels) is not always phosphorus limited, indicating the need to control for effects of both nutrients.

Alternatively, if there is no clear evidence for either N or P limitation, multiple regression using both TP and TN may provide significant relationships, provided that the nutrient variables are not themselves significantly correlated. A multivariate relationship can be visualised graphically using a scatter plot of TN v TP with lines defining the pairs of TN and TP concentrations that would predict the EQR G/M boundary of 0.6 (Figure 4-2). See Appendix 1 Section A4.3.1 for further information.

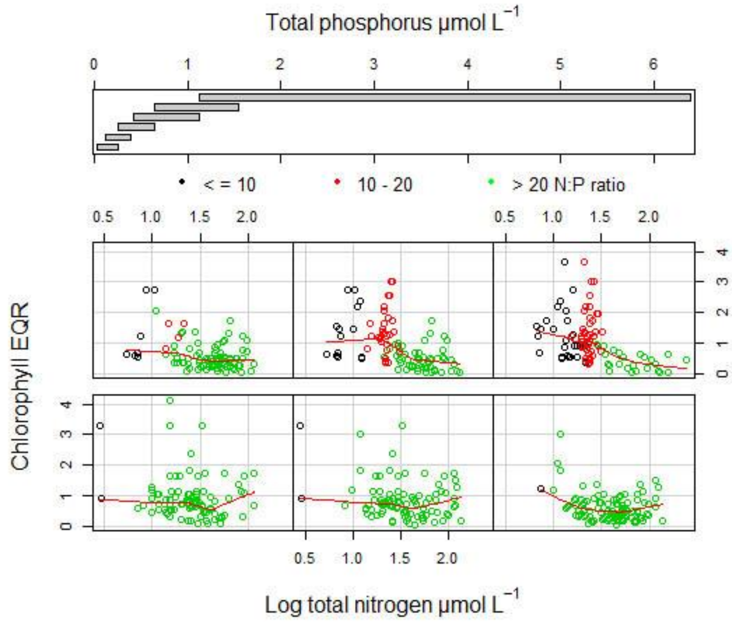
4.4. Ranges of nutrient boundary values associated with broad river and lake types

In step 5 we suggest that the boundary values obtained should be compared with any existing WFD boundary values or with values obtained from the analysis of GIG data sets. For freshwaters, details of these values have been made available in reports (boundary values: Phillips and Pitt, 2015; analysis of GIG data sets: Phillips et al., 2016) and are summarised in the following two sections. For transitional and coastal waters more details of the analysis of GIG data sets are provided in Section 4.5.

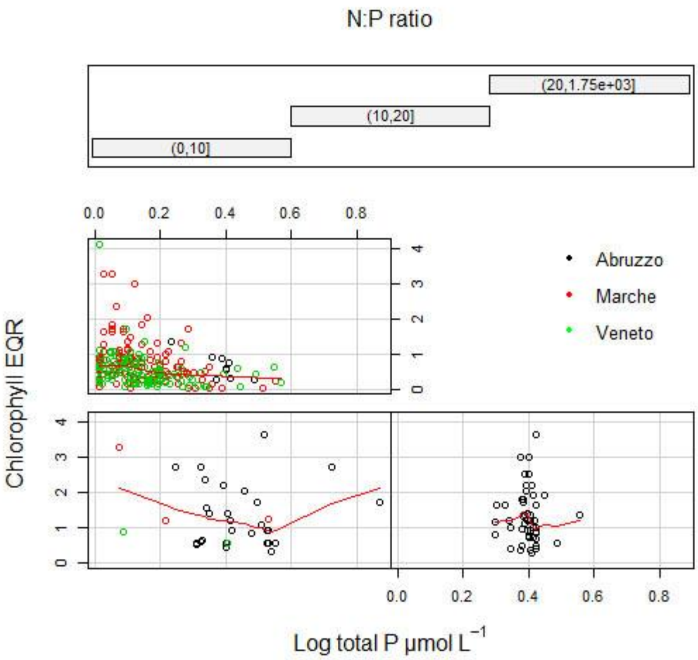
It is important that similar types are compared, and for freshwaters one approach is to use the recently developed European broad typology (Lyche-Solheim et al., 2015), or alternatively the IC typology. In freshwaters, where boundary values calculated using the toolkit fall beyond the ranges predicted from the analysis of GIG data (option 1) or outside the upper and lower quartiles of the boundary values used by other countries (option 2), then further validation of the boundaries may be required.

Option 1: results of analyses of GIG data sets

Available data from the IC exercise were analysed by Phillips et al. (2016) to provide an indication of the ranges of nutrient boundary values for lakes and rivers that could be produced from pressure response relationships. Two ranges of values are presented; the 'most likely' taken from the minimum and maximum value predicted from the different regression and categorical approaches and a 'possible' range taken from the maximum and minimum of the upper and lower quartiles of the regression residuals. Additional data for lakes and rivers were collated during this project to supplement the results in Phillips et al. (2016), although few adequate relationships were found. Details of these relationships for alpine lakes and an overview of river data are presented in Appendix 2. The current results are provided in Table 4-1 (TP in lakes), Table 4-2 (soluble phosphorus in rivers) and Table 4-3 (TN in rivers). Be aware, when making comparisons with the data in these tables, that the uncertainty of some of the relationships, particularly for rivers, produces substantial ranges of 'possible' boundary values. The reported values were not subject to the validation checks proposed in this document; they are provided for comparison to allow those carrying out similar analysis to place their results into a wider context.



(a)



(b)

Figure 4-1: Relationship between EQR for chlorophyll and (a) total nitrogen (log_{10}) across different ranges of total phosphorus concentration (concentrations increase from bottom left to top right panels) and (b) total phosphorus (log_{10}) across different ranges of the N:P molar ratio (0-10; 10-20; > 20). Examples using data from Italy, coastal waters in the Adriatic (MED II).

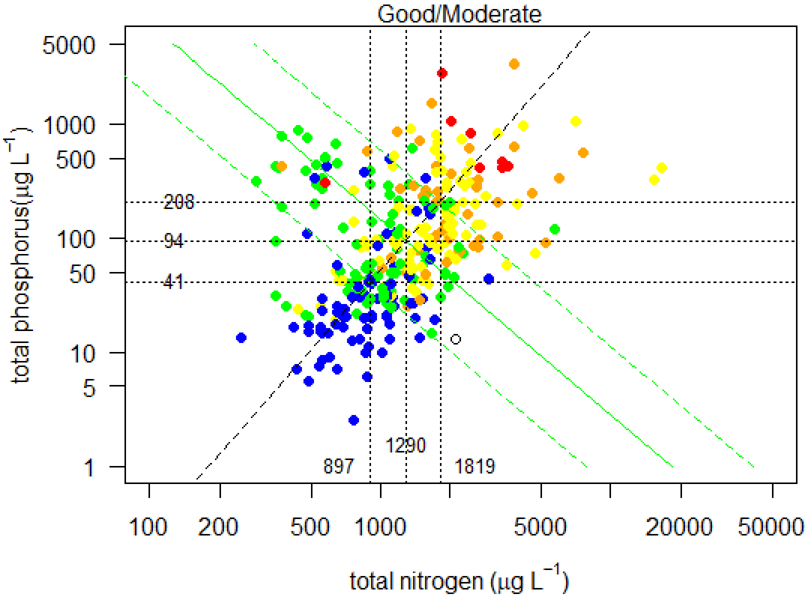


Figure 4-2: Relationship between mean TN and TP for very shallow calcareous lakes. Points coloured by phytoplankton class, dotted line marks the mean N:P ratio, green lines show contours that would predict the good/moderate EQR boundary of 0.6 (dotted green lines show uncertainty band). Horizontal and vertical dotted lines show potential pairs of TN and TP boundaries together with their uncertainty ranges.

Table 4-1: Range of TP boundary values for broad lake types across Europe (where appropriate data were available).

	Broad type	IC type	r^2	Metric	G/M (TP µg L^{-1})				H/G (TP µg L^{-1})			
					most likely range		possible range		most likely range		possible range	
1	Very large deep stratified	L-AL3										
		L-N2b										
2	Lowland siliceous	L-N2a	0.37	Phytoplankton	11	22	9	31	8	10	6	13
		L-N2b	0.37	Phytoplankton	8	15	7	20	6	8	5	10
		L-N1	0.81	Phytoplankton	18	20	15	23	11	12	9	15
3	Lowland calcareous/mixed, stratified	L-AL3										
		L-CB1	0.55	Phytoplankton	35	44	28	61	22	32	15	37
4	Lowland, calcareous/mixed very shallow unstratified	L-CB2	0.68	Phytoplankton	45	70	35	122	32	35	22	55
		L-EC1	0.15	Phytoplankton	Regression not significant							
5	Lowland, organic and siliceous	L-N3a	0.61	Phytoplankton	17	24	14	31	11	14	9	16
		L-N8a	0.80	Phytoplankton	26	28	20	38	14	19	11	23
6	Lowland, organic and calcareous/mixed											
7	Mid altitude, siliceous	L-N5										
8	Mid altitude, calcareous/mixed	L-AL4	0.72	Phytoplankton	22	27	19	37	9	14	8	17
		L-AL3	0.63	Phytoplankton	11	15	10	20	5	8	4	8
9	Mid altitude organic and siliceous	L-N6a	0.41	Phytoplankton	14	31	10	44	10	15	8	21
10	Mid altitude organic and calcareous/mixed				Not available							
11	Highland, siliceous				Not available							
12	Highland, calcareous/mixed				Not available							
13	Mediterranean, small-large siliceous (incl. reservoirs)	L-M5/7			Not available							

	Broad type	IC type	r^2	Metric	G/M (TP µg L^{-1}) most likely range	G/M (TP µg L^{-1}) possible range	H/G (TP µg L^{-1}) most likely range	H/G (TP µg L^{-1}) possible range
14	Mediterranean, small-large calcareous/mixed (incl. reservoirs)	L-M8			Not available			
15	Mediterranean, very small				Not available			

Table 4-2: Range of soluble P ([8]) boundary values for broad river types across Europe (where data were available).

	Broad type	IC type	r^2	Metric	G/M (P µg L^{-1}) predicted range		G/M (P µg L^{-1}) possible range		H/G (P µg L^{-1}) predicted range		H/G (P µg L^{-1}) possible range	
1	Very large rivers (all Europe)	R-L1 and R-L2	0.357	Phytobenthos	40	56	27	117	16	39	8	39
			Quantile	Invertebrates	n/a	n/a	40	170	n/a	n/a	16	105
2	Lowland, siliceous, medium-large											
3	Lowland, siliceous, very small-small	R-C1	0.480	Macrophytes	32	45	18	98	11	22	5	37
			0.490	Phytobenthos	31	62	16	126	10	17	7	36
			0.500	Combined	20	36	12	72	8	12	6	18
4	Lowland, calcareous/mixed, medium-large				Not available							
5	Lowland, calcareous/mixed, very small-small				Not available							
6	Lowland, organic and siliceous				Not available							
7	Lowland, organic and calcareous/mixed				Not available							
8	Mid-altitude, siliceous, medium-large				Not available							
9	Mid-altitude, siliceous, very small-small	R-C3	0.400	Macrophytes	48	128	25	128	11	18	5	50
			0.430	Phytobenthos	34	86	22	124	13	25	7	45

([8]) Details of the form of soluble P were not available for these IC data sets at the time of analysis. It is most likely that they were reactive ortho-phosphate.

	Broad type	IC type	r^2	Metric	G/M (P µg L^{-1})				H/G (P µg L^{-1})			
					predicted range		possible range		predicted range		possible range	
			0.500	Combined	25	46	17	93	6	13	2	27
10	Mid-altitude, calcareous/mixed, medium-large				Not available							
11	Mid-altitude, calcareous/mixed, very small-small				Not available							
12	Mid-altitude, organic and siliceous				Not available							
13	Mid-altitude, organic and calcareous/mixed				Not available							
14	Highland (all Europe), siliceous including organic				Not available							
15	Highland (all Europe), calcareous/mixed				Not available							
16	Glacial rivers (all Europe)				Not available							
17	Mediterranean, lowland, medium-large, perennial				Not available							
18	Mediterranean, mid-altitude, medium-large, perennial				Not available							
19	Mediterranean, very small-small, perennial				Not available							
20	Mediterranean, temporary/intermittent streams				Not available							

Table 4-3: Range of total nitrogen boundary values for broad river types across Europe (where data were available).

	Broad type	IC type	r^2	Metric	G/M (TN mg L^{-1})				H/G (TN mg L^{-1})			
					predicted range		possible range		predicted range		possible range	
1	Very large rivers (all Europe)	R-L1 and R-L2	0.236	Phytobenthos	1.6	2.5	1.1	3.3	0.5	1.0	0.2	1.4
2	Lowland, siliceous, medium-large				Not available							
3	Lowland, siliceous, very small-small	R-C1	0.480	Macrophytes	1.4	3.5	0.5	9.5	0.2	0.9	0.1	2.3
			0.490	Phytobenthos	1.9	4.6	0.9	12.3	0.6	0.9	0.2	2.3
			0.540	Combined	1.0	2.1	0.2	5.3	0.2	0.6	0.1	1.1

	Broad type	IC type	r^2	Metric	G/M (TN mg L^{-1})				H/G (TN mg L^{-1})			
					predicted range		possible range		predicted range		possible range	
4	Lowland, calcareous/mixed, medium-large				Not available							
5	Lowland, calcareous/mixed, very small-small				Not available							
6	Lowland, organic and siliceous				Not available							
7	Lowland, organic and calcareous/mixed				Not available							
8	Mid-altitude, siliceous, medium-large				Not available							
9	Mid-altitude, siliceous, very small-small	R-C3	0.490	Macrophytes	1.3	6.0	0.8	12.0	0.5	0.7	0.2	1.3
			0.530	Phytobenthos	1.4	3.8	0.9	8.3	0.7	0.9	0.3	2.2
			0.540	Combined	0.9	2.2	0.6	5.0	0.4	0.4	0.2	0.9
10	Mid-altitude, calcareous/mixed, medium-large				Not available							
11	Mid-altitude, calcareous/mixed, very small-small				Not available							
12	Mid-altitude, organic and siliceous				Not available							
13	Mid-altitude, organic and calcareous/mixed				Not available							
14	Highland (all Europe), siliceous including organic				Not available							
15	Highland (all Europe), calcareous/mixed				Not available							
16	Glacial rivers (all Europe)				Not available							
17	Mediterranean, lowland, medium-large, perennial				Not available							
18	Mediterranean, mid-altitude, medium-large, perennial				Not available							
19	Mediterranean, very small-small, perennial				Not available							
20	Mediterranean, temporary/intermittent streams				Not available							

Option 2: summary of existing Member State boundaries (excluding those set using expert opinion and the distribution of nutrients in all water bodies approach)

Table 4-4: Range of national good/moderate TP boundary values (µg L^{-1}) for lakes, reported by Member States where method used was not expert judgement or from distribution in all water bodies and summary metric was mean or median, derived using data from Phillips and Pitt (2015). Min or max replace quartiles when N ≤ 3, or single value shown.

	Lake broad type	Lower 25th quartile (min)	Median	Upper 75th quartile (max)	Number of national boundary values
1	Very large, deep stratified lakes	(15)	18	(25)	3
2	Lowland, siliceous	11	18	23	29
3	Lowland, calcareous/mixed, stratified	26	39	55	27
4	Lowland, calcareous/mixed, very shallow, unstratified	33	60	62	30
5	Lowland, organic and siliceous	21	24	28	20
6	Lowland, organic and calcareous/mixed	23	24	60	5
7	Mid-altitude, siliceous	13	16	34	21
8	Mid-altitude, calcareous/mixed	18	30	34	23
9	Mid-altitude, organic and siliceous	15	18	21	8
10	Mid-altitude, organic and calcareous/mixed		22		1
11	Highland, siliceous	9	16	28	15
12	Highland, calcareous/mixed	10	15	20	3
13	Mediterranean, small-large siliceous (incl. reservoirs)	24	42	51	6
14	Mediterranean, small-large calcareous/mixed	20	23	29	13
15	Mediterranean, very small		26		1

(Countries include AT, DK, EE, FI, FR, HR, IT, LT, NL, NO, PL, PT, SE and UK.)

Table 4-5: Range of national good/moderate TN boundary values (mg L^{-1}) for lakes, reported by Member States where method used was not expert judgement or from distribution in all water bodies and summary metric was mean or median, derived using data from Phillips and Pitt (2015). Min or max replace quartiles when N ≤ 3, or single value shown.

Lake broad type		Lower 25th quartile (min)	Median	Upper 75th quartile (max)	Number of national boundary values
1	Very large, deep stratified lakes	(0.50)	0.55	(0.60)	2
2	Lowland, siliceous	0.48	0.48	0.50	8
3	Lowland, calcareous/mixed, stratified	0.95	1.63	2.00	11
4	Lowland, calcareous/mixed, very shallow, unstratified	0.95	1.15	1.48	8
5	Lowland, organic and siliceous	0.65	0.66	0.70	5
6	Lowland, organic and calcareous/mixed	(0.90)	1.10	(1.30)	2
7	Mid-altitude, siliceous	0.43	0.43	0.71	6
8	Mid-altitude, calcareous/mixed				
9	Mid-altitude, organic and siliceous		0.55		2
10	Mid-altitude, organic and calcareous/mixed				
11	Highland, siliceous	0.25	0.25	0.48	6
12	Highland, calcareous/mixed				
13	Mediterranean, small-large siliceous (incl. reservoirs)				
14	Mediterranean, small-large calcareous/mixed				
15	Mediterranean, very small				

(Countries include DK, FI, EE, LT, NL, NO, PL and PT)

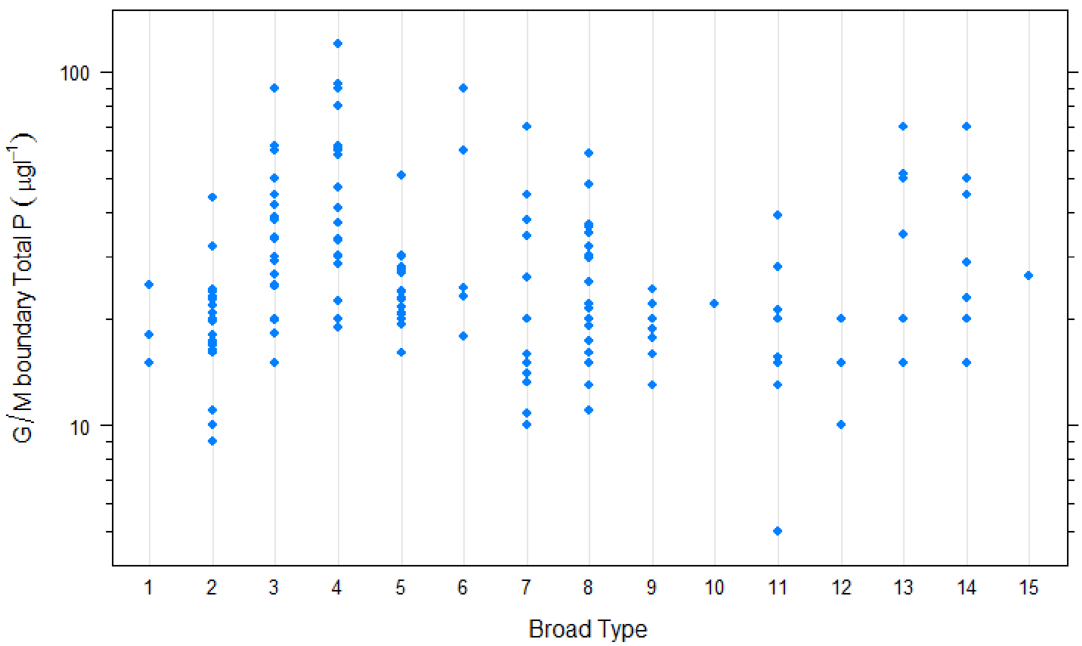


Figure 4-3: Range of national good/moderate boundary TP values for broad lake types where method used was not expert judgement or from distribution in all water bodies (redrawn from data collated by Phillips and Pitt, 2015).

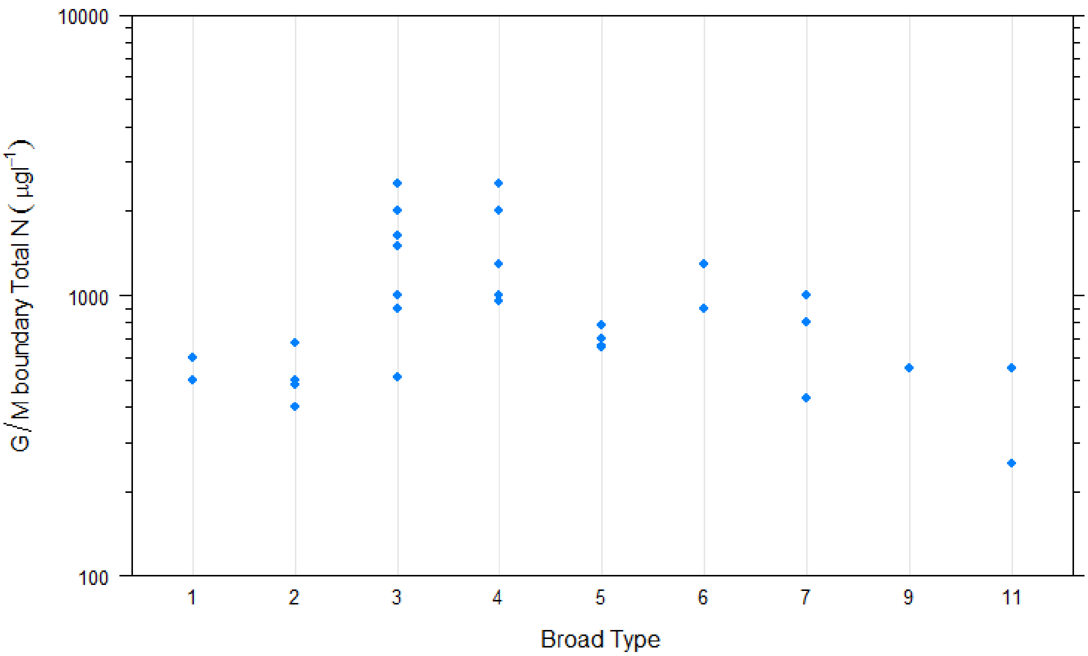


Figure 4-4: Range of national good/moderate TN boundary values for broad lake types where method used was not expert judgement or from distribution in all water bodies (redrawn from data collated by Phillips and Pitt, 2015).

Table 4-6: Range of national good/moderate TP (µg L^{-1}) boundary values for rivers, reported by Member States where method used was not expert judgement or from distribution in all water bodies and summary metric was mean or median, derived using data from Phillips and Pitt (2015). Min or max replace quartiles when N ≤ 3, or single value shown. Countries included: CY, CZ, FI, HR, LT, NL, NO and SE.

Broad type		Lower 25th quartile (min)	Median	Upper 75th quartile (min)	Number national boundary values
1	Very large rivers (all Europe)	70	140	200	7
2	Lowland, siliceous, medium-large	22	25	28	5
3	Lowland, siliceous, very small-small	19	28	91	14
4	Lowland, calcareous/mixed, medium-large	70	110	140	18
5	Lowland, calcareous/mixed, very small-small	37	70	110	13
6	Lowland, organic and siliceous	29	32	34	13
7	Lowland, organic and calcareous/mixed	33	40	50	9
8	Mid-altitude, siliceous, medium-large	(14)	15	(17)	2
9	Mid-altitude, siliceous, very small-small	15	15	15	6
10	Mid-altitude, calcareous/mixed, medium-large	50	50	60	8
11	Mid-altitude, calcareous/mixed, very small-small	45	50	60	9
12	Mid-altitude organic and siliceous	20	20	23	6
13	Mid-altitude organic and calcareous/mixed		28		2
14	Highland (all Europe), siliceous, incl. organic	8	8	17	6
15	Highland (all Europe), calcareous/mixed	30	30	30	5
16	Glacial rivers		no results		
17	Mediterranean, lowland, medium-large, perennial	60	60	85	3
18	Mediterranean, mid-altitude, medium-large, perennial		60		1
19	Mediterranean, very small-small, perennial	60	80	102	4
20	Mediterranean, temporary/intermittent streams		165		2

Table 4-7: Range of national good/moderate TN (mg L^{-1}) boundary values for rivers, reported by Member States where method used was not expert judgement or from distribution in all water bodies and summary metric was mean or median, derived using data from Phillips and Pitt (2015). Min or max replace quartiles when N ≤ 3, or single value shown. Countries included: FI, HR, LT, NL and NO.

	River broad type	Lower 25th quartile (min)	Median	Upper 75th quartile (min)	Number national boundary values
1	Very large rivers (all Europe)	1.3	2.5	2.5	7
2	Lowland, siliceous, medium-large		0.8		1
3	Lowland, siliceous, very small-small	0.48	0.74	2.3	10
4	Lowland, calcareous/mixed, medium-large	2.3	2.5	3	11
5	Lowland, calcareous/mixed, very small-small	2.15	2.3	2.45	7
6	Lowland, organic and siliceous	0.65	0.78	0.9	5
7	Lowland, organic and calcareous/mixed		0.78		1
8	Mid-altitude, siliceous, medium-large			no results	
9	Mid-altitude, siliceous, very small-small	0.43	0.43	0.43	4
10	Mid-altitude, calcareous/mixed, medium-large	1.5	1.5	1.5	3
11	Mid-altitude, calcareous/mixed, very small-small	1.45	1.5	1.75	3
12	Mid-altitude organic and siliceous		0.55		2
13	Mid-altitude organic and calcareous/mixed			no results	
14	Highland (all Europe), siliceous, incl. organic	(0.25)	0.55	(0.55)	3
15	Highland (all Europe), calcareous/mixed			no results	
16	Glacial rivers			no results	
17	Mediterranean, lowland, medium-large, perennial	(1)	1	(1. 7)	3
18	Mediterranean, mid-altitude, medium-large, perennial	1	1	1	1
19	Mediterranean, very small-small, perennial	1	1.25	1.5	4

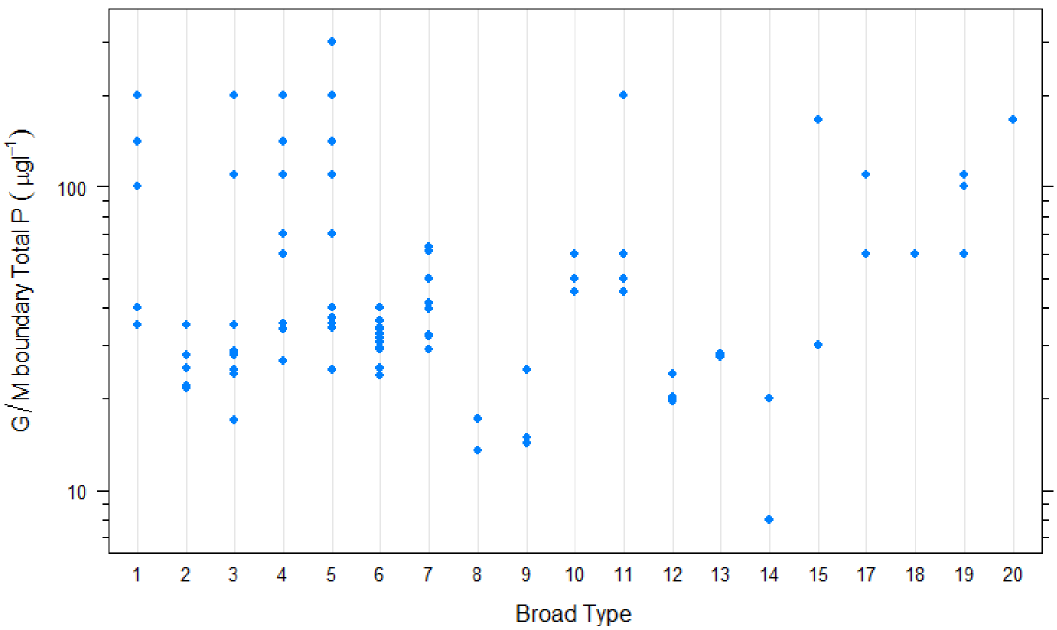


Figure 4-5: Range of national good/moderate boundary TP values for broad river types where method used was not expert judgement or from distribution in all water bodies (redrawn from data collated by Phillips and Pitt, 2015)

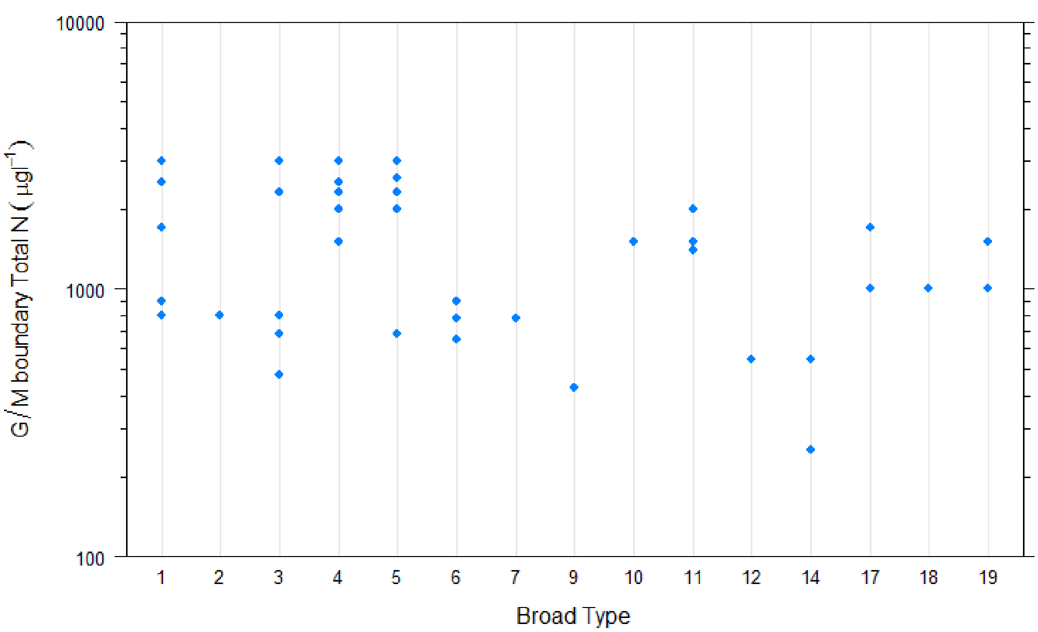


Figure 4-6: Range of national good/moderate TN boundary values for broad river types where method used was not expert judgement or from distribution in all water bodies (redrawn from data collated by Phillips and Pitt, 2015).

4.5. Ranges of nutrient standards associated with coastal and transitional water types

For transitional water (TRW) and coastal water (CW) categories, available GIG data sets have been analysed using the toolkit procedures in order to obtain type-specific nutrient boundary values. Most of these analyses are described in Appendix 4, with a few examples presented in this report (e.g. TRW NEA11 common type). The nutrient boundaries estimated for GIG data sets using the approaches recommended in this document should be compared with the nutrient boundaries reported by Member States within similar types (Dworak et al., 2016), and/or with other existing boundary values, for example adopted by the regional sea conventions.

Where a mismatch is found between the boundary values predicted using methods in the toolkit and those reported by Member States and/or adopted within the regional sea conventions, then further consideration of the validity of the boundaries may be required (see Chapter 6). A summary of these boundary values per common typology is presented in the following subsection.

For Member States using this document in the future, their existing national boundary values or the ones created using the toolkit should be compared with values reported by other countries in similar types and with values obtained from the analysis of the GIG data sets.

Results of analyses of GIG data sets in transitional waters and coastal waters

Data from the biological quality element phytoplankton available from the WFD IC exercise were analysed by Teixeira and Salas (2017) to provide an indication of the ranges of nutrient boundary values for common water types in coastal and transitional waters (CTRW) that could be produced from pressure–response relationships (this report is provided as Appendix 4). Phytoplankton data, mostly chlorophyll *a* (Chl*a*), the metric used for IC, was used to establish pressure–response relationships with nutrients. The nutrients analysed for this exercise were TN, TP and DIN, depending on the water systems. For these nutrients, two ranges of values are presented: the 'most likely' taken from the minimum and maximum value predicted from the different regression and categorical approaches and a 'possible' range taken from the maximum and minimum of the upper and lower quartiles of the regression residuals. Where other approaches have been tested (e.g. bivariate linear regression, logistic binomial regression or quantile regression), the results are also presented along with derived nutrient boundaries.

This section identifies the CTRW data sets available and the respective EQR boundaries produced during IC which were used for this exercise (Table 4-8). It also shows which data sets have been analysed for this exercise, summarises the major constraints found in each data set (Table 4-9 TRW and Table 4-10 CW), and indicates the results currently available from analysing these data sets using the toolkit. Finally we present the proposed nutrient boundaries (Table 4-11 and Table 4-12) derived from the analyses undertaken, and compare them with the Member State nutrient boundaries, where available (Tables 4-13 and 4-14).

The relationships found using linear regression approaches in coastal lagoons were acceptable (Table 4-9), although few adequate relationships were found for the remaining transitional waters data sets or for coastal waters using this approach (Table 4-10). The CTRW data analysed revealed that univariate linear regression approaches might not be robust for these water categories and alternative approaches, for example multivariate regression or categorical methods, might need to be considered (Teixeira and Salas, 2017). However, several common CW types have not been analysed for this exercise, in particular for the north-east Atlantic (NEA) GIG (see Appendix 4 for details).

Another limitation while deriving common nutrient boundaries at the type level was the use of non-normalised data from different Member States within common IC types. Therefore, some of the values presented within each type are based on data from a single Member State and may not apply to the whole common type.

Further testing with transitional waters data sets from the NEA GIG common type NEA11 has been used to illustrate how the guidance should be applied in situations where linear regression methods are unlikely to produce meaningful results (see Section 7.2). There will be cases where the biota shows a poor relationship with nutrients due to the presence of other pressures and/or environmental factors, often not covered by the data sets, and which are controlling phytoplankton response to nutrients.

The NEA11 TRW example also demonstrates how to overcome some practical issues of combining all the Member State data sets available within a type, by, for example, normalising intercalibrated EQRs. This will compensate for insufficient gradient coverage and/or small numbers of samples within a country, and allow more meaningful nutrient boundaries for the whole type to be produced.

Details of the relationships found using earlier versions of the toolkit (Table 4-9 and Table 4-10) are presented in Appendix 4, which includes pressure–response relationships for TRW data (coastal lagoons and estuaries) in the Baltic, NEA and Mediterranean, and for CW data in the Baltic and Mediterranean. Where results from the toolkit were not adequate, the use of other approaches, as for example those suggested in Chapter 5 of this document, may be required.

A summary of selected boundaries derived from the most plausible results (i.e. significant and/or meaningful) obtained for each common type, using the toolkit, are presented in Tables 4-11 (transitional waters) and 4-12 (coastal waters). See Appendix 4 for a detailed discussion of the results presented here.

Table 4-8: WFD EQR IC boundaries for BQE phytoplankton used in this exercise (adopted from European Commission, 2018).

Water category	Common type	Country	Data sets	EQR boundaries			
				H/G	G/M	M/P	P/B
TRW	BALBT1	LT	ds1	0.83	0.57	0.39	0.29
		PL	ds2	0.77	0.61	0.5	0.4
	MEDpolyCL	IT/GR	ds3	0.78	0.51	—	—
		FR	ds4	0.71	0.39	—	—
	NEA11	NL/UK/IE	ds25	0.8	0.6	—	—
		FR	ds26	0.67	0.393	—	—
		ES	ds27	0.67	0.37	—	—
		PT	ds28	0.667	0.467	—	—
CW	BALBC4	LV	ds5	0.82	0.67	0.33	0.23
		EE	ds6	0.83	0.67	0.33	0.23
	BALBC5	LV	ds7	0.65	0.39	0.33	0.2
		LT	ds8	0.87	0.6	0.28	0.21
	MEDI	IT	ds9	0.85	0.62	—	—
	MEDII Adriatic	IT	ds10	0.81	0.6	—	—
	MEDII Tyrrhenian	IT	ds11	0.84	0.62	—	—
	MEDIIIE	GR/CY	ds12	0.66	0.37	—	—
	NEA1-26A	FR	ds13	0.76	0.33	—	—
		IE	ds14	0.82	0.6	—	—
		ES/NO	ds15	0.67	0.33	—	—
		UK	ds16	0.8	0.6	—	—
	NEA1-26B	FR	ds17	0.67	0.44	—	—
		UK*south*	ds18	0.82	0.63	—	—
		NL/UK*north*	ds19	0.8	0.6	—	—
		BE	ds20	0.8	0.67	—	—
	NEA1-26C	DK/DE	ds21	0.67	0.44	—	—
	NEA1-26E	SP/PT*sUpW*	ds22	0.67	0.44	—	—
		PT*UpW*	ds23	0.88	0.49	—	—
	NEA3-4	DE/NL	ds24	0.8	0.6	—	—

Table 4-9: List of TRW data sets available for analysis, with indication of specific combination of countries with unique IC EQR boundaries and nutrient data available. Summary of the results for the toolkit analysis performed in each data set; results correspond to univariate linear regression type II (LR), if not mentioned otherwise: BvR — bivariate linear regression; LQR — linear quantile regression; AQR — additive quantile regression. Figures in red indicate where $r^2 < 0.36$ and/or not significant.

GIG	Common type	Country	Dataset	Nutrient tested	Complete obs N	R^2	Pearson *r*	*P* value	Notes	Toolkit version
Transitional waters										
Baltic	BT1	Lithuania	ds1	TN	25	**0.41**	0.641	<0.001		vs3
				TP	24	**0.432**	0.657	<0.001		vs3
				N+P BvR	23	**0.556**	0.745	<0.001		vs3
		Poland	ds2	TN	13	**0.86**	0.927	<0.001	Check interaction TN-TP	vs3
				TP	25	**0.209**	0.457	0.022	Check interaction TN-TP	vs3
				N+P BvR	23	**0.088**	0.296	0.154		vs3
Mediterranean	polyhaline CL	Italy/Greece	ds3	TN	12	**0.778**	0.882	<0.001	H/G/M range not covered; predictions outside data range	vs3
				TP	15	**0.603**	0.777	<0.001		vs3
		Analysis for:		DIN					*Not performed*	
		France	ds4	TN	13	**0.642**	0.801	<0.001	Collinearity TN-TP	vs3
				TP	14	**0.868**	0.932	<0.001	Collinearity TN-TP	vs3
		Analysis for:		DIN					*Not performed*	
North-east Atlantic	NEA11	Netherlands/ UK/ Ireland	ds25	DIN	98	**--**	--	--	EQS G/M class overlap	vs6c
		France	ds26	DIN	15	*excluded*			*insufficient gradient coverage; invert trend*	vs6c
		Spain	ds27	DIN	55	**--**	--	--	EQS G/M class overlap	vs6c
		Portugal	ds28	DIN	7	**--**	--	--	very low n; insufficient gradient coverage	vs6c
		Merged ds *normalised EQRs*	ds: 25, 27, 28	DIN	160	**0.210**	-0.458	<0.001	Wedge shape; EQS G/M class overlap	vs6c
				DIN AdQR	160(70th)			<0.001		vs6c

Table 4-10: List of CW data sets available for analysis, with indication of specific combination of countries with unique IC EQR boundaries and nutrient data available. Summary of the results for the toolkit analysis performed in each data set; results correspond to univariate linear regression type II (LR), if not mentioned otherwise: BvR — bivariate linear regression; LQR — linear quantile regression; AQR — additive quantile regression. Figures in red indicate where r^2 < 0.36 and/or not significant.

GIG	Common type	Country	Dataset	Nutrient tested	Complete obs N	R^2	*Pearson r*	*P* value	Notes	Toolkit version
Coastal waters										
Baltic	BC4	Latvia	ds5	TN	79	**0.284**	0.533	<0.001	Wedge shape; predictions outside data range	vs3
				TP	81	**0.308**	0.555	<0.001	Wedge shape	vs3
		Estonia	ds6	TN	22	**0.756**	0.87	<0.001	EQS G/M class overlap	vs3
				TP	40	**0.26**	0.501	<0.001	Wedge shape	vs3
	BC5	Latvia	ds7	TN	98	**0.48**	0.693	<0.001	Wedge shape; EQS G/M class overlap	vs3
				TP	98	**0.257**	0.507	<0.001	Wedge shape; EQS H/G class overlap	vs3
		Lithuania	ds8	TN	61	**0**	0.013	0.919	No trend in data	vs3
				TP	61	**0.214**	0.462	<0.001	Wedge shape	vs3
Mediterranean	MEDI	Italy	ds9	TN	82	**0.098**	0.314	0.004	Wedge shape	vs3
				TP	83	**0.043**	0.208	0.059	Wedge shape	vs3
		Analysis for:	*NO3*	*NO2*	*NH4*	*PO4*	*Si*		*Not performed*	
	MEDII Adriatic	Italy	ds10	TN	316	**0.217**	0.466	<0.001	Wedge shape	vs3
				TN (LQR)	332 (80th)	**0.615**	0.785		EQR range >>1	R
				N+P (BvR)	294	**0.228**	0.477	<0.001	H/G boundaries inversion	vs3
				TP	309	**0.066**	0.258	0.6	No trend in data	vs3
		Analysis for:	*NO3*	*NO2*	*NH4*	*PO4*	*Si*		*Not performed*	
	MEDII Tyrrhenian	Italy	ds11	TN	228	**0.088**	0.297	<0.001	Wedge shape	vs3
				TP	228	**0.005**	0.068	0.307	Wedge shape	vs3
		Analysis for:	*NO3*	*NO2*	*NH4*	*PO4*	*Si*		*Not performed*	
	MEDIIIE	Greece/ Cyprus	ds12	NO3		**0.132**	0.363	<0.001		vs3
		Analysis for:	*NH4*	*NO2*	*PO4*				*Not performed*	
North-east Atlantic	NEA1-26A	France	ds13	DIN	45				Merged	
		Ireland	ds14	DIN	41				Merged	
		Spain/ Norway	ds15	DIN	30				Insufficient gradient coverage; merged	
		UK	ds16	DIN	11				Insufficient gradient coverage; merged	
		Merged ds *normalised EQRs*	ds: 13-16	DIN	113				Inverted wedge shape	vs6c
	NEA1-26B	France	ds17	DIN	4	--	--	--	Very low N	
		UKsouth	ds18	DIN	40				*Not performed*	
		Netherlands/ UKnorth	ds19	DIN	18				*Not performed*	
		Belgium	ds20	DIN	3	--	--	--	Very low N	
	NEA1-26C	Denmark/ Germany	ds21	DIN	8	--	--	--	H/G/M range not covered; Very low N	
	NEA1-26E	Portugal/ Spain	ds22	DIN	25				*Not performed*	
		Portugal	ds23	DIN	6	--	--	--	Very low N	
	NEA3-4	Germany/ Netherlands	ds24	DIN	14				*Not performed*	

Table 4-11: Predicted nutrient boundaries for transitional waters common types, from the significant or most adequate approaches for each data set (toolkit Excel vs3 or vs6c and/or R scripts). Nutrient boundaries are expressed in the units used by each Member State in their national assessments. Values in red indicate that underlying relationships are not significant or that there are no robust data to support boundary positions.

Transitional waters common type and methods		Nutrient boundaries		Nutrient boundaries	
BAL BT1					
Lithuania		**TN**	µg L^{-1}	**TP**	µg L^{-1}
Regression methods (OLS and type II):		**H/G**	**G/M**	**H/G**	**G/M**
Most likely boundary predicted		1020	1224	73	89
	range	928-1084	1218-1298	66-78	88-90
	possible range	845-1187	1122-1333	61-86	82-99
Regression methods (bivariate TN+TP):					
Most likely boundary predicted		1014	1218	71	89
	range	939-1073	1128-1298	65-77	81-96
Categorical methods:					
Average adjacent class upper and lower quartiles				74	84
Average adjacent class median		1101	1206	74	85
75th quartile of class		1168	1235	72	85
Mismatch of biological v nutrient class (Excel)		960	1240	67	83
Poland		**TN**	µg L^{-1}	**TP**	µg L^{-1}
Regression methods (OLS and type II):		**H/G**	**G/M**	**H/G**	**G/M**
Most likely boundary predicted		948	1072		
	range	940-956	1071-1073		
	possible range				
Categorical methods:					
Average adjacent class upper and lower quartiles		662	1022		
Average adjacent class median		662	1022		
75th quartile of class		400	923		
Mismatch of biological v nutrient class (Excel)		600	900	88	101
MED Polyhaline Coastal Lagoons					
Italy/Greece		**TN**	µg L^{-1}	**TP**	µg L^{-1}
Regression methods (OLS and type II):		**H/G**	**G/M**	**H/G**	**G/M**
Most likely boundary predicted		1039		27	47
	range	1031-1049		25-28	44-53
	possible range	840-1176		17-38	25-97
Categorical methods:					
Average adjacent class upper and lower quartiles		1103		23	63
Average adjacent class median		1077		23	66
75th quartile of class		840	1463	28	25
Mismatch of biological v nutrient class (Excel)		870		21	97
France		**TN**	µg L^{-1}	**TP**	µg L^{-1}
Regression methods (OLS and type II):		**H/G**	**G/M**	**H/G**	**G/M**
Most likely boundary predicted		261	587	18	42
	range	216-304	582-594	17-19	42-42
	possible range	132-432	362-929	14-23	23-55
Categorical methods:					
Average adjacent class upper and lower quartiles		364	565	20	35
Average adjacent class median		374	559	21	39
75th quartile of class		432	470	21	23
Mismatch of biological v nutrient class (Excel)		225	570	19	28

Table 4-11 (cont.)

Transitional waters common type and methods	Nutrient boundaries		Nutrient boundaries
NEA 11 (NL, UK, IE, SP, PT)	**DIN**	µM	
Regression methods (OLS and type II):	**H/G**	**G/M**	
Most likely boundary *predicted*	36	62	
range	(14-43)	(61-72)	
possible range	5-79	23-278	
Additive quantile regression method (rqss):			
70th percentile	68	212	
Categorical methods:			
Average adjacent class upper and lower quartiles	49	80	
Average adjacent class median	47	82	
75th quartile of class	62	107	
Mismatch of biological v nutrient class (Excel)	50	72	
Mismatch of biological v nutrient class (R scripts)	53	75	
range	(47-59)	(66-83)	
Logistic binomial regression (*prob = 0.5*)	45	80	

Table 4-12: Predicted nutrient boundaries for coastal waters common types, from the significant or most adequate approaches for each data set (toolkit Excel vs3 or vs6c and/or R scripts).

Coastal waters common type and methods	Nutrient boundaries			
BAL BC4				
Latvia	**TN**	µM	**TP**	µM
Categorical methods:	**H/G**	**G/M**	**H/G**	**G/M**
average adjacent class upper and lower quartiles	29	30.4	0.68	0.72
Average adjacent class median	29.4	29.4	0.70	0.77
75th quartile of class	28.5	28.5	0.65	0.78
Mismatch of biological v nutrient class (Excel)	25.5	25.5	0.53	0.62
Estonia	**TN**	µM	**TP**	µM
Regression methods (OLS and type II):	**H/G**	**G/M**	**H/G**	**G/M**
Most likely boundary predicted	22.7	24.8		
range	22.4-22.9	24.8-24.8		
possible range	20.4-25.2	22.6-28.5		
Categorical methods:				
Average adjacent class upper and lower quartiles	25.2	27	0.69	0.76
Average adjacent class median	24.8	27.8	0.64	0.67
75th quartile of class	22.2	28.5		0.82
Mismatch of biological v nutrient class (Excel)	22.5	23.5	0.48	0.55
BAL BC5				
Latvia	**TN**	$\mu g\ L^{-1}$	**TP**	$\mu g\ L^{-1}$
Regression methods (OLS and type II):	**H/G**	**G/M**	**H/G**	**G/M**
Most likely boundary predicted	312	368		
range	292-327	366-370		
possible range	260-360	330-410		
Categorical methods:				
Average adjacent class upper and lower quartiles	332	353	20	21
Average adjacent class median	331	347	20	22
75th quartile of class	339	375		
Mismatch of biological v nutrient class (Excel)	320	340	18	21

Table 4-12 (cont.)

Coastal waters common type and methods	Nutrient boundaries			
BAL BC5				
Lithuania	**TN**	µg L^{-1}	**TP**	µg L^{-1}
Categorical methods:	**H/G**	**G/M**	**H/G**	**G/M**
Average adjacent class upper and lower quartiles			21	25
Average adjacent class median	388	409	20	25
Mismatch of biological v nutrient class (Excel)	190	285	13	23
MED I (Italy)	**TN**	µM	**TP**	µM
Categorical methods:	**H/G**	**G/M**	**H/G**	**G/M**
Average adjacent class upper and lower quartiles			0.58	0.72
Average adjacent class median	26	43	0.62	0.72
75th quartile of class			0.71	0.81
Mismatch of biological v nutrient class (Excel)	24	39	0.57	0.76
MED II Adriatic (Italy)	**TN**	µM	**TP**	µM
Linear quantile regression:	**H/G**	**G/M**	**H/G**	**G/M**
80th percentile	97	197		
Categorical methods:				
Average adjacent class upper and lower quartiles	23.7	33.2		
Average adjacent class median	24.8	33.2		
75th quartile of class	28.6	43.4		
Mismatch of biological v nutrient class (Excel)	23	30	0.35	0.58
MED II Tyrrhenian (Italy)	**TN**	µM	**TP**	µM
Categorical methods:	**H/G**	**G/M**	**H/G**	**G/M**
Average adjacent class upper and lower quartiles	15.4	19.6	0.22	0.46
Average adjacent class median	13.8	18	0.15	0.20
75th quartile of class	16.8	28.3	0.40	0.87
Mismatch of biological v nutrient class (Excel)	15.3	20	0.34	0.42
MED III (Greece, Cyprus)	**NO_3**	µM		
Categorical methods:	**H/G**	**G/M**		
75th quartile of class	0.99	1.25		
Mismatch of biological v nutrient class (Excel)	0.74	1.06		

The national nutrient boundaries adopted by Member States (Table 4-13) for the CTRW common types analysed in this report should be compared with the nutrient boundaries predicted by the toolkit analysis.

Some Member States have defined national subtypes within broad common types included in the WFD IC data set (e.g. UK or FR for TRW NEA11). In other cases, Member States have not reported H/G nutrient boundaries, but the reference conditions (RC) set for the nutrients. For this reason, comparison and correspondence of results derived from the approaches tested (Tables 4-11 and 4-12) with national nutrient boundaries should be interpreted on a case-by-case basis considering, in particular, the subtypes included in the IC data set of Member States (indicated by an asterisk in Tables 4-13 and 4-14).

Table 4-13: Member States' national nutrient boundaries adopted for transitional waters common types. Comparison of Member States' national boundaries with results produced using the toolkit requires caution for data sets marked *.

Water category	Common type	Country	Data set	Nutrient	Units	Notes	MS nutrient boundaries H/G (*Ref Cond*)	MS nutrient boundaries G/M
TRW	BAL BT1	LT	ds1*	TN	µg L^{-1}	*a* N CL salinity 0.5 - 5 *b* N CL salinity < 0.5 *c* Plume CL salinity < 2 *d* Plume CL salinity 2 - 4 *e* Plume CL salinity > 4	940 - 1080 *a,c* (RC < 750) 950 - 1070 *b* (RC < 760) 430 - 670 *d* (RC < 330) 130 - 250 *e* (RC < 100)	1090 - 1230 *a,c* 1080 - 1170 *b* 680 - 810 *d* 260 - 400 *e*
				TP	µg L^{-1}	a N CL salinity 0.5 - 5 b N CL salinity < 0.5 c Plume CL salinity < 2 d Plume CL salinity 2 - 4 e Plume CL salinity > 4	60 - 80 *a,c* (RC < 47) 61 -79 *b* (RC < 48) 37 - 53 *d* (RC < 29) 15 - 26 *e* (RC < 11)	81 - 136 *a,c* 80 - 130 *b* 54 - 84 *d* 27 - 33 *e*
		PL	ds2	TN	µg L^{-1}	salinity 0.5 - 5; winter mean	650	980
				TP	µg L^{-1}	salinity 0.5 - 5; winter mean	*not available*	120
	MED polyCL	IT/GR	ds3	TN	*not available*			
				TP	*not available*			
				DIN	µg L^{-1}	salinity < 30	*not available*	420 (ca 30 µM)
		FR	ds4	TN	µM		50	75
				TP	µM		2	3
	NEA11	NL	ds25	DIN	*not available*			
		UK	ds25*	DIN	µM	*a* **Mean** DIN at Clear waters: SPM > 10, Salinity 25 **99th percentile:** *b* Intermediate waters: SPM 10 - 100, SPM midpoint: 55 *c* Medium turbidity waters: SPM range: 100 - 300, SPM mid-point 200 *d* Very turbid water SPM < 300	(RC <20)	30 *a* 70 *b* 180 *c* 270 *d*
		IE	ds25	DIN	*not available*			
		FR	ds26	DIN	µM	normalised DIN salinity 33	20	29 (NEA26A) 33 (NEA26B)
		SP	ds27	DIN	*not available*			
		PT	ds28	DIN	*not available*			

Table 4-14: National nutrient boundaries adopted for coastal waters common types, for comparison with toolkit results. Comparison of national boundaries with BPG results requires caution for data sets marked *.

Water category	Common type	Country	Data set	Nutrient	Units	Notes	MS nutrient boundaries H/G (Ref Cond)	MS nutrient boundaries G/M
CW	BAL BC4	LV	ds5	TN	*not available*			
				TP	*not available*			
		EE	ds6	TN	µM	salinity 4 - 6; summer mean	19.2	23.7
				TP	µM	salinity 4 - 6; summer mean	0.4	0.5
	BAL BC5	LV	ds7	TN	*not available*			
				TP	*not available*			
		LT	ds8*	TN	µg L^{-1}	salinity 5 - 18; summer mean	130 - 250 (RC < 100)	260 - 400
				TP	µg L^{-1}	salinity 5 - 18; summer mean	15 - 26 (RC < 11)	27 - 33
	MEDI	IT	ds9*	TN	*not available*			
				TP	µM	salinity 20 - 37; winter mean	(RC < 0.24)	0.6
	MEDII Adriatic	IT	ds10	TN	*not available*			
				TP	µM	salinity 33 - 38; winter mean	(RC < 0.23)	0.37
	MEDII Tyrrhenian	IT	ds11	TN	*not available*			
				TP	µM	salinity 33 - 38; winter mean	(RC < 0.26)	0.54
	MEDIIIE	CY	ds12*	NO3	µM	salinity > 37.5; winter mean; (0.0091 mg L-1)	(RC < 0.14)	0.15
		GR	ds12	NO3	µM	annual mean; (0.023 mg L-1)		0.36
	NEA1-26A	FR	ds13	DIN	µM	normalised DIN salinity 33	20	33
		IE	ds14	DIN	mg L^{-1}	salinity 34.5, winter and summer median		0.25
		SP	ds15	DIN	*not available*			
		NO	ds15	DIN	*not available*			
		UK	ds16*	DIN	µM	winter *a* Mean DIN at Clear waters: SPM >10, Salinity 32 99th percentile: *b* Intermediate waters: SPM 10 - 100, SPM midpoint: 55 *c* Medium turbidity waters: SPM range:100 - 300, SPM mid-point 200 *d* Very turbid water SPM < 300	(RC < 12 *a*)	18 *a* 70 *b* 180 *c* 270 *d*
	NEA1-26B	FR	ds17	DIN	µM	normalised DIN salinity 33	20	29
		UK*south*	ds18*	DIN	same as for ds16 UK			
		UK*north*	ds19*	DIN	same as for ds16 UK			
		NL	ds19	DIN	*not available*			
		BE	ds20	DIN	*not available*			
	NEA1-26C	DK	ds21	DIN	*not available*			
		DE	ds21	DIN	*not available*			
	NEA1-26E	SP	ds22	DIN	*not available*			
		PTsUpW	ds22	DIN	*not available*			
		PTUpW	ds23	DIN	*not available*			
	NEA3-4	DE	ds23	DIN	*not available*			
		NL	ds24	DIN	*not available*			

5. Alternative approaches

The approaches described in Chapters 2 to 4 allow nutrient boundaries to be established, so long as a number of conditions are fulfilled (see Section 2.4). There will be situations (particularly in coastal and marine environments) where this is not the case and some options for establishing nutrient boundaries for these circumstances are outlined below. The options are: detecting change points in historical time series (see Section 5.1), nutrient load models (see Section 5.2), site-specific predictions (see Section 5.3), or deriving boundaries from other spatially linked water categories (see Section 5.5). In addition, a method used to derive nutrient thresholds in Catalonian coastal waters is described (see Section 5.4). These methods were contributed by Member States and Helcom and are optimised for local conditions. **These have been included to indicate the range of options available for situations where the toolkit is not effective and, with the exception of the use of historical time series, have not been subject to the same strict evaluation as the methods described in earlier sections.**

5.1. Detecting change points in historical time series

Marine waters pose a particular challenge when trying to establish nutrient boundaries. This is because the approach proposed in this document requires intercalibrated biological quality elements as the starting point for deriving nutrient boundaries, and marine waters are outside the scope of the WFD. Nutrient concentrations are, nonetheless, mandatory criteria for assessment of descriptor 5 'eutrophication' in the MSFD (see Commission Decision 2017/848). This decision also requires that nutrient boundaries set beyond coastal waters be consistent with the WFD and that they be established through regional or subregional cooperation. Helcom has already applied a Baltic-Sea-wide approach to derive common targets for nutrient concentrations as well as other eutrophication parameters (Helcom, 2013); however, other regional seas conventions are yet to follow.

In marine waters eutrophication can often be best evaluated using physicochemical parameters such as nutrient and oxygen concentrations and Secchi depth. Of the biological quality elements, a direct relationship with nutrients can often only be demonstrated for phytoplankton. Macrophytes are unlikely to occur due to the great depth of marine waters and the relationship between nutrients and macrozoobenthos is indirect and confounded by other human pressures (mainly bottom trawling and other activities causing physical damage to the seafloor).

A further challenge in the open Baltic Sea is that the whole area is highly eutrophic and there are no sites in good or high status. This means that several of the statistical criteria outlined in Section 2.4 are not fulfilled and it is not possible to use the methods outlined in this document. An alternative approach for deriving nutrient boundaries was applied here, using a long-term time series of nutrients and nutrient-sensitive parameters (chlorophyll *a*, Secchi depth and oxygen) extending back to the pre-eutrophic era (around 1900) in order to find 'break points' that could be related to alterations in ecological functioning (Helcom, 2013). The approach assumes that ecosystems can cope with some human activities and pressures, but only to a certain extent. Once this capacity has been exceeded, ecological effects become pronounced and the system collapses. Hence the target should be set at a level well below the point at which such a break point occurs in the time series.

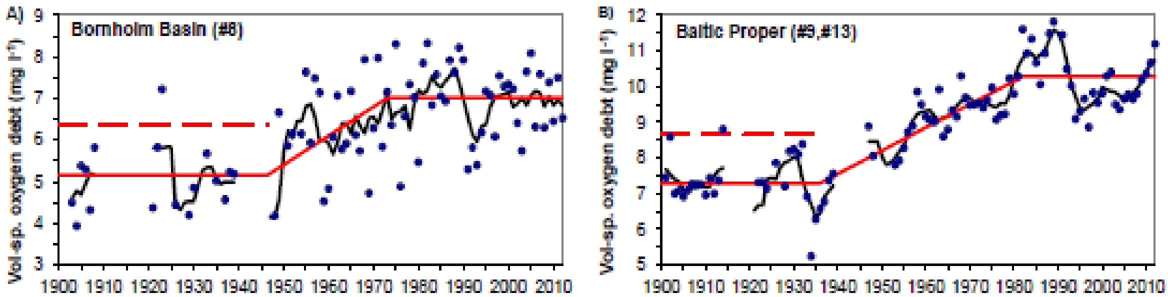


Figure 5-1*: Change point detection analysis for volume-specific oxygen debt, adjusted for variations in physical forcing using nitrogen as an input. Similar results were obtained using phosphorus. The dashed lines mark the upper 95th percentile of the oxygen debt distributions for the first period (pre-eutrophication) (Helcom, 2013). Oxygen debts above this 95 % confidence interval represent a significant departure from the natural variation within the two basins and hence the dashed line is used for setting the good/moderate boundary for this parameter.*

Volume-specific oxygen debt (defined as the difference between measured and full oxygen saturation in the water column) showed a distinct alteration (at around 1940) that could be linked to nutrient concentrations in order to define targets (G/M thresholds) (Figure 5-1). The period around 1900 was considered to have no more than minor anthropogenic disturbance and, in order to take account of the natural variation of oxygen debt in this period, any deviation above the 95 % confidence interval of the average debt was assumed to constitute a significant departure from this natural variation and was used to set the G/M boundary.

In situations where there was a gradual change in response to nutrients over time with no distinct 'break points', targets were based on the concept of an acceptable deviation from reference conditions (Helcom, 2013). In the open Baltic Sea other eutrophication parameters (nutrients and chlorophyll *a*) showed this type of gradual response although, in these cases, the time series only extended back to the 1960s or 1970s. In this situation ensemble hindcast modelling was used to estimate nutrient and chlorophyll *a* concentrations for 1900 using three coupled physical-biogeochemical models, Baltsem, Mikeecolab and ERGOM. The concentrations from the earliest data period were then compared with estimates from the simulation models (1900 scenario) to infer potential targets (Helcom, 2013).

5.2. Modelling approaches estimating nutrient loads compatible with good status

A number of Member States are using various types of hydrodynamic and biogeochemical models to establish nutrient boundaries particularly in the marine environment where, as we have seen, the spatial models described in the earlier chapters of this document often do not work.

Denmark

Denmark has not yet established nutrient boundaries for coastal and marine waters. However, a project with the aim of deriving nutrient boundaries from an existing modelling system developed to support the implementation of the WFD started in October 2017. This modelling system is used to calculate the maximum allowable input (MAI) of nutrients, which should not be exceeded in order to maintain and/or achieve GES in Danish coastal waters as required by the WFD.

The modelling system consists of a combination of dynamic mechanistic models (including hydrodynamic and biogeochemical processes) that cover larger Danish coastal waters and selected estuaries and statistical models that cover other smaller estuaries. Overall, the models cover approximately 90 % of Danish coastal water body area — equivalent to 70 % of the entire Danish catchment area. The models are validated on extensive monitoring data from Danish coastal waters.

Monitored land-based loadings of nitrogen and phosphorus constitute an important input to the development of both the statistical and mechanistic models. The models also consider nutrient loadings from the entire Baltic Sea as well as from atmospheric deposition.

The modelling system is used to calculate MAI for all Danish coastal water bodies. The calculations are based on knowledge of the current (2007-2012) status, the G/M boundary for intercalibrated BQEs and model-based relations between nitrogen loading and indicator values. The focus is therefore on loadings rather than concentrations, while the nutrient boundaries project aims at deriving model-based concentrations of TN, TP, DIN, DIP and bottom oxygen as chemical elements supporting GES. Project results are expected in early 2018.

For more information see Erichsen et al. (2017).

Germany

In Germany, the catchment area model Moneris (MOdelling Nutrient Emissions in RIver Systems) was used to model nutrient inputs from German rivers entering the North Sea and Baltic Sea in 1880. The concentrations of TN were, on average, 1.6 mg L^{-1} for North Sea rivers and 1.5 mg L^{-1} for Baltic Sea rivers (Schernewski et al., 2015). The year 1880 was chosen because it represents the period before industrialisation and agricultural intensification and because there was evidence that macrophytes were still abundant in German coastal waters of the North Sea and Baltic Sea. Following this, the historic nutrient inputs (including atmospheric nitrogen inputs) were used as input values for the ERGOM-MOM ecosystem model in the Baltic Sea. The ERGOM-MOM model was also run with current nutrient inputs and the relative difference between the historic and the recent model simulation was analysed for the resulting nutrient concentrations (Schernewski et al., 2015). Background concentrations for nutrients were calculated by multiplying recent concentrations with the factor gained from the two model simulations. Where the model had weaknesses (e.g. underestimation of nitrogen fixation) expert judgement was incorporated to set the nutrient background conditions.

United Kingdom

The dynamic Combined Phytoplankton and Macroalgae (dCPM) model (Aldridge, 2010a, 2010b) is used to manage inorganic nutrients in UK transitional and coastal waters (Ireland has also adapted the model for their own use). The dCPM model treats a water body as a single well-mixed box with direct nutrient inputs from rivers and point source discharges along with exchanges of nutrients and chlorophyll with coastal waters. The model then determines daily phytoplankton and macroalgal production within the box. The model has been updated in recent years to include multiple boxes linked together in a flexible configuration. Each box represents a different portion of the water body and can have its own characteristics such as depth, area available for macroalgal growth, and light attenuation. Nutrients and phytoplankton are exchanged between the boxes with the outer-most box having the only direct exchange with the coastal zone. Most recently, the model has been updated to include a freshwater source of chlorophyll.

The main outputs from the model are:

- average summer and winter nutrient concentrations;
- average summer and winter chlorophyll concentrations and macroalgal biomass; and
- an indication of factors limiting primary production (light, N, P or space). Space only applies to macroalgae and relates to the availability of suitable growth habitat within the estuary.

Nutrient loads are calculated as the sum of the loads from the rivers and streams (and any sewage treatment works) flowing into the estuary or coastal water. Monitoring data from within the estuary (winter and summer nutrients, chlorophyll) are also needed to verify the model accuracy and to reproduce the observations. Likewise tidal exchange at the estuary mouth is calibrated by winter and summer salinity monitoring data. The calibrated models can be used to predict the effect on primary production (phytoplankton and macroalgae) of a range of reductions to both direct and off-shore nutrient loadings. These are then used to set management targets on a case-by-case basis.

Norway

Rivers flowing through clay deposits in lowland areas of Norway are often turbid with soft sediments that are resuspended during high flow periods. The BQE responses to phosphorus in these rivers have exceptionally high variability and low biodiversity due to the poor light penetration and unstable, fine sediments. Dose-response relationships are extremely weak and not fit for boundary setting. Most of these rivers are also found in agricultural areas, so data from reference rivers are scarce.

However, TP and soluble phosphorus data from a few forested catchments with clay deposits (assumed to be only minimally impacted by agriculture and other human nutrient emissions) have been used to construct a model predicting reference value for TP as a function of the proportion of clay in the catchment (Borch, 2008, Greipsland et al., 2017). The clay proportion was estimated from geological maps with fine spatial resolution. The G/M boundary was set at twice the reference value (Table 5-1).

Table 5-1: Reference values and G/M boundaries for total phosphorus in rivers in clay-dominated catchments.

Proportion of clay in catchment	Estimated ref. cond. for TP ($\mu g\ L^{-1}$)	G/M boundary for total P ($\mu g\ L^{-1}$)	G/M boundary for total P (EQR)
20 %	20	40	0.5
30 %	25	50	0.5
40 %	30	60	0.5
50 %	40	80	0.5

Greater North Sea (Belgium, Denmark, France, Germany, Netherlands, Norway, Sweden and United Kingdom)

For the Greater North Sea the approaches outlined in this document will not work since chlorophyll-a good/moderate boundaries applied in offshore waters have been derived partly from historic nutrient concentrations and are not intercalibrated. These show large differences even in ecologically homogenous areas due, at least in part, to the use of different approaches by the countries bordering the Greater North Sea. As for the Baltic Sea (see Section 5.1), other biological quality elements either do not occur or are less suitable to derive nutrient boundaries and an alternative approach has been sought. The Directorate-General for Environment is currently funding the joint monitoring programme of the eutrophication of the North Sea with satellite data (JMP Eunosat), which uses a modelling approach to establish G/M boundaries for nutrient concentrations and chlorophyll *a* to fulfil obligations set by the OSPAR Commission and MSFD. In the first step the E-HYPE model will be used to estimate coherent natural background levels for nutrient loads in the major inputs (from rivers and the Baltic Sea) to the North Sea. E-HYPE is a tool that estimates river discharges and nutrient loads for all European rivers, based on land use, population density, elevation, rainfall, etc. In a second step natural background concentrations of chlorophyll *a* in coastal and marine waters will be derived by ensemble modelling based on local box models or multiple regression models that estimate chlorophyll concentrations based on local nutrient and light conditions. The models will be validated for present nutrient conditions with chlorophyll observations from satellite data and in situ data and available primary productivity data. The project will propose reference values and G/M boundaries for nutrient concentrations and chlorophyll *a* for the Greater North Sea, as well as suitable assessment areas. First results are expected in early 2018.

Other multivariate and modelling approaches

A number of methods have been suggested which, at present, have no official endorsement from Member States but which may have potential for the future. These include the use of Bayesian networks (Fernandes et al., 2012) and the application of Principal Components Analysis on a data set including nutrient (nitrate, nitrite, ammonia and phosphate) and chlorophyll concentrations to derive a eutrophication index (Primpas et al., 2010). Such methods may be used to either establish boundaries or validate those set by other means.

5.3. Site-specific predictions

Further possibilities for developing nutrient thresholds arise when the variables which were used to determine ‘type’ (e.g. alkalinity and altitude in freshwaters) are treated as independent variables in models spanning several ‘types’ from which site-specific thresholds can be obtained. A development from the use of bivariate regressions is to include extra predictor variables into the models from which thresholds are obtained. In the United Kingdom, for example, river phosphorus standards were computed from models which used the alkalinity and altitude of the site, along with the biological EQR (macrophytes and phytobenthos combined, in this case) to set standards (UK TAG, 2013).

The first step was to predict the concentration of phosphorus expected if a site were at ‘reference condition’ — an estimate of the natural condition of the site. The prediction used values of alkalinity and altitude to represent key geological and geographic factors that determine a site’s natural phosphorus concentrations. The next step was to calculate the ratio between the estimated ‘natural’ phosphorus concentration and the concentration actually measured at the site (this is, in effect, a phosphorus ‘EQR’). A regression equation was then developed to describe the link between the biological data (the lower class of macrophytes and diatoms) and these phosphorus ratios. Provided a site’s alkalinity and altitude are known, the equation can be used to estimate the likely ranges of phosphorus concentrations at the site associated with each biological status class. Finally, the equation was rearranged and used to calculate the most likely phosphorus concentration at the midpoint of each biological class. As an example the most likely concentrations for the midpoints of the five biological classes for a particular pair of values of alkalinity and altitude are shown in Figure 5-2 as small shapes at the centre of coloured horizontal lines. The lines represent ranges in the estimates of the phosphorus concentrations predicted by the regression model at the mid-point of the biological class. The ‘EQR’ values on the Y axis represent the degree of disturbance to the biology compared with near undisturbed conditions. The results are for a lowland, high alkalinity river in the United Kingdom.

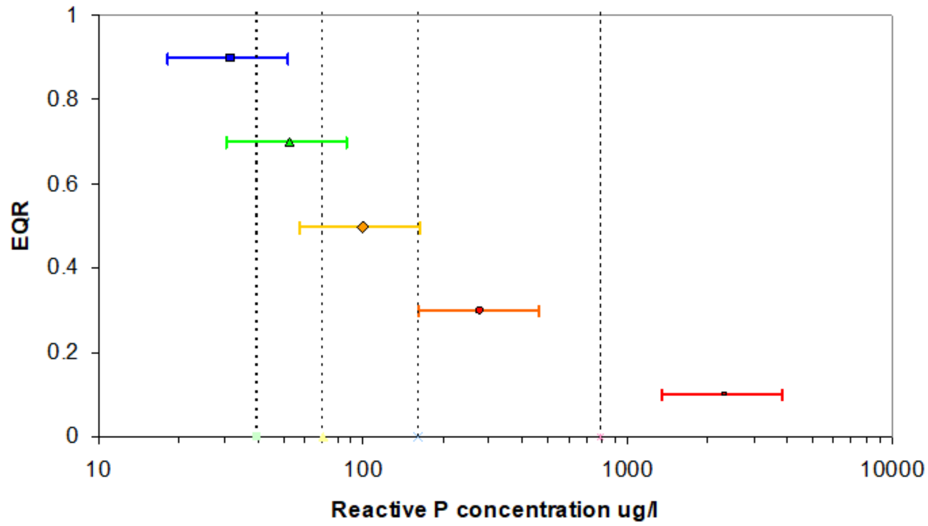


Figure 5-2: The relationship between reactive P concentration and EQR (minimum of macrophytes and phytobenthos) for a lowland high alkalinity river in England.

The recommended phosphorus standards are set at the midway point of the overlapping error bars. This position represents a concentration at which there is equal statistical confidence (P = 0.5) of the biology being in adjacent classes. The class boundaries are the vertical dotted lines in the example illustrated in Figure 5-3 with the corresponding EQR values marked as a cross. For any site, the phosphorus concentrations at these boundaries are calculated using the following equation:

Boundary P concentration =
$10^{((1.0497 \times \log_{10}(EQR) + 1.066) \times (\log_{10}(\text{reference condition RP}) - \log_{10}(3{,}500)) + \log_{10}(3{,}500))}$.
$R^2 = 0.333$ $P < 0.001$ $N = 573$

where:

EQR = ecological quality ratio (macrophytes and phytobenthos)

Reference condition RP = phosphorus concentration expected at reference condition, calculated as:

Reference condition RP = $10^{(0.454 (\log_{10} alk) - 0.0018 (altitude) + 0.476)}$

where:

Alk = alkalinity (mg L^{-1} $CaCO_3$)

Altitude = height above sea level (metres)

A benefit of the approach described here is that it does not rely on dividing rivers up into types. By using the alkalinity and altitude of the site concerned, the method derives phosphorus standards that are, in principle, specific to each point in a river. In contrast, most of the other approaches specify a single boundary applicable to the continuum of waters within a type. On the other hand, care is needed when applying such models in regions where calcium carbonate or related materials (‘lime’) are applied to agricultural land (or to mitigate acidification in low alkalinity rivers), as this may raise the alkalinity of the receiving water and indirectly influence the phosphorus target. In theory, the natural alkalinity of a river could be modelled but this has not yet been incorporated into this assessment scheme.

5.4. Establishment of nutrient boundary values for coastal waters in Catalonia

The FAN and FLU indices method assesses the physicochemical state of coastal waters and allows nutrient boundary values to support GES to be established. This method is based on a distinctly different process to establish these values than those described elsewhere in this document. Rather than using nutrient and BQE data simultaneously, it assesses the physicochemical state of coastal waters and then it relates this to the BQE. Nutrient boundary values are then established from this relationship. The approach considers several dissolved inorganic nutrients concentrations and their stoichiometry at the same time rather than focusing on a single nutrient, as is the case when applying the toolkit.

The FAN and FLU indices method was developed using the physicochemical database of the National Catalan Coastal Water Monitoring Programme. The data are representative of the north-west Mediterranean and comprise 20 102 records from 268 sampling stations collected between 1994 and 2014. A factorial analysis performed with this database revealed that the main pressures impacting coastal waters are continental influences (CI), which are related to gradients of dissolved inorganic nutrients, and freshwater content (inverse of salinity). An assessment of the physicochemical state of coastal waters based on the CI yielded results nearly equivalent (correlation of 0.93; Table 5-1) to those obtained with the Trophic Index (TRIX) of Vollenweider et al. (1998). A further rotation applied to the factorial analysis revealed that CI is divided into two distinct gradients: levels of dissolved inorganic ammonium, phosphate, and nitrite define a gradient of urban influences while levels of dissolved inorganic silicate, and nitrate as well as the freshwater content, represent a gradient of freshwater influences or fluviality. The former is considered to reflect urban influences and the latter natural continental pressures on coastal waters (although freshwater influences are partly related to nitrate enrichment from agricultural sources).

These gradients of urban and freshwater influences were the basis for the development of the FAN and FLU indices. The FAN index is scaled into five categories of water quality (high, good, moderate, poor, and bad) and the FLU index into five categories of fluviality (very low, low, medium, high, and very high). The combined results provide a final assessment of the CI reaching coastal waters (urban, fluvial, mixed, or none) and, therefore, an assessment of their physicochemical state. The indices can be applied using data from inshore (0-200 m from the shore) or offshore (> 200 m from the shore) waters or both. The procedure, equations, and boundaries to apply the FAN and FLU indices together with detailed information on the method are available in Flo (2017).

The FAN and FLU indices were used to establish the nutrient boundary values and their ranges to support GES in Catalan coastal waters. First, the physicochemical state of the Catalan coast was assessed. Second, the relationship between CI and chlorophyll *a* concentration (Chl*a*; µg L^{-1}), a proxy of phytoplankton biomass, was explored. Chl*a* increased with increasing CI, as higher concentrations of dissolved inorganic nutrients are available for phytoplankton growth (Table 5-2). The FAN and FLU indices were also linked to Chl*a* (Table 5-2), although not as strongly as CI (sum of the FAN and FLU indices). Similar relationships were found between the FAN and FLU indices and Chl*a* when only inshore or offshore data were taken into account (not shown). Chl*a* increased with increasing FAN and FLU values, as phytoplankton growth is independent of the continental sources of dissolved inorganic nutrients. This is supported by the lack of any correlation between Chl*a* and any of the five dissolved inorganic nutrients included in the data set (all values < 0.40; not shown). In contrast to TRIX, FAN and FLU do not include Chl*a* within their equations, so reducing circularity. Finally, the nutrient boundary values and their ranges were established. The 25th, 50th, and 75th percentiles of each nutrient were calculated for each quality and fluviality category and, afterwards, the mean between corresponding percentiles of neighbouring categories were calculated to obtain the values.

Table 5-2: Correlations and linear regressions between CI, Vollenweider's Trophic Index (TRIX), the FAN and FLU indices, and Chl*a* levels calculated with a data set from the Catalan coast (1994-2004; N sampling stations = 166; N data = 5.967). All correlations and linear regressions are significant.

Variable 1	Variable 2	Correlation	Linear regression (LR)	LR r^2
CI	TRIX	0.93	TRIX = 4.00 + 1.14 * CI	0.87
Chl*a*	CI	0.76	Chl*a* = 2.19 + 2.34 * CI	0.59
Chl*a*	FAN index	0.65	Chl*a* = 2.01 + 2.15 * FAN	0.42
Chl*a*	FLU index	0.49	Chl*a* = 2.02 + 1.52 * FLU	0.24

5.5. Setting thresholds to protect downstream water bodies

Although the primary role of boundaries described in this document is to support ecological status in the water body under consideration, almost all water bodies are connected to others and, therefore, boundaries that are intended to protect an upstream water body may, in some circumstances, be insufficient to protect the water bodies into which this flows. Member States have approached this in a variety of ways. In some cases, consideration of downstream water bodies is an integral part of the boundary setting process. In others, regulatory boundaries are established for a water body but these may be overridden by more stringent management targets set to protect downstream water bodies. The underlying science is, however, similar, and four examples of how Member States consider downstream water bodies have been included.

Norway

Some river types may be more resilient than lakes into which they flow and therefore have higher nutrient boundaries. In such circumstances, it may be possible to use phosphorus retention in a lake to derive an appropriate boundary for rivers that flow into this. This approach can also be used for river types where other approaches are difficult to use, as well as for checking whether existing river boundaries are sufficiently stringent to protect downstream lakes.

A number of models have been used to estimate phosphorus retention in lakes, from simple empirical models of the Vollenweider type to advanced lake-specific dynamic models (Brett and Benjamin, 2008). For lakes in good or better status the internal load from the sediments is assumed to be minor, permitting a simplified assessment of phosphorus retention using the difference between inflowing and outflowing TP concentration. Using the data set compiled by Brett and Benjamin, the mean TP retention was found to be 40 % of the phosphorus in the inflowing rivers. As this data set also included some more eutrophic lakes with internal loading, the mean phosphorus retention in lakes in good or better status is assumed to be slightly higher, ca. 50 %. In humic lakes phosphorus retention may be less than in clear lakes, due to adsorption onto humic substances and, consequently, lower bioavailability and less sedimentation of phosphorus (Jones et al., 1988). A rough estimate of mean phosphorus retention in humic lakes in good or better status is assumed to be 30 %. For very large and deep lakes with very long retention time, the phosphorus retention will probably be higher than for stratified lakes, and can be roughly assumed to have a mean value closer to 70 %. In contrast, very shallow, unstratified lakes often have considerably shorter retention time, and also more resuspension of sediments, and therefore are likely to have lower phosphorus retention than stratified lakes, with a mean value roughly estimated at 20 %.

The boundary for a river that flows into a lake, therefore, could be taken as the boundary for the lake itself (derived using methods described in Chapter 4) plus the fraction of TP that is retained by the lake. Examples for broad lake types (see Lyche-Solheim et al., 2015) are given in Table 5-3.

Germany

In the Baltic Sea and the North Sea, coastal waters are susceptible to eutrophication and are often not sufficiently protected by nutrient standards applied in rivers, particularly in the case of nitrogen. Hence, Germany has set a management target of 2.8 mg L^{-1} for TN for all individual German rivers (Elbe, Weser, Ems, Eider) entering the North Sea and a target of 2.6 mg L^{-1} for German rivers entering the Baltic Sea in order to manage eutrophication in coastal and marine waters. These target values are measured at the freshwater-marine border. For the River Rhine which flows through Germany but enters the sea in the Netherlands the management target applies at the border between Germany and the Netherlands. The management target is laid down in the German Surface Water Ordinance and is, consequently, legally binding. If the target is reached, it is assumed that this will enable coastal waters to achieve GES according to the WFD and marine waters to achieve good status with respect to eutrophication under the MSFD. In order to help water managers achieve nitrogen reduction requirements for upstream waters the catchment model Moneris (see Section 5.2.2) was used. Based on this, nitrogen reduction requirements for catchments have been set (Figure 5-3) and are used as a basis for planning measures in the River Basins Management Plans. For example, at the freshwater-marine border nutrient reduction requirements have been calculated considering freshwater discharge. They range between 30 % for the river Weser and 49 % for the Schlei/Trave.

Table 5-3. Estimated river standards for total phosphorus based on lake standards for broad types (Table 4-4) and mean phosphorus retention for each type (see text above).

	Broad lake type	Lake mean depth (m)	P retention (mean) (% of inflow TP)	TP standard for lake type		TP standard for corresponding river type	
				25th	75th	25th	75th
				percentiles			
2	Lowland, siliceous, shallow (stratified)	3-15	50	11	23	22	46
3	Lowland, calcareous, shallow (stratified)	3-15	50	26	55	52	110
4	Lowland, calcareous, very shallow	< 3	20	33	62	41	78
5	Lowland, humic, shallow (stratified)	3-15	30	21	28	30	40
7	Mid-altitude, siliceous, shallow (stratified)	3-15	50	13	34	26	68
8	Mid-altitude/highland calcareous, shallow (stratified)	3-15	50	18	34	36	68
11	Highland, siliceous, shallow (stratified)	3-15	50	9	28	18	56
12	Highland, calcareous, shallow (stratified)	3-15	50	12	18	24	36

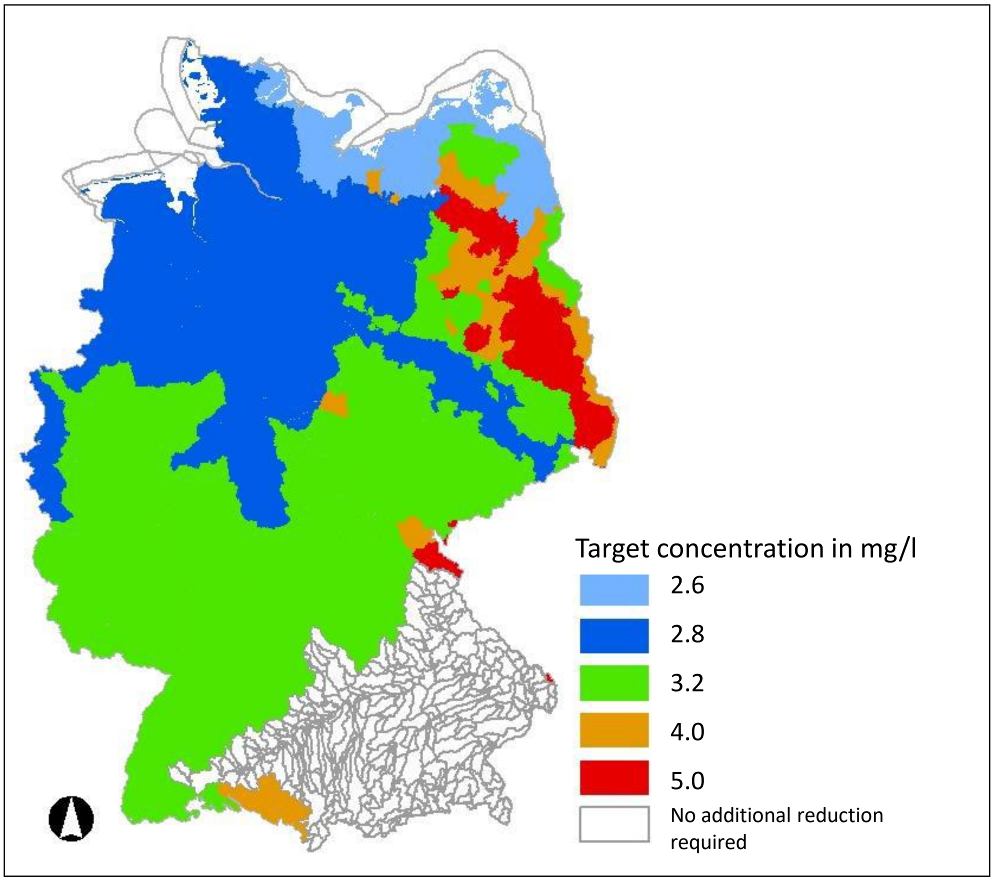


Figure 5-4: Mean annual target concentration for total nitrogen (in mg L^{-1}) that is required in the catchments of German rivers in order to achieve good status with respect to eutrophication in coastal and marine waters of the North Sea and Baltic Sea. At the time that this work was performed no nutrient reduction targets had been set for the Black Sea, into which the Danube catchment drains (LAWA 2016) (but see 5.5.2 for recent developments).

Nutrient reduction targets for the German parts of the Danube catchment ‘are set to ensure that the Black Sea ecosystems could recover to conditions similar to those observed in the 1960s, representing river loads under low pressures’ (Ibisch et al., 2016). According to a recent European Topic Centre on Inland, Coastal and Marine Waters report these target loads correspond to ca. 250 kt/yr DIN, ca. 300 kt/yr TN and ca. 20 kt/yr TP (Adam Kovacs, ICPDR Secretariat, pers. comm. 2016 in Ibisch et al., 2016). The target loads are based on the Moneris model (Behrendt and Zessner, 2005; Malagó et al., 2015, Ibisch et al., 2016).

Netherlands

The standard for DIN in coastal waters of the Netherlands was determined by minimising the number of observations where nutrient concentrations would be in good status while the status for phytoplankton would be moderate or worse. The Netherlands suggest using the DIN winter mean target concentration at salinity 30 (33 µM for the G/M boundary) and natural background concentrations of DIN in Atlantic Ocean water (salinity > 35) to calculate a target concentration for winter means of DIN in rivers. There is an inverse linear relation between salinity and nutrient concentrations during the winter months due to conservative mixing (OSPAR, 2013). Winter mean DIN and TN concentrations in the rivers Rhine and Meuse showed a strong linear correlation over the years 1975-2007 (Theo Prins, pers. comm.). Using this linear correlation, the winter mean DIN standard for rivers could be converted into a TN concentration of 2.8 mg/L TN (arithmetic annual mean, based on weekly (Meuse) or bi-weekly (Rhine) sampling). This target value is in line with that suggested by Germany for German rivers entering the North Sea.

France

The occurrence of episodic blooms of toxic dinoflagellates has focused attention on eutrophication of the Seine Bight. In order to better understand the relationships between these processes and human activity in the Seine watershed, two modelling studies were performed with the following results.

- The first model (Cugier et al., 2005) describing nutrient (nitrogen, phosphorus, silica) transfer processes at the scale of the whole Seine Basin (Riverstrahler) allows human activity (agricultural practices, waterscape management, urban wastewater management, etc.) to be related to fluxes delivered to the sea. The models were validated by their ability to reproduce observed trends of inter-annual variations of nutrients delivered by the Seine during the last 50 years, as well as the response of the marine system in terms of diatoms and dinoflagellate development, for which data are available from 1976 to 1984 for the former and from 1987 to 1997 for the latter. The results show that dry years, where silica inputs show a deficit with respect to nitrogen and phosphorus, are those where summer blooms of dinoflagellates are particularly pronounced. Various scenarios of human activity in the watershed have been simulated by the two models, including a reconstitution of the ‘pristine’ state, a historical state corresponding to the situation at the end of the 18th century, as well as several scenarios corresponding to the present situation with alternative policies of wastewater nitrogen and/or phosphorus treatment.

- The second study (Passy et al., 2016) also coupled a riverine and marine model (the Seneque/Riverstrahler and the ECO-MARS3D, respectively) in the Bay of Seine to explore the effects of two scenarios of watershed management. The first scenario, focused on an upgrade of wastewater treatment plants, decreased phosphorus fluxes by 5 to 35 kg phosphorus km^{-2} yr^{-1} on average over the 2000-2006 period, depending on the watershed, and would lead to a threefold reduction in the concentration of dinoflagellates in the adjacent coastal zone. The second scenario envisaged adoption of organic farming in all agricultural areas of the basin. Although not realistic, this showed the theoretical best outcomes. With this, nitrogen fluxes decreased by almost 50 %, and the dinoflagellate blooms generally and *Dinophysis* spp. blooms in particular were reduced by a factor of 20 to 40. Diatoms, the main primary producers in the bay which sustain the marine food web, are not significantly affected by this scenario.

The other concern in France is the annual occurrence of massive green macroalgal blooms along the Brittany coastline since the 1970s. Perrot et al. (2014) modelled *Ulva* proliferation using a two-dimensional model by coupling hydrodynamic and biological models (called 'MARS-Ulves'). Calibration of the biological model was based mainly on the seasonal variation of the maximum nitrogen uptake rate ($V_{max}N$) and the half-saturation constant for nitrogen (K_N) to reproduce the internal nutrient quotas measured in situ for each site. In each bay, model predictions were in line with observed algal coverage converted into biomass. A numerical tracking method was implemented to identify the contribution of the rivers that empty into the study bays, and scenarios of decreased nitrate concentration in rivers were simulated. Results from numerical nitrogen tracking highlighted the main nitrogen sources of green tides and also showed that each river contributed locally to green tides. In addition, dynamic modelling showed that the nitrate concentrations in rivers must be limited to between 5 and 15 mg L^{-1}, depending on the bay, to reduce *Ulva* biomass at the coast by half. The three-step methodology developed in this study (analysing total DIN flux from rivers, tracking nitrogen sources in Ulva and developing scenarios for reducing nitrogen) provides qualitative and quantitative guidelines for stakeholders to define specific nitrogen reduction targets for better environmental management of water quality.

More information on eutrophication in France can be found in Pinay et al. (2017).

6. Validation

It is recommended that boundaries be tested using independent data in order to ensure that they will protect the ecosystem of interest. The focus in this document has been on the use of primary producers to establish standards, as these are the organisms that have the most direct link with inorganic nutrients. However, the consequences of eutrophication extend to all trophic levels and other BQEs (benthic invertebrates and fish) can be used to validate boundaries. Validation should focus on protecting key ecological functions and ecosystem services (European Commission, 2009). A further straightforward and easy check of any boundary is to compare it with values used by neighbouring countries for similar water body types, or with the values reported in Sections 4.4 and 4.5 of this report. Agreement does not automatically validate a boundary, but major differences should provide the impetus to re-examine the data to ensure that any values produced are robust. Some options for validating nutrient boundaries are listed in Table 6-1.

Validation steps depend upon a good understanding of the ecosystems under consideration and will, as a result, vary between regions and water body types. In general, a hypothesis based on known cause-effect relationships needs to be established, and then tested. For example: the Arctic char (*Salvelinus alpinus*) is a salmonid associated with some deep lake types in northern and western Europe. It feeds in the hypolimnion and has been shown to decline in lakes that are subject to eutrophication, as the hypolimnion becomes progressively deoxygenated (Winfield et al., 2008). This generates a testable hypothesis: that lakes with healthy Arctic char populations should have nutrient concentrations consistent with high or good status. If such lakes have nutrient concentrations consistent with moderate status, it is possible that the standard is too precautionary; if Arctic char are confined to lakes where the nutrient concentrations should support high status then the opposite may be true.

Free et al. (2016) used data on the distribution of charophytes to test nutrient standards developed by methods similar to those described in Chapter 4. Although charophytes are included in macrophyte assessment methods, they have a value as indicators of high quality habitats, protected under the habitats directive, and therefore it is important to ensure that nutrient standards are sufficient to preserve such taxa. In this study, G/M boundaries produced by phytoplankton, phytobenthos and macrophytes ranged from 16-30 µg L^{-1} TP while a separate analysis of charophytes suggested that a more stringent standard of 10-15 µg L^{-1} TP may be necessary if populations of charophytes in marl lakes are to be preserved (Figure 6-1). Another concern is to set a nutrient standard so as to prevent a lake switching into an alternative stable state. Marl lakes with high populations of charophytes and unique phytobenthos communities can lower phosphorus levels through co-precipitation and this function can be lost with excessive eutrophication. Allowing eutrophication to reach a level where internal ‘cleansing’ processes are disabled will result in higher long-term costs for restoration.

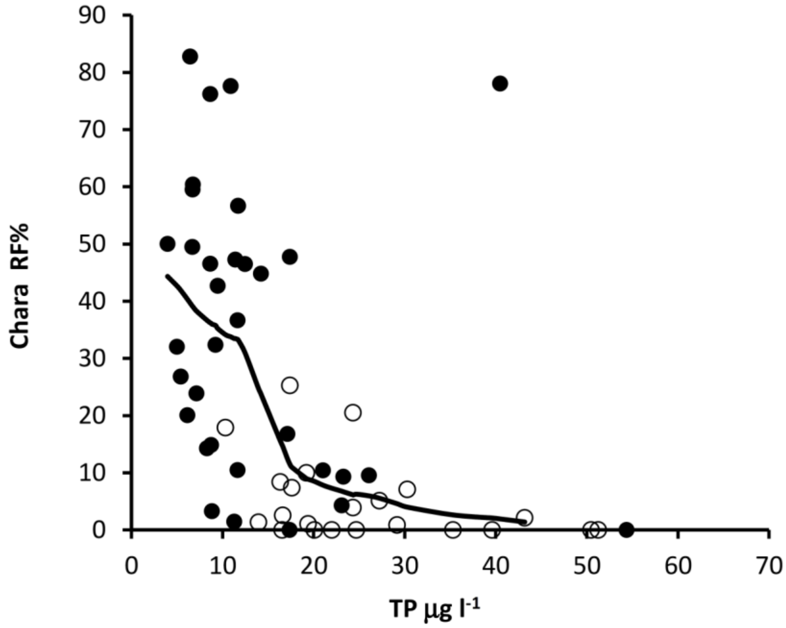


Figure 6-1: Relationship between total phosphorus and the relative frequency of *Chara* spp. (the most frequently-encountered charophyte) in marl lakes with alkalinity > 100 mg L^{-1} $CaCO_3$ (n = 58). Closed and open circles represent lakes that are above and below the good/moderate boundary for macrophytes, respectively. A LOESS smoothed line is fitted to the data (from Free et al., 2016).

Validation based on organisms that are included within BQEs from which the standards were derived risks circularity and needs to be approached with caution. This particular approach works because Free et al. (2016) draw on a broader understanding of the role of charophytes in shallow calcareous lakes. By contrast, Norway has used Average Score Per Taxon, their river invertebrate metric, for an independent check on the standards set using phytobenthos. A non-linear regression line of invertebrate data intersected the G/M boundary at very similar TP concentration (16-17 µg L^{-1}) to that obtained using phytobenthos (Figure 6-2).

Another option, tested on invertebrates in German rivers, is to analyse taxon-specific physicochemical change points using Threshold Indicator Taxon Analysis (TITAN: Baker and King, 2010), a development from indicator species analysis (Dufrêne and Legendre, 1997). Sundermann et al. (2015) used this approach to test current supporting element standards in Germany and noted a need to tighten these in some cases. A similar multi-taxa analysis (gradient forest, Ellis et al., 2012), independent of biological indices, has been applied to phytoplankton and fish data in French lakes (Roubeix et al., 2016 and 2017). This method gives the possibility to consider the response of a given BQE to several environmental variables in a single analysis. It is a useful tool for checking the ecological relevance of some nutrient thresholds.

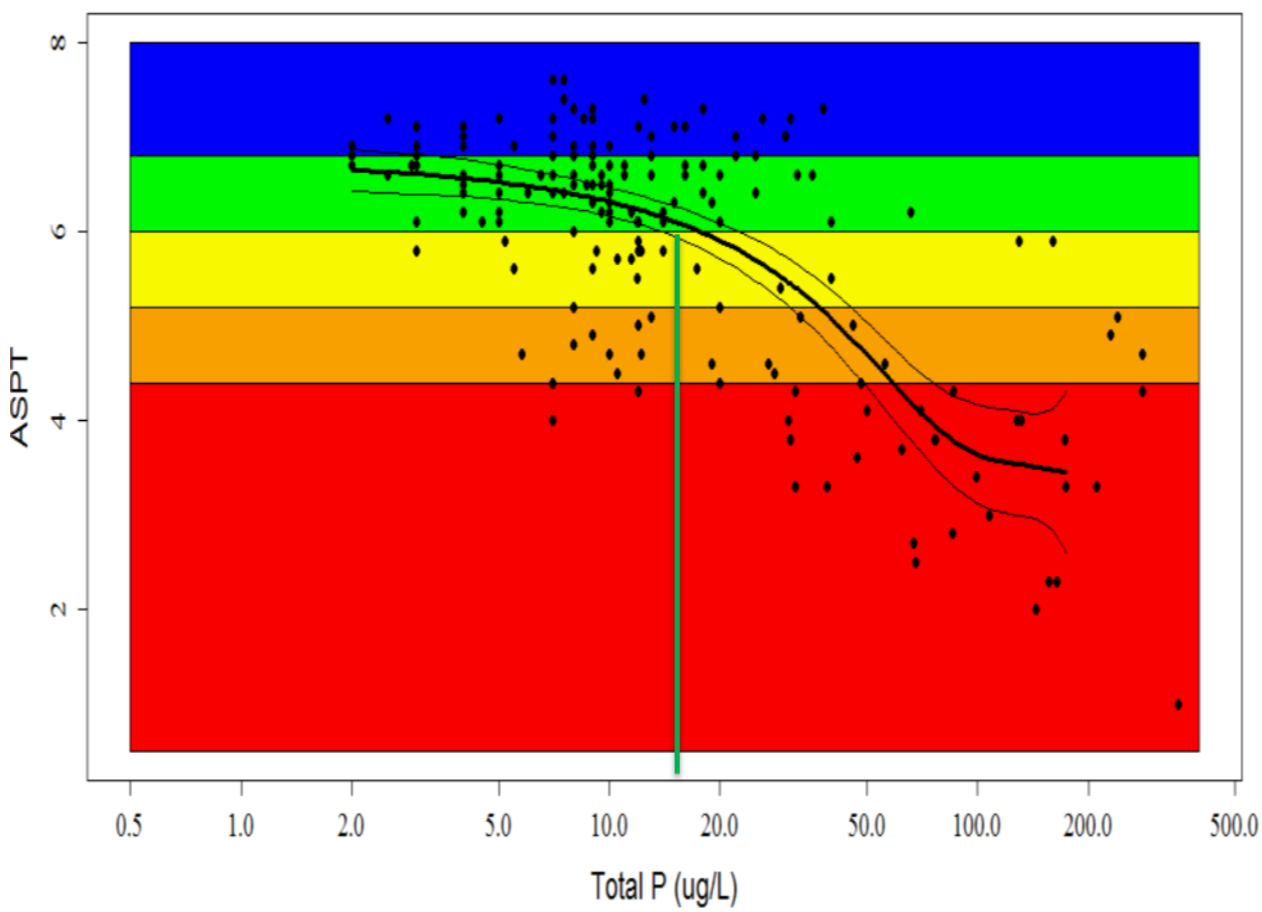


Figure 6-2: Nonlinear regression of the ASPT index for benthic invertebrates vs total phosphorus, based on data from Norwegian low-alkalinity rivers compiled by T. Bækken, regression done by T.E. Eriksen and J. Moe, NIVA. The G/M intercalibrated boundary for ASPT is 6.0. Green line indicates G/M boundary for total phosphorus, using the lower confidence band of the regression.

Further validation may come from palaeoecological sources, particularly where fossil assemblages have been used to infer past phosphorus concentrations. As in Section 5.1, the focus should be on detecting change points in distributions that can be related to changes in composition or functioning. There are many good examples from lakes, but it may also be possible, under some circumstances, to extend this approach to CTRW (see Clarke et al., 2006).

Table 6-1: Options for validating nutrient thresholds in different water categories.

Indicator	Rationale
Lakes	
Charophytes	Sensitive key macrophytes in high-/good-status lakes with threshold response to nutrients, indicating the position of the G/M boundary (Penning et al., 2008).
Arctic char (*Salvelinus alpinus*)	Distribution coincides with lakes with nutrient concentrations at or below good-status boundary. The same principle applies to other salmonids in lakes; however, char are particularly appropriate as they require an oxygenated hypolimnion in which to feed.
Cyanobacterial blooms	Absent or rare when TP is at or below good-status boundary. Cyanobacteria biomass should be below the WHO low risk threshold for cyanotoxins (corresponding to 1-2 mg L^{-1}, Carvalho et al., 2013).
Palaeoecology	Diatom-inferred TP in era before industrialisation/agricultural intensification is at or below good-status boundary.
Rivers	
Cladophora and other persistent filamentous algae	Low cover at sites with nutrient concentrations at or below good-status boundary.
Presence of salmonid fish	In rivers where salmonid fish are expected, these should be present and in good condition when nutrient concentrations are below the boundary for GES.
Evidence of fish kills	No reports of fish kills associated with night-time anoxia, when nutrient concentrations are at or below good-status thresholds.
Transitional waters	
Palaeoecology	Diatom-inferred TN or TP in era before industrialisation/agricultural intensification is at or below good-status boundary.
Coastal waters	

Indicator	Rationale
Palaeoecology	Diatom-inferred TN or TP in era before industrialisation/agricultural intensification is at or below good-status boundary.
All water body types	
Analysis of taxon-specific change points	Statistical technique using Threshold Indicator Taxon Analysis (TITAN).

7. Use of nutrient boundaries in ecological assessment and management

Both the MS Excel and R versions of the toolkits give an array of numbers for potential boundary values for each water body type (minor differences in output between the two toolkits will reflect rounding conventions within the software and can be discounted). The final question that needs to be addressed is which values are appropriate for use as a nutrient boundary. The array of choices can be reduced in two ways.

- Statistical criteria: if the relationship between the explanatory and response variables has relatively low predictive power (i.e. $r^2 < 0.36$), the categorical approaches may be better. This is particularly appropriate if there are a limited number of points at either of the extremes of the range of the pressure variable to ensure that the regression is securely 'anchored' (although good coverage for the two classes either side of the boundary in question is still necessary).
- Regulatory criteria: the two options outlined in Section 2.2 were either to set precautionary boundaries, but not acting unless the biological boundary was also exceeded, or to set a less-precautionary boundary, but acting as soon as this was breached, regardless of the state of the biology. The former option might lead to the adoption of the lower confidence interval and the latter to use of the predicted boundary (i.e. the point at which the regression line crosses the biological status class boundary).

The MS Excel and R toolkits have been used on two different test data sets and the output of each can now be evaluated using these criteria.

7.1. Deciding how to select suitable phosphorus boundaries for a shallow high-alkalinity lake type, an example for broad type 3 lakes

The r^2 for the relationship between the BQE and TP is 0.544; this exceeds the limit of 0.36 recommended for confident use of type II regression. This allowed us to calculate several potential boundary values to be produced, using all possible methods (Table 7-1). The predicted G/M boundary values range from 46 to 61µg L^{-1}. We cannot say that any of these numbers are 'correct', the pros and cons of each approach have been discussed above. However, in management terms, this is a relatively narrow range of values and within the range of uncertainty associated with the data. The value of 55 µg L^{-1} was obtained using type II regression ([9]) which, for reasons discussed above, also has several theoretical advantages. This approach provides the best estimate of the nutrient concentration at the biology G/M boundary and is thus likely to minimise the mismatch between biology and chemistry. If, instead, a more precautionary boundary was to be set (i.e. using 25th percentiles), then the range is 34-40 µg L^{-1}, again, with type II regression providing the value between the two OLS regression extremes. Finally, if mitigation measures were triggered as soon as the phosphorus standard was breached, regardless of the state of the biology, then the 75th percentile might be more appropriate, as the number of instances where action was taken when biology was still in good status would be minimised. The range of options is 65-82 µg L^{-1} and,

([9]) Ranged major axis regression, the preferred type II method again, type II regression provides an intermediate value. In each case, the uncertainties in the data are such that some judicious rounding of the values could be undertaken.

These values can be compared with the categorical approaches. The preferred type II regression predicted G/M boundary value of 55 µg L^{-1} is very similar to the binomial regression estimate at a probability of 0.5 and the values estimated from the box plots. The minimisation-of-mismatch approach produced slightly lower values (49-53 µg L^{-1}). These results highlight that where there is a relatively strong relationship between biology and nutrients, as for example with lake phytoplankton, all the methods produce similar results.

The choice of boundary value is, as can be seen from this brief summary, very much influenced by the objective of the standard (see Table 3-1). Is it to protect most water bodies, which would favour the use of the lower percentile, or to ensure that we allow biology to dictate status, leading to selection of the upper percentile to minimise false negatives? It is also important, in this discussion, to consider ecosystem services and, in particular, to test the outcomes against the distributions of species of particular economic or conservation interest, and against the likelihood of undesirable disturbances both in the water body under consideration and in those downstream.

Table 7-1: Summary of predicted boundaries for broad type 3 (L-CB1) lake ['Excel example'], using the different statistical methods outlined above (model r^2 = 0.544, see Appendix 2 for more details).

		Predicted	25th percentile	75th percentile
G/M	Model 1 OLS(Phyto on TP)	61	40	82
	Model 2 OLS (TP on Phyto)	46	34	65
	Model 4 Ranged Major Axis	55	39	70
	Categorical (adj. Quartiles)	52		
	Categorical (median of class)	51		
	Categorical (75th Quartile)	54		
	Mismatch	50		
	Binomial logistic regression P = 0.5	56		
H/G	Model 1	27	18	37
	Model 2	30	22	42
	Model 4	28	20	36
	Categorical (adj. Quartiles)	28		
	Categorical (median of class)	33		
	Categorical (75th Quartile)	27		
	Mismatch	30		
	Binomial logistic regression P = 0.5	30		

7.2. Establishing nutrient thresholds in transitional waters: a case study using common type NEA11 (north-east Atlantic transitional water category)

Introduction

This section applies the steps described in Chapters 3 and 4 to a real situation in transitional waters. The example presented here uses phytoplankton IC data from the common IC type NEA11 in transitional waters. This is a very broadly defined type consisting of estuaries entering the NEA. As relationships within national data sets are all relatively weak, there is a good case for combining data to produce a single data set spanning several countries and the additional screening steps associated with that process are also explained here.

The estuaries within this type are likely to differ on features that could influence the outcome of the phytoplankton and nutrient relationships across this type. However, this data set contains few parameters that might account for such differences, so this example describes the best way to derive nutrient boundaries based on the information available. This means that the values obtained will need to be validated with independent data. This step is, however, out of scope for the present exercise.

The data set includes mean winter DIN and EQR based on chlorophyll *a* from six Member States: France (FR); Ireland (IE); Netherlands (NL); Portugal (PT). Spain (SP) and United Kingdom (UK) Note that Chl*a* values are expressed as the national chlorophyll *a* metric calculated in a 6-year period.

Step 1: data check

The data set should span at least four ecological quality classes and show a linear relationship for at least H, G, M with $r^2 > 0.36$.

The WFD IC exercise resulted in different EQR boundaries across Member States within this type (see Table 4-8), therefore EQRs have been normalised (nEQR) (using toolkit template *TKit_Normalise.xlsx*) to be able to pull all data sets together and derive nutrient boundaries for this common water type (Table 4-13). The identity of each data set (ds) has been preserved in order to be able to check for their effect on the results.

Data covers all five ecological status classes (Figure 7-1), with DIN concentrations slightly decreasing towards higher quality classes, although there is some overlap, particularly below good status. Some data sets appear to benefit from being combined: Portugal (ds28), for example, does not cover enough of the nutrient gradient on its own to establish a relationship with nutrient concentration (Figure 7-2a). DIN concentrations in the French data (ds26), however, show an inverse trend to that expected (Figure 7-2b and Figure 7-3), and have been excluded from the analysis.

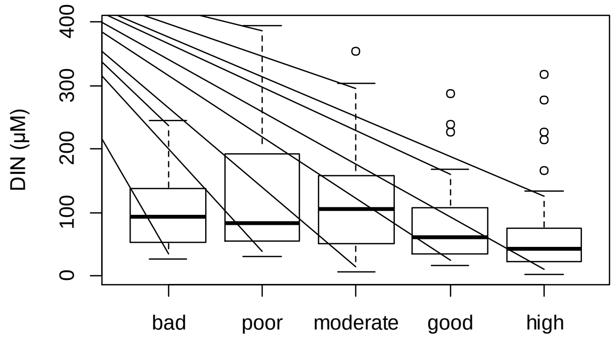
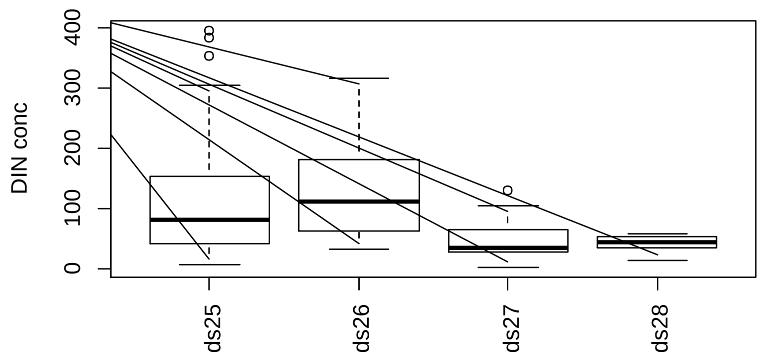


Figure 7-1: DIN concentrations (µM) between status classes for all national data sets combined. Outliers were identified in the combined data set at DIN concentrations of 353, 288, 227, 240, 317, 227, 276, 166, and 214 µM.

(a)



(b)

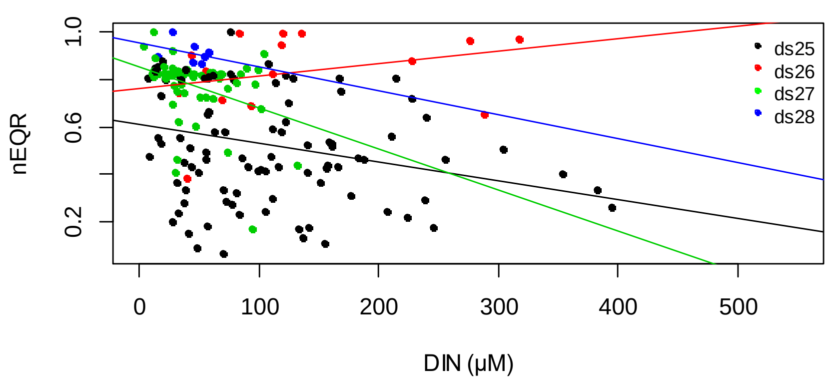


Figure 7-2: (a) Dissolved inorganic nitrogen concentrations (µM) range in each data set (ds) within common type NEA11: ds25 (Netherlands, United Kingdom, Ireland), ds26 (France), ds27 (Spain), and ds28 (Portugal); outliers were identified at DIN concentrations of 353; 394; 382 and 131 µM. (b) Relationship between nutrient concentrations and normalised EQRs, trend lines shown for each data set within NEA11.

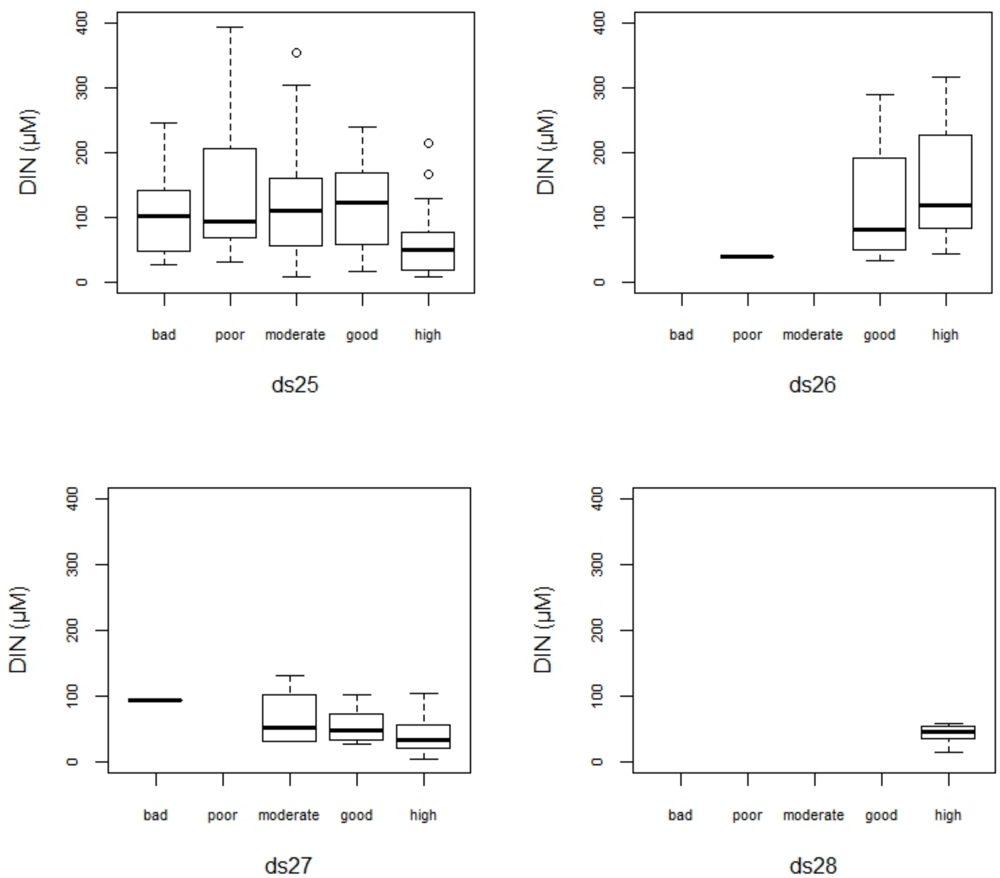
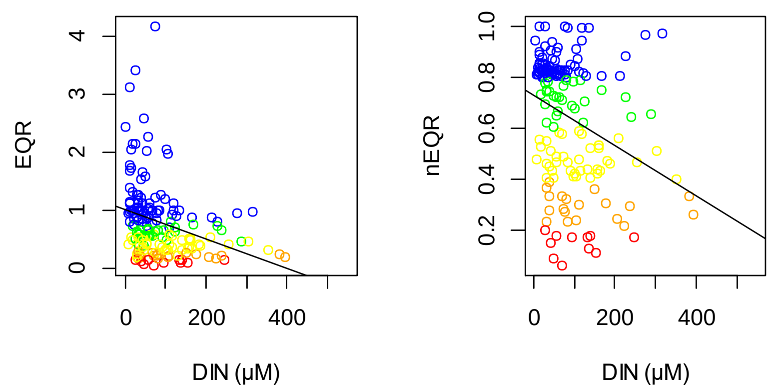


Figure 7-3: DIN concentrations (µM) between status classes for each data set (ds) separately: ds25 (Ireland, Netherlands, United Kingdom), ds26 (FR), ds27 (ES) and ds28 (PT).

The regressions show a weak relationship between EQR and DIN, with r = -0.308 and r = -0.310, respectively for EQRs and nEQRs. The plots (Figure 7-4, top) show that the regression is very weak, suggesting that pressures other than DIN are contributing to the decrease in EQR and are influencing the distribution of data, resulting in a wedge-shaped distribution. In this case other approaches such as quantile regression, or the use of an upper-quantile categorical (e.g. 75th), or binomial logistic regression, will be considered.

Removing the French data set (ds26) increases the strength of the relationship slightly: r = -0.347 and r = -0.431, respectively for EQRs and nEQRs (Figure 7-4, bottom) but data are still very scattered. The overlap between status classes decreased slightly, in particular for the range of interest (Figure 7-5).



When removing ds26:

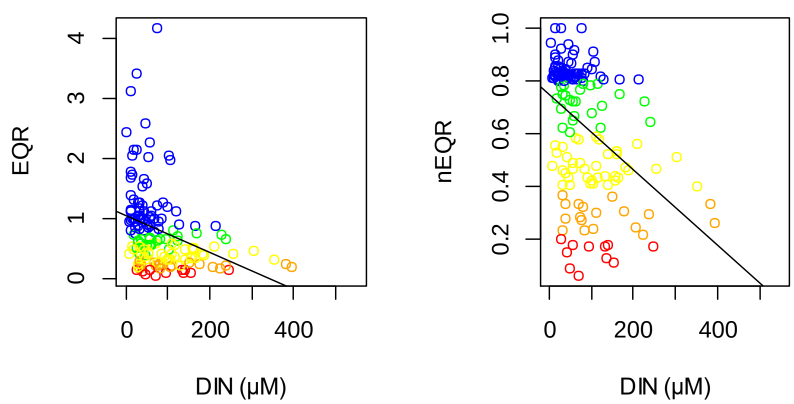


Figure 7-4: Scatterplot of the relationship of nutrient dissolved inorganic nitrogen concentrations (µM) with intercalibrated EQRs (left) and normalised nEQRs (right), including observations from all data sets in the NEA11 Type (top) and removing data set from France (ds26) (bottom); colour by status class.

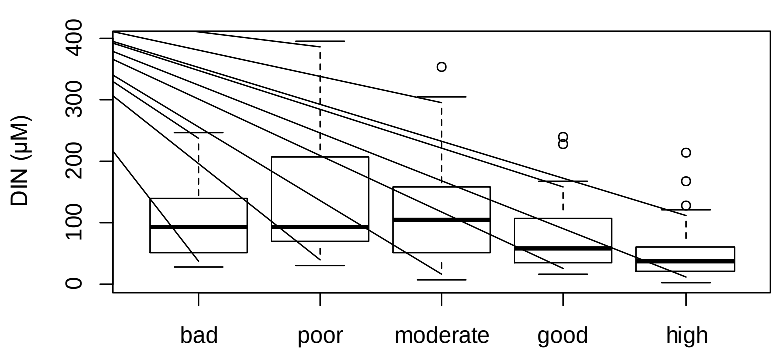


Figure 7-5: DIN concentrations (µM) between status classes for NEA11 data sets after ds26 had been removed. Outliers identified at DIN concentrations of: 353, 227, 240, 166, 214 and 128 µM.

Box 6: Conclusions from data check

Data removal

Ds26 will be removed from this analysis as it presents a pattern opposite to that expected, differing from all other data sets, and there are no additional explanatory variables available in this data set that can help explain this.

EQR data use

Normalised EQRs will be used for regression-based approaches and ecological status classes will be used for categorical approaches.

Outlier removal

Those outliers highlighted will be considered for their effect and eventual removal, but not removed *a priori*.

Gradient coverage

Full range of disturbance covered when data sets are combined within common type. The slight overlap of classes within the interval of interest (H/G/M) will be tested for significance of differences between nutrient concentrations across classes.

Strength of the relationship

Weak relationship ($r^2 < 0.36$) between biological and nutrient data.

Data shape

Scattered wedge-shaped data distribution, indicating that other pressures besides DIN are affecting phytoplankton and contributing to EQR decrease.

Data exploration indicates that linear regression methods (OLS and type II regression) may not be suitable for deriving nutrient boundaries for this common type using the available data set, therefore quantile regression and categorical approaches will also be considered and compared.

Most of the graphs shown for the data exploration can be performed using the R scripts templates available with toolkit vs 6c (*TKit_check_data.R* and *TKit_CoPlot.R*).

Step 2: statistical analysis

Excel toolkit analysis showed that the correlation between DIN and nEQR based on chlorophyll *a*, though significant ($P < 0.001$), was lower than $r < 0.6$ ($r = -0.458$) and thus caution is needed when interpreting the predicted nutrient boundaries for DIN for this type. Some critical data points have been highlighted, essentially data points with the highest DIN concentrations observed (> 238 μM); however, their removal had no significant influence on the regression outputs. The outliers were kept in the data set because: (a) other approaches will be tested; and (b) they may not be real outliers as they correspond to the tail of the distribution, eventually covering the full gradient of disturbance, and potentially have an important influence on regression outputs.

Significant differences between the nutrient concentrations in adjacent quality classes were only observed between high and good status classes and the difference between good and moderate status was not significant (non-parametric Wilcoxon Rank Sum Test P = 0.052). Boundary values between good and moderate status calculated using categorical methods based on the quantiles (Table 4-11) must be treated with extreme caution.

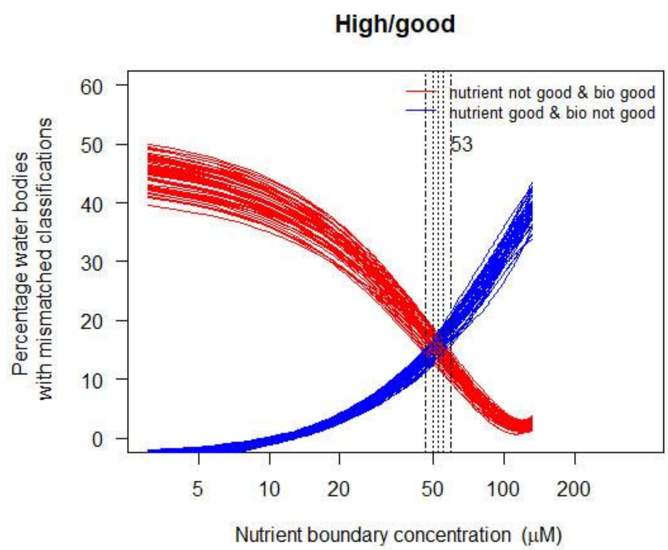
Quantile regression could be more appropriate for the type of data distribution observed in this data set. Another possibility is the minimisation-of-mismatch method, as this is the least sensitive to outliers and nonlinear relationships. This method does, however, require a significant correlation which, as mentioned above, is not the case around the G/M boundary.

Because the Excel toolkit outputs provide no estimate of uncertainty for the minimisation-of-mismatch-of-class method, a bootstrap approach was included in the R scripts (*Tkit_mismatch3_HG.R* and *Tkit_mismatch3_GM.R*). Using this approach (Figure 7-6) the mean estimated H/G boundary for DIN is 52.5 µM, within a range of 47-59 µM, with a total mismatch classifications rate of 30 %, ranging from 28-34 %. For G/M the mean estimated boundary is 74.5 µM, which is within the range of 66-83 µg L^{-1} reported in Table 4-11. At this point the total mismatch of classifications is 28 % and lies within the range of 24-34 %. Although not shown here, a sufficient number of iterations has been used to achieve convergence (details in Appendix 1 Section A5.2.2).

Binomial Logistic regression is also included in the toolkit (*TKit_LogisticRegHigh.R* and *TKit_LogisticRegGood.R*) and, as this is the most reliable categorical method, it should be used when linear modelling is not appropriate. However, like other methods, boundary estimates will be unreliable if other pressures are operating. As this is often the case in estuaries, where multiple pressures are encountered frequently, the results must be interpreted with caution.

The binomial logistic regression of DIN on biology (nEQRs), for both the H/G and the G/M range are presented in Figure 7-7. Nutrient boundary estimates are presented for a 50 % probability of being in moderate or worse status for the G/M boundary, or in good or worse for the H/G boundary (Table 7-2), but nutrient values at lower and higher probability thresholds (25 % and 75 %) are also presented, which provide precautionary and non-precautionary values.

At this stage, due to the lack of environmental information for predicting relevant features and potential subtypes across the NEA11 broad type, but also due to other pressures, and a need to accommodate uneven gradient coverage across Member State data sets, quantile regression offers a possible alternative. Quantile regression is not included in the current toolkit version, but some testing has been done which is included here. For this NEA11 data set a higher quantile has been adopted, for coping with the influential role of potential unknown stressors (or environmental features) in the shape of the data, as this is the only portion of the data set possible to be modelled at this stage. The univariate model fitted indicated that the maximum DIN levels (using the 0.7 quantile) that could still support H/G and G/M ecological status correspond to nutrient values of 68 and 212 µM respectively (Figure 7-8 and Table 7-2). However, the 95 % confidence intervals obtained for the G/M boundary are too wide indicating that, at such nutrient concentrations, EQRs ranging from 0.39 to 0.81 could be expected. Boundaries derived from a high quantile may be appropriate when pressures other than nutrients are downgrading the biological status. However, such boundaries are not precautionary and indicate a high risk of negative effects on the biota at these values. In the case of NEA11 these pressures might be operating together with environmental factors to control phytoplankton growth dynamics, which makes it more difficult to disentangle the most relevant factors. The nutrient boundaries indicated here should therefore be taken with caution until additional environmental factors are considered and further guidance on quantile selection is developed.



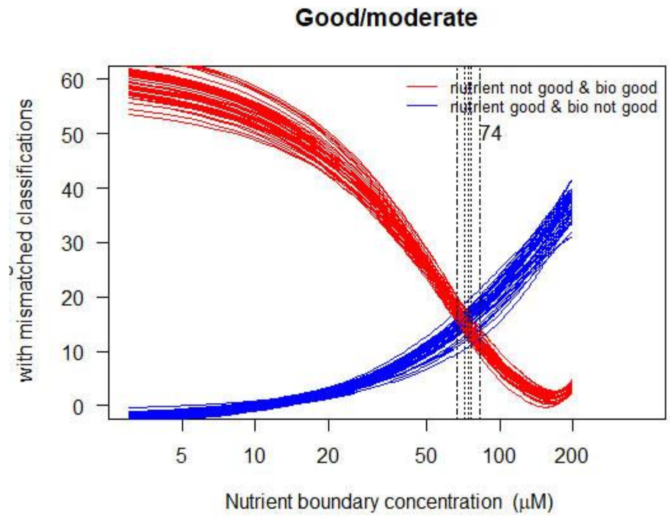
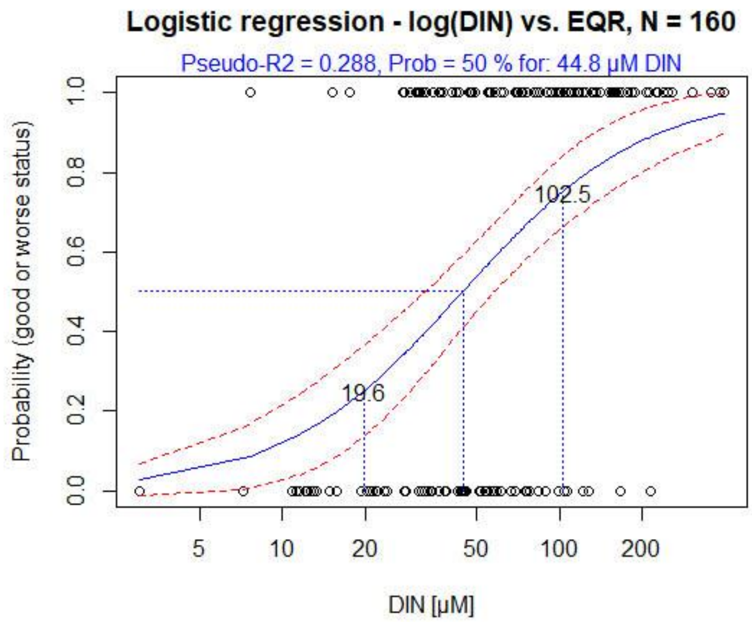


Figure 7-6: Relationship between percentage of misclassified records when biological and nutrient classifications are compared. Vertical lines mark the range of crossover points where the misclassification is minimised, together with the mean nutrient concentration. (each line shows a sub-sample of the data set selected at random).

(a)



(b)

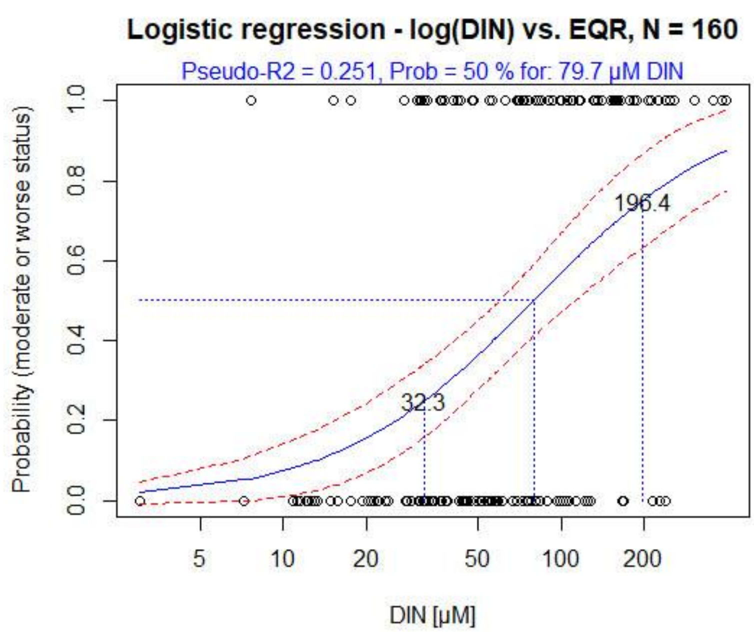


Figure 7-7: Binomial logistic regression of DIN on probability of being (a) good or worse status and (b) moderate or worse status (normalised EQRs used). Lines show potential boundary values at different probabilities of being (a) good or worse status and (b) moderate or worse.

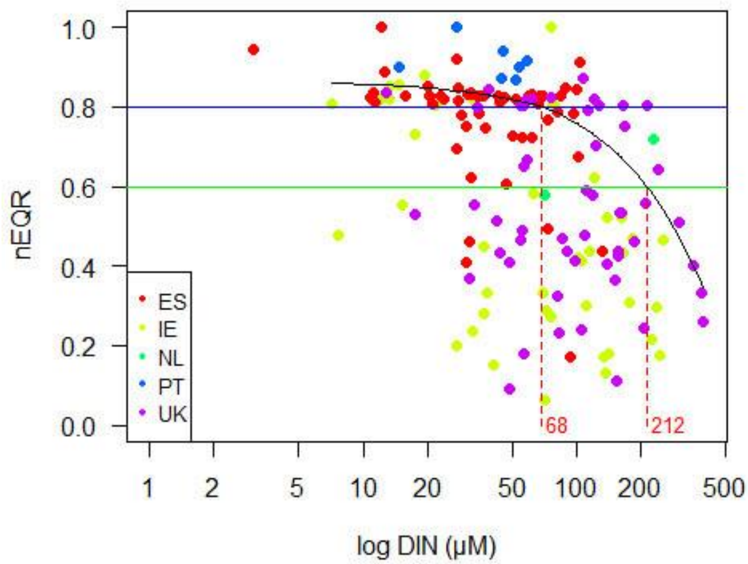


Figure 7-8: Quantile regression fit at the 70th quantile (additive quantile regression smoothing rqss using quantreg R package by Koenker) for nEQR v DIN (µM) in the NEA11 common type (data from the Spain, Ireland, Netherlands, Portugal, and the United Kingdom. Horizontal lines indicate EQR boundaries at H/G and G/M, and vertical lines the nutrient boundaries respectively for high/good and good/moderate, at this quantile.

NEA11 is a broad common type that includes estuaries from eight Member States (see Commission Decision 2018/229/EU), encompassing a wide geographic range and great natural variability. Such variability is reflected in different national nutrient boundaries adopted by Member States for different systems within type NEA11. For example, the United Kingdom has defined four sets of boundaries within NEA11 depending on water transparency, from clear to very turbid waters; while France has defined two sets of boundaries depending on whether NEA11 estuaries are in the Channel and Atlantic or in the North Sea (Table 4-13). The boundaries for DIN in NEA11 suggested by the different approaches tested are summarised in Table 7-2, and can be compared with the national nutrient boundaries for this type.

The analysis of the available IC data has shown that the difference between good and moderate status was not significant, which compromises the robustness of the results obtained for most of the tested approaches. For the H/G boundary, however, values suggested by the categorical approaches and quantile regression may be considered for guidance. The categorical and quantile regression results obtained with this IC data set (Table 7-2) indicate G/M boundaries in line with UK boundaries adopted for medium to very turbid waters, but do not seem adequate to protect clearer waters. The boundaries are also not in line with the more stringent H/G or G/M French boundaries in either of their NEA11 national subtypes. These results reinforce the need to account for additional environmental factors when creating data sets for establishing nutrient boundaries across common types, in order to accommodate within type natural variability, particularly for broadly defined types across Europe. A mixed data set covering a wide gradient of more to less turbid systems would mask the relationship between nutrients and phytoplankton in clearer waters, as turbidity would control phytoplankton growth allowing for good ecological quality values to be attained at higher nutrient values than would be expected for example in non-turbid systems. This is the reason for the widely scattered data observed, and emphasises the need to evaluate and interpret all values produced using the toolkit.

Table 7-2: Predicted boundary values for DIN (µM) in common type NEA11 (n = 160), derived from the most adequate approaches for this data set (Excel toolkit (vs 6c) and/or R scripts). Results from regression and categorical methods are presented; those in red need to be taken with caution. 'Possible range' refers to values derived from the interquartile range of the residuals.

Boundaries NEA11		DIN µM	
Regression methods (OLS and type II):		**H/G**	**G/M**
Most likely boundary	predicted	36	62
	range	(14-43)	(61-72)
	possible range	5-79	23-278
Additive quantile regression method (rqss):			
70th percentile		68	212
Categorical methods:			
Average adjacent class upper and lower quartiles		49	80
Average adjacent class median		47	82
75th quartile of class		62	107
Mismatch of biological v nutrient class (*Excel toolkit*)		50	72
Mismatch of biological v nutrient class (*R scripts*)		53	75
	range	(47-59)	(66-83)
Logistic binomial regression (*prob = 0.5*)		45	80

7.3. Why is it challenging to establish nutrient concentrations that would support good ecological status?

The CIS eutrophication guidance document recognised that there would be difficulties in establishing appropriate boundary values (paragraph 139 stated 'In setting the value of the (nutrient) standard, Member States are faced with a choice of selecting values anywhere on the spectrum between close to the G/M boundary for either most sensitive or the least sensitive water body in the type*')* due to lack of scientific understanding of causal links (see paragraph 187). As a result Chapter 4 of the CIS classification guidance document proposes a checking procedure 'designed to ensure that the type-specific values established for the general physicochemical quality elements are no more or no less stringent than required by the WFD'. The implication of this advice was that the mismatch between biology and nutrient concentration was a result of inadequate data and understanding of causal links and that this could be solved as more data became available. However, from the results of the work carried out to produce this document, it is clear that the statistical relationship between the supporting nutrient elements and ecological status will always be uncertain. Uncertainty can be reduced by using adequate data sets and categorising water bodies into similar types, but the complexity of the ecological interactions involved will always result in a variation in biological status (EQR) at any nitrogen or phosphorus concentrations for any water body. It is thus essential that we understand and manage this variation as we both set and use boundaries for environmental management.

At a more practical level it is also important to be aware of the different sources of uncertainty and how these are expressed in the outputs from the toolkit.

Sources of uncertainty

7.3.1. Data uncertainty

The toolkit uses a data set, which we assume is made up of:

- similar types of water body, such that their response to nutrients will be similar;
- adequate data, that captures the spatial and temporal variation, including the part of the gradient where biology responds linearly to nutrients;
- reliable analytical procedures for both explanatory and response variables, particularly the use of an intercalibrated biological metric.

Data that fulfil these criteria should be available from national monitoring programmes.

7.3.2. Statistical approach to uncertainty

Given reliable data, we use statistical procedures to develop a type-specific relationship that allows the nutrient concentration at a particular biological status, typically at EQR boundary values, to be determined. The conceptual model underpinning this assumes that increases in nutrient concentration 'cause' a decrease in ecological status. However, as we are using real-time monitoring data, rather than a controlled experimental design, the values used to generate predictions cannot be assumed to be error free. Thus, relying on the minimisation of deviations of EQR from a regression line is likely to underestimate the slope of the 'true' relationship. The converse is also true: it is very unlikely that our measurement of EQR is error free, so fitting a line where we minimise the deviations along the nutrient axis will overestimate the slope of the 'true' relationship. This is why we propose the use of type II regression where deviations are minimised along both EQR and nutrient axis, in effect at an angle of 90 degrees to the best fit line (orthogonal).

The certainty of each of the regression lines is assessed by r^2, and the standard error of their slopes can also be determined. In reality we also do not know which of our models predicts the best approximation to the 'true' regression line, so we present the predictions from each of the OLS regressions to represent the range of the most likely values, with the type II model prediction as the best single estimate.

Thus, depending on the scatter within the data and the choices made when selecting the data used in the models (outliers, linear range, etc.), a range of predicted regression lines can be produced from which the 'most likely' boundary concentrations has to be determined.

Similar estimates are also produced from the categorical box plot methods; these are not influenced by the uncertainty of statistical modelling, and each method produces a single estimate. However, the estimate depends upon factors such as the relative number of water bodies in each biological class and the width of that class. Thus, these categorical estimates are also uncertain, only representing the value for the particular data set. Uncertainty can be estimated by fitting a binary logistic model, or by the use of bootstrapping when estimating misclassification rates but, as for the simpler methods, the results are dependent on the reliability of the biological status classification.

Bringing all these results together results in a range of 'most likely' boundary values. The extent of the range reflects our statistical uncertainty. Data where r^2 values are low will have higher uncertainty — some data sets may be so uncertain it is impossible to make a prediction.

7.3.3. Model uncertainty

The regression models provide our best estimate of the 'average' response for water bodies in the data set. Individual water bodies will fall above or below that line. This may be because the water body mean EQR or nutrient concentration is itself uncertain and does not reflect its true value (data uncertainty) or that environmental conditions or other pressures combine to give a unique outcome for each water body. Hence, the result is that a boundary established from the average behaviour of a water body type will often be too high or too low in comparison to its biological status for any specific water body of that type.

We express this uncertainty in the toolkit using the interquartile range of the residuals of the regression models and use this to predict a further range of boundary values, the 'possible range'. The magnitude of the possible range depends on the quality of our conceptual model. For example, in mesotrophic deep lakes phytoplankton biomass is normally highly dependent on phosphorus and thus the relationship between phytoplankton EQR and TP is normally very good ($r^2 > 0.65$). Conversely, in rivers phytobenthos and macrophytes will respond to many other pressures and be subject to other influences such as grazing or shade. Thus, while ecological dogma assumes that nutrients are a major pressure influencing the ecological structure and function that the WFD tools attempt to measure as 'ecological status', the reality is that simple pressure response models are unlikely to be adequate and our predictions of nutrient boundaries will have very large uncertainty bands. Until it is possible to improve our conceptual models and develop statistical models that incorporate a wider range of variables we need to recognise and manage this variation when we set boundary values for management.

Most likely boundary value

Given the uncertainty described above, decisions about the values to select need to be made. The boundary values derived from the regression best fit lines and most of the categorical methods (not the method based on the 75th quantile of class) all represent the 'most likely' boundary value (i.e. the average nutrient concentration that occurs when the BQE was at the G/M boundary. When nutrient concentrations are lower than this value, a typical water body should be at or above the BQE boundary. If the classifications based on biology and nutrient concentration for a large number of water bodies were compared, then there would probably be some water bodies with different classifications. The number with different classifications would depend on the responsiveness of the BQE to nutrients and how successful the typology was for reducing variability, but there should be no bias in either direction. The likelihood of achieving good status with the mean nutrient concentration as the boundary would be 50 %.

Most protective boundary value

If the boundary is to be used to establish the nutrient concentration that will ensure that the majority of water bodies within a type are at good status, then a value predicted using one of the lower quantiles of the linear regression residuals, a lower quantile of a quantile regression, or a lower probability value from binary logistic regression should be used. Such a boundary should support good status.

Box 7: Pros and cons of selecting the most likely boundary value (mean regression line)

- Selecting a boundary derived from the mean estimate of the best-fit regression line would ensure that half of water bodies would have good status for the nutrient sensitive BQEs if the nutrient boundary was achieved.
- There would be a low misclassification rate with most water bodies having a nutrient assessment in the same status class as the nutrient-sensitive biology.
- There would be a medium risk of downgrading a water body using the WFD 'worst of any' approach to assessment.

This has the disadvantage that for any particular water body there is a greater risk that the nutrient concentration would be exceeded, and thus not be considered to be supporting GES, even though the biological assessment still categorised the water body as being at good status. Thus, for a group of water bodies there would be a bias against the biological assessment. This approach would be highly protective, but could result in unnecessary downgrades of status using the 'one out all out' rule.

Most certain that biology dictates status

If avoiding unnecessary downgrades due to the supporting element is the primary requirement, then an upper quantile of the regressions should be used. If a regression has not been possible, then an upper quantile of the distribution of nutrients in water bodies at good status, or higher probability value of logistic regression, could provide a similar value. As for the previous scenario, the misclassification rate would be higher, but the bias would be in the opposite direction. On average, only 25 % of water bodies would be classified as not being at good status based on nutrients when their biological status was good.

In the case of an inverted wedge, where other pressures are probably reducing biological status, then only an upper quantile can be determined. This would be a nutrient concentration which would be relatively certain of causing a downgrade of biological status.

Box 8: Pros and cons of selecting a protective boundary value (lower percentiles)

- Selecting a boundary derived from the lower uncertainty estimate of the best-fit regression line would ensure that most water bodies would have good status for the nutrient sensitive BQEs if the nutrient boundary was achieved.
- A protective boundary would minimise the risk for undesirable secondary effects of eutrophication, including a low risk for negative impacts on ecosystem function and services.
- A protective boundary would counteract negative synergistic effects of other pressures, as well as effects of climate change, which would otherwise pose a risk for deterioration.
- A protective boundary would have a higher probability of protecting downstream water bodies from eutrophication.
- There would be a higher misclassification rate with the nutrient assessment being less than good more often than the biology.
- There would be a greater risk of downgrading a water body using the WFD 'worst of any' approach to assessment (increase of type I error, i.e. false positives, incorrectly assuming a nutrient negative effect on BQE when there is none).

Box 9: Pros and cons of selecting a relaxed/lenient boundary value (upper percentiles)

- Selecting a boundary derived from the upper uncertainty estimate of the best-fit regression line would ensure that only a minority of water bodies would have good status for the nutrient sensitive BQEs if the nutrient boundary was achieved.
- If the boundary is used only for status assessment, another boundary would be needed to plan nutrient-reduction measures to restore those water bodies where the BQEs are less than good.

- An upper percentile boundary would not protect against the risk for undesirable secondary effects of eutrophication, and would pose a risk for negative impacts on ecosystem function and services, as well as on downstream water bodies.
- There would be a higher misclassification rate with the biology assessment being less than good more often than the nutrient assessment (increase of type II error, i.e. false negatives, incorrectly assuming no nutrient negative effect on BQE when in fact there is).

Interaction of nitrogen and phosphorus

When nitrogen and phosphorus are highly correlated, it is impossible to determine which nutrient is responsible for determining biological status. However, the very shallow lake example (7.1) which included lakes from the eastern continental GIG, showed evidence of lakes where low nitrogen occurred in lakes that were in good biological status despite very high phosphorus concentrations. When these lakes were included, a regression using phosphorus did not produce significant relationships. Phosphorus boundary values based on other very shallow lakes (broad type 4) would appear to be much too low for these lakes and cause substantial misclassification. A solution might be to place these lakes into a different lake type, perhaps accounting for the low nitrogen and high phosphorus as a result of higher continental temperature accelerating biogenic cycling. Another option would be to establish phosphorus and nitrogen boundaries using the data set for the broad type, but to use a nutrient combination rule that required both nitrogen and phosphorus boundaries to be exceeded before the nutrient supporting quality element was not at good status. This would deal with nitrogen or phosphorus limited biota, minimise the risk of unnecessary downgrades while still establishing boundary values more typical of the broad type. It also has the advantage that there is no need to determine in advance which lakes, or other water categories, are nitrogen or phosphorus limited, a feature that might change over time. It would also allow for any synergistic interaction of nutrients, a topic of growing interest in the literature.

A pragmatic approach has been adopted in Norway where nitrogen can occasionally becoming limiting in highly eutrophic water bodies, especially in late summer. TN boundaries have, therefore, been derived for different types of rivers and corresponding lake types, aiming for use in classification when nitrogen limitation is indicated (TN:TP < 20 and low N:P combined with nitrate-N and ammonium-N below detection limit) and to protect downstream coastal waters.

TN boundaries for lakes were set in two different ways:

- using the lower 25th percentile of the regression of phytoplankton nEQR versus TN for different lake types (Figure 7-9a);
- using the type-specific TP boundaries to infer the corresponding TN boundary from a regression model spanning all lake types (Figure 7-9b).

In practice, the two approaches produce very similar boundary values so either can be used. The type-specific TN boundaries are the same for lakes as for rivers of corresponding types (in terms of altitude, alkalinity and humic content), because nitrogen retention is minor in most Norwegian lakes, which are mostly well oxygenated and, as a result, denitrification is negligible.

7.4. Establishing nutrient boundaries in the presence of other stressors

The challenge presented by multiple stressors has been a recurring theme in this document, although no satisfactory solution has been achieved.

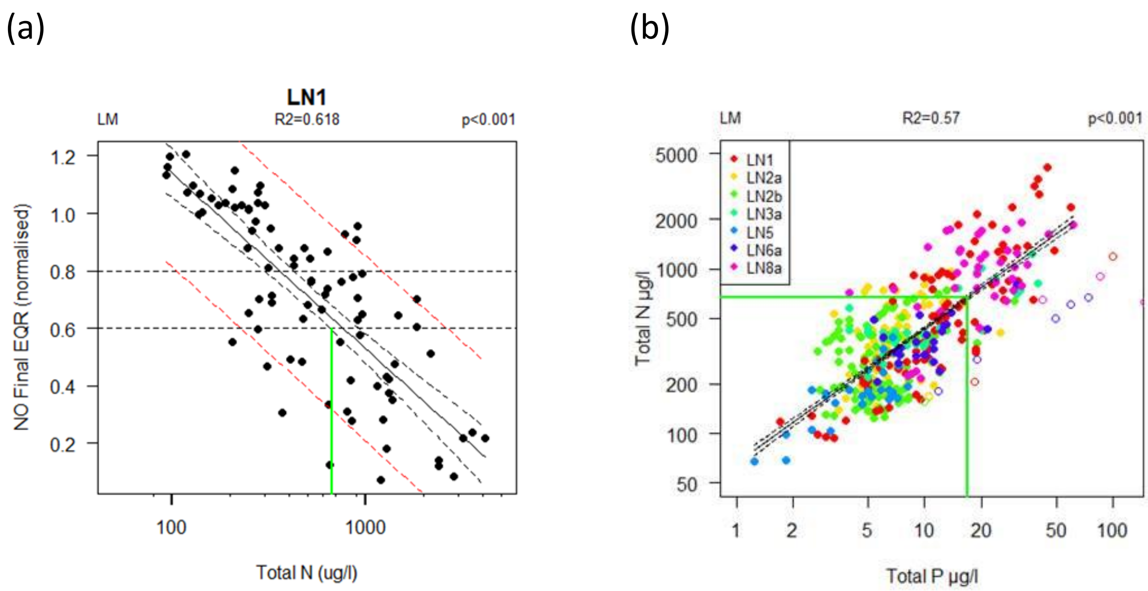


Figure 7-9: (a) Linear regression (r2 = 0.62) of the Norwegian nEQR for phytoplankton versus TN in lakes of NGIG type L-N1 (lowland, moderate alkalinity, stratified, clear). The green line shows the total-N boundary using the lower 25th confidence limits of the regression (675 µg L-1); (b) regression of TN vs TP for Norwegian lakes across all NGIG types (r2 = 0.57). Green lines are total-P (17 µg L-1) and Total-N (675 µg L-1) boundary for the type L-N1. The open circles below the hatched line are data that were excluded from the regression.

Situations where the scatter of biological status against nutrients is in the form of a wedge are assumed to indicate situations where another stressor (potentially, more than one) overrides the nutrient-biology relationship. In such situations, a low EQR may arise either due to nutrient pressure, to these other pressures or to interactions between them. In these cases, it is not appropriate to fit a type I or type II regression line and either quantile regression (Section 2.4.3) or use of an appropriate upper percentile from a categorical approach (Section 2.4.4) would be the preferred solution. An alternative might be to fit a binary logistic regression and estimate boundary values from a higher probability value, although it would be important to determine that the logistic model was significant. The upper quantile shown in Figure 2-2a, for example, is the highest concentration of nutrients that is associated with good status in the data set and many other status outcomes are possible at the same level of enrichment. However any of these methods will, inevitably, predict values that are relatively lenient. Such values are, at best, interim targets that ensure that inappropriate investment in nutrient reduction strategies is minimised, but are far from ideal as means of ensuring long-term sustainable water resources.

Example approaches

Germany's approach has been to define so-called orientation values, which represent the highest values of a physicochemical parameter that is consistent with GES (Müller et al., 2017). This means that GES is very unlikely to be achieved if this value is violated, even if all other supporting elements are at levels that should support GES. In essence, this involves preliminary assembly and analysis of the data set similar to the steps described in Section 4.2.1, followed by categorical analysis based on box and-whisker plots (Section 4.2.4). The steps are:

- establishment of 'threshold values' as outer whiskers of either good status or high and good status combined (see Figure 7-10);
- determining that there is a significant difference between classes using a Kruskal-Wallis test (when five classes are used) or Wilcoxon test (if just two groups of classes are used);
- testing that there is a relationship between the pressure and the outer whiskers using linear regression ($r^2 \geq 0.8$).

These steps are repeated for each river type, after which results are scrutinised and, if necessary, values obtained for similar types are harmonised (bearing in mind factors such as limited data availability for some less common types, and other factors). At this stage, expert judgement is used to check that orientation values produced are consistent with other evidence for the effect of each particular physicochemical parameter on the biota.

A different approach has been adopted in France. Here, the limitations of a simple linear analysis in multi-stressor situations is acknowledged and, instead, data points are envisaged as occupying an area under a curve, where the curve represents the highest value of an indicator at that point along the nutrient (Figure 7-11; this is conceptually identical to the reasoning behind the 'wedge', described above).

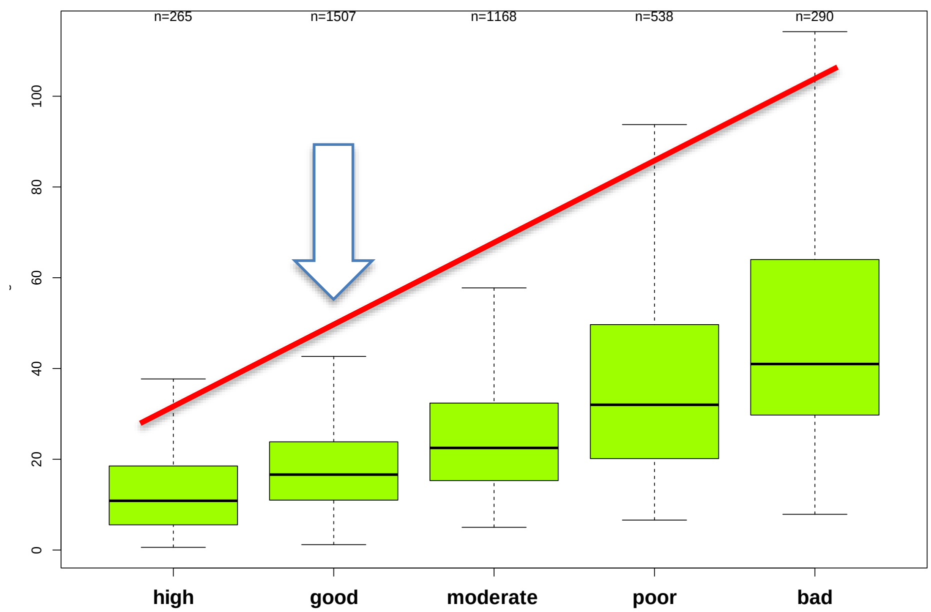


Figure 7-10: Example for the determination of threshold values as upper bounds for good ecological status.

Dodds et al (2010) listed some methods for identifying nonlinear and threshold relationships including the Cumulative Frequency Distributions (Paul and McDonald, 2005) which can be used to define thresholds derived from ecological states. This method detects changes in expected distributions, and can be combined with analytical methods to determine differences among distributions (e.g. detection limit of the method) and acceptable levels of change (e.g. region with 95 % certainty of change). This method identifies a concentration (with an associated confidence interval) for a physicochemical parameter where it appears that it becomes a limiting factor for a given BQE. Before reaching this threshold, the value of the index will be the result of interactions amongst a set of stressors while , beyond this point, the best possible state is capped due to the limiting effect (direct or indirect) of the parameter under consideration.

This conditional probability approach is similar to the minimisation of mismatch but relies on nonparametric deviance reduction to determine the change point (Paul and McDonald, 2005). This is a two-step process: the derivation of empirical cumulative frequency (i.e. successive calculations by adding the values one to one) followed by identification of the change point on this curve by nonparametric deviance reduction. This approach determines the dividing point for splitting the data into two groups, resulting in the largest reduction in the deviance in the data (Qian et al., 2003). When the data are divided into two groups, the sum of the deviance for the two subgroups is always less than or equal to the deviance for the entire data set. The threshold is identified as the point where the split in the data minimises the deviance (Figure 7-12). Uncertainty in the deviance reduction change point (10 and 90 percent confidence intervals) can be estimated from the empirical percentiles for the bootstrap distribution by resampling.

While the point of lowest deviance can generally be determined, it is dangerous to infer biological significance without further investigation. A three-stage checking procedure was therefore developed. The first step consists of visual checks of graphs to ensure that results are convincing; the second step is a post hoc analysis based on taxon lists to check ecological consistency (with an-ANOVA-like test based on Bray-Curtis distance matrix); and the third step involves considering the rate of false negative defined by the method (when rivers are at GES but nutrients are less than good). A score ranging from one to three is allocated to each of these three criteria and their sum makes it possible to estimate the level of confidence of the threshold under consideration.

This study has been tested with French river data in order to detect thresholds of the physicochemical elements supporting the biological elements beyond which the achievement of the values specified for the biological quality elements in terms of EQR is no longer ensured (Ferreol and Bougon, in prep).

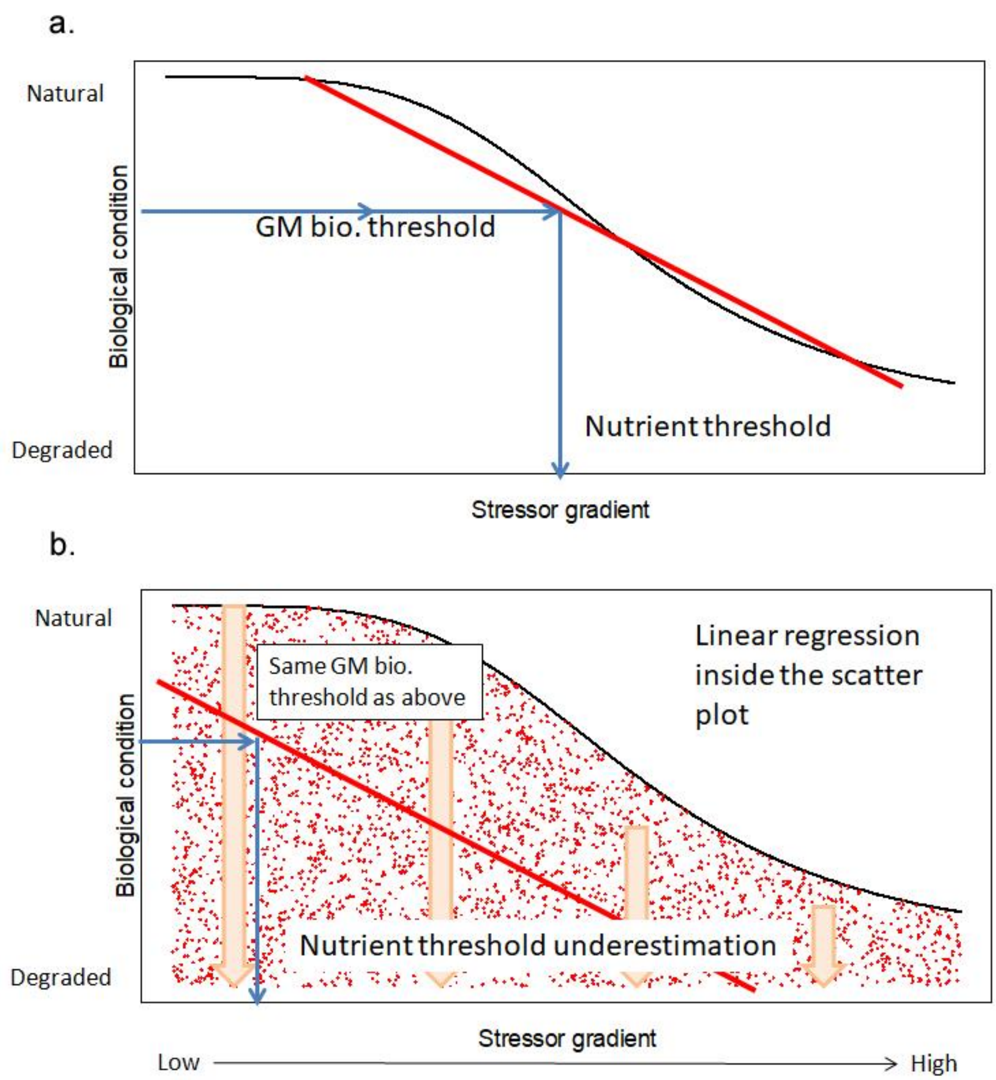


Figure 7-11: (a) Dose-response relationship with single stressor gradient; (b) Dose-response relationship along multi-stress gradient.

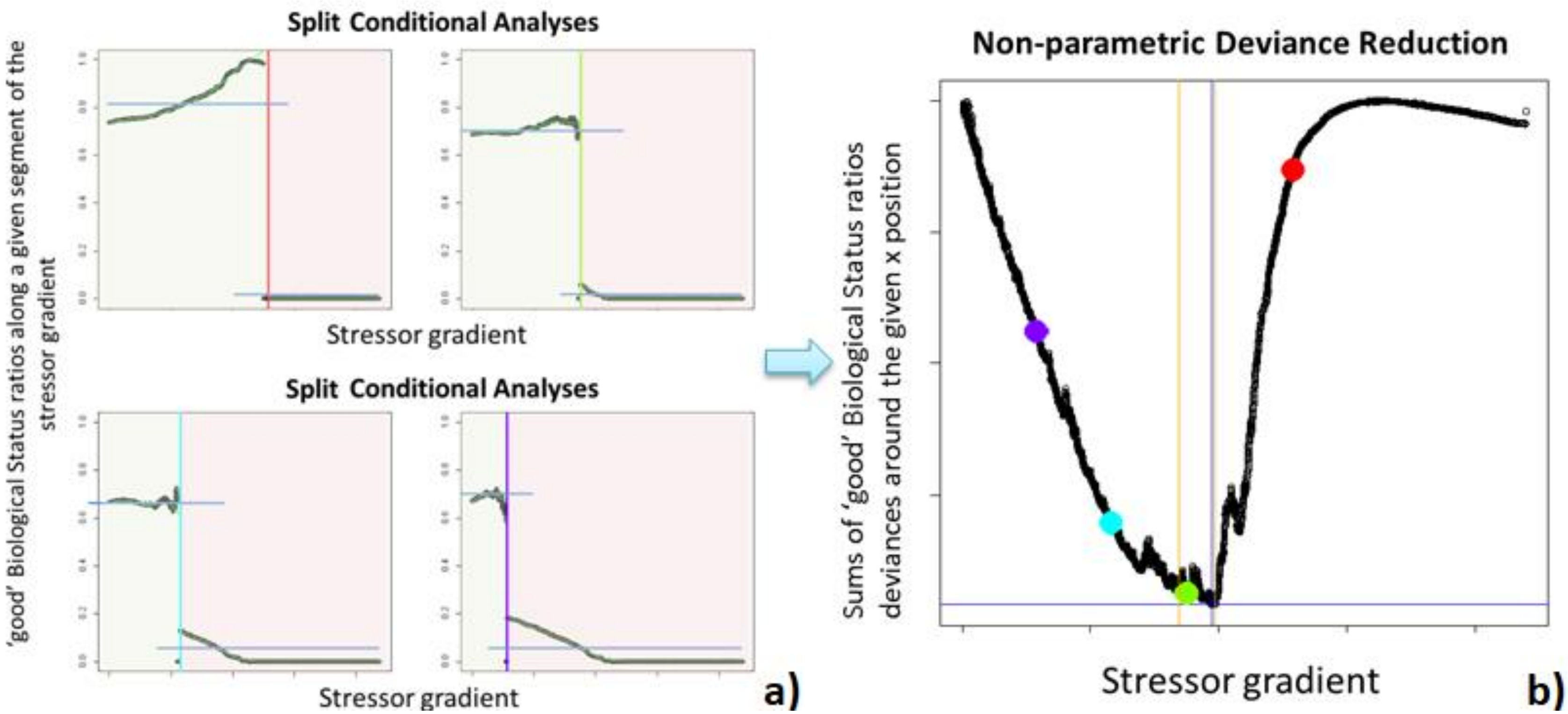


Figure 7-12: (a) Example of sums of the deviance before (green zone) and after (red zone) a given value of the extent levels of one physicochemical parameter. The sum of the deviance is calculated for the two subgroups separately. For the red zone, the sum starts at the right part and ends at the breakpoint; and for the green zone, the sum starts at the breakpoint and ends at the left part.; (b) The sum of the deviances are then used on one graphic and the point of minimum deviance (vertical blue line) is defined as the minimum of the empirical curve. The confidence interval is represented by the vertical orange lines.

7.4.2 Regulatory response

From a regulatory point of view, stressors that confound a nutrient–biology relationship fall into two categories.

- Cases where there is potential to minimise the effect of the confounding stressor by regulation, in which case, the ‘true’ nutrient–biology relationship should gradually be revealed over time, permitting nutrient boundaries to be reviewed. In rivers that are subject to organic pollution the impact of nutrients is conflated with that of high concentrations of fine particulate organic material, low oxygen concentrations, ammonium toxicity, etc. Regulators have become aware of the role of nutrients as an independent stressor as the impact of these other factors have declined in many parts of Europe due primarily to better sewage treatment.
- Cases where the effect of the confounding stressor cannot be minimised by regulation. Hydromorphological alteration of lowland rivers is a good example of this type of situation, with amelioration possible in some, but far from all, circumstances. In this case, a lenient nutrient target established by quantile regression is not a realistic long-term solution.

In the second case, extensive hydromorphological alteration may lead to designation of a ‘heavily modified water body’, with the objective of good ecological potential rather than good ecological status. In practice, however, the need to manage nutrients remains, in order to ‘maintain the functioning of the ecosystem …’ (see WFD Annex V, Section 1.2.5). If it is not possible to do this using the intercalibrated biological metrics due to interactions between stressors, then alternatives must be sought. Other parts of this document have emphasised the need to have a clear understanding of good ecological status before embarking on attempts to derive nutrient targets and, similarly, having a clear vision of good ecological potential is an option for developing nutrient targets in the presence of hydromorphological stressors. Guidance for determining good ecological potential is in preparation which will give more details on general chemical requirements.

A pragmatic option could be to develop targets that minimise secondary effects in order to protect or enhance ecosystem function and services. Focusing attention on ‘undesirable disturbances’ such as the frequency of harmful algal blooms or fish kills caused by deoxygenation, for example, rather than on the usual constituents of ecological assessment metrics, may give insights that allow thresholds established by quantile regression to be improved. Other options described in Chapters 5 (Alternative Approaches) and 6 (Validation) may also be appropriate.

It should also be emphasised that some alterations to channel morphology and connectivity, as well as climate change, may combine to make a given system more sensitive to nutrients than would be the case under more natural conditions (warming, alterations to flow and removal of shade, for example, combining to promote high algal/plant biomass and greater risk of secondary effects). Under such circumstances, a standard that is more stringent than that predicted for an equivalent ‘natural’ river may be necessary to offset the consequences of such alterations.

The MARS project output *Recommendations on how to best assess and mitigate impacts of multiple stressors in aquatic ecosystems* offers more guidance on how to manage multiple pressures (Schinegger et al., 2018).

7.5. 'Weight of evidence' approach

The United Kingdom uses a categorical 'weight of evidence' approach to determining eutrophication risk in both freshwaters and transitional and coastal waters. The transitional and coastal water framework is presented here as an example of how different strands of evidence can be combined to aid decision-making. This considers information about compliance with the WFD DIN standard GES (as a measure of exposure pressure) along with information on eutrophication-related impacts based primarily on the WFD classification tools for phytoplankton and opportunistic macroalgae and their G/M status boundaries. These are the biological elements most sensitive to nutrient pressure. The latest WFD classifications are the main source of data, although results for previous years were also considered. A matrix is used to enable results for nutrients, WFD impact indicators and wider evidence to allow an expert evaluation of the likelihood of eutrophication impacts.

This is a four-step process:

Step 1 — Assess nutrient pressure using WFD classification results.

The most recent classification results are used to assess whether:

- DIN data indicate no problem (good status or better);
- DIN data indicate a possible problem (moderate status or worse).

Where there are no DIN data for a water body expert judgement is used to decide the status is expected to be the same as that for adjacent water bodies. A precautionary approach is taken at present and few water bodies have been grouped at this stage. This is a step that could be considered in the future.

Step 2 — Assess evidence of a potential eutrophication problem using WFD classification results for the primary biological indicator.

Classification data for both phytoplankton and opportunistic macroalgae are not available for all water bodies. Expert judgement is used to decide which BQE is expected to be the most responsive to elevated nutrients in each water body. In general opportunistic macroalgae are usually the most responsive BQE in water bodies with suitable intertidal areas for the algae to attach and grow, while phytoplankton is the most responsive element in other water bodies.

The most recent opportunistic macroalgae and phytoplankton classification results are used to assess whether the following are true.

- The most responsive BQE does not indicate a problem (good status or better).
- One or both of the primary biological elements indicate a possible problem (moderate status or worse).
- It is not clear whether there is a problem because the BQE is close to the G/M boundary or the status is variable or there is some evidence of a changing trend.
- It is not clear whether there is a problem due to limited data for the most responsive element. WFD opportunistic macroalgae classifications are considered to be uncertain if there are less than 2 years of data. WFD phytoplankton classifications are considered to be

uncertain if there are data for only one metric, or if samples need to be collected in more years or months. Data from 3 years are required for coastal waters, with records from at least 9 months in each year. Five years of data are required for estuaries, with records from at least 10 months in each year.

- There are no data for the most responsive quality element.

Step 3 — Assess further supporting evidence.

If data for phytoplankton or opportunistic macroalgae indicate there is potential for eutrophication to occur as nutrients appear to be fuelling at least some excess plant growth, more evidence is required to evaluate whether secondary effects are occurring or likely. This could include indications that algal growth is causing, for example: dissolved oxygen problems, algal scums, entrained and/or overwintering macroalgae or effects on invertebrates/seagrass. Some of this evidence may come from local investigations carried out to investigate WFD moderate status classifications. Evidence collected for other purposes (e.g. habitats directive and urban waste water treatment directive assessments) may be used, along with evidence collected by other organisations. This new evidence should improve certainty when considering appropriate measures.

Step 4 — Bring the evidence together to assign a level of certainty of eutrophication to each water body.

A matrix (Figure 7-12) is used to assign one of the following outcomes to each coastal and estuarine water body:

- very certain there is a problem;
- quite certain there is a problem;
- uncertain there is a problem (statistical) — this may be because biological status is near the G/M boundary or there is high variability in the data or we are not yet sure if quality is improving or deteriorating;
- uncertain there is a problem (data) — this applies in situations where biological data are limited;
- certain there is not a problem;
- not assessed.

Any anomalies in this process will lead to either an investigation or a re-consideration of standards.

-

		Biological evidence					
		No data on most responsive QE	Most responsive QE is good status	Insufficient data on most responsive QE	Most responsive QE varies in status or is close to good/moderate boundary	Most responsive QE fails good status	Most responsive QE fails and there is further supporting evidence of adverse effects from the elevated plant growth
DIN evidence	DIN is good status, or assumed to be good having grouped waterbodies	Certain there is not a problem	Certain there is not a problem	Certain there is not a problem	Uncertain there is a problem (statistical)		
	No DIN data	Not assessed	Certain there is not a problem	Uncertain there is a problem (data)	Uncertain there is a problem (statistical)	Quite certain there is a problem	
	DIN status is uncertain or close to the good moderate boundary	Not assessed	Certain there is not a problem	Uncertain there is a problem (data)	Uncertain there is a problem (statistical)	Quite certain there is a problem	
	DIN is moderate status	Not assessed	Certain there is not a problem	Uncertain there is a problem (data)	Uncertain there is a problem (statistical)	Quite certain there is a problem	Very certain there is a problem

Figure 7-13: Decision matrix used as part of the weight of evidence process for evaluating eutrophication risk in United Kingdom transitional and coastal waters

7.6. Wider considerations

Finally, it is important to remind ourselves that the WFD makes it clear that the purpose of establishing boundary values for nutrients is to ensure the functioning of the ecosystem, not simply to ensure that BQEs achieve good or better status ('nutrient concentrations do not exceed the levels established *so as to ensure the functioning of the ecosystem* and the achievement of the values specified for the biological quality elements', WFD Annex V, Section 1.2)

By using pressure response relationships to determine equivalence between a biological G/M boundary and nutrient concentration we are assuming that our biological indicators do indeed reflect ecological function. The intercalibrated WFD biological methods are our current best available assessments of ecological status; however they are not perfect, and it is important to use the wider ecological literature alongside empirical analysis. For example, it is widely reported that the response to phosphorus in lakes by phytoplankton reaches a plateau at concentrations above 100 µg L^{-1} (Phillips et. al., 2008). Although nutrient responses in rivers are more difficult to quantify, significant ecological changes do occur in rivers at similar concentrations (Mainstone, 2010). Thus, further consideration of whether the functioning of the ecosystem is still ensured might be needed where G/M boundary values are substantially higher than these levels, particularly where they are derived from weak relationships or where multiple pressures exist.

8. References

Aldridge J.N., Tett, P., Painting, S.J., Capuzzo, E. and Mills, D.K. (2010a). The dynamic Combined Phytoplankton and Macroalgae (CPM) Model: User Guide. Contract C3290 Report, Environment Agency.

Aldridge J.N., Tett, P., Painting, S.J., Capuzzo, E. and Mills, D.K. (2010b). The dynamic Combined Phytoplankton and Macroalgae (CPM) Model: Technical Report. Contract C3290 Report, Environment Agency.

Andersen, J.H. and Conley, D.J. (eds) (2006). Eutrophication in coastal Ecosystems: selected papers from the second International Symposium on Research Management of Eutrophoication in Coastal ecosystems, 20-23 June 2006, Nyborg, Denmark.

Andersson, A., Hajdu, S., Haecky, P., Kuparinen, J. and Wikner, J. (1996). Succession and growth limitation of phytoplankton in the Gulf of Bothnia (Baltic Sea). Marine Biology 126:791-801.

Anon. (1982). Eutrophication of waters: monitoring, assessment and control. Organisation for Economic Cooperation and Development Environmental Directorate, Paris, 154 pp.

Baker, M.E., King, R.S. (2010). A new method for detecting and interpreting biodiversity and ecological community thresholds. Methods in Ecology and Evolution 1: 25-37.

Behrendt, H. and Zessner, M. (2005). Point and diffuse nutrient emissions and loads in the transboundary Danube River Basin — II. Long-term changes. River Systems 16: 221-247.

Berland, B. R., Bonin, D. and Maestrini, S.Y. (1980). Azote ou phosphore? Considérations sur le paradoxe nutritionnel de la Mer Méditerranée (Nitrogen or phosphorus? Considerations on the nutritional paradox of the Mediterranean Sea) . Oceanologica Acta 3: 135-142.

Bethoux, J.P., Morin, P., Madec, C. and Gentili, B. (1992). Phosphorus and nitrogen behaviour in the Mediterranean sea. Deep-Sea Res 39: 1 641-1 654

Birk, S. and Hering, D. (2008). A new procedure for comparing class boundaries of biological assessment methods: a case study from the Danube Basin. Ecological Indicators 60: 71-83.

Birk, S., Willby, N.J., Kelly, M.G., Bonne, W., Borja, A., Poikane S. and van de Bund, W. (2013). Intercalibrating classifications of ecological status: Europe's quest for common management objectives for aquatic ecosystems. Science of the Total Environment 454-455: 490-499.

Black Sea Commission (2008). State of the Environment of the Black Sea (2001-2006/7). Edited by Temel Oguz. Publications of the Commission on the Protection of the Black Sea Against Pollution (BSC) 2008-3, Istanbul, Turkey, 421 pp.

Bonin, D.J., Bonin, M.C. and Berman, T. (1989). Mise en evidence expérimentale des facteurs nutritifs limitants de la production du micro-nanoplancton et de l'ultraplankton dans une eau côtiere de la Méditerranée orientale (Haïfa, Israël). (Experimental demonstration of the nutrients limiting production of micro-nanoplankton and ultraplankton in coastal waters of the eastern Mediterranean (Haifa, Israel). Aquatic Science 51: 129-152.

Borch, H., Eggestad, H.O., Engebretsen, A., Skarbøvik, E., Lyche-Solheim, A. and Tjomsland, T. (2008). *Klassegrenser for næringssalter i leirvassdrag* (Class boundaries for nutrients in naturally turbid rivers in catchments with clay sediments). Chapter 3 and Annex D in Lyche Solheim, A., Berge, D., Tjomsland, T., Kroglund, F., Tryland, I., Schartau, A.K., Hesthagen, T., Borch, H., Skarbøvik, E., Eggestad, H.O. and Engebretsen, A. (2008). *Forslag til miljømål og klassegrenser for fysisk-kjemiske parameter I innsjøer og elver, inkludert leirvassdrag og egnethet for brukerinteresser* (Suggestion for environmental objectives and class boundaries for physicochemical parameters in lakes and rivers, including rivers in clay-influenced river basins and suitability for user interests. Supplement to Guidance on classification of ecological status) (in Norwegian with English summary). NIVA report 5708: 77 pp.

Brett, M. T. and Benjamin, M. M. (2008). A review and reassessment of lake phosphorus retention and the nutrient loading concept. Freshwater Biology 53: 194-211.

Burson, A., Stomp, M., Akil, L., Brussaard, C.P.D. and Huisman, J. (2016). Unbalanced reduction of nutrient loads has created an offshore gradient from phosphorus to nitrogen limitation in the North Sea. Limnology and Oceanography. 61: 869-888.

Cade, B.S. and Noon, B.R. (2003). A gentle introduction to quantile regression for ecologists. Frontiers in Ecology and the Environment 1: 412-420.

Carstensen, J., Sanchez-Camacho, M., Duarte, C., Krause-Jense, D. and Marbàa, N. (2011). Connecting the Dots: Responses of coastal ecosystems to changing nutrient concentrations. Environmental Science & Technology 45: 9 122-9 132.

Carvalho, L., McDonald, C., Hoyos, C., Mischke, U., Phillips, G., Borics, G., Poikane, S., Skjelbred, B., Lyche-Solheim, A., Van Wichelen, J. and Cardoso, A. C. (2013). Sustaining recreational quality of European lakes: minimizing the health risks from algal blooms through phosphorus control. Journal of Applied Ecology 50: 315-323.

Clarke, A.L., Weckström, K., Conley, D.J., Anderson, N.J., Adser, F., Andrén, E., de Jonge, V.N., Ellegaard, M., Juggins, S., Kauppila, P., Korhola, A., Reuss, N., Telford, R.J. and Vaalgamaa, S. (2006). Long-term trends in eutrophication and nutrients in the coastal zone. Limnology and Oceanography 51: 385-397.

Cugier, P., Billen, G., Guillaud, J.F., Garnier, J. and Ménesguen, A. (2005). Modelling the eutrophication of the Seine Bight (France) under historical, present and future riverine nutrient loading, Journal of Hydrology, Special issue on nutrient mobility within river basins. Journal of Hydrology 304: 381-396.

Dafner, E.V., Boscolo, R. and Bryden, H.L. (2003). The N:Si:P molar ratio in the Strait of Gibraltar. Geophysical Research Letters 30: 1 506-1 509.

Dent, C.D., Cumming, G.S. and Carpenter, S.R. (2002). Multiple states in river and lake ecosystems. Philosophical Transactions of the Royal Society Series B: 357: 635-645.

Devlin, M. and Bonne, W. (2016). NEA GIG Intercalibration Report. Phytoplankton element. Common type NEA 1/26 (Available at CIRCA).

Dodds, W.K., Clements, W.H., Gido, K., Hilderbrand, R.H. and King, R.S. (2010). Thresholds, breakpoints, and nonlinearity in freshwaters as related to management. Journal of the North American Benthological Society 29: 988-997.

Dufrêne, M. and Legendre, P. (1997). Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecological Monographs 67: 345-366.

Dworak, T., Berglund, M., Haider, S., Leujak, W. and Claussen, U. (2016). A comparison of European nutrient boundaries for transitional, coastal and marine waters. Working Group on Ecological Status.

European Environment Agency (2012). Nutrients in transitional, coastal and marine waters. https://www.eea.europa.eu/data-and-maps/indicators/nutrients-in-transitional-coastal-and-3/assessment. (date accessed: 2.10.2018)

Ellis, N., Smith, S.J. and Pitcher, C.R. (2012). Gradient forests: calculating importance gradients on physical predictors. Ecology 93:156-168.

Erichsen, A.C., Timmermann, K., Christensen, J.P.A., Kaas, H., Markager, S. and Møhlenberg, F. (2017). Development of models and methods to support the Danish River Basin Management Plans. Scientific documentation. Aarhus University, Department of Bioscience and DHI. https://mfvm.dk/fileadmin/user_upload/MFVM/Miljoe/AU_DHI_Development_of_models_and_methods_to_support_the_establishment_of....pdf. (date accessed: 2.10.2018)

Estrada, M. (1996). Primary production in the Northwestern Mediterranean. Scientia Marina 60:55-64.

European Commission (2009). Guidance document on eutrophication assessment in the context of European water policies — Common implementation strategy for the water framework directive (2000/60/EC) — Guidance document No 23. Office for Official Publications of the European Communities, Luxembourg.

European Commission (2017). Commission Decision (EU) 2017/848 of 17 May 2017 laying down criteria and methodological standards on good environmental status of marine waters and specifications and standardised methods for monitoring and assessment, and repealing Decision 2010/477/EU. Official Journal of the European Union, L 145: 43-74.

European Commission (2018). Commission Decision of 12 February 2018 establishing, pursuant to Directive 2000/60/EC of the European Parliament and of the Council, the values of the Member State monitoring system classifications as a result of the intercalibration exercise and repealing Commission Decision 2013/480/EU. Official Journal of the European Union, L 47: 1-91.

Feld, C.K., Segurado, P. and Gutiérrez-Cánovas, C. (2016). Analysing the impact of multiple stressors in aquatic biomonitoring data: a cookbook with applications in R. Science of the Total Environment 573: 1320-1339.

Fernandes, J.A., Kauppila, P., Uusitalo, L., Fleming-Lehtinen, V., Kuikka, S. and Pitkänen, H. (2012). Evaluation of reaching the targets of the water framework directive in the Gulf of Finland. Environmental Science and Technology 46: 8220-8228.

Ferreol, M. and Bougon, N. (in prep.) Identification des seuils physico-chimique en soutien au bon état écologique des cours d'eau. (Identification of physicochemical thresholds to support good ecological status in rivers)

Fiala, M., Cahet, G., Jacques, G., Neveux, J. and Panousse, M. (1976). Fertilisation de communautés phytoplanktoniques, I. Cas d'un milieu oligotrophe: mediterranée nord-occidentale (Fertilization of phytoplankton communities, I. The case of an oligotrophic environment: North-West Mediterranean). . Journal of Experimental Marine Biology and Ecology 24: 151-163.

Fielding, A. H., and Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24: 38-49.

Flo, E. (2017). Opening the black box of coastal inshore waters in the NW Mediterranean Sea: environmental quality tools and assessment. PhD. 372 pages. https://www.tdx.cat/handle/10803/461378 (date accessed: 2.10.2018)

Free, G., Tierney, D., Little, R., Kelly, F., Kennedy, B., Plant, C., Trodd, W., Wynne, C., Caroni, R. and Byrne, C. (2016). Lake ecological assessment metrics in Ireland: relationships with phosphorus and typology parameters and the implications for setting nutrient standards. Biology and Environment: Proceedings of the Royal Irish Academy 116B: 191-204.

Giovanardi, F., Precalli, R. and France, J. (2016). MED GIG Intercalibration Report. Phytoplankton element (Croatia, Italy and Slovenia working document). Available at CIRCA.

Granéli, E., Wallström, K., Larsson, U., Granéli W. and Elmgren, R. (1990). Nutrient limitation of primary production in the Baltic Sea area. Ambio 19: 142-151.

Greipsland, I., Barneveld, R. and Skarbøvik, E. (2017). *Multiparameteranalyse av feltkarakteristika og vannkjemi i leirvassdrag. Underlag for fastsettelse av miljømål i henhold til vannforskriften* (in Norwegian) (Multiparameter analysis of field characteristics and water chemistry in naturally turbid rivers in catchments with clay sediments. Basis for setting environmental objectives according to the WFD). NIBIO Report 3(110) 2017, 27 pp.

Guillaud, J.-F., Aminot, A., Delmas, D., Gohin, F., Lunven, M., Labry, C., Herbland, A. (2008). Seasonal variation of riverine nutrient inputs in the northern Bay of Biscay (France), and patterns of marine phytoplankton response. Journal of Marine Systems 72: 309-319.

Gunderson, A. R., Armstrong, E. J. and Stillman, J. H. (2016). Multiple stressors in a changing world: the need for an improved perspective on physiological responses to the dynamic marine environment. Annual Review of Marine Science 8: 357-378.

Harper, D.M. (1992). Eutrophication of Freshwaters. Chapman and Hall, London.

Harris, G.P. and Heathwaite, A.L. (2012). Why is achieving good ecological outcomes in rivers so difficult? Freshwater Biology 57: 91-107.

Helcom (2009). Eutrophication in the Baltic Sea — An integrated thematic assessment of the effects of nutrient enrichment and eutrophication in the Baltic Sea region. Baltic Sea Environment Proceedings No 115B.

Helcom (2013). Approaches and methods for eutrophication target setting in the Baltic Sea region. Baltic Sea Environment Proceedings No 133, 138 pp.

Helcom (2017). Total nitrogen. Helcom core indicator report. Online. ISSN 2343-2543. http://www.helcom.fi/baltic-sea-trends/indicators/total-nitrogen-(tn)/contributors-and-references/ (date accessed: 21.8.2017).

Hering, D., Carvalho, L., Argillier, C., Beklioglu, M., Borja, A., Cardoso, A.C., Duel, H., Ferreira, T., Globevnik, L., Hanganu, J., Hellsten, S., Jeppesen, E., Kodeš, V., Solheim, A.L., Nõges, T., Ormerod, S., Panagopoulos, Y., Schmutz, S., Venohr, M. and Birk, S. (2014). Managing aquatic ecosystems and water resources under multiple stress — An introduction to the MARS project. Science of the Total Environment 503-504: 10-21.

Hoepffner N. and Palialexis A. (eds) (2015). Review of the Commission Decision 2010/477/EU concerning MSFD criteria for assessing Good Environmental Status: Descriptor 5. JRC Technical Report (no. 201606230625).

Hurlbert, S.H. (1984). Pseudoreplication and the design of ecological field experiments. Ecological Monographs 54: 187-211.

Ibisch, R., Austnes, K., Borchardt D., Boteler, B., Leujak, W., Lukat, E., Rouillard, J., Schmedtje, U., Lyche-Solheim, A., Westphal, K. (2016). European assessment of eutrophication abatement measures across land-based sources, inland, coastal and marine waters. European Topic Centre on Inland, Coastal and Marine Waters, UFZ, Magdeburg, Germany.

Jacques, G., Cahet, G., Fiala, M. and Panousse, M. (1973). *Enrichissement de communautés phytoplanctoniques néritiques de méditerranée nord occidentale (in French)*. (Enrichment of neritic phytoplankton communities of the north-west Mediterranean)
Journal of Experimental Marine Biology and Ecology 11: 287-295.

Jones, R.I., Salonen, K. and De Haan, H. (1988). Phosphorus transformations in the epilimnion of humic lakes: abiotic interactions between dissolved humic materials and phosphate. Freshwater Biology 19: 357-369.

Karl, D.M. (1999). A sea of change: biogeochemical variability in the North Pacific Subtropical Gyre. Ecosystems 2: 181-214.

Karl D.M. (2000). Phosphorus, the staff of life. Nature (London) 406: 31-32.

Karl, D.M., Hebel, D.V., Björkman, H. and Letelier, R.M. (1998). The role of dissolved organic matter release in the productivity of the oligotrophic North Pacific Ocean. Limnology and Oceanography 43: 1270-1286.

Koenker, R. (2011). Additive models for quantile regression: model selection and confidence bandaids. Brazilian Journal of Probability and Statistics 25: 239-262.

Koenker, R. (2016). Package 'quantreg'.
https://cran.r-project.org/web/packages/quantreg/quantreg.pdf
(date accessed: 15.8.2017).

Krom, M.D., Kress, N. and Benner, S. (1991). Phosphorus limitation of primary productivity in the eastern Mediterranean Sea. Limnology and Oceanography 36: 424-432.

Labry, C., Herbland, A. and Delmas, D. (2002). The role of phosphorus on planktonic production of the Gironde plume waters in the Bay of Biscay. Journal of Plankton Research 24: 97-117.

Lazzari, P., Solidoro, C., Salon, S. and Bolon, G. (2016). Spatial variability of phosphate and nitrate in the Mediterranean Sea: a modeling approach. Deep Sea Research Part 1: Oceanographic Research Papers 108: 39-52.

LAWA (2016). *Empfehlung zur Übertragung flussbürtiger, meeresökologischer Reduzierungsziele ins Binnenland* (in German) (Recommendation on the transfer of river-borne, marine biological reduction targets inland). LAWA AO, 19pp.

Legendre, P. (2013). Model II regression user's guide, R edition. (A tutorial within R session, R package lmode2) http://cran.r-project.org/web/packages/lmodel2/vignettes/mod2user.pdf. (date accessed: 7.7.2015).

Legendre, P. and Legendre, L.F. (2012). Numerical ecology. Elsevier.

Lignell, R., Seppälä, J., Kuuppo, P., Tamminen, T., Andersen T. and Gismervik, I. (2003). Beyond bulk properties: Responses of coastal summer plankton communities to nutrient enrichment in the Northern Baltic Sea. Limnology & Oceanography 48: 189-209.

Lyche-Solheim, A., Persson, J., Austnes, K., Moe, J., Kampa, E., Stein, U., Feher, J., Kristensen, P. (2015). Freshwater Ecosystem Assessment: Cross-walk between the water framework directive and habitats directive types, status and pressures, ETC/ICM technical report 2/2015. European Topic Centre on Inland, Coastal and Marine Waters, Magdenurg.
https://icm.eionet.europa.eu/ETC_Reports/FreshwaterEcosystemAssessmentReport_201509 (date accessed: 2.10.2018).

Maberly, S., King, L., Dent, M., Jones, R. and Gibson, C. (2002). Nutrient limitation of phytoplankton and periphyton growth in upland lakes. Freshwater Biology 47: 2 136-2 152.

Maberley, S.C., King, L., Gibson, C.E., May, L., Jones, R.I., Dent. M.M. and Jordan, C. (2003). Linking nutrient limitation and water chemistry in upland lakes to catchment characteristics. Hydrobiologia 506-509: 83-91.

Malagó, A., Venohr, M., Gericke, A., Vigiak, O., Bouraoui, F., Grizzetti, B. and Kovacs, A. (2015). Modelling nutrient pollution in the Danube River Basin: a comparative study of SWAT, MONERIS and GREEN models. JRC Technical Report EUR 27676 EN, doi:10.2788/156278.

Margalef, R. (1963). *El ecosistema pelágico de un área costera del Mediterráneo occidental.* (in Spanish). (The pelagic ecosystem of a coastal area of the western Mediterranean) *Memorias de la Real Academia de Ciencias y Artes de Barcelona* 35: 1-48

McArdle, B. H. (2003). Lines, Models, and Errors: Regression in the Field. Limnology and Oceanography 48: 1363-1366.

Mainstone, C. P. (2010). An evidence base for setting nutrient targets to protect river habitat http://naturalengland.etraderstores.com/NaturalEnglandShop/NERR034
(date accessed: 12.11.2010).

Moss, B., Jeppesen, E., Søndergaard, M., Lauridsen, T.L. and Liu, Z. (2013). Nitrogen, macrophytes, shallow lakes and nutrient limitation: resolution of a current controversy? Hydrobiologia 710: 3-21.

Müller, A., Halle, M. and Bellack, E. (2017). *Schwellenwerte und biologische Indikatoren für physikalisch-chemische Parameter in Fließgewässern*. (Threshold values and biological indicators for physicochemical parameters in running waters) Wasser und Abfell 06/2017: 25-30.

Nausch, M., Nausch, G. and Wasmund, N. (2004). Phosphorus dynamics during the transition from nitrogen to phosphate limitation in the central Baltic Sea. Marine Ecology Progress Series 266: 15-25.

Nõges, P., Argillier, C., Borja, Á., Garmendia, J.M., Hanganu, J., Kodeš, V., Pletterbauer, F., Sagouis, A. and Birk, S. (2016). Quantified biotic and abiotic responses to multiple stress in freshwater, marine and ground waters. Science of the Total Environment 540: 43-52.

Ockenden, M., Deasy, C.E., Benskin, C., Beven, K.J., Burke, S., Collins, A.L., Evans, R., Falloon, P.D., Forber, K.J., Hiscock, K.M., Hollaway, M.J., Kahana, R., Macleod, C.J.A., Reaney, S.M., Snell, M., Villamizar, M., Withers, P., Zhou, J. and Haygarth, P.M. (2016). Changing climate and nutrient transfers: evidence from high temporal resolution concentration-flow dynamics in headwater catchments. Science of the Total Environment 548-549: 325-339.

Officer, C.B and Ryther, J.H. (1980). The possible importance of Silicon in marine eutrophication. Marine Ecology Progress Series 3: 83-91.

OSPAR (2009). Second Integrated Report on the Eutrophication Status of the OSPAR Maritime Area. Oslo-Paris Convention, London.

OSPAR (2013). Common Procedure for the Identification of the Eutrophication Status of the OSPAR Maritime Area. Oslo-Paris Convention, London, Reference number: 2013-8, 66 pp.

Page, T., Heathwaite, A.L., Moss, B., Reynolds, C., Beven, K.J., Pope, L. and Willows, R. (2012). Managing the impacts of nutrient enrichment on river systems: dealing with complex uncertainties in risk analysis. Freshwater Biology 57: 108-123.

Palialexis, A., Hoepffner, N., Aigars, J., Axe, P., Costea, F., Fleming-Lehtinen, V., Giovanardi, F., Hart, V., Leujak, W., Magaletti, E., Malcolm, S., Poikane, S., Poje, M., Sever, M., Pons, C., Precali, R., Ruiter, H., Van De Bund, W., Van Den Berg, M., Yebra Mora, L., Ysiak-Pastuszak, E. and Devreker, D. (2016). Report of the JRC's Descriptor 5 workshop to support the review of the Commission Decision 2010/477/EU concerning MSFD criteria for assessing Good Environmental Status, JRC Technical Report EUR 27814 EN, Publications Office of the European Union, Luxembourg, doi: 10.2788/130847.

Passy, P., Le Gendre, R., Garnier, J., Cugier, P., Callens, J., Paris, F., Billen, G., Riou, P. and Romero, E. (2016). Eutrophication modelling chain for improved management strategies to prevent algal blooms in the Bay of Seine. Marine Ecology Progress Series 543: 107-125.

Paul, J.F. and McDonald, M.E. (2005). Development of empirical, geographically specific water quality criteria: a conditional probability analysis approach. Journal of the American Water Resources Association 41: 1 211-1 223.

Penning, W.E., Mjelde, M., Dudley, B., Hellsten, S., Hanganu, J., Kolada, A., van den Berg, M., Poikane, S., Phillips, G., Willby, N. and Ecke, F. (2008). Classifying aquatic macrophytes as indicators of eutrophication in European lakes. Aquatic Ecology 42: 237-251.

Perrot, T., Rossi, N., Menesguen, A. and Dumas, F. (2014). Modelling green macroalgal blooms on the coasts of Brittany, France to enhance water quality management. Journal of Marine Systems 132: 38-53.

Phillips, G., Pietiläinen, O.P., Carvalho, L., Solimini, A., Lyche Solheim, A. and Cardoso, A. (2008). Chlorophyll–nutrient relationships of different lake types using a large European dataset. Aquatic Ecology 42: 213-226.

Phillips, G., Birk, S., Bohmer, J., Kelly, M. and Willby, N. (2016). The use of pressure response relationships between nutrients and biological quality elements as a method for establishing nutrient supporting element boundary values for the water framework directive. Unpublished report for DG Environment.

Phillips, G. and Pitt, J. (2015). A comparison of European freshwater nutrient boundaries used for the water framework directive: Report to Ecostat, October 2015. https://circabc.europa.eu/w/browse/58a2363a-c5f1-442f-89aa-5cec96ba52d7 (date accessed: 2.10.2018)

Piggott, J.J., Townsend, C.R. and Matthaei, C.D. (2015). Reconceptualizing synergism and antagonism among multiple stressors. Ecology and Evolution 5: 1538-1547.

Pinay, G., Gascuel, C., Ménesguen, A., Souchon, S., Le Moal, M., Levain, A., Etrillard, C, Moatar, F., Pannard, A. and Souchu, P. (2017). *L'eutrophisation: manifestations, causes, conséquences et* *prédictibilité* (Eutrophication: manifestations, causes, consequences and prediction). *Synthèse de l'Expertise scientifique collective* CNRS — Ifremer — INRA — Irstea (France), 148 pp.

Pitkänen, H. and Tamminen, T. (1995). Nitrogen and phosphorus as production limiting factors in the estuarine waters of the eastern Gulf of Finland. Marine Ecology Progress Series 129: 283-294.

Poikane, S., Birk, S., Böhmer, J., Carvalho, L., de Hoyos, C., Gassner, H., Hellsten, S., Kelly, M., Solheim, A.L., Olin, M., Pall, K., Phillips, G., Portielje, R., Ritterbusch, D., Sandin, L., Schartau, A.-K., Solimini, A.G., van den Berg, M., Wolfram, G. and van de Bund, W. (2015). A hitchhiker's guide to European lake ecological assessment and intercalibration. Ecological Indicators 52: 533-544.

Primpas, I., Tsirtsis, G., Karydis, M. and Kokkoris, G.D. (2010). Principal component analysis: Development of a multivariate index for assessing eutrophication according to the European water framework directive. Ecological Indicators 10: 178-183.

Qian, S.S., King, R.S. and Richardson, C.J. (2003). Two statistical methods for the detection of environmental thresholds. Ecological Modelling 166: 87-97.

R Development Core Team (2006). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org

Ramírez, T., Cortés, D., Mercado, J.M., Vargas-Yañez, M., Sebastián, M. and Liger, E. (2005). Seasonal and interannual dynamics of DIN and DIP in the NW Alboran Sea. Geophysical Research Abstracts 7: 04315.

Rinaldi, A. (2014). Fioriturі algali In Adriatico. Il bacino padano-adriatico tra sviluppo e scienza (Algal blooms in the Adriatic. The Padano-Adriatic basin between development and science) Editrice La Mandragora.

Roubeix, V., Danis, P.-A., Feret, T. and Baudoin, J.-M. (2016). Identification of ecological thresholds from variations in phytoplankton communities among lakes: contribution to the definition of environmental standards. Environmental Monitoring and Assessment 188: 246.

Roubeix, V., Daufresne, M., Argillier, C., Dublon, J., Maire, A., Nicolas, D., Raymond, J.-C. and Danis, P.-A. (2017). Physico-chemical thresholds in the distribution of fish species among French lakes. Knowledge and Management of Aquatic Ecosystems 418: 41.

Schernewski, G., Friedland, R., Carstens, M., Hirt, U., Leujak, W., Nauscha, G., Neumann, T., Petenati, T., Sagert, S., Wasmund, N. and von Weber, M. (2015). Implementation of European marine policy: New water quality targets for German Baltic waters. Marine Policy 51: 305-321.

Schindler, D.W. (2012). The dilemma of controlling cultural eutrophication in lakes. Proceedings of the Royal Society of London Series B 279: 4322-4333.

Schindler, D.W., Hecky, R.W., Findlay, D.W., Stainton, N.P., Parker, B.R., Paterson, M.J., Beaty, K.G., Lyng, M. and Kaisan, S.E.M. (2009). Eutrophication of lakes cannot be controlled by reducing nitrogen input: Results of a 37-year whole-ecosystem experiment. Proceedings of the National Academy of Science 105: 11254-11258.

Schinegger, R., Schülting, L., Schmutz, S., Solheim, A.L., Birk, S., Feld, C., Hering, D., Kuijper, M., Chrzanowski, C., Buijse, T., Borja, A., Venohr, M., Globevnik, L. (2018). MARS Recommendations on how to best assess and mitigate impacts of multiple stressors in aquatic ecosystems. http://www.mars-project.eu/files/download/recommendations/MARS_Recommendations.pdf

Schneider, S., Hilt, S., Vermaat, J.E. and Kelly, M. (2016). The forgotten ecology behind ecological status evaluation: re-assessing the roles of aquatic plants and benthic algae in ecosystem functioning. Progress in Botany 78: 285-304.

Smith, R. J. (2009). Use and misuse of the reduced major axis for line-fitting. American Journal of Physical Anthropology 140: 476-86.

Sokal, R. R. and Rohlf, F. J. (1995). Biometry: the principles and practice of statistics in biological research. WH Freeman & Co, San Francisco.

Sundermann, A., Leps, M., Leisner, S. and Haase, P. (2015). Taxon-specific physic-chemical change points for stream benthic invertebrates. Ecological indicators 57: 314-323.

Tamminen, T. and Andersen, T. (2007). Seasonal phytoplankton nutrient limitation patterns as revealed by bioassays over Baltic Sea gradients of salinity and eutrophication. Marine Ecology Progress Series 340: 121-138.

Teixeira, H. and Salas, F. (2017). The use of pressure response relationships between nutrients and biological quality elements as a method for establishing nutrient supporting element boundary values for the water framework directive: Coastal and transitional waters. (available as Appendix 4 of this CIS Guidance document)

Thingstad, T.F., Krom, M.D., Mantoura, R.F.C., Flaten, G.A.F., Groom, S., Herut, B., Kress, N., Law, C.S., Pasternak, A., Pitta, P., Psarra, S., Rassoulzadegan, F., Tanaka, T., Tselepides, A., Wassmann, P., Woodward, E. M. S., Wexels Riser, C., Zodiatis, G. and Zohary T. (2005). Nature of phosphorus limitation in the ultraoligotrophic Eastern Mediterranean. Science (New York) 309: 1 068-1 071.

Tsirtis, G.E. (1995). A simulation model for the description of a eutrophic system with emphasis on the microbial processes. Water Science and Technology 32: 189-196.

Tyrrell, T. (1999). The relative influences of nitrogen and phosphorus on oceanic primary production. Nature (London) 400: 525-531.

UK TAG (2013). A revised approach to setting water framework directive phosphorus standards. UKTAG; October 2012.
http://www.wfduk.org/resources%20/revised-approach-setting-wfd-phosphorus-standards
(date accessed: 2.10.2018).

Vinebrooke, R.D., Cottingham, K.L., Norberg, J., Scheffer, M., Dodson, S.I., Maberly, S.C. and Sommer, U. (2004). Impacts of multiple stressors on biodiversity and ecosystem functioning: the role of species co-tolerance. Oikos 104: 451-457.

Vollenweider, R.A. (1968). Water Management Research; Scientific Fundamentals of Eutrophication of Lakes and Flowing Waters, with particular reference to Nitrogen and Phosphorus as factors in Eutrophication. Technical Report DAS/CSI/68.27, Organisation for Economic Cooperation and Development, Paris.

Vollenweider, R.A., Giovanardi, F., Montanari, G. and Rinaldi, A. (1998). Characterization of the trophic conditions of marine coastal waters with special reference to the NW Adriatic Sea: Proposal for a trophic scale, turbidity and generalised water quality index. Environmetrics 9, 329-357.

Wagenhoff, A., Townsend, C.R., Phillips, N. and Matthaei, C.D. (2011). Subsidy-stress and multiple-stressor effects along gradients of deposited fine sediment and dissolved nutrients in a regional set of streams and rivers. Freshwater Biology 56: 1916-1936.

Whitton, B.A. and Neal, C. (2010). Organic phosphate in UK rivers and its relevance to algal and bryophyte surveys. Annales de Limnology 47: 1-8.

Winfield, I.J., Fletcher, J.M. and James, J.B. (2008). The Arctic Charr (*Salvelinus alpinus*) populations of Windermere, UK: population trends associated with eutrophication, climate change and increased abundance of roach (*Rutilus rutilus*). Environmental Biology of Fish 83: 25-35.

Zohary T. and Robarts, R.D. (1998). Experimental study of microbial P limitation in the eastern Mediterranean. Limnology and Oceanography 43: 387-395.

Annex: List of contributors

8.1. Steering group

Ulrich Claussen	Federal Environment Agency (UBA), Dessau-Roßlau, Germany
Anne Lyche Solheim	NIVA, Norway
Jo-Anne Pitt	Environment Agency, UK
Sandra Poikane	European Commission, Joint Research Centre, Sustainable Resources Directorate
Fuensanta Salas	European Commission, Joint Research Centre, Sustainable Resources Directorate
Marcel van den Berg	Rijkswaterstaat Water, Verkeer & Leefomgeving, Netherlands

8.2. Drafting group, development of statistical toolkit and Shiny app

Martyn Kelly	Bowburn Consultancy, UK
Wera Leujak	Federal Environment Agency (UBA), Dessau-Roßlau, Germany
Geoff Phillips	University of Stirling, UK
Heliana Teixeira	University of Aveiro, Portugal
Gábor Várbíró	Danube Research Institute, Hungary

8.3. Others who have contributed text and information

Mike Best	Environment Agency, UK
Nolwenn Bougon	Agence Française pour la Biodiversité
Anne Daniel	IFREMER, France
Karin Deutsch	Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft, Austria
Eva Flo	Institut de Ciències del Mar, CSIC, Barcelona, Catalunya, Spain
Gary Free	Environmental Protection Agency, Ireland
Martin Halle	Umweltbüro Essen, Germany
Theis Bødker Jensen	Ministry of Environment and Food of Denmark
Theo Prins	Deltares, Netherlands