Adjustment of the Netherlands WFD-assessment methods for fish in lakes

Contents

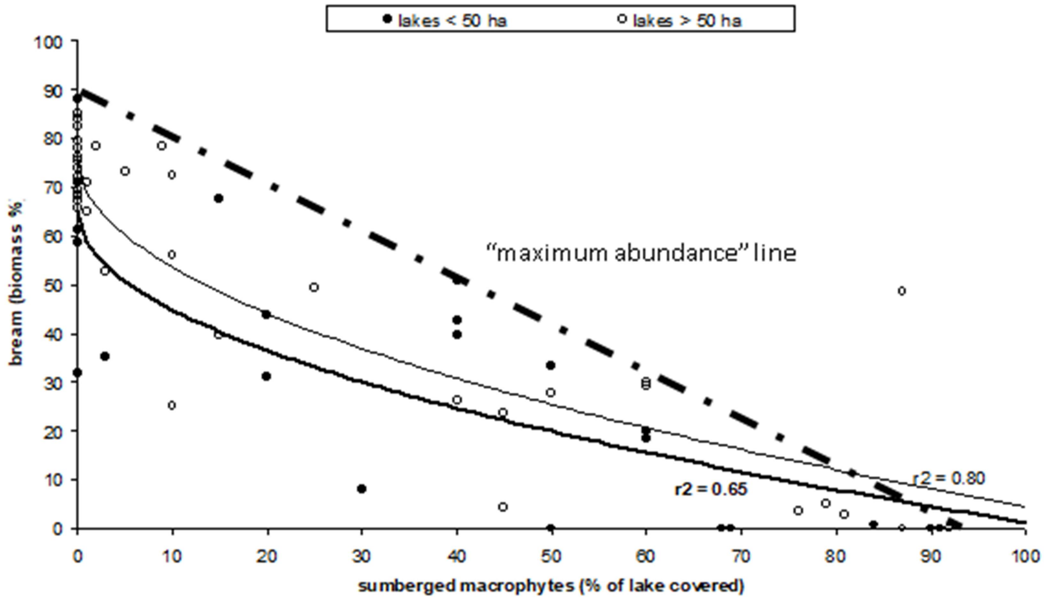
1. Introduction

In 2017, the Dutch EU Water Framework Directive (WFD)-assessment method for fish in lakes was intercalibrated with the methods of the Czech Republic, Denmark, Germany, Estonia, Lithuania and Poland. The methods of Belgium, the United Kingdom and Latvia could not yet be intercalibrated because they were not yet ready or did not meet the basic requirements. Intercalibration for fish was not an easy process, but was successfully completed by comparing the standards of the Member States on the basis of a common metric based on 'human pressures'. Completing the intercalibration procedure means that adjustments to the current (applicable) standards or new standards must be tested against the requirements of the EU-intercalibration, in accordance with EU-WFD guidance document no. 30 (European Union, 2015).

The Dutch WFD assessment methods for fish in lakes were developed in the period 2003-2005, with the knowledge and data available at the time (Klinge, *et. al.,* 2004, Jaarsma *et. al.,* 2007). Since then, the methods have remained largely unchanged. The most important modification that has been implemented is the removal of the metric "number of species", which turned out to be very sensitive for sampling effort (N.G. Jaarsma & M. Klinge, 2012). There has been relatively little discussion about the choice of the remaining four metrics that are used to assess the lake fish community for most lakes. For one specific lake (type) a fifth metric has been developed to assesses the fishery intensity. See chapter 3 for a short overview of the Dutch assessment system for fish in lakes.

Although the metric itself was broadly accepted, the class boundaries of the metric "% weight of bream" were subject to debate, especially for the "good" and higher (reference) classes but also for the lower classes. The class boundaries were derived based on the available data and knowledge at the time, but were generally found to be very / too strict. This can partly be explained by the fact that the class boundaries were based on the assumption of complete cover (100%) of a lake with submerse vegetation under reference conditions. However, this is debatable since the Dutch WFD assessment method for macrophytes in lakes is based on a coverage of submerse vegetation of 45 to 100% in the "very good" class and 25 to 45% in the "good" class (van der Molen *et al.,* 2016). This means that the assumed reference situation for fish differs substantially from that of the macrophytes. A higher proportion of bream is appropriate for such a situation, see Figure 1.

Figure 1. Relationship between the coverage of the lake area with submerse vegetation and the relative biomass of bream, as was used at the time of deriving class boundaries for the metric.



Therefore, an evaluation of the class boundaries for the "% weight of bream" was considered necessary. More in general; the far greater availability of good quality data made it possible to carry out a more extensive evaluation of the methods for fish in lakes. This was also considered desirable. The evaluation has focused on the two metrics in the Dutch system that are mainly associated with eutrophication ("% weight of bream" and "% weight of roach+perch to all eurytopic species").

2. Methods

The main approach for the study consisted of the following steps:

- Collecting an up to date dataset for fish in Dutch lakes and collecting a dataset with human pressures for these locations;
- Analysing the fish stocks in relation to pressures and evaluation of the current assessment methods;
- Proposal for adaptions to the current assessment methods;
- Selection of new metrics and class boundaries for the new assessment methods, based on the relations to pressures and the requirements of the intercalibration;
- Comparing the new EQRs to the old EQRs and consultation of fish experts from the water authorities.

The study was performed in 2018. An extensive description of the project, corresponding methods and results is reported in Jaarsma & Klinge (2018).

3. Selection of indicators and relation with pressures

The current (2012) Dutch assessment methods for fish in lakes consist of four metrics, as described in Ritterbusch *et. al.* (2017a). Each of these metrics is indicative for a specific pressure. The metrics and pressures are (% W = weight percentage of total fish stock):

- % W Abramis brama: eutrophication
- % W (roach+perch)/eurytopic species: eutrophication
- % W phytophilic: shoreline degradation/water level regulation
- % W low oxygen tolerant: shoreline degradation/water level regulation

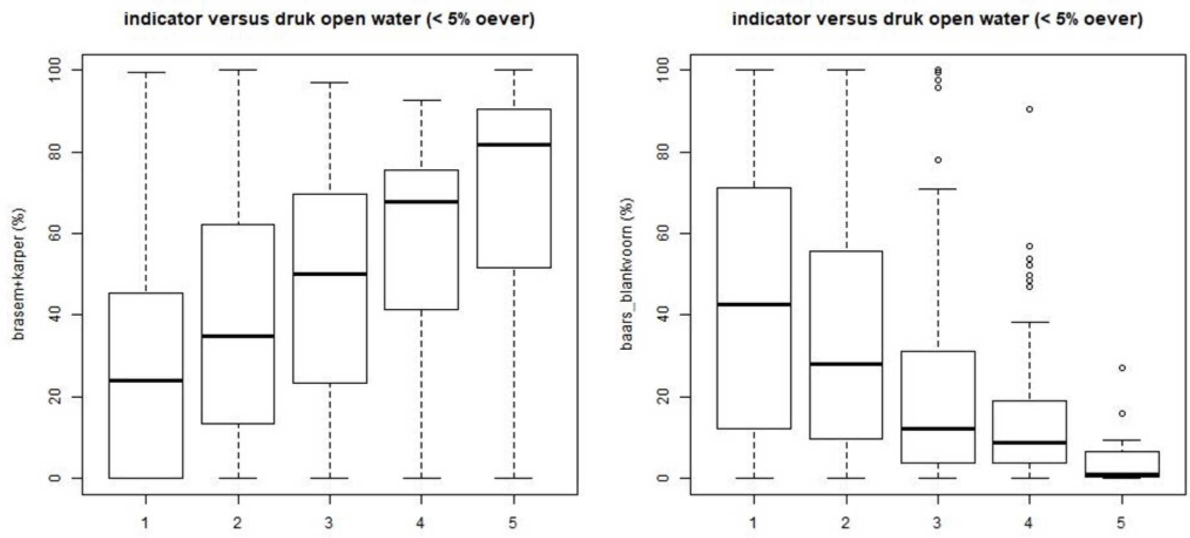
Only for one specific type of lakes (lakes larger than 10.000 ha) also fishery intensity is assessed based on the population structure of pikeperch:

- % W pikeperch > legal length limit: fishery intensity

In the proposed new assessment methods only the first metric (% W Abramis brama) is changed and replaced by (% W Abramis brama + Cyprinus carpio). This was done because data analysis showed that the relative abundance of bream + carp has a better relation with eutrophication than bream alone. For this metric and the other metric for eutrophication (% W (roach+perch)/eurytopic species) new class boundaries have been derived, based on the relation with pressures (eutrophication) in the dataset.

In order to assess the level of eutrophication, a pressure index has been constructed similar to the one that has been developed in the intercalibration exercise (Ritterbusch *et. al.* 2017b). The index is based on total-P, total-N, chlorophyll-a and secchi disk depth of the lakes in the dataset. The index has five classes (1=high, 2=good, 3=moderate, 4=poor and 5=bad), based on the class boundaries of these parameters in the Dutch WFD classification system (boundaries differ between lake types). Figure 2 shows the range of values for both metrics under evaluation, for each of the pressure classes.

Figure 2. Box-plot of the percentage of bream + carp (left) and the percentage of (roach+perch)/eurytopic species in relation to the pressure index (1=high, 2=good, 3=moderate, 4=poor and 5=bad), based on the data of Dutch lakes > 50 ha.



4. Intercalibration requirements

A modified assessment method, as proposed in this case, has to be tested against the intercalibration-requirements following rules in the EU-guidance (European Union, 2015). This guidance applies to two specific situations:

a) completely new methods which have not previously been subject to intercalibration but for which an exercise has already been completed for the BQE and GIG in question, and
b) methods which were part of a completed exercise but have since been revised in some way.

The latter (b) applies (the Dutch method was revised), in this case the requirements are in short:

1. check correlation between “old” and “new” method. If r^2 >= 0.8 check boundaries (next step), if r^2 < 0.8 then carry out feasibility check:
    a. to qualify for intercalibration the revised method should have a correlation (r) with the common metric >= 0.5 (r^2 >= 0.25);
    b. the slope of the regression (revised method vs common metric) should be between 0.5 and 1.5.
2. check if the boundaries H-G and G-M of the revised method are higher (i.e. more precautionary) than the old method. If this is true, the intercalibration “fitting procedure” is accomplished. If not, follow the procedure for fitting new classification methods.

In our case the correlation of the revised method with the old method was high enough (r^2 > 0.8), but the class boundaries for H-G and G-M were less strict than those of the already intercalibrated method. This means that the procedure for fitting new classification methods (a) had to be applied. The procedure for fitting new classification methods uses the following criteria:

- sufficient correlation (r^2> = 0.25) of the new metric with the intercalibration metric (common metric);
- sufficiently distinctive: slope of the regression line with the intercalibration metric (common metric) between 0.5 and 1.5
- not too lenient: class boundaries for good-very good and for moderate-good must lie within a specified bandwidth, based on the relation with the common metric (“boundary bias” check).

The proposed adjustments have been (iteratively) tested with the dataset used for the intercalibration (Ritterbusch *et. al.,* 2017b). In the next paragraphs it is shown that:

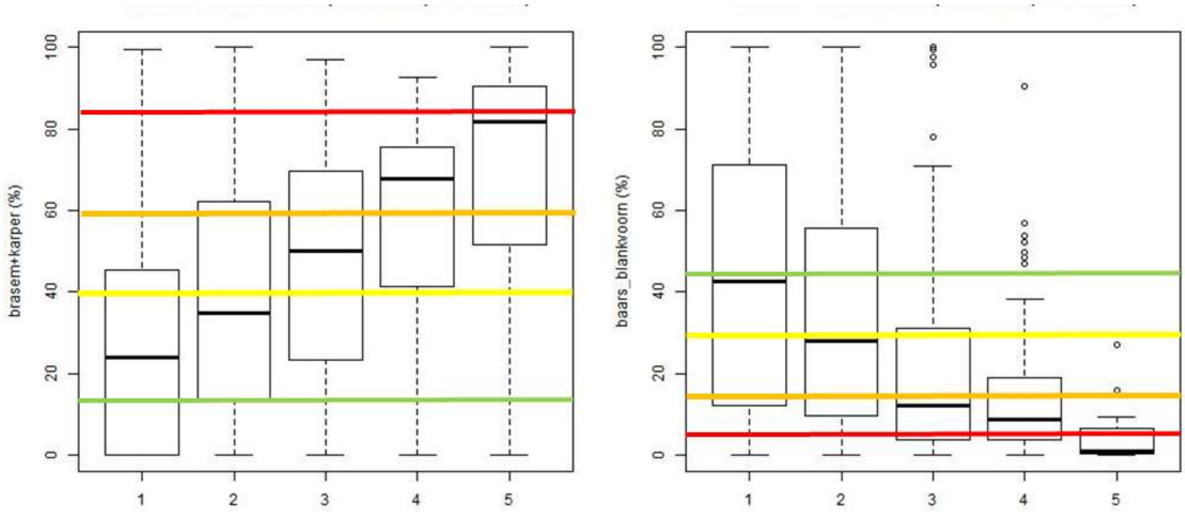
- both the correlation with the common metric (TAPI) and the slope of the regression line amply meet the intercalibration requirements (see 4.2);
- the joint effect of the proposed changes falls within the requirements imposed on the "boundary bias" (see 4.1 and 4.3).

4.1. Class boundaries

An important step in the intercalibration procedure is the harmonization of class-boundaries. Therefore the proposed class boundaries have to be tested against the intercalibration requirements, by again going through the boundary harmonization process that was originally carried out during the intercalibration. This shows whether the proposed class-boundaries of the new assessment methods fit within the determined range of the common metric (+/- 0,25 class width). This is done by calculating the ‘boundary bias’, which is the deviation of the proposed class boundaries from the average class boundaries on the common metric of all member states participating in the intercalibration.

The class boundaries we propose here were derived iteratively; they were initially based on the relation with the pressures (Figure 2) and then tested against the requirements of the EU-intercalibration. This showed that the H-G and G-M boundaries for bream+carp had to be slightly adapted (made stricter) to be able to meet the EU-WFD requirements. For practical reasons, we adjusted the class boundaries of the indicators rather than the EQRs for the H-G and G-M boundaries. In the Netherlands, at all times these boundaries are at EQR = 0.8 and 0.6 respectively.

Figure 3. Box-plot of the percentage of bream + carp (left) and the percentage of (roach+perch)/eurytopic species in relation to the pressure index (1=high, 2=good, 3=moderate, 4=poor and 5=bad), based on the data of Dutch lakes. In the figure the proposed class boundaries are shown, where green = boundary H-G, yellow=G-M, orange=M-P and red =P-B.



4.2. Relation with common pressure-index (TAPI)

During the intercalibration exercise, a data set was compiled by the participating member states and a multi-pressure index (TAPI) was developed. In 2017, the fish metrics for freshwater lakes of 7 member states were successfully intercalibrated by comparing EQR scores with the common multi-pressure index (Poikane, *et. al.,* 2017). The multi-pressure index is based on indicators for eutrophication (indicators: total-P and chlorophyll-a) and habitat degradation (indicators: shore modification, habitat loss and lake use intensity). The structure of the index is explained in Table 4-1.

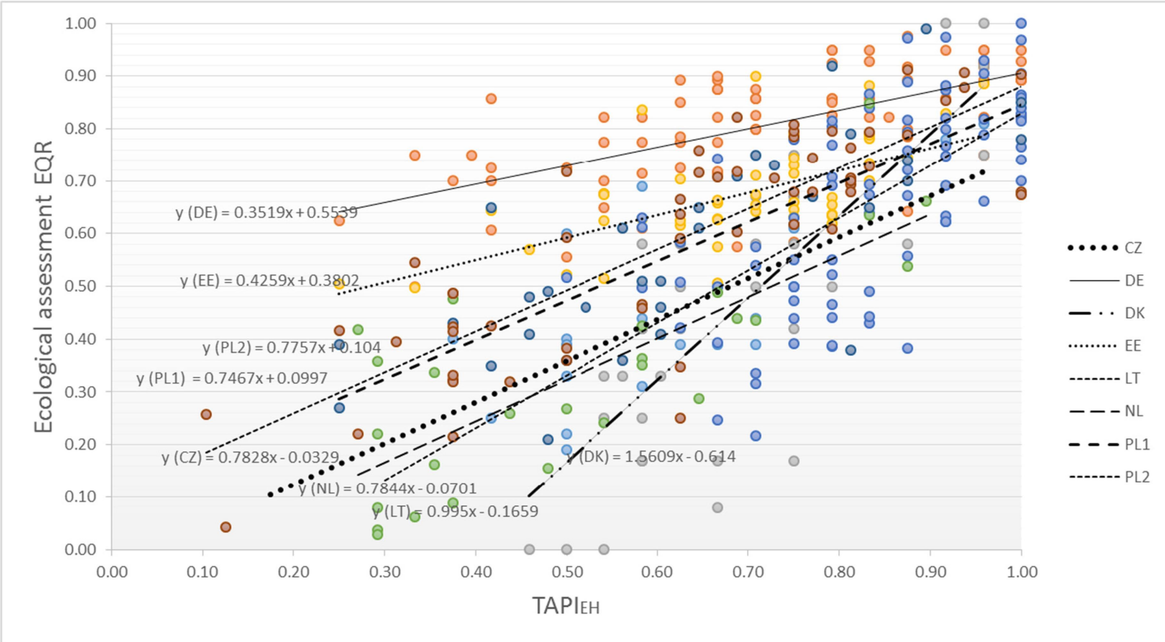
Table 4-1. TAPI index for the intercalibration of fish in lakes within the Central Baltic Intercalibration Group (Poikane, *et. al.,* 2017).

Scoring criteria for TAPI metrics (for other metrics see Tables S2 and S3, Supporting information). P – polymictic lakes, S – stratified lakes, D – deep stratified lakes with max depth > 30 m.

	TAPI metric	5 points least disturbed	4 points minor impact	3 points major impact	2 points strong impact	1 point extreme impact
50 %	Eutrophication					
	Chl-*a* ($\mu g\ L^{-1}$)	<11 (P) <6 (D, S)	11–21 (P) 6–10 (D, S)	21–52 (P) 10–26 (D, S)	52–215 (P) 26–104 (D, S)	>215 (P) >104 (D, S)
	TP spring	<32 (P)	32–45 (P)	45–100 (P)	100–200 (P)	>200 (P)
	TP summer ($\mu g\ L^{-1}$)	<25 (D, S)	25–32 (D, S)	32–45 (D, S)	45–100 (D, S)	>100 (D, S)
50 %	Hydromorphological alterations and lake use					
	Shore modification	≤10%	11–30%	31–50%	51–70%	>70%
	Habitat loss	Natural/increased	All habitats	1–3 habitats missing	4–6 habitats missing	>6 habitats missing
	Lake use intensity	Low (bath, boat, sail)	–	Intense (motorboat, ships, dive)	–	Very intense

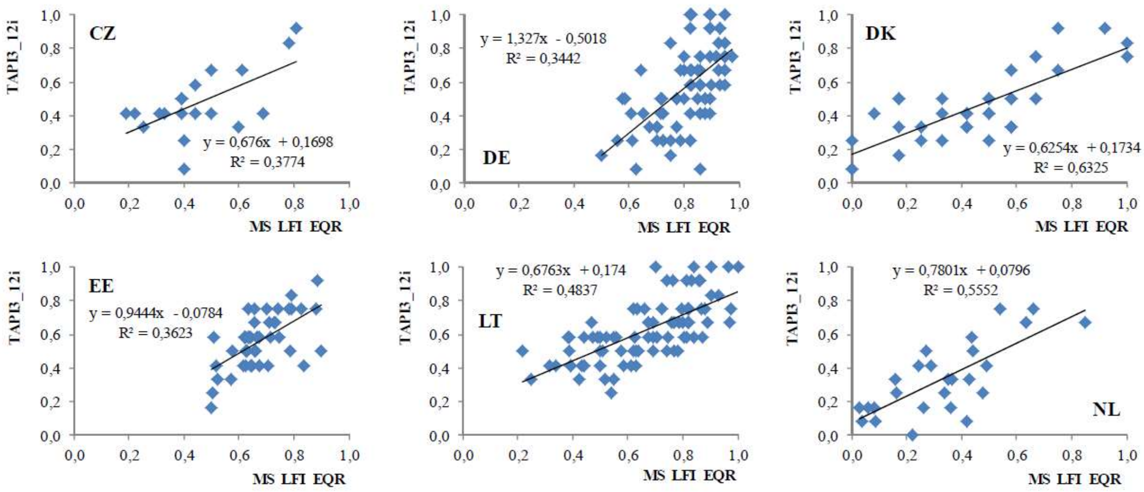
It appeared that the Dutch methods showed a very good correlation with the developed multi-pressure index (TAPI), but that the Dutch EQR score was systematically lower than that of most other member states in the relevant intercalibration group (see Figure 4). This was corrected in the intercalibration exercise, by adjusting class boundaries for MS that were too lenient.

Figure 4. Relation between the member states EQR-values with the common multi-pressure index (Poikane, et. al., 2017).



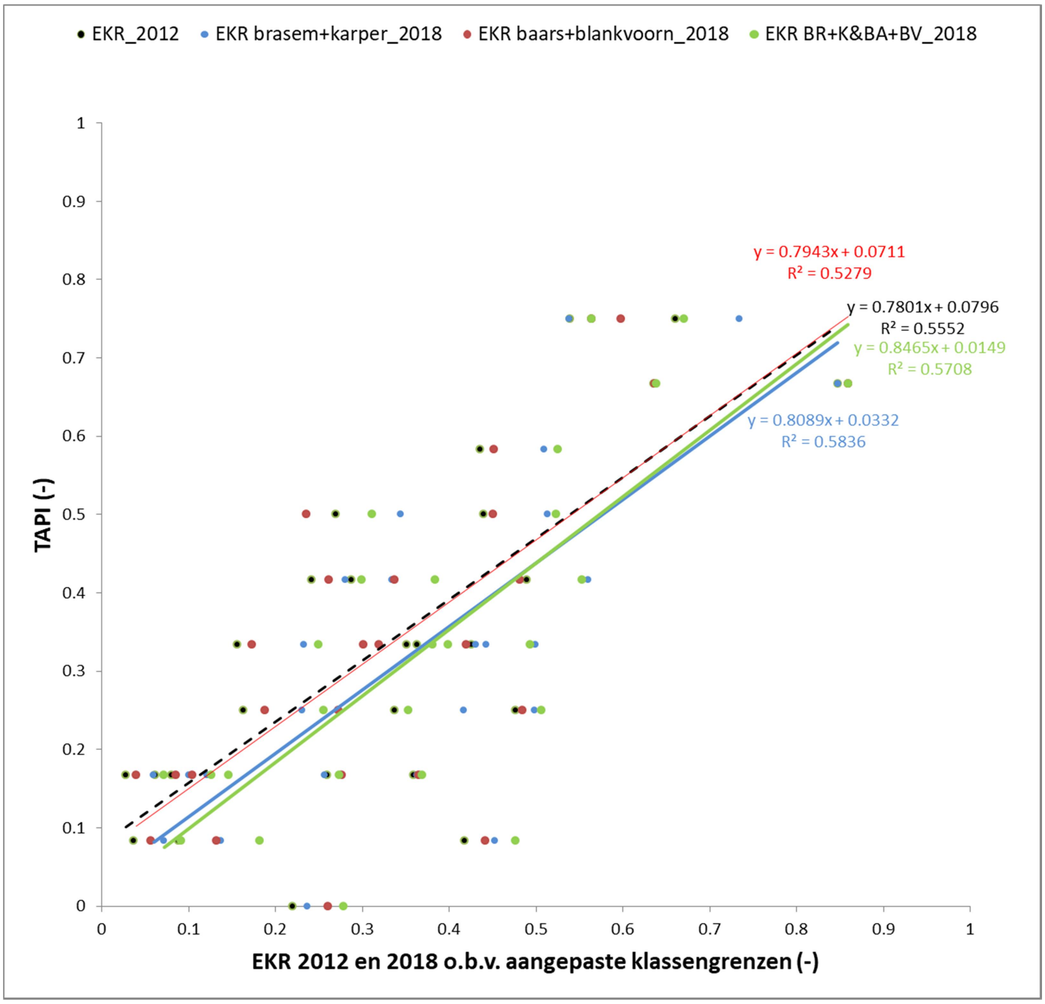
For each of the member states the TAPI scores were plotted against the EQRs and tested against the criteria as mentioned above (regression line: $r^2 > 0.25$ and slope between 0.5 and 1.5). Figure 5 shows the member states for which the acceptance criteria for intercalibration are met.

Figure 5. Regression lines for the score of national measures (x-axis) and the score of the TAPI index (y-axis) of the member states that meet the acceptance criteria for intercalibration (Ritterbusch*, et al.,* 2017).



The EQR-values for the proposed adjustments of the Dutch methods have been calculated and tested against the acceptance requirements for intercalibration, using the intercalibration data set (see Figure 6). This shows that, in broad terms, not much has changed compared to the current (2012) methods. The correlation with the TAPI is even slightly higher for the adjusted methods and the slope of the regression line is also somewhat higher (closer to 1). This applies both to the individual adjustments and to both adjustments combined. In that respect, the adjusted methods even show a better fit with the TAPI.

Figure 6. Regression lines for the EQR score of Dutch method and proposed adjustments * and the TAPI index.



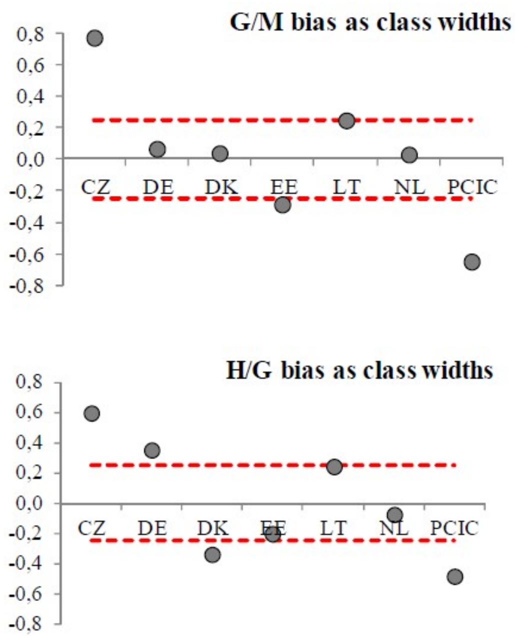
* Explanation of abbreviations:

- EKR_2012 = current method.
- EKR brasem+karper_2018 = metric % W bream replaced by bream+carp and class boundaries adjusted.
- EKR baars+blankvoorn_2018 = class boundaries adjusted for the metric % W (roach+perch)/eurytopic species.
- EKR BR+K&BA+BV_2018 = both adjustments combined.

4.3. Boundary bias

The boundary bias check has to assess whether the class boundaries meet the IC-criteria. For this purpose, a procedure has been developed for intercalibration in which the boundary bias is determined. This is the deviation of the class boundaries from the very good-good and good-moderate classes of a member state, compared to the "average" of all participating member states. The boundary bias is expressed as a fraction of the class width. Figure 7 shows the boundary bias (initial situation, before adaptation of class boundaries) for the member states that met the requirements for acceptance. For each member state it is indicated how the boundaries between good and moderate and between very good and good differ from the average, expressed in class widths. The figure shows that in both cases the Netherlands is close to the average. A positive value means that the method is on the strict side, a negative on the flexible side. The allowed bandwidth is +/- 0.25 class width.

Figure 7. Boundary bias expressed in class widths, initial situation, for those Member States that met the acceptance criteria for intercalibration (Ritterbusch, et. al., 2017b).



Adjustments to the methods must be tested against the criteria for the boundary bias again. For this, the initial situation must be assumed (i.e. the situation as in Figure 7). The reference (the 0 line in Figure 7) is determined by the average of the TAPI scores, corresponding to the boundaries between good and moderate and very good and good for the methods of all Member States. The new (proposed) class boundaries are set against this. For this purpose, the average TAPI score for the class boundaries was taken from the original boundary bias determination, and the boundary bias was recalculated for both the current (2012) method and the proposed adjustments. Table 4-2 and Figure 8 show the results.

Firstly, the calculated boundary bias was checked for the current (2012) method, which exactly matches the value from the intercalibration report (Ritterbusch, et. al., 2017), which shows that both the data used and the calculation method are the same. The boundary bias was then calculated for the proposed changes, for each metric separately and for both combined. This shows that the changes all fall within the limits of +/- 0.25 class boundary. They therefore meet the intercalibration requirements.

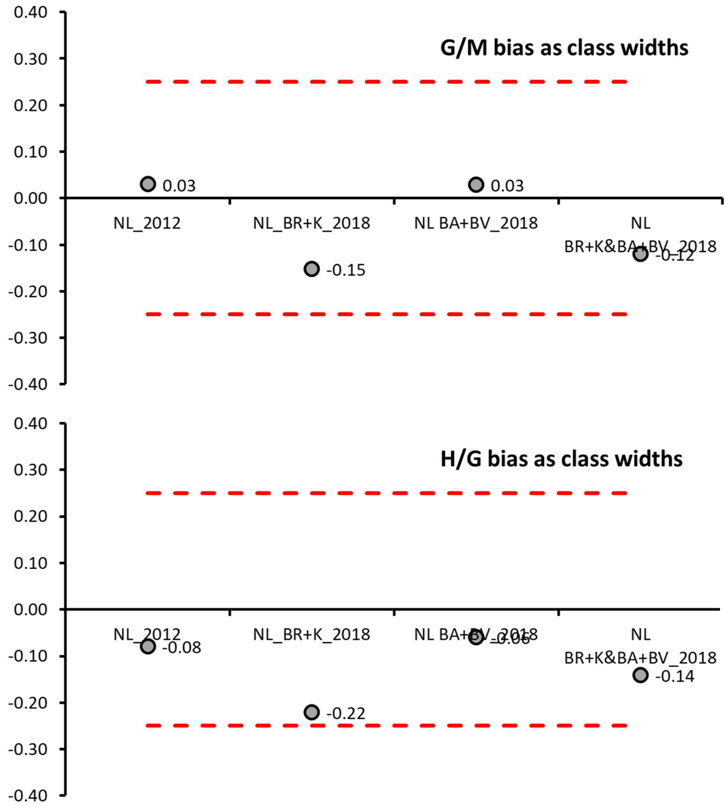
Table 4-2. Calculation of the boundary bias expressed in class widths, relative to the average TAPI score in the initial situation (AVG_MS), for the current method (NL_2012) and for the proposed adjustments*.

score Common Metric (TAPI)	AVG_MS	NL_2012	NL_BR+K_2018	NL BA+BV_2018	NL BR+K&BA+BV_2018
H-G	0.72	0.70	0.68	0.71	0.69
G-M	0.54	0.55	0.52	0.55	0.52
class width					
H-G		0.16	0.16	0.16	0.17
G-M		0.16	0.16	0.16	0.17
Bias H-G		-0.01	-0.04	-0.01	-0.02
Bias G-M		0.00	-0.02	0.00	-0.02
Bias H-G in CW		-0.08	-0.22	-0.06	-0.14
Bias G-M in CW		0.03	-0.15	0.03	-0.12

boundary bias calculation

- perform a linear regression (y = ax + b) of the TAPI on the EQR (Figure 6), determine a and b;
- calculate the TAPI value for the good-moderate limit (G-M, EQR = 0.6) and the very good-good limit (H-G, EQR = 0.8);
- calculate the width of the classes; this is the same for all classes and can be calculated a * 0.2 (EQR) or by TAPI with EQR = 0.8 - TAPI with EKR = 0.6;
- the boundary bias is the difference between the calculated TAPI for the new method and the average TAPI of all member states (AVG_MS). This is calculated at EQR = 0.6 (G-M) and EQR = 0.8 (H-G);
- finally, the boundary bias is divided by the class width, so that it is expressed in terms of the number of class widths

*Figure 8. Boundary bias expressed in class widths relative to the initial situation, for the current method (NL_2012) and for the proposed adjustments *.*



* Explanation of abbreviations:

- NL_2012 = current method.
- NL_BR+K_2018 = metric % W bream replaced by bream+carp and class boundaries adjusted.
- NL_BA+BV_2018 = class boundaries adjusted for the metric % W (roach+perch)/eurytopic species.
- BR+K&BA+BV_2018 = both adjustments combined.

5. Results and discussion

The class boundaries of the metric "% bream" were evaluated in several ways (Jaarsma & Klinge, 2018). This showed that the current class boundaries do not match the quality classes of the pressure gradient based on eutrophication parameters; the metric is too strict. This also corresponds to the aforementioned relationship of the percentage of bream (+carp) with the % submerse vegetation and the corresponding quality classes in the Dutch assessment system for macrophytes. Both approaches arrive at comparable class boundaries, these are therefore included as the proposed new boundaries in this report (see table 5-1). We consider this to be more than sufficient grounds for adjusting the metric.

Table 5-1. Proposed class boundaries for the adapted metrics for fish in freshwater lakes (water types M14, M20, M21, M23 and M27)

	Class boundaries	ref_upper	H-G	G-M	M-P	P-B	ref_lower
metric	**EQR**	1	0.8	0.6	0.4	0.2	0
proportion bream+carp (%)		5	15	40	60	85	100
proportion roach+perch (%)		60	45	30	15	5	0

We chose to use the "% bream + carp", analogous to the metrics for ditches and canals in the Netherlands. Although carp only rarely reaches high biomass in the larger lakes, this leads to better results and more uniformity in the assessment methods. Furthermore, the analysis seems to indicate that for bream+carp, a distinction in class boundaries between lake types is not necessary for lakes > 50 hectares. This means that the metric is better substantiated than so far (both in terms of functioning and with regard to class boundaries), is more uniform (no distinction between lake types) and the criticism of class boundaries is met.

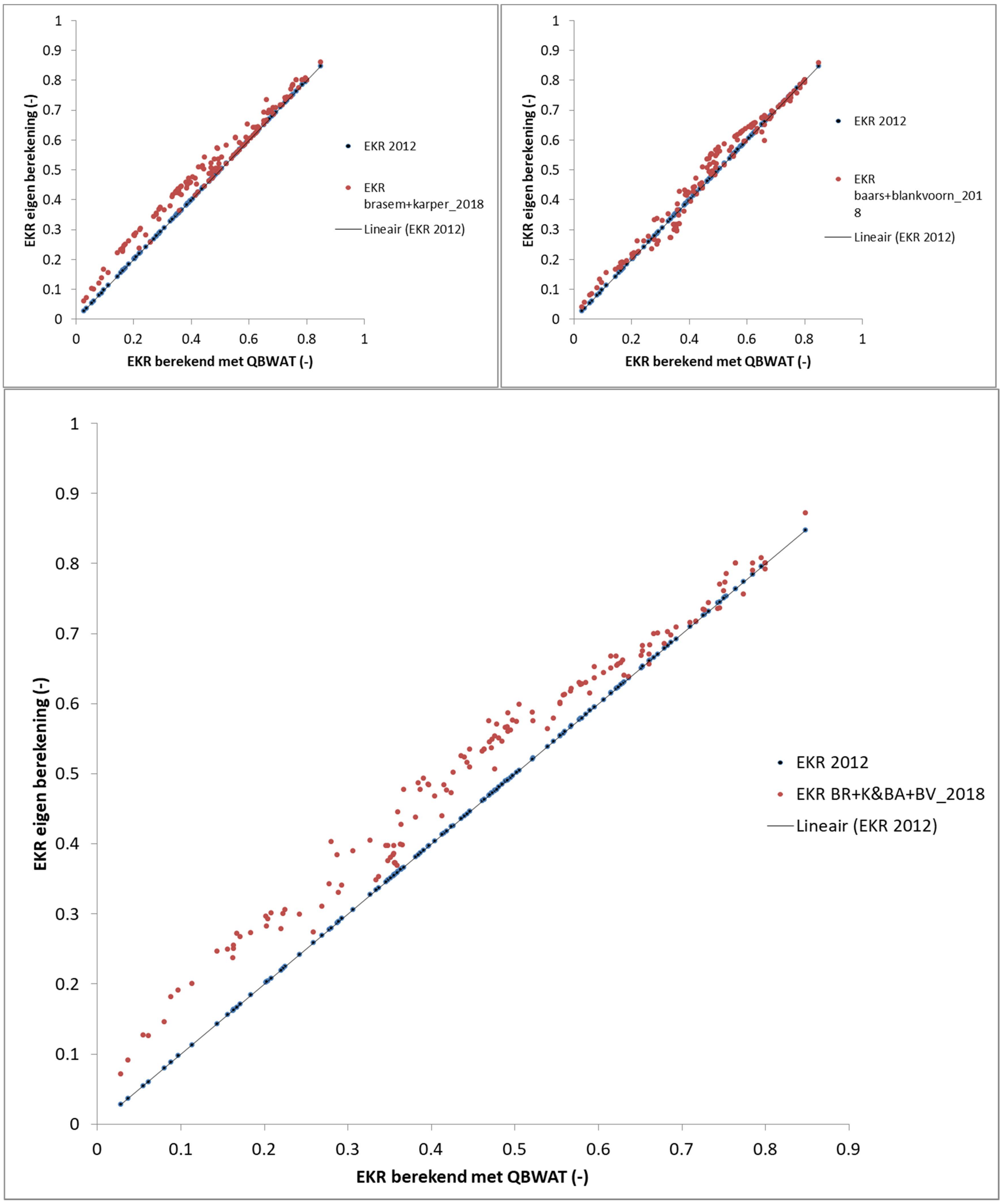
The class boundaries for the metric "% perch + roach" were evaluated at the same time. This led to the proposed slightly adjusted class boundaries, whereby, as with the previous metric, the distinction between lake types was abandoned.

With the proposed adjustments, the assessment methods for fish in lakes will become somewhat less strict, but also more distinctive between lakes with different quality. The width of the class “bad” has decreased and the class boundaries are more even spaced. This means that for the lower classes in particular, a change in the fish community is more likely to be noticeably reflected in a change in the EQR-score. This is important with respect to the effect of measures taken to improve lake quality.

Figure 9 shows that the adjustments mainly have an effect in the lower EQR-ranges, in other words: lakes with lower scores will often score somewhat better (order of magnitude around 0.1 EQR). This meets the main criticism on the methods: that the metric % W bream was too strict.

Finally, it has been shown that the proposed changes also fall within the criteria of the EU-intercalibration. Therefore, the fish experts of the water boards in the Netherlands have decided to adopt the proposed changes.

Figure 9. Overall effect of the proposed adjustments based on the data from the intercalibration data set (multiple water types and years). On the horizontal (x-axis) the EQR scores of the current methods (NL_2012). On the y-axis the EQR scores of the proposed adjusted methods (top left bream + carp adjustment, top right roach+perch adjustment, below both metrics adjusted).



6. References

European Union, 2015. Procedure to fit new or updated classification methods to the results of a completed intercalibration exercise. Guidance Document No. 30. Technical Report 2015–085.

Jaarsma, N., 2012. Aanpassingen KRW-Maatlatten M-typen. STOW114-8. Witteveen+Bos in opdracht van Rijkswaterstaat Waterdienst.

Jaarsma, N., M. Klinge & R. Pot (red.) 2007. Achtergronddocument Vissen. Expertgroep vissen.

Klinge, M., J. Backx, M. Beers, B. Higler, N. Jaarsma, Z. Jager, J. Kranenbarg, J. de Leeuw, F. Ottburg, M. van der Ven & T. Vrieze, 2004. Achtergronddocument referenties en maatlatten voor vissen. www. stowa.nl.

Sandra Poikane, David Ritterbusch, Christine Argillier, Witold Białokoz, Petr Blabolil, Jan Breine, Nicolaas G. Jaarsma, Teet Krause, Jan Kubečka, Torben L. Lauridsen, Peeter Nõges, Graeme Peirson, Tomas Virbickas, 2017. Response of fish communities to multiple pressures: Development of a total anthropogenic pressure intensity index. Science of The Total Environment. Volume 586. Pages 502-511.

D. Ritterbusch, C. Argillier, J. Arle, W. Białokoz, J. Birzaks, P. Blabolil, J. Breine, H. Draszkiewicz-Mioduszewska, N. Jaarsma, I. Karottki, T. Krause, J. Kubečka, T. Lauridsen, M. Logez, A. Maire, A. Palm, G. Peirson, M. Říha, J. Szlakowski, T. Virbickas, S. Poikane, 2017a. Water Framework Directive Intercalibration: Central-Baltic Lake Fish fauna ecological assessment methods; Part A: Descriptions of fish-based lake assessment methods. European Union, JRC, Ispra.

D. Ritterbusch, C. Argillier, J. Arle, W. Białokoz, J. Birzaks, P. Blabolil, J. Breine, H. Draszkiewicz-Mioduszewska, N. Jaarsma, I. Karottki, T. Krause, J. Kubečka, T. Lauridsen, M. Logez, A. Maire, A. Palm, G. Peirson, M. Říha, J. Szlakowski, T. Virbickas, S. Poikane, 2017b. Water Framework Directive Intercalibration: Central-Baltic Lake Fish fauna ecological assessment methods; Part B: Development of the intercalibration common metric; Part C; Intercalibration. European Union, JRC, Ispra.

Van der Molen, D.T., Pot, R., Evers, C.H.M., Van Herpen, F.C.J., and Van Nieuwerburgh, L.L.J. (2016). Referenties en maatlatten voor natuurlijke watertypen voor de kaderrichtlijn water 2015-2021 (Amersfoort: STOWA).