



Institut für Binnenfischerei e.V. (IfB)
Potsdam-Sacrow
Im Königswald 2, 14469 Potsdam

**INTERKALIBRIERUNG UND FORTENTWICKLUNG
DER FISCHBASIERTEN SEEN- UND
FLIEßGEWÄSSERBEWERTUNG
GEMÄß EU-WASSERRAHMENRICHTLINIE**

Projekt-Nr. O 5.11

Endbericht

Mitarbeit: Dr. David Ritterbusch (Verfasser)
Dr. Michael Schubert
Dr. Uwe Brämick

Im Januar 2011

INHALT - BERICHT

Ziel des Vorhabens und Übersicht der geplanten Tätigkeiten in den vertraglich festgelegten Arbeitsbereichen.....	5
Praxistest Seen.....	7
Allgemeines.....	7
SITE-Ansatz	8
Stand der Arbeiten Ende 2010	8
Verbesserungen des Verfahrens	8
Testbewertungen und Leistungsfähigkeit.....	9
Alpiner Bereich	9
Norddeutschland.....	9
Zusammenfassung	10
Bewertungsvorlage.....	11
TYPE-Ansatz	12
Stand der Arbeiten Ende 2010	12
Verbesserungen des Verfahrens	12
Testbewertungen und Plausibilität	13
Entwicklungsdatensatz	13
Testdatensatz	13
Stabilität des Verfahrens	14
Bewertungsvorlage.....	14
Anwendung von SITE oder TYPE.....	15
Vorbereitung der Veröffentlichung	15
Interkalibrierung und GIG Leitung	16
Central/Baltic GIG	16
Stand der Arbeiten Ende 2010	16
Fortschritte der Tätigkeiten, Berichterstattung und Kommunikation	16
Entwicklung und Prüfung eines common metric	17
Entwicklung einer alternativen Option zur Interkalibrierung	18
Alpine GIG.....	21
Stand der Arbeiten Ende 2010	21
Tätigkeiten und Ergebnisse 2011	21
Nutzen für den wasserrechtlichen Vollzug	22
Zusammenfassung und weiterer Forschungsbedarf	23
Literatur.....	24
Anhang	27

INHALT - ANHANG

Die Anhänge behandeln spezielle Fragestellungen und sind nicht mit ausgedruckt. Sie sind Bestandteil der vollständigen Version als Datei und dort zu finden.

Sitzung 1/2011 des VDFF-AK „Fischereiliche Gewässerzustandsbewertung“ am 19./20.April 2011 in Würzburg: Protokollauszug Top 1 Seebewertung.....	27
Sitzung 2/2011 des VDFF-AK „Fischereiliche Gewässerzustandsbewertung“ - am 25./26. Oktober 2011 in Dresden: Protokollauszug Seebewertung.....	30
CB LakeFish Milestone 6 report	31
Development of the common metric	45
Suitability of the common metric	50
C/B Intercalibration - harmonization of good/moderate class boundaries.....	58

ZUGEHÖRIGE DATEIEN

- Gesamtdokument einschließlich der Anhänge
- Excel-Bewertungsvorlage SITE Version
- Excel-Bewertungsvorlage TYPE Version

ZIEL DES VORHABENS UND ÜBERSICHT DER GEPLANTEN TÄTIGKEITEN IN DEN VERTRAGLICH FESTGELEGTEN ARBEITSBEREICHEN

Die Wasserrahmenrichtlinie (WRRL 2000) fordert die Herstellung eines guten ökologischen Zustandes bzw. Potenzials in den Oberflächengewässern der EU. Dazu sind Bewertungen mit biologischen Qualitätselementen zur Einschätzung des aktuellen ökologischen Zustandes erforderlich. Zeigt die Biologie des Gewässers einen mäßigen, unbefriedigenden oder gar schlechten Zustand an, so muss durch geeignete Maßnahmen ein guter Zustand hergestellt werden. Im vorliegenden Bericht wird das Qualitätselement FISCHE IN SEEN behandelt.

Im Rahmen des hier beschriebenen LfP-Projektes und der Vorgängerprojekte O 02.09, O 14.09 und O 14.10 wurde die Bearbeitung des Qualitätselementes Fische in Seen für Deutschland durchgeführt. Zum einen gehört dazu die Entwicklung und ständige Verbesserung eines Verfahrens zur Bewertung des ökologischen Zustandes von Seen anhand der Fische. Zum anderen beinhalten die Vorhaben die Beteiligung am europaweiten Interkalibrierungsprozess, durch den die internationale Vergleichbarkeit der Bewertungsergebnisse gewährleistet werden soll.

In Fortführung der bis Ende 2010 erzielten Ergebnisse enthielt die Projektbeschreibung vom 02.12.2010 als Anlage des Förderantrags folgende Schwerpunkte:

Seenbewertung Deutschland: Fortentwicklung

- Weitere Datensammlung und Testbewertung des deutschen Bewertungsvorschlags
- Defizitanalyse und ggf. Modifikation
- Prüfung der Übertragbarkeit auf andere Wasserkörper
- Vorbereitung der Publikation der Ergebnisse für den deutschsprachigen Raum
- Teilnahme an Treffen zur Koordination

Seenbewertung EU: Interkalibrierung und Leitung Central-Baltic GIG

- Fortführung der Erarbeitung eines gemeinsamen Bewertungssystems
- Berichtserstattung und Austausch mit den Mitgliedsstaaten der C/B-GIG
- Harmonisierung der Bewertungssysteme für Deutschland, die CB-GIG und ggf. Abgleich mit dem WISER-Projekt
- Berichtserstattung für das JRC oder andere europäische Koordinierungsstellen
- Teilnahme an Treffen mit Präsentation der Ergebnisse

Die Ende des Jahres 2010 beigefügte Projektbeschreibung enthielt Tätigkeiten im Bereich der **Fließgewässerbewertung mit Fischen**. Im Rahmen der Projektplanung sollten das nationale fischbasierte Bewertungssystem für Fließgewässer (fiBS) verbessert und die deutschen Tätigkeiten im Rahmen der Interkalibrierung abgedeckt werden. Durch geänderte Bedingungen (Anstellung eines Projektpartners, Übernahme der Berichterstellung durch das JRC¹) konnten diese Tätigkeiten im Rahmen der dienstlichen Aufgaben der vorgesehenen Projektpartner durchgeführt werden. Die genannten Leistungen werden nicht mehr als Bestandteil des Projektes O 5.11 bearbeitet und vorliegend auch nicht dargestellt.

Die Fortentwicklung von fiBS wird an der Fischereiforschungsstelle Baden-Württemberg durchgeführt. Die aktuelle Version ist fiBS 8.0.6a. Die Bewertungsvorlage und das Handbuch sind auf der Webseite der Fischereiforschungsstelle Baden-Württemberg öffentlich

¹ Joint Research Center: die EU-Koordinierungsstelle für den Interkalibrierungsprozess in Ispra, Italien

zugänglich². Auch zukünftige überarbeitete Versionen des fiBS werden dort zur Verfügung gestellt.

Die Interkalibrierung der Fließgewässerbewertung mit Fischen ist unter deutscher Beteiligung abgeschlossen. Die Ergebnisse sind als Milestone 6 im „Communication & Information Resource Centre“ des JRC zugänglich³. Die Ergebnisse der Interkalibrierung werden noch von Gremien der EU geprüft und voraussichtlich im April 2012 als „Technical Report“ veröffentlicht.

Da die Ergebnisse der entsprechenden Tätigkeiten kostenfrei zugänglich sind oder sein werden, ist durch den Wegfall der Bearbeitung im Projekt O 05.11 nicht mit anderweitig entstehenden Kosten für die Länder oder die Öffentlichkeit zu rechnen.

² <http://www.landwirtschaft-mlr.baden-wuerttemberg.de/servlet/PB/menu/1296703/index.html>

³ http://circa.europa.eu/Public/irc/jrc/jrc_eewai/library?l=/download-3&vm=detailed&sb=Title

PRAXISTEST SEEN

ALLGEMEINES

Im Laufe des Jahres 2011 wurde nach weiteren Verbesserungen des deutschen Verfahrens ein anwendungsreifer Bearbeitungsstand erreicht. Die Entwicklung des Verfahrens ist abgeschlossen, weitere Verbesserungen können nur auf der Basis neuer Bewertungen als Ergebnis der Anwendung durchgeführt werden.

Das deutsche Verfahren zur Bewertung des ökologischen Zustandes von Seen anhand der Fische beinhaltet zwei Ansätze:

- Beim Gewässer-spezifischen Ansatz (SITE) wird die Fischgemeinschaft modelliert. Die Grundlage für die Ermittlung des Referenz-Zustandes können historische Aufzeichnungen, Fischereistatistiken, Literaturquellen und Expertisen sein. Für die Modellierung des gegenwärtigen Zustandes können Probebefischungen, aber auch Daten der Berufsfischerei, Literaturquellen, Fischereistatistiken und Expertisen herangezogen werden. Die Beschreibung der Fischgemeinschaften erfolgt mit artspezifischen Häufigkeitsklassen.
- Beim Typ-spezifischen Ansatz wird auf der Basis einer Typisierung und entsprechenden typbasierten Referenzzuständen bewertet. Die aktuelle Fischgemeinschaft wird auf der Basis von Fängen mit benthischen Multimaschen-Stellnetzen nach CEN 14757 bewertet. Der Ansatz ist anhand von natürlichen Seen des norddeutschen Tieflandes entwickelt worden und dementsprechend auf die Ökoregion 14 begrenzt.

Die Anforderungen an ein deutsches fischbasiertes Bewertungssystem für Seen und eventuelle Verbesserungsvorschläge des jeweiligen Bearbeitungsstandes wurden laufend zwischen Vertretern des Instituts für Binnenfischerei und verantwortlichen Experten diskutiert. Hierbei sind besonders die Vorschläge des VDFF AK „Fischereiliche Gewässerzustandsbewertung“ (im weiteren AK) als beratendes Gremium relevant. In den zweimal jährlich stattfindenden Sitzungen wurden die Schwerpunkte der Forschung und Entwicklung festgelegt.

Das hier beschriebene Projekt O 05.11 ist das Folgeprojekt des ebenfalls vom LfP geförderten Projektes O 14.10. Der Stand der Arbeiten Ende 2010 wird in jedem Abschnitt kurz vorangestellt. Details zur Bearbeitung im letzten Jahr können dem entsprechenden Bericht entnommen werden (RITTERBUSCH et al. 2011).

SITE-ANSATZ

Stand der Arbeiten Ende 2010

Der SITE-Ansatz wurde im Jahr 2010 weiter entwickelt. Mit Hilfe entsprechender Excel-Vorlagen wurden Testbewertungen durchgeführt, allerdings noch in geringem Umfang. Für den alpinen Bereich zeigten sich hervorragende Plausibilitäten. Bei der Anwendung auf süddeutsche und österreichische Seen stimmten die Bewertungsergebnisse und die durch Experten eingeschätzten ökologischen Zustände der Seen überein. Es konnten acht deutsche Seen bewertet werden. Für norddeutsche Seen zeigten sich weniger plausible Ergebnisse (ebenfalls acht Seen). Weitere Testbewertungen waren erforderlich, um die Ursachen zu identifizieren und zu beheben.

Verbesserungen des Verfahrens

Das SITE-Verfahren wurde im Zusammenhang mit den Testbewertungen verbessert. Auf eine Beschreibung der metrics und der Berechnungsroutinen wird hier verzichtet. Im Prinzip sind sie im Bericht zum LfP-Projekt O 02.09 zu finden (RITTERBUSCH & BRÄMICK 2010). Im laufenden Bearbeitungszeitraum wurden unter Anderem folgende Änderungen vorgenommen:

- Die Klassengrenzen für die Zuordnung von Statusklassen zu EQR-Werten wurden angepasst.
- Die Berechnung und Bewertung der Abundanz von Präferenzgruppen wurde angepasst.
- Im metric „Reproduktion“ werden besetzte, nicht reproduzierende Arten nun als aktuell fehlend angesehen. Dadurch führt ein Besatz fehlender Arten nicht mehr zu mehrfachen Abwertungen.
- Weitere Fischarten und ihre Präferenzklassen wurden in das Verfahren integriert.

Die Anpassung erfolgte durch Prüfung von Bewertungsergebnissen und zugehörigen Erwartungswerten für den ökologischen Zustand auf Basis von Experteneinschätzungen. Insgesamt sind die Korrekturen so umfangreich, dass vorhergehende Plausibilitätsprüfungen weitgehend hinfällig sind und alle Seen neu bewertet wurden. Da sich die Änderungen im Wesentlichen im Bereich der mäßig/unbefriedigend/schlechten Bewertung auswirken, sind hierdurch die Ergebnisse der alpinen Bewertung weniger betroffen als die des norddeutschen Bereiches.

Testbewertungen und Leistungsfähigkeit

Alpiner Bereich

Bisher konnte das SITE-Verfahren auf acht deutsche Seen im alpinen Bereich angewandt werden. Die Bewertungsergebnisse sind plausibel. In sieben der acht bewerteten Fälle stimmen die durch die Fischgemeinschaft angezeigten ökologischen Zustandsklassen mit einer Experteneinschätzung überein. In einem Fall weicht die Bewertung um eine Klasse vom erwarteten Zustand ab, bleibt dabei aber im Bereich der sehr guten/guten Zustandsklassen. Das SITE-Verfahren führt zu keiner Fehlindikation von Maßnahmebedarf.

Tab. 1: Bewertung des ökologischen Zustandes für 8 süddeutsche Seen anhand der Fische mit dem SITE-Verfahren und Experteneinschätzung des Gewässerzustandes (Experte: M. Schubert).

Lake	SITE EQR	SITE Klasse	Klasse Experte
Chiemsee	0,69	3	3
Starnberger	0,75	4	4
Walchensee	0,75	4	4
Königssee	0,81	4	4
Ammersee	0,88	4	4
Boden-/Obersee	0,94	5	5
Wörthsee	0,94	5	4
Brombachsee	1,00	5	5

Auf Grundlage der bisherigen Daten zeigt das SITE Verfahren eine sehr gute Eignung zur Bewertung des ökologischen Zustandes alpiner Seen.

Norddeutschland

Zur Prüfung der Leistungsfähigkeit des SITE-Verfahrens können die Bewertungen des Fischbestandes mit einer Einschätzung des Gewässerzustandes verglichen werden. Dabei lassen sich die Vergleiche auf die Kategorien sehr gut/gut und mäßig/unbefriedigend reduzieren. Es wird zwischen Abwesenheit und Vorhandensein eines Maßnahmebedarfs zur Wiederherstellung eines mindestens guten ökologischen Zustandes unterschieden. „Schlechte“ Gewässer fehlen im Datensatz, sowohl für die Fischbewertung als auch für die Einschätzung des Seestatus‘. Das SITE-Verfahren beruht auf der sorgfältigen Modellierung der Referenzzustände und der aktuellen Fischgemeinschaften auf der Grundlage möglichst detaillierter Informationen. Eine Bewertung erfordert die gewissenhafte und individuelle Untersuchung jedes einzelnen Gewässers. Sie ist damit nur durch Experten mit guter Kenntnis der Gewässergegebenheiten, der lokalen Fischgemeinschaften und der historischen Gegebenheiten möglich. Diese Voraussetzungen begrenzen den zur Verfügung stehenden Testdatensatz.

Für 118 natürliche Seen in Brandenburg liegen die modellierten Fischgemeinschaften für die Referenz und die aktuelle Fischgemeinschaft als Ergebnis früherer Untersuchung des IfB vor (ZAHN & BORKMANN 2006, 2007, 2008). Davon wurden **28 Brandenburger Seen** neu mit der aktuellen SITE-Version 2.11 bewertet. Zur Einschätzung des Gewässerzustandes lagen folgende Informationen vor: See-Steckbriefe des Umweltministeriums (MUGV 2010), Belastungsdaten zu Trophie, Verbau und Nutzungsintensität (MEHNER et al. 2004), ergänzt

durch Erhebungen des Verfassers (RITTERBUSCH, unveröff.) sowie Experten-Einschätzungen von ZAHN & BORKMANN. Damit war eine gute Grundlage für die Plausibilitätsprüfung der SITE-Ergebnisse gegeben. Das SITE-Verfahren liefert zufriedenstellende Ergebnisse für die 28 getesteten Gewässer Brandenburgs. In etwa 70 % der Fälle wird ein nach Expertise vorhandener Maßnahmenbedarf angezeigt bzw. es wird analog zur Experteneinschätzung kein Maßnahmenbedarf durch die Fischbewertung angezeigt. In etwa 14 % der Fälle wird entgegen der Experteneinschätzung ein Maßnahmenbedarf durch das SITE-Verfahren angezeigt. In den verbleibenden 16 % wird kein Maßnahmenbedarf angezeigt, obwohl dieser nach Expertise vorhanden wäre.

Für **20 norddeutsche Großseen** mit Flächen über 1.000 ha wurden die Referenzen im Rahmen eines früheren BMBF-Projektes modelliert (BRÄMICK et al. 2008; MEHNER et al. 2004). Die Informationen wurden zur Neubewertung mit der SITE-Version 2.11 genutzt. Die Prüfung der Plausibilität erfolgte anhand einer Einschätzung des Gewässerzustandes auf Basis von Belastungsdaten (s. o.), Einschätzungen anhand von Literaturquellen (NIXDORF et al. 2004), sowie Expertisen aus den Bundesländern. Für die Bewertung der norddeutschen Großseen ist die Übereinstimmung von Fischbewertung und Experteneinschätzung mit 55 % deutlich schlechter. Hierbei ist zu beachten, dass die entsprechenden Modellierungen knapp 10 Jahre alt sind und sich damit die Zusammensetzung der aktuellen Fischgemeinschaft und der Gewässerzustand geändert haben können. Zudem zeigten Überprüfungen, dass die damaligen Modellierungen mit anderen inhaltlichen Schwerpunkten erstellt wurden und es dadurch in einigen Fällen zu Unstimmigkeiten bei der Übertragung auf die aktuelle SITE Version kommt.

Für **19 Seen in Schleswig-Holstein** wurden die SITE-Bewertungsvorlagen vom Verfasser auf der Basis von Literaturdaten vorläufig ausgefüllt. Für 16 Gewässer lagen grobe Modellierung der Referenz-Fischgemeinschaft sowie Ergebnisse von Befischungen mit Zugwadern (BÖTTGER 2005, 2006) und Elektrobefischungen (STAAS et al. 2006) vor. Für 3 weitere Seen wurden Daten aus dem o.g. BMBF-Projekt genutzt. Die Vorlagen wurden an Herrn M. Brunke (LLUR SH, AK Fische) zur Kontrolle, Ergänzung und Plausibilitätsprüfung übermittelt und in seinem Auftrag fachgutachterlich geprüft. Die Ergebnisse lagen Ende Dezember 2011 vor (NEUMANN 2011). Die Prüfung ergab, dass sich Belastungen nicht ausreichend im Ergebnis des SITE-Ansatzes niederschlagen. Eine Analyse dieser Ergebnisse und ihrer Ursachen durch den Verfasser war im Rahmen des hier dargestellten Vorhabens nicht mehr möglich. Es ist aber wahrscheinlich, dass zukünftige Modifikationen des SITE-Verfahrens die Plausibilität der Bewertungsergebnisse im norddeutschen Raum erhöhen können.

Zusammenfassung

Insgesamt weist das SITE Verfahren eine gute Eignung zur Bewertung von Seen anhand der Fischgemeinschaft auf. Die Ergebnisse sind überwiegend plausibel. Dabei treten regionale Unterschiede auf. Im alpinen Bereich Deutschlands wird eine sehr hohe Plausibilität ohne fehlende oder falsche Identifikation von Maßnahmenbedarf erreicht. Im norddeutschen Gebiet ist die Leistungsfähigkeit des Verfahrens etwas reduziert. Es erscheint sinnvoll, das Verfahren zur Anwendung für norddeutsche Gewässer weiterhin fortlaufend zu verbessern und so die Plausibilität zu erhöhen.

Die Plausibilität der Ergebnisse hängt von einer sorgfältiger Modellierung und Beachtung der spezifischen Besonderheiten der bewerteten Seen und der vorkommenden Arten ab. Eine fachliche Prüfung der Bewertungsergebnisse ist erforderlich.

Bewertungsvorlage

Die aktuelle Version der Bewertungsvorlage ist 2.11. Bei der Prüfung des Verfahrens stellte sich heraus, dass die Umsetzung der Verfahrensvorgaben in die Excel-Bewertungsvorlagen mit erheblichen Fehlern behaftet war. Im Vergleich zum Stand Ende 2010 wurden unter Anderem folgende Verbesserungen vorgenommen:

- Die Klassengrenzen wurden korrigiert.
- Die Berechnung und Bewertung der Abundanz von Präferenzgruppen wurden berichtigt.
- Der metric „Reproduktion“ fließt nun in die Gesamtbewertung ein. Der metric wurde modifiziert.
- Zusätzliche Fischarten und ihre Präferenzklassen wurden der Vorlage zugefügt.
- Berechnungsfehler durch unvollständige Referenzen wurden korrigiert.
- Die Benutzerfreundlichkeit der Vorlage wurde erhöht.

Bei der Anwendung und Bedienung der Bewertungsvorlage durch Dritte ergaben sich keine Probleme (NEUMANN 2011).

TYPE-ANSATZ

Stand der Arbeiten Ende 2010

Für den TYPE-Ansatz wurde nachgewiesen:

- dass unbelastete (Referenz-)Seen anhand der Fischbewertung einen guten bis sehr guten ökologischen Zustand zugeordnet bekommen,
- dass das Verfahren über das Spektrum möglicher ökologischer Zustände plausible Ergebnisse liefert und
- dass die Bewertungsergebnisse signifikant mit Belastungsparametern aus den Bereichen Eutrophierung und Störung korrelieren.

Die genannten Zusammenhänge wurden anhand des Datensatzes geprüft, der zur Entwicklung des Verfahrens genutzt wurde (75 Seen). Die Anwendung des Verfahrens auf 12 „normale“ natürliche Seen, 6 besondere Seen (saline Seen oder Flusseen) und 3 stark anthropogen geprägte Seen zeigte plausible Ergebnisse. Die Expertisen zur Einschätzung des ökologischen Gewässerzustandes und die Bewertungsergebnisse des Fisch-Verfahrens stimmten weitgehend überein.

Verbesserungen des Verfahrens

Der Schwerpunkt der Tätigkeiten 2011 lag in der Anwendung auf weitere Befischungsdaten zur Erhöhung des Testumfangs. Das TYPE-Verfahren wurde im Zusammenhang mit den Testbewertungen laufend verbessert und soweit möglich vereinfacht. Auf eine Beschreibung der metrics und der Berechnungsroutinen wird hier verzichtet. Im Prinzip sind sie im Bericht zum LfP-Projekt O 02.09 zu finden (RITTERBUSCH & BRÄMICK 2010). Unter Anderem wurden folgende Änderungen vorgenommen:

- Die Klassengrenzen der den ökologischen Statusklassen zugeordneten EQR-Werte wurden angehoben und eine höhere Plausibilität erreicht. Es werden weniger sehr gute Bewertungsergebnisse erreicht, was den überaus geringen Anteil gänzlich unbeeinflusster Seen in Deutschland spiegelt.
- Es ist im Verfahren nun möglich, zwischen dem Verlust von Arten durch anthropogene Einflüsse und dem natürlichen Fehlen von Arten zu unterscheiden. Natürlicherweise fehlende Arten fließen nicht mehr abwertend in das Ergebnis ein.
- Folgende Arten werden nicht mehr im metric [Arten der benthischen Netze] berücksichtigt: Quappe (vernachlässigbar) und Wels (Fang eher zufällig, dann aber ggf. großer Einfluss). Der Karpfen wurde als funktioneller Indikator hinzugefügt.
- Folgende Arten werden nicht mehr im metric [benthivore Arten] berücksichtigt, da ihr Einfluss vernachlässigbar ist: Gründling, Schlammpeitzger und Steinbeißer. Der Karpfen wurde als funktioneller Indikator hinzugefügt.

Testbewertungen und Plausibilität

Entwicklungsdatensatz

Die Fortentwicklung des TYPE-Verfahrens hat nicht zu größeren Änderungen der berechneten EQR-Werte geführt. Die im Bericht zum Vorgängerprojekt (RITTERBUSCH et al. 2011) dargestellten Zusammenhänge zwischen dem TYPE-Fischbewertungsverfahren und Parametern der Intensität abiotischer Belastungen sind weiterhin gültig. Sie werden nicht erneut dargestellt.

Ein Fisch-Bewertungsergebnis wird als plausibel eingeschätzt, wenn es mit einer Experten-Bewertung der gesamten Belastungsintensität am entsprechenden See übereinstimmt. Durch die geänderten Klassengrenzen für die Zuordnung von Statusklassen zu EQR-Werten hat sich die Plausibilität der Ergebnisse verbessert. Insbesondere der Status „sehr gut“ wird in der Fischbewertung seltener erreicht. Für den Datensatz von 70 Seen, die für die Entwicklung des Verfahrens genutzt wurden, stimmen Fisch-Bewertungen und Experten-Erwartung in 70 % der Fälle exakt überein. Bei einer Reduzierung der Vergleiche auf die maßnahmenrelevanten Kategorien (sehr gut/gut) und (mäßig/unbefriedigend) werden 14 % der Seen falsch bewertet, wobei alle Fehlbewertungen zu gut sind. Ein gewisser Anteil zu guter Bewertungen liegt „in der Natur der Sache“, da Fische als oberste Stufe der Nahrungspyramide des Sees ein erhebliches Potenzial haben, störende Einflüsse auszugleichen.

Die Leistungsfähigkeit des TYPE-Verfahrens wurde im Rahmen des Vorhabens weiter verbessert und kann inzwischen als sehr gut eingeschätzt werden.

Testdatensatz

Für die Prüfung des Verfahrens an einem unabhängigen Datensatz standen die Ergebnisse von 43 Befischungskampagnen an natürlichen Seen bei Verwendung von Multimaschen-Stellnetzen nach CEN-Standard zur Verfügung. Die im Bericht zum Vorgängerprojekt O 14.10 bewerteten Gewässer werden in die nachfolgende Darstellung einbezogen. Für die Gewässer lagen Experteneinschätzungen des ökologischen Gewässerzustandes vor. Fisch-Bewertungen und Experten-Erwartung stimmen auch hier in 70 % der Fälle exakt überein. Bei einem Vergleich in den Kategorien (sehr gut/gut) und (mäßig/unbefriedigend) werden 16 % der Seen falsch bewertet, wobei wieder alle Fehlbewertungen zu gut sind.

Die Anwendung auf den Testdatensatz bestätigt die hohe Plausibilität der Bewertung mit dem TYPE-Verfahren.

Die Anwendung des TYPE-Verfahrens auf besondere natürliche Seen ergab überwiegend plausible Ergebnisse. Zu den besonderen natürlichen Seen gehören Gewässer mit erhöhter Salinität oder Tideeinfluss in Küstennähe, Flussseen oder Seen mit besonders intensivem fischereilichen Einfluss durch Besatz. In dieser Gruppe lagen Daten zu 10 Gewässern vor, für alle wurde mit dem TYPE-Verfahren ein Maßnahmenbedarf identifiziert. Damit werden die genannten Besonderheiten durch das TYPE-Verfahren in Form einer schlechten Bewertung erkannt. Es ist jedoch nicht möglich zwischen natürlichen Besonderheiten (Salinität) und anthropogenen Belastungen (Verbau, evtl. intensiver Besatz) zu unterscheiden. Die genannte Gruppe der besonderen Seen ist in sich heterogen, für jede einzelne Besonderheit liegen sehr wenige Daten vor. Es wird empfohlen, die Plausibilität des TYPE-Verfahrens verstärkt zu prüfen, wenn Seen bekanntermaßen besonderen natürlichen Belastungen mit Wirkung auf die Fischfauna ausgesetzt sind.

Für anthropogene oder anthropogen stark überprägte Seen (AWB und HMWB: anthropogenic or heavily modified water bodies im Sinne der WRRL) lagen Daten zu 4 Gewässern vor. In drei Fällen stimmten die Fisch-Bewertungsergebnisse mit den Experteneinschätzungen des

ökologischen Gewässerzustandes überein. In einem Fall wurde jedoch ein Maßnahmenbedarf angezeigt, der nicht durch Experteinschätzungen gestützt wurde. Die Datenlage ist zu gering um eine Aussage über die Plausibilität des TYPE-Verfahrens zu treffen. Zudem lässt sich der ökologische Zustand (oder das ökologische Potenzial) dieser Gewässer kaum fachlich einschätzen. Der Fischbestand der Seen ist stark besatzgeprägt und unterliegt besonders in Tagebauseen starken Änderungen. Laut Beschluss des AK ist die fischbasierte Bewertung von Seen im Bereich der HMWB oder AWB derzeit nicht möglich (Anhang S. 27 ff.).

Stabilität des Verfahrens

Für 9 natürliche Seen des norddeutschen Tieflandes lagen die Ergebnisse von zwei bis drei Befischungen nach CEN-Standard aus verschiedenen Jahren vor. Für eines dieser Gewässer ergab sich zwischen den einzelnen Jahren eine Abweichung des TYPE-Bewertungsergebnisses um eine Statusklasse. Die Abweichung war vergleichsweise unbedeutend, da in beiden Fällen ein Maßnahmenbedarf resultierte (Abweichung mäßig - unbefriedigend). Ein natürlicher See wurde sieben Mal befischt, die entsprechenden Bewertungen waren sechs Mal „mäßig“ und ein Mal „gut“.

Insgesamt kann das TYPE-Verfahren angesichts der hohen Variabilität von Fischbestandsuntersuchungen als sehr stabil bewertet werden. Die Datenlage ist allerdings gering, so dass die Variabilität der Bewertungsergebnisse weiter geprüft werden sollte.

Bewertungsvorlage

Die aktuelle Version der Bewertungsvorlage ist 2.00. Im Vergleich zum Stand Ende 2010 wurden unter Anderem folgende Verbesserungen vorgenommen:

- Es ist nun möglich, die Häufigkeiten der relevanten Fischarten einzeln einzugeben. Metrics, die Arten zusammenfassen, werden durch die Vorlage berechnet und bewertet. Dadurch ist die Überführung von Rohdaten in die Excel-Vorlage vereinfacht. Vorberechnungen per Hand sind nicht mehr nötig.
- Alle Verbesserungen des Verfahrens wurden in die Vorlage integriert (siehe S. 12).
- Zur Abfrage von Interkalibrierungsdaten wurde die Berechnung des entwickelten common metric in die Bewertungsvorlage eingebaut (vergl. Abschnitt „Interkalibrierung und GIG Leitung“).
- In der Vorlage wird nun auf fehlende oder nicht plausible Eingaben hingewiesen.
- Durch die Trennung von Daten-Eingabe und Bewertungs-Ausgabe wurde die Übersichtlichkeit erhöht.

Die Vorlage für die Nutzung des deutschen Verfahrens wurde von den Vertretern anderer EU-Staaten (Dänemark, Niederlande, Polen, Estland, Litauen) mit den jeweiligen nationalen Daten ausgefüllt. Dabei traten keine Verständnisprobleme auf. Die Vorlage ist für fachlich versierte Nutzer problemlos nutzbar.

ANWENDUNG VON SITE ODER TYPE

Beide Verfahrensansätze liefern überwiegend plausible Bewertungsergebnisse. Sie weisen jedoch eine regional unterschiedliche Leistungsfähigkeit auf, so dass die Wahl des Verfahrens nicht beliebig ist.

Im alpinen Bereich, d.h. in den Bundesländer Bayern und Baden-Württemberg ist das SITE-Verfahren anzuwenden. Auf der Grundlage der bisher bewerteten Seen ergeben sich in allen Fällen plausible Ergebnisse. Die Anwendung und Plausibilitätsprüfung des Verfahrens in alpinen Seen der Länder Österreich und Italien unterstützt die sehr gute Tauglichkeit des Verfahrens für diese geographische Region. Das SITE-Verfahren fließt auch in die Interkalibrierung in der alpinen GIG ein. Das TYPE-Verfahren ist im süddeutschen Bereich nicht nutzbar. Es fehlen die entsprechenden Ergebnisse aus CEN-Standardbefischungen und es ist nicht absehbar, dass sie erhoben werden können. Zudem sind die typspezifischen Referenzbedingungen an einem Datensatz von norddeutschen Seen entwickelt worden und nur mit erheblichen Modifikationen auf andere geographische Regionen übertragbar. Die Datenlage zur Erarbeitung der notwendigen Modifikationen liegt nicht vor.

In Norddeutschland (Seen der Ökoregion 14) ist das TYPE-Verfahren zu nutzen. Wie an einem Testdatensatz von Brandenburger Seen gezeigt wurde, zeigt es für Norddeutschland plausible Ergebnisse als das SITE-Verfahren. Die Ergebnisse sind damit vertrauenswürdiger und unabhängiger von der abschließenden Expertenprüfung. Das SITE-Verfahren sollte daher nur zum Einsatz kommen, wenn eine CEN Multimaschen-Befischung aus zwingenden Gründen nicht möglich ist. Das ist derzeit in Schleswig-Holstein der Fall, wo das SITE-Verfahren auf Tauglichkeit geprüft wird (NEUMANN 2011).

VORBEREITUNG DER VERÖFFENTLICHUNG

Für die öffentliche Nutzbarkeit der Verfahren zur Bewertung des ökologischen Zustandes von Seen anhand der Fische sind entsprechende Darstellungen der Vorgehensweise notwendig. Für die Verfahren liegen Beschreibungen der wissenschaftlichen Grundlagen, der Berechnungsschritte, der Leistungsfähigkeiten sowie der fortlaufend durchgeführten Änderungen vor. Zudem wurden Excel-Bewertungsvorlagen entwickelt und verbessert, die eine sofortige Anwendung beider Verfahrensansätze ohne vertiefte Kenntnis der fachlichen und mathematischen Grundlagen ermöglichen. SITE und TYPE-Ansatz sind somit auf dem aktuellen Bearbeitungsstand durch Dritte nutzbar.

Grundsätzlich wurden die Verfahren bis zum Ende des vorliegend beschriebenen Vorhabens fortlaufend verbessert. Eine abschließende Verfahrensbeschreibung konnte daher nicht erstellt werden. Auch sind Bewertungsvorlagen sowie Beschreibungen zwar vom Verfasser erhältlich, aber nicht öffentlich in Form eines Downloads oder einer Drucklegung zugänglich. Es erscheint daher notwendig, die nun anwendungsreifen Bewertungssysteme zusammenfassend darzustellen und an geeigneter Stelle zu publizieren. Für das TYPE-Verfahren insgesamt sowie für das SITE-Verfahren zur Anwendung in Süddeutschland ist ein entsprechender Bearbeitungsstand erreicht. Über die Notwendigkeit von Modifikationen des SITE-Verfahrens zur Anwendung auf norddeutsche Seen muss noch entschieden werden, z.B. von Seiten des AK.

INTERKALIBRIERUNG UND GIG LEITUNG

CENTRAL/BALTIC GIG

Stand der Arbeiten Ende 2010

Ende 2010 gab es in der Central/Baltic GIG nur ein fertiggestelltes und angewandtes Bewertungssystem (Niederlande). Eine Interkalibrierung im Sinne eines Vergleichs der Ergebnisse von Bewertungssystemen war damit unmöglich. Daher wurde im Vorjahr versucht, auf der Basis von Befischungsergebnissen deutscher Seen ein Verfahren zu entwickeln, das für alle Mitgliedsstaaten der Interkalibrierungsgruppe anwendbar ist. Der Versuch scheiterte jedoch. Die weite geographische Ausdehnung der GIG und die damit einhergehend natürlicherweise stark variierenden Fischgemeinschaften machen die Nutzung eines einheitlichen Systems unmöglich.

Fortschritte der Tätigkeiten, Berichterstattung und Kommunikation

Der Verfasser war als Vertreter Deutschlands und als Leiter der Central/Baltic GIG mit der Koordination von Tätigkeiten und der Berichterstattung beauftragt. In diesem Zusammenhang sind die Milestone Reports an das JRC zu nennen, in denen die vorhandenen Verfahren beschrieben, ihre Übereinstimmung mit den Anforderungen der WRRL geprüft, die Interkalibrierbarkeit abgeschätzt sowie die entsprechenden Pläne dargestellt werden sollen. Drei solcher Reports wurden im Jahr 2011 verlangt: Milestone 4 (März), Milestone 5 (Juni) und Milestone 6 (Oktober). Der aktuellste Milestone 6 ist im Anhang beigelegt (ab S. 31).

Im Laufe des Jahres 2011 haben zahlreiche Mitgliedsstaaten nationale Verfahren entwickelt. Neben den Niederlanden besitzen nun auch Dänemark, Deutschland, Estland, Frankreich, Litauen und Polen potenziell interkalibrierbare Verfahren. Außer den Niederlanden befischen alle Mitgliedstaaten mit Multimaschen-Stellnetzen nach dem CEN 14757- Verfahren (CEN 2005), teilweise allerdings in modifizierter Variante. Die Interkalibrierbarkeit des Niederländischen Verfahrens ist durch Parallelbefischungen gewährleistet. Trotz der inzwischen stark zugenommenen Anzahl von Verfahren konnte die Interkalibrierung nicht in Phase II, d.h. in 2011 abgeschlossen werden. Die Entwicklung der nationalen Verfahren der Mitgliedsstaaten wurde erst in der zweiten Jahreshälfte abgeschlossen, z. T. erst nach dem Termin für den Abschlussbericht. Zudem gelang es nicht, einen geeigneten common metric als Vergleichswert für die Bewertungsergebnisse zu identifizieren. Insgesamt zeigten sich die Vorgaben des Guidance Documents zur Vorgehensweise bei Interkalibrierung (CIS 2011) als wenig anwendbar.

Neben der Zusammenstellung der Fortschritte in den Mitgliedsstaaten wurden fortlaufend Anfragen des JRC zum Fortgang und zur Planung des Interkalibrierungsprozesses beantwortet. Die Mitgliedsstaaten wurden bei der Entwicklung von nationalen Verfahren beraten. Das im Rahmen des WISER Projektes entwickelte Bewertungsverfahren wurde an deutschen Daten geprüft und eine Plausibilitätsabschätzung mit Verbesserungsvorschlägen übermittelt.

In Kooperation mit der Leitung der alpinen GIG wurden Standpunktpapiere zur Reaktion von Fischgemeinschaften auf einzelne Belastungen (s. S. 80) und zur Tauglichkeit der WRRL-Qualitätskriterien „Alter“ und „sensitive Arten“ im Rahmen der Bewertung von Seen anhand der Fische verfasst. Die Standpunkte wurden in einer ECOSTAT-Sitzung diskutiert und sinngemäß akzeptiert.

Entwicklung und Prüfung eines common metric

Als Ergebnis eines Treffens der Mitgliedsstaaten Ende 2010 erwies sich der Vorschlag für ein gemeinsames System als ungeeignet. Es wurden jedoch metrics vorgeschlagen, die nach Meinung der anwesenden Experten für eine Bewertung der Fischgemeinschaften der gesamten GIG geeignet sind. Diese waren:

- 1) Einheitsfang in benthischen Netzen
- 2) benthische Arten, die auf eine Eutrophierung reagieren
- 3) Masseanteil Blei + Plötze in benthischen Netzen
- 4) Sensitive Arten des Uferbereichs

Als Reaktion auf diese Vorschläge wurde durch den Verfasser in seiner Funktion als GIG-Leiter ein reduziertes Bewertungsverfahren auf der Basis der vier genannten metrics entwickelt. Das Verfahren sollte als „kleinster gemeinsamer Nenner“ zum Abgleich der Bewertungsergebnisse in der Interkalibrierung dienen. Es entspricht dem common metric, der in der entsprechenden Richtlinie als entscheidende Voraussetzung für die Anwendung von zwei von drei Interkalibrierungs-Optionen gefordert wird (CIS 2011). Auf Anregung der auf dem Treffen anwesenden Experten wurde zur Festsetzung der Klassengrenzen die gemeinsame Datenbank genutzt. Diese wurde vom CEMAGREF zur Verfügung gestellt. Die Beschreibung der Verfahrensentwicklung des common metric ist stichwortartig im Anhang beigefügt (ab S. 45, in Englisch).

Die Berechnung des common metric wurde in die Excel-Vorlage zur Bewertung nach dem deutschen Verfahren eingebaut. Anschließend wurde die Vorlage an alle Mitgliedsstaaten versendet, die ihre Fischbestände mit Multimaschen-Stellnetzen nach oder in Anlehnung an die europäische Standardprozedur (CEN 2005) befischen. Die Bearbeiter wurden gebeten, die entsprechenden Informationen zu Befischungsergebnissen einzutragen. Rückmeldungen kamen aus Dänemark, Estland, Holland, Litauen, Tschechien und Polen (Frankreich konnte das zeitaufwändige Ausfüllen der Vorlagen nicht leisten). Zur Plausibilitätsprüfung wurden die Bewertungsergebnisse der nationalen Verfahren sowie Experteneinschätzungen der Gewässerzustände abgefragt. Mit den so gesammelten Daten wurde geprüft:

- Ob das nationale System plausibel ist (Vergleich der ökologischen Statusklassen der Expertenbewertung der Seen und des nationale Fisch-Bewertungssystems)
- Ob der common metric plausible Ergebnisse liefert (Experte für Seen - common metric für Fisch)
- Ob die deutsche Methode plausible Ergebnisse liefert (Experte für Seen - deutsches Verfahren für Fisch)

Die Ergebnisse der nationalen Fischbewertungsmethoden zeigen überwiegend gute Zusammenhänge zu den nationalen Experteneinschätzungen. Der common metric und das deutsche Verfahren haben durchweg schlechtere, teilweise auch keine erkennbaren Zusammenhänge zu den Expertenbewertungen. Dadurch wird die Annahme unterstützt, dass ein gemeinschaftliches, GIG-weites Bewertungssystem nicht zielführend ist.

Zudem wurde geprüft, ob signifikante Zusammenhänge zwischen den Ergebnissen der nationalen Fisch-Bewertungssysteme und dem common metric bestehen. Signifikante Korrelationen sind die Voraussetzung der Nutzbarkeit des common metric im Interkalibrierungsprozess. Testweise wurden auch die Korrelationen zu den Ergebnissen des deutschen Verfahrens geprüft. Zum Vergleich wurden die EQR-Werte herangezogen. Die Zusammenhänge zeigen, dass der common metric zur Unterstützung der Interkalibrierung nutzbar wäre, allerdings nur in einem begrenzten Bereich direkter Nachbarstaaten (Dänemark, Deutschland, Holland).

Eine genauere Darstellung der Prüfkriterien, der Ergebnisse und der Konsequenzen ist im Anhang zu finden (ab S. 50, in Englisch). Insgesamt muss die Interkalibrierbarkeit der Bewertungssysteme auf Basis eines common metrics als begrenzt eingeschätzt werden. Zudem setzt die Eichung an einem common metric voraus, dass eine Einigung über sog. benchmark sites erzielt wurde. Dies sind Gewässer, die einem auf Basis der Belastungen gemeinsam definierten ökologischen Zustand entsprechen. Eine solche Einigung wurde im Interkalibrierungsprozess nicht erzielt.

Entwicklung einer alternativen Option zur Interkalibrierung

Im September/Oktober 2011 stellte sich die Situation so dar, dass im Laufe des Jahres zwar zahlreiche neue Bewertungssysteme entwickelt worden waren, dass eine Interkalibrierung aber ohne Korrelationen zwischen Fisch-metrics und Belastungsparametern, ohne benchmark-sites und weitgehend ohne common metric durchgeführt werden musste. Damit waren (und sind) die vorgeschlagenen Interkalibrierungsoptionen des offiziellen Guidance (CIS 2011) nicht anwendbar. Zudem wird die direkte Anwendbarkeit des Guidance auf das Qualitätselement „Fische in Seen“ von zahlreichen Bearbeitern einschließlich dem Verfasser kritisch gesehen. Wesentlicher Kritikpunkt ist die Forderungen nach möglichst hoch korrelierten Beziehungen zwischen einzelnen Belastungen und Fisch-metrics. Hierdurch würde man im Endeffekt direkte und präzise Messungen von Belastungen durch indirekte Messungen von Fischbestandsmerkmalen ersetzen. Gemäß der Wasser-Rahmenrichtlinie soll aber festgestellt werden, ob sich Belastungen ökologisch auswirken. Das ist nicht gleichbedeutend mit der Messung von Belastungen.

Unter Beachtung der spezifischen Leistungen der Fisch-Bewertungssysteme wurde eine alternative Interkalibrierungsoption entwickelt. Hierbei waren zwei Dinge entscheidend: Erstens können Merkmale der Fischgemeinschaft (metrics) einen schlechten ökologischen Zustand der Fischgemeinschaft und damit des Gewässers anzeigen. Der Umkehrschluss ist aber unzulässig; fehlen Hinweise auf Störungen in den metrics der Fischgemeinschaft ist das nicht gleichbedeutend mit einem guten Zustand des Sees. Der ökologische Zustand von Gewässern mit „guter“ Fischgemeinschaft ist unbekannt; er kann sowohl gut als auch schlecht sein. Hauptursache ist die Resilienz von Ökosystemen, d. h. die Wirkung von störenden Einflüssen wird dynamisch ausgeglichen. Zweitens ist die „wirkliche“ Anzahl degradierter Gewässer im Datensatz unbekannt und nicht feststellbar. Mängel im ökologischen Zustand können, müssen aber nicht durch die Fischgemeinschaft angezeigt werden. Experteneinschätzungen „ihrer“ Gewässer durch die nationalen Bearbeiter sind eine wertvolle Interkalibrierungsgrundlage, aber nicht objektiv quantifizierbar. Grenzwerte für abiotische Parameter, die ökologischen Zustandsklassen entsprechen, können nicht festgelegt werden (sonst wäre die Fischbewertung nicht mehr nötig).

Der alternative Interkalibrierungsvorschlag beruht auf den metrics des deutschen Verfahrens und des common metrics, die über ausgefüllte Excel-Vorlagen von den Mitgliedsstaaten der GIG zur Verfügung gestellt wurden. Das Verfahren beschränkt sich auf den Vergleich der Übergänge zwischen den Statusklassen gut und mäßig. Die grundlegenden Schritte des Verfahrens sind: 1) die Identifikation von Seen bei denen die Fische einen mäßigen oder schlechteren ökologischen Zustand anzeigen und 2) die Prüfung, ob diese Seen in den nationalen Verfahren auch einen mindestens mäßigen Zustand erhalten.

1. Identifikation von Seen, bei denen die Fische einen mäßigen oder schlechteren ökologischen Zustand anzeigen
 - Für die meisten Seen lagen Expertenbewertungen des ökologischen Zustands vor. Die Experteneinschätzungen wurden zu zwei Kategorien zusammengefasst. Die Kategorie Experte:GUT umfasst den sehr guten und guten ökologischen Zustand nach Expertise, Experte:SCHLECHT umfasst den mäßigen, unbefriedigenden oder schlechten ökologischen Zustand nach Expertise.
 - Für viele Fisch-metrics sind Grenzwerte erkennbar, die Seen im Zustand Experte:SCHLECHT identifizieren, ohne dass Seen im Zustand Experte:GUT einbezogen sind (Abb. 1). Die Fisch-metrics jenseits der Grenzwerte zeigen einen garantiert degradierten ökologischen Zustand anhand der Fischgemeinschaft an (metric: SCHLECHT). Für diese Seen sind sich alle Experten einig, dass sie sich im schlechten ökologischen Zustand befinden.
 - Die Seen, die mit metric: SCHLECHT identifiziert werden, sind ein Teildatensatz aller Seen, die sich in einem schlechten ökologischen Zustand befinden. Durch die Kombination mehrere metrics wird die Anzahl erkannter Seen im degradierten ökologischen Zustand gemäß der Fischgemeinschaft erhöht.

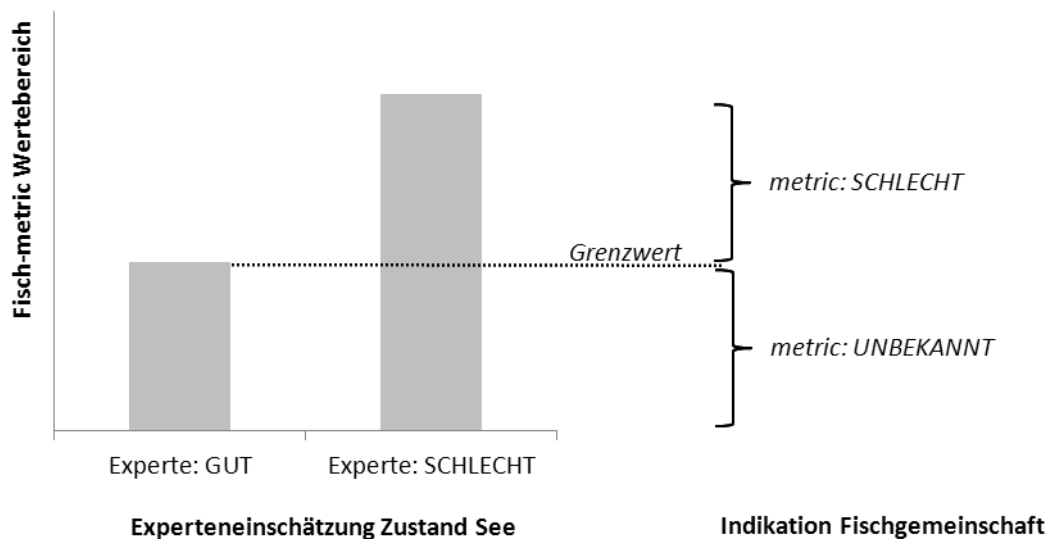


Abb. 1: Prinzip der Identifikation von Seen im schlechten ökologischen Zustand mit Fisch-metrics. Für die meisten Metrics gibt es Grenzwerte, ab denen ein Gewässer nach einhelliger Expertenmeinung im schlechten ökologischen Zustand ist. Unterhalb dieser Grenzwerte kommen die Werte des Fisch-metrics sowohl in Seen mit guter als auch mit schlechter Expertenbewertung vor, der Zustand des Gewässers ist in diesem Bereich anhand des metrics nicht festzustellen (metric:UNBEKANNT).

2. Prüfung der nationalen Verfahren
 - Mit der beschriebenen Vorgehensweise wurde anhand von Fischgemeinschaftsmerkmalen ein Datensatz von Seen ermittelt, der sich nach einhelliger Expertenmeinung im degradierten ökologischen Zustand befindet.
 - An die nationalen Verfahren wurde die Anforderung gestellt, dass alle nach einhelliger Expertenmeinung degradierten Seen in der nationalen Bewertung höchstens einen mäßigen Zustand zugeschrieben bekommen

Im Ergebnis erfüllen die meisten nationalen Verfahren diese Kriterien. In einigen Fällen sind jedoch Anhebungen der EQR-Entsprechungen für den Übergang von guter zu mäßiger Zustandsklasse notwendig.

Die alternative Interkalibrierungsoption wurde schriftlich als Entwurf dargelegt und den Mitgliedsstaaten Mitte November zur Prüfung und Kommentierung übersandt (Entwurf im Anhang ab S. 58, in Englisch). Die Mitgliedsstaaten haben den Vorschlag sinngemäß akzeptiert, aber Unklarheiten in der Darstellung bemängelt. Nach Überarbeitung soll der Vorschlag an das JRC übermittelt werden. Eine abschließende Entscheidung liegt demnach nicht vor. Die im Anhang dargelegten Inhalte sind daher als Tätigkeitsbeleg zu werten und kein Interkalibrierungsergebnis.

ALPINE GIG

Stand der Arbeiten Ende 2010

Im Dezember 2010 fand in Scharfling (Österreich) ein erstes Treffen zum Vergleich der Bewertungsergebnisse der deutschen und österreichischen Ansätze zur Seebewertung mit Fischen statt. Im Ergebnis zeigte sich, dass der deutsche SITE Ansatz im alpinen Bereich hervorragend mit den Expertenbewertungen der süddeutschen Seen korrespondiert. Zudem stimmten die vergleichenden Bewertungen von acht Seen mit deutschem SITE Ansatz und mit dem österreichischen Verfahren weitestgehend überein.

Tätigkeiten und Ergebnisse 2011

Im März 2011 fand in Scharfling ein Interkalibrierungstreffen statt, in dem neben Österreich und Deutschland (vertreten durch Bearbeiter des vorliegenden Berichtes) auch Slowenien und Frankreich vertreten waren. Zu diesem Zeitpunkt waren nur das österreichische und das deutsche Verfahren anwendungsreif. Die Zahl der im direkten Vergleich bewerteten Gewässer wurde erhöht. Weiterhin stimmten die Ergebnisse zwischen den beiden Ansätzen hervorragend überein. Die Datenlage für die Seen aus Frankreich und Slowenien war jedoch für die Modellierung eines gesicherten Referenzzustandes nicht ausreichend.

Im weiteren Verlauf des Jahres trat die Teilnahme von Frankreich und Slowenien in den Hintergrund. Es wurde jedoch möglich, italienische Gewässer in den Vergleich einzubeziehen und mit den Verfahren der deutschsprachigen Länder zu bewerten. So konnte die Interkalibrierung über den Direktvergleich Österreich-Deutschland ausgedehnt werden.

Bis Ende 2011 wurden 5 italienische, 8 deutsche und 13 österreichische Seen mit den Verfahren der Mitgliedsstaaten bewertet. Im Resultat zeigt sich eine weiterhin hervorragende Übereinstimmung zwischen deutschem und österreichischem Verfahren (Abb. 2). Das italienische Verfahren bewertet etwa eine Klasse besser, daher wurde vorgeschlagen die Klassengrenzen zu modifizieren. Plausibilitätsprüfungen anhand der Fischgemeinschaften unterstützen dieses Ergebnis.

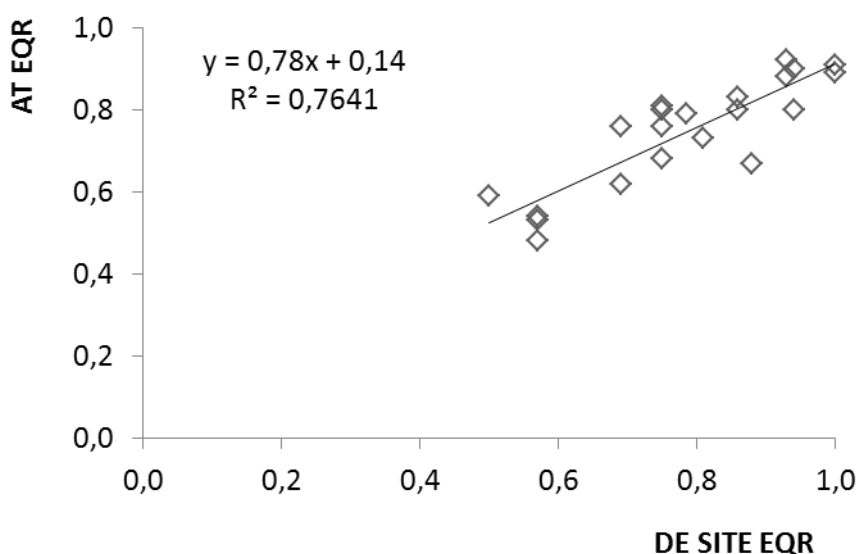


Abb. 2: Korrelation zwischen den EQR-Werten des deutschen SITE-Verfahrens und des österreichischen Verfahrens zur Bewertung des ökologischen Zustandes von Seen anhand der Fische.

Von den aktiven Mitgliedern der Alpinen Gruppe wird die Interkalibrierung als abgeschlossen betrachtet, d. h. die Vergleichbarkeit der Seebewertung in den Mitgliedstaaten ist gewährleistet. Entsprechende Berichte und Präsentationen der GIG-Leitung Österreich wurden jedoch nicht vollständig akzeptiert (diese Tätigkeiten sind nicht Bestandteil des vorliegend beschriebenen Projektes). Wesentlicher Kritikpunkt war der mangelnde Beleg von Belastungskorrelationen als Vorgabe des Intercalibration Guidance (CIS 2011). Die aktiven Mitglieder Österreich, Italien und Deutschland lehnen die Kritik aus fachlicher Sicht deutlich ab. Dennoch versucht die GIG-Leitung, den formellen Anforderungen gerecht zu werden und die Interkalibrierung innerhalb der Phase II offiziell abzuschließen. Im Rahmen des Vorhabens wurden die Anfragen der GIG Leitung bearbeitet und testweise weitere Gewässer aus Österreich und Italien mit dem deutschen Verfahren bewertet.

NUTZEN FÜR DEN WASSERRECHTLICHEN VOLLZUG

Die im vorliegenden Bericht beschriebenen Tätigkeiten beinhalten:

- Die Entwicklung und Optimierung eines deutschen Systems zur Bewertung des ökologischen Zustandes von Seen mit Fischen. Eine solche Bewertung wird von der Wasserrahmenrichtlinie für alle EU Mitgliedsstaaten gefordert.
- Die Vertretung Deutschlands im europäischen Interkalibrierungsprozess. Die Beteiligung aller Mitgliedsstaaten in der europäischen Harmonisierung der Bewertungsergebnisse ist ebenfalls durch die Wasser-Rahmenrichtlinie verpflichtend vorgegeben.
- Die Leitung der Central/Baltic GIG durch Deutschland gewährleistet den Interkalibrierungsprozess, da sich kein anderer Mitgliedsstaat zur Übernahme dieser Funktion in der Lage sah. Die Leitung erlaubt die Vertretung deutscher Interessen und Standpunkte im Gesamtprozess. Wichtig ist neben der Präsenz auch die fachliche Begründung von Verspätungen durch die besonderen Anforderungen des Qualitätselementes Fische in Seen.

Das Projekt dient unmittelbar der Durchführung wasserrechtlicher Maßgaben und ihres Vollzugs.

ZUSAMMENFASSUNG UND WEITERER FORSCHUNGSBEDARF

Im Rahmen des hier beschriebenen Vorhabens wurde die Bewertung von Seen mit Fischen erheblich verbessert und hat einen anwendungsreifen Zustand erreicht. Dabei ist nicht nur ein nationales Verfahren entwickelt worden, auch dessen Vergleichbarkeit im europäischen Kontext wurde gezeigt. Folgende Ergebnisse des Vorhabens wurden erzielt:

Seenbewertung Deutschland: Fortentwicklung

Beide Ansätze des deutschen Verfahrens wurden durch Anwendung auf Neudaten geprüft und erheblich verbessert. Die beiden Ansätze SITE-Verfahren und TYPE-Verfahren decken die Notwendigkeiten zur fischbasierten Seebewertung im Bundesgebiet ab. Sie sind für natürliche Seen nutzbar. Die Übertragbarkeit des Verfahrens auf anthropogene oder stark anthropogen beeinflusste Seen ist derzeit nur an sehr wenigen Seen geprüft, zudem gibt es Unklarheiten bei der Interpretation der Bewertungsergebnisse.

Seenbewertung EU: Interkalibrierung und Leitung Central-Baltic GIG

Es ist nicht möglich, alle Seen der geographischen Interkalibrierungsgruppe Central/Baltic mit einem gemeinsamen Verfahren WRRL-konform zu bewerten. Nationale Systeme erzielen jeweils wesentlich plausiblere Ergebnisse. Trotz der schwierigen Situation der Seebewertung mit Fischen wurde ein vielversprechender Vorschlag zur Interkalibrierung entwickelt und von den Mitgliedsstaaten grundsätzlich befürwortet. Die Interkalibrierung in der alpinen Gruppe wurde abgeschlossen. Offizielle Bestätigungen stehen für beide Gruppen noch aus. Der aus der WRRL resultierende Verpflichtung Deutschlands zur Entwicklung eines fischbasierten Verfahrens zur Bewertung des ökologischen Zustandes von Seen wurde nachgekommen und an den entsprechenden Stellen darüber berichtet.

Trotz der guten Ergebnisse sind noch wichtige Aspekte unbearbeitet und machen eine Fortführung der Bearbeitung erforderlich. Eine deutliche Verbesserung der Verfahren auf Basis der vorhandenen Daten ist nicht zu erwarten. In Anbetracht der guten Plausibilitäten können die Verfahren als anwendungsreif angesehen werden. In Absprache mit dem AK sind die Schwerpunkte der weiteren Tätigkeiten auf die Fortentwicklung des nationalen Verfahrens zu setzen und folgende Ziele anzustreben:

- Praxistest der Verfahren, d.h. Anwendung auf Seen der Bundesländer. Damit wird den Anforderungen der WRRL nachgekommen, zudem lassen sich neue Erkenntnisse und Daten zur Verbesserung des Verfahrens gewinnen.
- Dazu ist zunächst eine Abschätzung der zu bewertenden Gewässer in den Bundesländern nötig und eine Kostenabschätzung sinnvoll.
- Ein Handbuch mit allgemein verständlicher Bewertungsanleitung sowie detaillierter Darstellung der wissenschaftlichen Hintergründe muss für den gegenwärtigen Stand verfügbar gemacht werden.
- Auf der Grundlage fortlaufender Testbewertungen sollten die Verfahren weiter geprüft werden.
- Modifikationen und Verbesserungen scheinen insbesondere für den SITE-Ansatz bei der Anwendung auf norddeutsche Gewässer weiter erforderlich.

Die im Vorhaben entwickelte alternative Interkalibrierungsoption sollte dazu führen, dass das deutsche Verfahren noch innerhalb der offiziellen Phase II als interkalibriert gelten kann. Im Februar 2012, nach Ablauf des Vorhabens, stellte sich aber heraus, dass der Vorschlag nicht mehr von JRC bzw. ECOSTAT akzeptiert wird, der Fortgang der Interkalibrierungstätigkeiten jedoch gewünscht wird.

LITERATUR

- ANDERSON, J. N., JEPPESEN, E. & SONDERGAARD, M. (2005): Ecological effects of reduced nutrient loading (oligotrophication) on lakes: an introduction. *Freshwater Biology* 50: 1589-1593.
- APPELBERG, M. (2000): Swedish standard methods for sampling freshwater fish with multi-mesh gillnets. Institute of Freshwater Research, Drottningholm.
- APPELBERG, M., BERGQUIST, B. C. & DEGERMAN, E. (2000): Using fish to assess environmental disturbance of Swedish lakes and streams-a preliminary approach. *Verhandlungen der Internationalen Vereinigung für Limnologie* 27: 311-315.
- BECK, M. W. & HATCH, L. K. (2009): A review of research on the development of lake indices of biotic integrity. *Environmental Reviews* 17: 21-44.
- BELPAIRE, C., SMOLDERS, R., AUWEELE, I. V., ERCKEN, D., BREINE, J., VAN THUYNE, G. & OLLEVIER, F. (2000): An index of biotic integrity characterizing fish populations and the ecological quality of Flandrian water bodies. *Hydrobiologia* 434: 17-33.
- BÖTTGER, T. (2005): Fischmonitoring in 12 schleswig-holsteinischen Seen. Gutachten Landesamt für Natur und Umwelt des Landes Schleswig-Holstein.
- BÖTTGER, T. (2006): Fischmonitoring in 15 schleswig-holsteinischen Seen. Gutachten Landesamt für Natur und Umwelt des Landes Schleswig-Holstein.
- BRÄMICK, U., DIEKMANN, M., LEMCKE, R. & MEHNER, T. (2008): Assessing shifts in fish assemblages of German large lakes by literature data and commercial catch statistics. *Fundamental and Applied Limnology* 171: 87-103.
- CAUSSE, S., GEVREY, M. & ARGILLIER, C. (2011a): Multimetric Index based on Fish Fauna for the assessment of the central-baltic lakes. System description CEMAGREF.
- CAUSSE, S., GEVREY, M., ARGILLIER, C. & PEDRON, S. (2011b): Fish indicators for ecological status assessment of lakes affected by eutrophication and hydromorphological pressures - final report under revision. CEMAGREF, WISER deliverable 3.4-4, Aix-en-Provence.
- CEN (2005): Water quality - sampling of fish with multi-mesh gillnets vom 08/2005. EN 14757 Seiten 36 ff.
- CIS (2003): River and lakes - Typology, reference conditions and classification systems - WG 2.3 REFCOND. Common Implementation Strategy for the Water Framework Directive (2000/60/EC) - Guidance Document, 10, Luxembourg.
- CIS (2011): Guidance document on the Intercalibration Process 2008-2011. Implementation Strategy for the Water Framework Directive (2000/60/EC) - Guidance Document, 14.
- DEGERMAN, E., HAMMAR, J., NYBERG, P. & SVARDSON, G. (2001): Human impact on the fish diversity in the four largest lakes of Sweden. *Ambio* 30: 522-528.
- DIEKMANN, M., BRÄMICK, U., LEMCKE, R. & MEHNER, T. (2005): Habitat-specific fishing revealed distinct indicator species in German lowland lake fish communities. *Journal of Applied Ecology* 42: 901-909.
- GARCIA, X.-F., DIEKMANN, M., BRÄMICK, U., LEMCKE, R. & MEHNER, T. (2006): Correlations between type indicator fish species and lake productivity in German lowland lakes. *Journal of Fish Biology* 68: 1144-1157.
- GASSNER, H. & WANZENBÖCK, J. (2005): Application of population size structure indices to Austrian whitefish (*Coregonus* sp.) stocks. *Advances in Limnology* 60: 377-384.
- GASSNER, H., TISCHLER, G. & WANZENBÖCK, J. (2003): Ecological integrity assessment of lakes using fish communities - Suggestions of new metrics developed in two austrian prealpine lakes. *International Review of Hydrobiology* 88: 635-652.
- GASSNER, H., WANZENBÖCK, J., ZICK, D., TISCHLER, G. & PAMMINGER-LAHNSTEINER, B. (2005): Development of a fish based lake typology for natural Austrian lakes >50 ha

- based on the reconstructed historical fish communities. *International Review of Hydrobiology* 90: 422-432.
- HELMINEN, H., KARJALAINEN, J., KURKILAHTI, M., RASK, M. & SARVALA, J. (2000): Eutrophication and fish biodiversity in Finnish lakes. *Verhandlungen der Internationalen Vereinigung für Limnologie* 27: 194-199.
- HOLMGREN, K. & APPELBERG, M. (2000): Size structure of benthic freshwater fish communities in relation to environmental gradients. *Journal of Fish Biology* 57: 1312-1330.
- JAARSMA, N. (2007): Description of references and metrics for fish in lakes in the netherlands (unpublished). Witteveen & Bos consulting engineers.
- JENNINGS, M. J., LYONS, J., EMMONS, E. E., HATZENBELER, G. R., BOZEK, M. A., SIMONSON, T. D., BEARD, T. D. & FAGO, D. (1999): Toward the development of an index of biotic integrity for inland lakes in Wisconsin. In: T.P. Simon (Ed.) *Assessing the sustainability and biological integrity of water resources using fish communities*, pp. 541-562. CRC Press LLC, Boca Raton.
- JEPPESEN, E., JENSEN, J. P., SONDERGAARD, M. & LAURIDSEN, T. L. (2005a): Response of fish and plankton to nutrient loading reduction in eight shallow Danish lakes with special emphasis on seasonal dynamics. *Freshwater Biology* 50: 1616-1627.
- JEPPESEN, E., JENSEN, J. P., SONDERGAARD, M., LAURIDSEN, T. & LANDKILDEHUS, F. (2000): Trophic structure, species richness, and biodiversity in Danish lakes: changes along phosphorus gradient. *Freshwater Biology* 45: 201-218.
- JEPPESEN, E., MEERHOFF, M., JACOBSEN, B. A., HANSEN, R. S., SØNDERGAARD, M., JENSEN, J. P., LAURIDSEN, T. L., MAZZEO, N. & BRANCO, C. W. C. (2007): Restoration of shallow lakes by nutrient control and biomanipulation—the successful strategy varies with lake size and climate. *Hydrobiologia* 581: 269-285.
- JEPPESEN, E., SONDERGAARD, M., JENSEN, J. P., HAVENS, K. E., ANNEVILLE, O., CARVALHO, L., COVENEY, M. F., DENEKE, R., DOKULIL, M. T., FOY, B., GERDEAUX, D., HAMPTON, S. E., HILT, S., KANGUR, K., KÖHLER, J., LAMMENS, E., LAURIDSEN, T. L., MANCA, M., MIRACLE, M. R., MOSS, B., NOGES, P., PERSSON, G., PHILLIPS, G., PORTIELJE, R., ROMO, S., SCHELSKE, C. L., STRAILE, D., TATRAI, I., WILLEN, E. & WINDER, M. (2005b): Lake responses to reduced nutrient loading - an analysis of contemporary long-term data from 35 case studies. *Freshwater Biology* 50: 1747-1771.
- LAUNOIS, L., VESLOT, J., IRZ, P. & ARGILLIER, C. (2011): Selecting fish-based metrics responding to human pressures in French natural lakes and reservoirs: towards the development of a fish-based index (FBI) for French lakes. *Ecology of Freshwater Fish* 20: 120-132.
- MEHNER, T., DIEKMANN, M., BRÄMICK, U. & LEMCKE, R. (2005a): Composition of fish communities in German lakes as related to lake morphology, trophic state, shore structure and human-use intensity. *Freshwater Biology* 50: 70-85.
- MEHNER, T., DIEKMANN, M., GARCIA, X.-F., BRÄMICK, U. & LEMCKE, R. (2004): Ökologische Bewertung von Seen anhand der Fischfauna. *Berichte des IGB* 21: 202.
- MEHNER, T., DIEKMANN, M., GARCIA, X.-F., BRÄMICK, U. & LEMCKE, R. (2005b): Möglichkeiten und Grenzen der Bewertung der ökologischen Qualität von Seen anhand der Fischfauna. In: C.K. Feld, S. Rödiger, M. Sommerhäuser & G. Friedrich (Eds.), *Typologie, Bewertung, Management von Oberflächengewässern*, pp. 137-150. E. Schweizerbart'sche Verlagsbuchhandlung, Stuttgart.
- MEHNER, T., HOLMGREN, K., LAURIDSEN, T. L., JEPPESEN, E. & DIEKMANN, M. (2007): Lake depth and geographical position modify lake fish assemblages of the European 'Central Plains' ecoregion. *Freshwater Biology* 52: 2285-2297.
- MUGV (2010): Steckbriefe: Seen der EG-Wasserrahmenrichtlinie. <http://www.mugv.brandenburg.de/cms/detail.php/bb2.c.546917.de>, abgerufen 23.08.2010.

- NEUMANN, M. (2011): Fischbiologische Bewertung von 22 schleswig-holsteinischen Seen: Ein Vergleich verschiedener Bewertungsentwürfe. Gutachten Büro Michael Neumann, Kiel.
- NIXDORF, B., HEMM, M., HOFFMANN, A. & RICHTER, P. (2004): Dokumentation von Zustand und Entwicklung der wichtigsten Seen Deutschlands. Umweltbundesamt Texte 05/04: 1056.
- PEDRON, S., DE BORTOLI, J. & ARGILLIER, C. (2010): Fish indicators for ecological status assessment of lakes affected by eutrophication and hydromorphological pressures - provisional report. CEMAGREF, WISER deliverable 3.4-4, Aix-en-Provence.
- PERSSON, L., DIEHL, S., JOHANSSON, L., ANDERSSON, G. & HAMRIN, S. F. (1991): Shifts in fish communities along the productivity gradient of temperate lakes - patterns and the importance of size-structured interactions. *Journal of Fish Biology* 38: 281-293.
- RITTERBUSCH, D. & BRÄMICK, U. (2010): Praxistest Seenbewertung sowie Interkalibrierung Seenbewertung für Fische. Endbericht Institut für Binnenfischerei e.V., O 2.09, Potsdam-Sacrow.
- RITTERBUSCH, D., BRÄMICK, U., SCHUBERT, M., SCHÜTZ, C. & DUBLING, U. (2011): Interkalibrierung der Fließgewässer- und Seenbewertung anhand der Fische, Praxistest Seenbewertung sowie Leitung der Central Baltic GIG für Seen. Endbericht Institut für Binnenfischerei e.V., O 14.10, Potsdam-Sacrow.
- SONDERGAARD, M., JEPPESEN, E., JENSEN, J. P. & AMSINCK, S. L. (2005): Water Framework Directive: ecological classification of Danish lakes. *Journal of Applied Ecology* 42: 616-629.
- STAAS, S., ROCHOL, F. & SCHARBERT, A. (2006): Vorbereitung eines Monitorings der Fischfauna größerer Seen des Landes Schleswig-Holstein gemäß EU-WRRL - Teil B: Begleitende Elektrofischungen. Untersuchung LimnoPlan, Erfstadt.
- TAMMI, J., LAPPALAINEN, A. & RASK, M. (2001): Using Swedish fish index FIX in assessing degradation of Finnish eutrophic lakes - what does fish community data tell about them? Nordic Council of Ministers, Copenhagen.
- TAMMI, J., APPELBERG, M., BEIER, U., HESTHAGEN, T., LAPPALAINEN, A. & RASK, M. (2003): Fish status survey of Nordic lakes: Effects of acidification, eutrophication and stocking activity on present fish species composition. *Ambio* 32: 98-105.
- WHITTIER, T. R. (1999): Development of IBI metrics for lakes in southern New England. In: T.P. Simon (Ed.) *Assessing the sustainability and biological integrity of water resources using fish communities*, pp. 563-582. CRC-Press.
- WRRL (2000): Richtlinie 2000/60/EG des Europäischen Parlaments und des Rates zur Schaffung eines Ordnungsrahmens für Maßnahmen der Gemeinschaft im Bereich der Wasserpolitik vom 23.10.2000. ABl. L 327.
- ZAHN, S. & BORKMANN, I. (2006): Bestandserhebungen der Fischfauna in ausgewählten Fließgewässern und Seen des Landes Brandenburg. Endbericht Institut für Binnenfischerei e.V., Potsdam.
- ZAHN, S. & BORKMANN, I. (2007): Bestandserhebungen der Fischfauna in ausgewählten Fließgewässern und Seen des Landes Brandenburg. Endbericht Institut für Binnenfischerei e.V., Potsdam.
- ZAHN, S. & BORKMANN, I. (2008): Bestandserhebungen der Fischfauna in ausgewählten Fließgewässern und Seen des Landes Brandenburg. Endbericht Institut für Binnenfischerei e.V., Potsdam.
- ZICK, D., GASSNER, H., FILZMOSER, P., WANZENBÖCK, J., PAMMINGER-LAHNSTEINER, B. & TISCHLER, G. (2006): Changes in the fish species composition of all Austrian lakes >50 ha during the last 150 years. *Fisheries Management and Ecology* 13: 103-111.

ANHANG

SITZUNG 1/2011 DES VDFF-AK „FISCHEREILICHE GEWÄSSERZUSTANDBEWERTUNG“ AM 19./20.APRIL 2011 IN WÜRZBURG: PROTOKOLLAUSZUG TOP 1 SEEBEWERTUNG

Sachstand und Praxistest Bewertungsverfahren

Es sind zwei Bewertungsansätze verfügbar: SITE (Bewertung auf Basis von Fangstatistiken der Berufsfischerei) und TYPE (Bewertung anhand Befischungen mit Multimaschennetzen nach CEN-Standard). Das Konzept TYPE wurde geprüft und ein Excel-Werkzeug erstellt; gemeinsam mit Österreich erfolgten erste vergleichende Bewertungen. Aus dem AK liegen bisher keine Rückmeldungen zum Praxistest vor. Es ist eine Überarbeitung der Excel-Vorlage geplant, diese sollte bei ev. geplanten Arbeiten abgewartet werden.

Praxistest SITE

Im alpinen Bereich (16 Seen) wurde eine sehr gute Übereinstimmung mit den Ergebnissen des österreichischen Systems festgestellt. Beim Vergleich beider Verfahren wurden allerdings nur drei Zustandsklassen (mäßig, gut, sehr gut) betrachtet; es fehlten Seen in schlechtem Zustand. Die Validierung erfolgte im Vergleich mit einer Experteneinschätzung der Seen. Bei den untersuchten 16 Seen trat hinsichtlich der korrekten Identifizierung des Maßnahmebedarfs nur ein kritischer Fehler auf. Als Problem stellte sich die Übersetzung der Abundanzklassen in konkrete Werte heraus; es wird um Vorschläge zur Lösung dieses Problems gebeten.

Im Tiefland (Central/Baltic) umfasste der Praxistest nur 8 Seen aus Brandenburg. Es wurden 5 kritische Fehler beobachtet, SITE arbeitet hier sehr unbefriedigend. SITE wurde für Großseen entwickelt und zeigt dort gute Ergebnisse. Da der Stichprobenumfang zu gering war, soll der Praxistest fortgesetzt werden. Es sollen genaue Maßgaben erarbeitet werden, für welche Gewässer SITE geeignet ist.

Als Fazit wurde festgestellt, dass SITE im alpinen Bereich gut anwendbar ist. Im Tiefland ist die Anwendbarkeit derzeit unklar, das Verfahren muss weiter geprüft werden.

Diskussion

- Umgang mit ausgestorbenen Arten: Diese sind nicht zur Indikation des aktuellen Status geeignet. Starke Auswirkungen auf die Bewertung resultieren jedoch in Fällen, wenn mit ausgestorbenen Arten auch ganze Gilden entfallen; der Bewertungsalgorithmus soll dahingehend überarbeitet werden.
- Qualität und Sicherheit der Experteneinschätzung als Referenz: In der Diskussion wurde die Nutzung anderer Wege zur Objektivierung als notwendig erachtet (z.B. andere Qualitätskomponenten, chem. Parameter?). Für den vorgestellten Praxistest wurde die Expertenmeinung erfragt, außerdem lagen für die alpinen Gewässer alle Informationen zu den übrigen Qualitätskomponenten vor.
- Einfluss der komplexen und variablen Fischgemeinschaften in Tieflandseen: Resultieren die festgestellten Probleme bei der Anwendung von SITE im Tiefland aus Fehlern bei der Referenzerstellung? Möglicherweise ist die Referenzsituation bei alpinen Gewässern besser definiert.
- Der vorgestellte Vergleich des deutschen und österreichischen Verfahrens ist noch nicht als Interkalibrierung anzusehen; es müssen weitere Verfahren einbezogen werden (in Vorbereitung: u.a. Frankreich).

Praxistest TYPE

Der Praxistest ergab für natürliche Seen überwiegend eine gute Übereinstimmung der Experteneinschätzung mit den Ergebnissen des TYPE-Konzeptes; hier arbeitet das Verfahren gut.

Die bisher verwendete Seentypisierung (3 Typen) lässt sich nicht auf Sonderfälle wie künstliche Gewässer (Stauseen, Talsperren, Tagebaurestseen) und Flusseen übertragen, hier ist TYPE bisher ungeprüft und somit derzeit nicht anwendbar. Für die Bewertung dieser sowie weiterer Sonderfälle (z.B. Salzeinfluss) fehlt bisher die Datengrundlage ganz bzw. ist unzureichend (Stauseen: N = 2; Flusseen: N = 2). Es sind daher weitere Arbeiten zur Entwicklung und Prüfung des Verfahrens notwendig.

Hinsichtlich der Interkalibrierung des Verfahrens fanden bisher lediglich erste Arbeiten statt; die eigentliche Interkalibrierung steht noch aus. Dies setzt jedoch die Entwicklung weiterer Verfahren in den Mitgliedsstaaten voraus. Da dazu entsprechende Arbeiten laufen bzw. geplant sind, erscheint eine Interkalibrierung perspektivisch möglich, kann aber 2011 nicht abgeschlossen werden. Laut ECOSTAT kann dabei auf den Parameter Altersstruktur verzichtet werden.

Obwohl derzeit die weitere Entwicklung unklar ist, sollte die Fischfauna Element der Seenbewertung bleiben. Die Fischfauna bildet die Gesamtbelastung des Gewässers ab und stellt eine integrative Messgröße dar, daher wird eine direkte Reaktion auf einzelne Belastungen nicht erwartet. Die bisherigen Ergebnisse zeigen eine Reaktion auf die Gewässertrophie, nicht jedoch auf andere Belastungen wie Störungen durch Nutzung und Uferverbau.

Diskussion

- Bewertung von Talsperren u.a. künstlichen Gewässern: Es wurde festgestellt, dass die Bewertung dieser Gewässer aufgrund der überwiegend durch Besatz geprägten Fischgemeinschaften und des unklaren Referenzzustandes sehr große Probleme aufwirft und eine Bewertung daher fragwürdig bzw. nicht möglich ist. Dies sollte ggf. in solchen Fällen klargestellt werden. In Nordrhein-Westfalen und Sachsen werden Talsperren daher derzeit nicht bewertet. In Bayern zählen Talsperren nicht zu den künstlichen (= „vom Menschen neu geschaffenen“), sondern zu den erheblich veränderten Wasserkörpern. Eine Bewertung erfolgt nur in HMWB-Seen der Überblicksüberwachung
- Trotz der og. Schwierigkeiten wird die grundsätzliche Berücksichtigung der Fischfauna in der Überblicksüberwachung künstlicher Standgewässer weiterhin als erstrebenswert und wichtig angesehen.
- TYPE und SITE erfordern ähnlich fiBS eine fachkundige Evaluierung und Qualitätssicherung des Bewertungsergebnisses
- Verzicht auf die Altersstruktur bei der Interkalibrierung: Es wurde festgestellt, dass diese Position von ECOSTAT (s.o.) nicht den Vorgaben der WRRL entspricht.

Zum Verhältnis von Fischerei und WRRL vertritt der AK folgende Position:

1. Die fischereiliche Bewirtschaftung (Fischerei und Besatz) ist grundsätzlich keine Maßnahme zur Erreichung des guten ökologischen Zustandes im Sinne der WRRL.
2. Einen Sonderfall stellen Gewässer dar, die allein aufgrund verschollener Fischarten das Ziel der WRRL verfehlen, sich ansonsten aber bereits in einem guten ökologischen Zustand befinden. Hier ist die Wiederansiedelung der verschollenen Arten grundsätzlich wünschenswert. Derartige Maßnahmen sollten sich hinsichtlich der Besatzeempfehlung an der Referenz orientieren.
3. Die fischereiliche Bewirtschaftung kann Einfluss auf das Bewertungsergebnis haben. Wenn im Ergebnis einer Einzelfallprüfung eine starke fischereibedingte Verzerrung der

Fischgemeinschaft festgestellt wird, sollte keine Bewertung vorgenommen werden. Dies gilt auch bei relevanten Effekten fischfressender Vögel wie z.B. Kormoran.

4. Bei Einhaltung einer guten fachlichen Praxis ist die Fischerei kein „pressure“ im Sinne der WRRL.

Intrakalibrierung und praktischer Einsatz von SITE und TYPE

Die Datengrundlage ist bisher unzureichend, es sind mehr Vergleiche erforderlich. Die Ergebnisse des Verfahrensvergleichs sind unbefriedigend, nur in zwei Fällen stimmten die Ergebnisse der beiden Verfahren überein. TYPE zeigte dabei tendenziell die bessere Übereinstimmung mit der Experteneinschätzung. Möglicherweise könnte das TYPE-Verfahren auch mit einem verringerten Befischungsaufwand („halber CEN_Standard“) auskommen; diese Frage ist jedoch noch offen. Für den Einsatz beider Verfahren sollte eine klare Handlungsanweisung erarbeitet werden. SITE arbeitet in Alpenseen sehr gut, der Einsatz im Tiefland ist jedoch problematisch. TYPE ist ungeeignet für Alpenseen, kann jedoch gut in natürlichen Seen des Tieflandes eingesetzt werden. Die Übertragbarkeit des Verfahrens auf Talsperren, Flusseen und andere Sonderfälle ist derzeit ungeklärt.

Diskussion

- Die routinemäßige Anwendbarkeit von TYPE erscheint aufgrund des sehr hohen methodischen Aufwandes und der hohen Kosten fraglich. In Schleswig-Holstein kann daher nur der SITE-Ansatz im operativen Monitoring verfolgt werden, und in der überblicksweisen Überwachung bestehen Schwierigkeiten bei der Akzeptanz der ansässigen Erwerbsfischer bzw. Fischereiberechtigten. In Bayern ist dieses Verfahren zudem aufgrund der verlangten Befischungen praktisch nicht genehmigungsfähig.
- Probleme ergeben sich für die Modellierung der Referenzen (SITE) aus der nicht durchgängig vorhandenen Datengrundlage.
- Es besteht weiterer Bedarf für die Entwicklung der Seenbewertung; daher sollte 2011 ein Antrag an die LAWA gestellt werden. Eine Bedarfsermittlung für die Bewertung der Fließgewässer mit fiBS ergab für die Fortführung der Interkalibrierung keinen Bedarf (Arbeiten sind abgeschlossen). Es wurde jedoch ein Projekt zur Vorbereitung einer zentralen Softwarelösung für eine datenbankgerechte Umsetzung von fiBS als sinnvoll angesehen. Für beide Inhalte soll daher ein gemeinsamer Antrag gestellt werden.

Arbeiten 2011

- Fortführung des Praxistests
- Einbeziehung weiterer Tieflandseen (SITE)
- Analyse der Variabilität des Ergebnisses z.B. in Abhängigkeit vom Befischungsaufwand.
- Überarbeitung der Excel-Vorlagen (eine neue Kurzdarstellung wird erst nach dem Abschluss der Arbeiten erstellt)
- Fortführung der Interkalibrierung
- Einreichung eines Projektantrages an die LAWA

SITZUNG 2/2011 DES VDFF-AK „FISCHEREILICHE GEWÄSSERZUSTANDSBEWERTUNG“ - AM 25./26. OKTOBER 2011 IN DRESDEN: PROTOKOLLAUSZUG SEEBEWERTUNG

Zum Zeitpunkt der Abgabe des vorliegenden Berichtes lag das Sitzungsprotokoll noch nicht vor. Der nachfolgende Auszug ist daher ein inoffizielles Sitzungsprotokoll des Verfassers:

Stand der Seebewertung - Diskussionspunkte

- Die Entwicklung des deutschen Verfahrens ist abgeschlossen.
- Die Plausibilitätsprüfung ist zufriedenstellend.
- Nächster Schritt in der Bewertung von Seen mit Fischen wäre der Praxistest, d.h. die Anwendung in den Bundesländern.
- Ob und in welchem Umfang die Anwendung stattfinden wird, entscheiden die Länder.
- Es sollte erarbeitet werden, welchen Mehrwert die Bewertung mit Fischen gegenüber der Bewertung mit anderen BQE liefert.
- Es sollte erarbeitet werden, welche Handlungshinweise aus dem Bewertungssystem abgeleitet werden können.

Kommentar D. Ritterbusch

Zu 5) Der Mehrwert der Fische liegt in der Anzeige eines Gesamtzustandes gegenüber einer eher definierten Trophieanzeige (Phytoplankton) oder Strukturanzeige (Benthos). Eigentlich stellt sich die Frage nicht, da die BQE nicht gegeneinander aufgewogen werden können.

Zu 6) Fische sind nicht als direkte Stellvertreter für Belastungsmessungen geeignet. Diese Vorgehensweise ist auch nicht im Sinne der WRRL. Handlungsempfehlungen lassen sich indirekt durch fachliche Analyse der Bewertungsergebnisse ableiten, z.B. sind Blei und Zander Eutrophierungsanzeiger, Barschanteile oder fehlende Arten wie Hecht und Rotfeder zeigen strukturelle Defizite. Überdies korrelieren die verschiedenen Belastungen stark miteinander (z.B. Eutrophierung, Uferverbau, Nutzungsintensität) so dass ihre Wirkungen nicht ohne weiteres getrennt werden können.

Interkalibrierung

- Die Interkalibrierung im Alpen Bereich wird als abgeschlossen betrachtet, dies ist aber noch nicht von der EU bestätigt.
- Die Interkalibrierung im CB GIG ist nicht abgeschlossen.
- In der CB GIG wurden mehrere Verfahren September/Oktober 2011 fertiggestellt und könnten jetzt interkalibriert werden.
- Die CB Interkalibrierung ist sehr schwierig, weil pressure-impact-Beziehungen fehlen, keine einheitliche Interpretation der Bedeutung von pressures vorliegt, dementsprechend benchmark-sites nicht existieren und common metrics bisher nicht zufriedenstellend sind.
- Dennoch wollen die Mitgliedsstaaten die Interkalibrierung fortführen.
- Ob die Interkalibrierung generell und speziell die Central Baltic-GIG Leitung durch Deutschland auch nach Phase II fortgeführt werden, ist Entscheidung der LAWA.
- Der AK befürwortet die Fortführung; die Interkalibrierung ist wichtiger Bestandteil der WRRL. Ggf. sollte gewartet werden, bis plausible Bewertungssysteme vorhanden sind.

Kommentar D. Ritterbusch: Wenn man die Interkalibrierung fortführen will, sollte eine Unterbrechung vermieden werden. Es ist unwahrscheinlich, dass die Leitung von einem anderen Land übernommen wird, wenn Deutschland die Leitung der CB GIG abgibt. Damit stehen dann auch Interkalibrierung und internationale Zusammenarbeit in Frage. Zudem ist fraglich, ob die Interkalibrierung wieder gestartet werden kann, wenn die Kommunikation einmal unterbrochen wird. Zumindest sind erhebliche Verzögerungen zu erwarten.

CB LAKEFISH MILESTONE 6 REPORT

Water category/GIG/BQE/ horizontal activity:	Lake / Central-Baltic / Fish /
Information provided by:	David Ritterbusch (Germany)

1. Organisation

1.1. Responsibilities

Indicate how the work is organized, indicating the lead country/person and **the list of involved experts of every country:**

DE	Germany (CB lead)	David Ritterbusch / Uwe Brämick
FR	France (cross-GIG coord.)	Christine Argillier / Simon Causse
BE-F	Belgium-Flanders	Jan Breine
CZ	Czech Republic	Marie Prchalova/Jan Kubecka (observational)
DK	Denmark	Torben Lauridsen
EE	Estonia	Teet Krause / Anu Palm
LT	Lithuania	Tomas Virbickas
LV	Latvia	Janis Birzaks
NL	Netherlands	Eddy Lammens / Nico Jaarsma / Bob Brederveld
PL	Poland	Hanna Draskiewicz-Mioduszevska, Witold Bialokoz
SK	Slovakia	Vladimír Kováč (observational)
UK	United Kingdom	Willie Duncan

1.2. Participation

Indicate which countries are participating in your group. Are there any difficulties with the participation of specific Member States? If yes, please specify:

All MS listed in 1.1 participate, except of Latvia, where contact was lost in early 2011. Czech Republic participates as an observational partner but actively contributes with data and advice. Slovakia is an observational member, too.

1.3. Meetings

List the meetings of the group:

2008: 31.03./01.04. 1st LakeFish meeting (Aix-en-Provence / France)
2008: 30.09./01.10. 2nd LakeFish meeting (Ranco / Italy)
2009: 22.09./23.09. 3rd LakeFish meeting (Drottningholm / Sweden)
2010: 30.11./01.12. 1st CB Lake Fish meeting (Berlin / Germany)

2. Overview of methods to be intercalibrated

Identify for **each** MS the national classification method that will be intercalibrated and the status of the method (national / intercalibrateable / under development / no method).

MS	status	expected for	sampling gear
NL	national	present	Trawl, seine, electrofishing
DE	intercalibrateable	2011	CEN / CENmod
DK*	intercalibrateable	2011	CEN
LT	intercalibrateable	2011	CENmod
EE	intercalibrateable	2012	CENmod
PL	intercalibrateable	2011	LFI+: Fisheries statistics
	u. d./ intercalibrateable	2013	CEN
BE	under development	2012	fyke + electrofishing
FR	intercalibrateable	2011	CEN
LV	under development	unclear	gillnets, trammel nets, statistics
CZ	no method	unclear	CEN, hydroacoustics, electrofishing
SK	no method	not expected	-
UK*	no method	unclear	

DK: expects a finalized and approved method within 2011. Method is intercalibrateable, but comparison could not be done in due time for Milestone 6. It is suggested to wait with the further intercalibration until the official confirmation is obtained.

UK: England and Wales are part of the CB GIG. The Irish LakeFish system and the CB methods will be checked for applicability. Gillnet fishing is not an option because of public relation issues.

Sampling gear was added to the table, because sampling is essential for comparability.

3. Checking compliance of national methods with the WFD requirements

Compliance criteria	Compliance checking conclusions
1. Ecological status is classified by one of five classes (high, good, moderate, poor and bad).	Yes for all MS
2. High, good and moderate ecological status are set in line with the WFD's normative definitions (boundary setting procedure)	ND: Yes for all MS BSP: No for most MS <u>See comment on pressures, p.</u>
3. All relevant parameters indicative of the biological quality element are covered. If parameters are missing, Member States need to demonstrate that the method is sufficiently indicative of the status of the QE as a whole.	Tax. composition: Yes for all MS Abundance: Yes for all MS Sensitive taxa: - Yes for BEF, EE, FR, LT, NL, PL - No for DE, DK - unknown for others Age: direct - no for all MS indirect - yes for most MS <u>See comment on age, p.</u>
4. A combination rule of parameters into assessment BQE is defined.	Yes for all MS
5. Assessment is adapted to intercalibration common types that are defined in line with the typological requirements of the WFD Annex II and approved by WG ECOSTAT	Yes: FR, NL No: BE WFD but not approved: other MS CB LakeFish typology is developed and adopted by DE, LT
6. The water body is assessed against type-specific near-natural reference conditions	Yes for most MS. FR uses site specific modelling.
7. Assessment results are expressed as EQRs	Yes for all MS.
8. Sampling procedure allows for representative information about water body quality/ ecological status in space and time	CEN and trawl fishing information is representative in space (whole lake). Temporal representativeness is under discussion/investigation Yes for time / space for Polish LFI+
9. All data relevant for assessing the biological parameters specified in the WFD's normative definitions are covered by the sampling procedure	Taxonomic composition: Yes Abundance: Yes Sensitive species: No for CEN, Yes for multiple gear (NL, PL) Age: indirectly, <u>see comment p.</u>
10. Selected taxonomic level achieves adequate confidence and precision in classification	Yes for all MS

Clarify if there are still gaps in the national method descriptions information and summarize the conclusions of the compliance checking:

Comprehensive English descriptions exist for some MS (DE, FR, NL). Short descriptions have been submitted by other MS. The descriptions will be made available at the CIRCA forum of the JRC as soon as they are in a publishable condition (folder Lakes-fish intercalibration/CB GIG/CB method descriptions).

The aim of the LakeFish assessment systems is to provide information on the ecological status of lakes based on a 5-step scale. General WFD compliance is present in the existing systems and will be given for systems to be developed. The aim of intercalibration process is to make the assessment results comparable. Minor deviations from WFD prerequisites do not hinder from intercalibration.

One deviation is the lack of a ‘real’ age metric in all national systems. In all CB systems, age is substituted by length or weight parameters. Scientific justification was provided by the CB GIG lead (see annex). Based on this, ECOSTAT decided to include national methods in the current intercalibration exercise even if they do not contain age structure metrics (VAN DE BUND et al. 2011).

4. Methods’ intercalibration feasibility check

4.1. Typology

Describe common intercalibration water body types and list the MS sharing each type

Common IC type	Type characteristics	MS sharing IC common type
POLY	Polymictic lakes	Most MS
STRAT	Stratified lakes	Most MS
DEEP	Stratified lakes > 30 m max. depth	Most MS

Conclusion:

A common CB IC typology has been developed (RITTERBUSCH et al. 2010)⁴ and was accepted for use in the CB intercalibration process (RITTERBUSCH 2010). IC Typology and national typologies are compatible. The IC typology covers most of the lakes.

Different typologies are used on the national level and FR developed a site-specific approach without typology. However, no difficulties in the comparison have appeared until now, that would have been caused by different typologies.

Only some special lake types can not be intercalibrated (naturally acidified, influenced by humic substances or marine water). The number of lakes assigned to these special types is too small to establish a systematic assessment of status.

⁴ Both documents are available on CIRCA/JRC - Lakes -fish intercalibration

4.2. Pressures

Describe the pressures addressed by the MS assessment methods.

MS	pressures addressed
BE-F	biotic integrity (habitat quality, water quality)
DE	general
DK	eutrophication
EE	eutrophication
FR	eutrophication, general degradation
LT	eutrophication
NL	eutrophication, water level regulation
PL	general and eutrophication
CZ, LV, SK, UK	-

Conclusion

Fish in lakes can be used to assess many pressures (eutrophication, acidification, deterioration, water level fluctation). Fish communities reflect the lake status as stipulated by the WFD. However, fish are a BQE that integrates pressures over time and space. They are less useful to describe individual pressure intensities (please see ‘Comment on the suitability ...’ in the Annex). For this reason, intercalibration processes can not rely on good pressure-impact correlations. Therefore, the future harmonization process might differ from the proposals in the IC Guidance. A scientific justification was provided by the leader of the Alpine and CB GIG (see annex). Based on these arguments, ECOSTAT decided to agree that MS will demonstrate a reaction of the total assessment result to the combined intensity of all pressures (if necessary, using expert judgment) and clearly document it in the reports (VAN DE BUND et al. 2011).

4.3. Assessment concept

Do all national methods follow a similar assessment concept?

Conclusion:

The methods are aware of the integrating indicating value of the fish community and assess the status of the fish community of the lake as a whole (including littoral, benthic and pelagic fish). All systems are based on the comparison of the current status with a reference condition. By now, the intercalibration seems feasible in terms of assessment concepts.

4.4. Sampling comparability

Conclusion:

Many MS follow the CEN 14757 multimesh gillnetting fishing procedure with randomized stratified setting of nets. The following methods are used

- DE, DK, EE, FR follow the CEN 14757 exactly (more or less)
- LT excluded small mesh sizes
- BE-F, NL and PL use other methods (like trawl, fishers statistics, see section 2.).

Data based on different gear is incomparable because:

- a) methods sample different habitats,
- b) active (trawl) and passive gears (nets) have different species-specific effectiveness,
- c) selectivity of gears to species or size-classes can not be converted to other gear,
- d) most systems work with percentages, deviation in one size/species impacts others,
- e) the evidence of some species is restricted to certain methods (e.g. littoral species).

4.5. Boundary setting / comparison and harmonization in common IC type

Boundaries were set at national level.

The boundary setting procedure is heterogeneous and differs between MS, lakes, or even metrics of individual systems. Some possibilities are:

- regression lines, discontinuities (NL, DE)
- definition of H/G boundary and consequent equidistant division (NL, FR)
- value distributions, discontinuities (DE)

In most cases, national expert judgment is included in the class boundary setting. There are no common agreements on abiotic parameters and threshold values representing H/G or G/M boundaries. In many cases pressure-impact relationships are of limited use to set class boundaries in LakeFish assessment (see comment).

More detailed descriptions will be provided in a more advanced state of intercalibration.

4.6 Conclusions

Typology, reaction on pressures and system concepts are of minor importance for the intercalibration process. Main problem is the incomparability of results obtained with different fishing gear. This requires the use of intercalibration option 2, i.e. the development of common metrics / pseudo common metrics. Until now, the search for a suitable common metric had no convincing result. The only way to apply a pseudo common metric is parallel sampling. This was done by NL and will be done by PL. However, results are not clear until now.

5. Collection of IC dataset

Describe data collection within the GIG.

The common dataset for the LakeFish IC process is hosted at the CEMAGREF (France, cross-GIG coordination). It is used as WISER database and includes data from European Member States that have done fishing according to CEN 14757 in natural lakes. For the CB-Gig the MS are: CZ, DE, DK, EE, FR, LT, LV. New data obtained throughout 2009/2010 is imported into the database. Only CEN 14757 data was summarized, so all data acceptance criteria are fulfilled. Concerning the IC process, a limited part of the dataset is really comparable. Reasons might be the lake type (e.g. the Czech Republic has anthropogenic reservoirs, but no natural lakes), modifications of the CEN procedure (exclusion of smaller mesh sizes by LT, two campaigns with ½ of nets instead of full standard for DE). The database includes physic-chemical and pressure data for the lakes. For detailed information on the data please refer to the host.

Within the GIG, there is a collection of an IC dataset with information on assessment results. It includes EQR values and status classes for both individual metrics and whole systems of the CB common metric, the German assessment method and the national systems. Additionally, expert's estimations of the ecological status of the lakes reflecting the total pressure intensity are collected.

MS	Data provided to the common CEMAGREF-database
Czech Republic	Yes - CEN 14757 (HMWB/AWB)
Denmark	Yes - CEN 14757
Estonia	Yes - CEN 14757
France	Yes - CEN 14757
Germany	Yes - CEN 14757
Latvia	Yes - CEN 14757
Lithuania	Yes - CEN 14757
Poland	CEN 14757 - not yet achievable
Netherlands	Some new CEN 14757, to be included in the DB soon
Belgium (Flan.)	No CEN 14757 data
Slovakia	No data
UK	CEN 14757 (with reduced effort)

6. Benchmarking: Reference conditions or alternative benchmarking

Clarify if you have defined

- common reference conditions → **No**
- or a common alternative benchmark for intercalibration → **Not yet**

Lakes in true reference condition concerning non biological parameters are absent or nearly absent in the CB GIG. Therefore most MS use other methods than abiotic reference sites to define the fish community in HIGH status (e.g. expert's definition, use of LDC sites, historical data, for France also WISER hindcasting method (CAUSSE et al. 2011b)).

There was a huge and enduring disagreement in the cross-GIG LakeFish experts group about the relevance of the REFCOND criteria for defining reference condition. The proposals of the WISER/CEMAGREF team were also discussed controversially. Therefore, an alternative solution was found using a modelling procedure that was not based on reference sites. In summary, both 'true reference' and LDC sites defined with abiotic parameters seem to be unsuitable to benchmark status equivalents for fish in lakes on a CB range because of their strong dependency on lake type and biogeographical location.

Additionally benchmarking with the use of a pressure-impact gradient is not a suitable intercalibration option for fish in lakes. They would be necessary for the 'official' IC spreadsheets, for example. Correlations between pressure intensities and fish community measures are weak. In lakes, Fish are mostly insensitive to individual pressures but react on the sum of different pressures (see comment on the suitability of fish to address specific pressures). This makes benchmarking even more difficult.

For future use in the IC process, expert judgments on the total ecological status are collected for use as a benchmarking system.

First experiences have shown that the comparison of assessment results with expert judgments of the ecological status of the corresponding lakes can be used to

- 1) estimate the reliability of assessment systems and
- 2) compare, evaluate and synchronize the results of different assessment systems.

This procedure was accepted by ECOSTAT (van de Bund et al. 2011), but the practical use is still under investigation.

7. Design and application of the IC procedure

7.1. Please describe the choice of the appropriate intercalibration option.

Which IC option do you plan to use?

IC Option 1 - same method:

The GIG lead tried to develop a common CB method during 2010. It was designed for MS applying the CEN 14757 multimesh method or slightly modified sampling. However, the results were not satisfying. The huge biogeographical differences within the GIG lead to the fact that individual metrics are not transferable to the whole range of MS, e. g. Bream is a reliably indicator of eutrophication in central CB regions, but naturally absent in the eastern part. The incomparability is even more pronounced for class boundaries (RITTERBUSCH 2010). Option 1 is not applicable.

IC Option 2 - Different data acquisition and numerical evaluation:

Most of the MS use different fishing methods (see section 2). Even modifications of a similar method lead to an incomparability of the results, both for absolute values per unit effort and for percentages (e.g. exclusion of small mesh sizes). Knowledge on the difference of methods is present, but there is no way to translate results of one sampling procedure into another. Ecological relevance and class boundaries of metrics are gear specific.

All MS have different procedures of fish data evaluation.

Principally, option 2 would be the appropriate choice of intercalibration. However, it was not possible to identify a suitable common metric until now.

Option 2 is not applicable.

IC Option 3 - Similar data acquisition, but different numerical evaluation:

A small subset of MS applies the CEN 14757 procedure in the standardized way (CZ, DE⁵, DK, EE, FR). First steps in the IC option 3 have been done. A common metric was proposed by Germany and was calculated for these countries. Comparisons of national assessment results with this common metric, the German system and expert's estimations of lakes status show different degrees of correlation. The relationships between national methods and common metric are prevalingly unacceptable. As things stand now, option 3 might be useful for some MS in geographical vicinity.

More recently, the index developed in the WISER project has reached a condition that it's suitability as common metric can now be tested. Unfortunately, this was reached too late to be included in the present Milestone 6 report.

Option 3 can become useful to intercalibrate within the central CB region: DE, DK, FR and later on PL. NL could be added via direct comparison using parallel sampling results of trawl and CEN fishing.

Option 3 could be applicable for a number of neighboured MS. The WISER results may further improve the use of option 3.

7.2. IC common metrics (if IC Options 2 or 3 are used)

Describe the IC common metric:

⁵ Using 2 times ½ CEN efforts is regarded as comparable, although not strictly following the standard.

There are three kinds of common metrics:

- a) single metrics selected of national methods or generated for the IC process
- b) common metric made of combined metrics
- c) a pseudo common metric, i.e. the mean assessment result of other methods

a) Was avoided because fish in lakes usually show weak correlations between parameters and pressures. It was expected that single metrics have no acceptable correlation to both national systems and pressure-intensities. LakeFish assessment systems always base on multiple metrics to provide acceptable results.

b) Following the discussion about the CB system, 4 suitable common metrics were selected. With this selection, a common metric mini-assessment system was developed for the CB GIG. The development was based on the WISER database. It was designed for the CEN 14757 sampling and based on the following metrics:

- WPUE
- Bream + Roach
- Benthic species reacting on eutrophication
- Species sensitive to the integrity of the shoreline

The common metric mini-system shows the same problems as the preceding complete system proposal: no reaction on pressures, limited applicability to metrics to the huge biogeographical range of fish communities, no comparability of EQR values and class boundaries between MS.

c) The pseudo common metric is applicable only, if multiple assessment systems can be applied to the same dataset. As already mentioned, the direct comparison is useful only for few MS in close vicinity.

7.3. Strategy

Status quo

The BQE Fishes in Lakes, Central Baltic GIG has achieved great progress during the intercalibration Phase II. Only one system with official status is present in the GIG (NL), but other intercalibrateable systems have been finalized recently (DE, DK, EE, FR, LT). Two additional systems will be ready for intercalibration in 2011 / 2012 (BE-F, PL).

The intercalibration process will not be finished within 2011. All member states have done considerable efforts to develop national assessment methods with the result that most of them expect to finalize their system within 2011/2012 (see section 2). Thus, the situation has greatly improved from only one finalized method at the beginning of the LakeFish IC Process in 2008 (The Netherlands) to currently roundabout six methods that can be compared.

Reasons for delay

Most of the CB methods were not ready to be compared until end of September 2011. Therefore, first comparisons could be done (see Annex 3), but real intercalibration processes in terms of class boundary comparisons and adjustments are still missing. The reasons for the delay are unclear. Despite its suitability to indicate the ecological status of lakes the BQE Fish is far behind in (nearly) all European MS. While other BQE started intercalibration in 2004, the CB LakeFish intercalibration started in 2008 with only one working system. The BQE fish fauna has some inherent properties that add further problems beside the delay of system development.

First of all, there seems to be focus on other Water Bodies/BQE in most MS. Manpower and money invested in LakeFish systems is small compared to other elements. Eventually the difficulties of LakeFish assessment have been foreseen and the initial focus of the WFD compliant system development was set to more easy/obvious biota. The existent datasets of fish are small, long time series are missing.

Methodological reasons are for example, that no agreement on the threshold values of pressures reflecting the reference conditions could be found. Not even the significance of pressures for fish was undisputed and the proposals of the REFCOND guidance were criticized. An important biological reason is the weak pressure-impact reaction of specific fish community traits. Fish are integrating in space and time and can avoid local or limited pressures. Pressure-impact relationships are the basis of most of the IC options proposed in the current guidance. Additionally, the resilience of the fish community is high. Therefore single metrics are unreliable indicators of pressures. To obtain a good performance, multimetric indices are needed, which are more stable if more metrics are used. This makes a single or reduced common metrics unusable. The huge geographical range of the CB GIG leads to very different fish communities, in some cases important indicator species are naturally missing.

We have to highlight that the existing typology are not able to take into account the entire natural variability of the natural environment of the CB lakes. As it is shown on the following figures (principal component analyses), there are large differences among the main environmental characteristics of the lakes in the different countries. As a consequence, the intercalibration process will probably be difficult because of the specificity of each national tool.

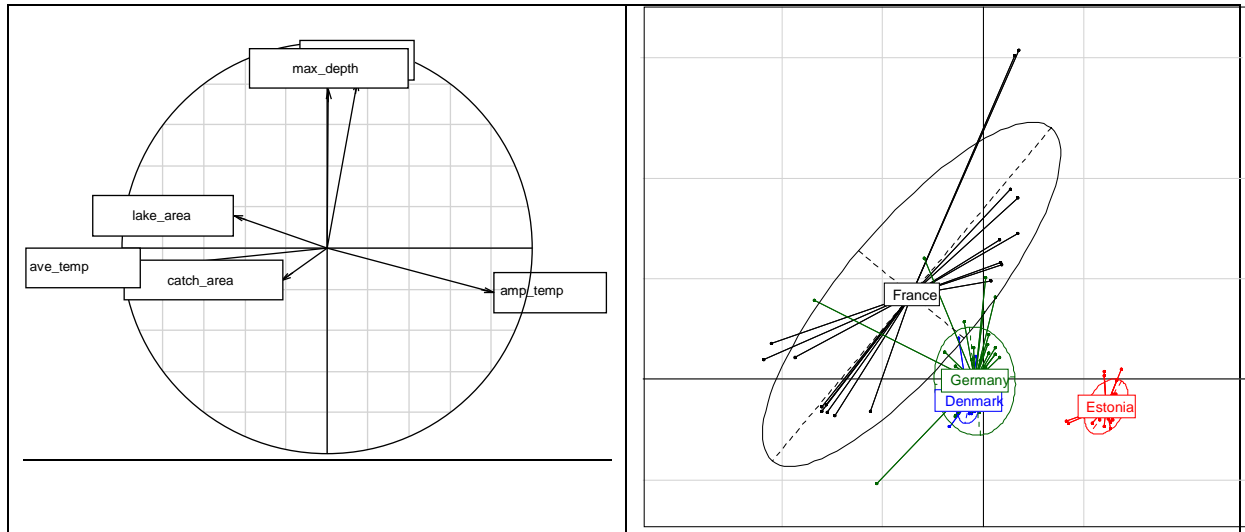


Fig. 1: Principal Component Analyses (PCA) performed on natural parameters of all CB lakes (a) Correlation circle for environmental parameters (amp_temp: temperature amplitude, ave_temp: average temperature, max_depth=maximum depth, altitude, lake_area and catch_area = catchment area); (b) Representation of lakes coordinates on the first factorial plan according to the country. For further explanations see WISER report (CAUSSE et al. 2011b).

Strategy up to now

Due to the lack of national systems in the MS of the Central Baltic GIG, the strategy for intercalibration focused on the development of a kind of common assessment.

First, this was done by proposing a common system (2010). The common system was modified and is now used as national German system; some metrics were adopted by other MS. At all, the system turned out to be too specific for use on a whole GIG level.

Afterwards, a mini common system (common metric) was developed.

First bilateral comparisons show, that the common metric has low correlations with the lake status for most MS. The common metric can not be used to assess the ecological status of Central/Baltic MS. The correlations of EQR values with the results of national systems are also weak. Apart from Germany they only correlate with NL expert judgment. However, a CEN-based system will not be used in The Netherlands.

In the following table, the “appl”-columns show if MS national system, CB common metric and DE system are applicable for use as national system (via comparison with expert judgment). The “corr”-columns indicate if there is a correlation to the MS national system (EQR values).

MS	appl. MS	appl. CB	appl. DE	corr CB	corr DE
CZ	-	ok	Not suitable	-	-
DE	ok	Not suitable	ok	ok	-
DK					
EE	limited	Not suitable	limited	Not suitable	ok
LT	ok	Not suitable	Not suitable	Not suitable	Not suitable
NL	ok	ok	ok	ok	ok (?)
PL	-	Not suitable	Not suitable	-	-

Both common system and common metric are not suitable for

- a) Assessment of the ecological status of most CB states or
- b) Comparison of the results of assessment systems.

Perspectives

The intercalibration of LakeFish assessment could not be finalized within Phase II/2011. The CB MS stress the importance of fish to assess the ecological status of lakes and great progress has been achieved in recent times. Most MS favour the straightforward continuation of the IC process if the results obtained in 2011 are promising. More systems can be expected in near future and the interruption of a dynamic process can undermine the successes obtained until now. Financing and manpower are estimated to be sufficient for the intercalibration purposes in most cases. However, no explicit decisions on the future of the LakeFish Intercalibration Process after Phase 2 have been made. Decisions will have to be made on the integration of Fish in the national lake assessment of the different MS. And on the future and way of realization of the IC process.

As shown in this Milestone report, it is unlikely that the options proposed by the IC guidance will fit the special requirements of Fishes in Lakes. Modifications of existing options or alternative pathways of intercalibration have to be found. Some approaches are:

“Normative” solutions

- a) **Agreement of expert’s judgment:** Fish in lakes have low reactions on the intensity of individual pressures but react on the total level of lakewide degradation. A common CB definition of relevant pressures, their combination to total pressure intensity and thresholds values for the status classes could be defined. Afterwards, the fish assessment could be calibrated against the total pressure intensity parameter. However, the experiences show that it is unlikely to find such a common agreement (see section 6.).
- b) **Precise normative definition of the status classes:** This would be a specification of the WFD normative definitions leading to a detailed description of a set of fish community parameters and unequivocal class boundary equivalents. Such an agreement is possible, but unlikely (see a).

Decrease of demands

- c) **The introduction of CB-subgroups:** The division of the CB GIG into geographical subgroups could lead to a higher similarity of fish communities and thus to a better fit of common metrics. However, the aim of the IC process should be to assure comparability in the largest geographical range possible.

- d) **Reduced common system and other combination of metrics:** As both common system and common metric failed to be reliable, I thought about mistakes in the combination of individual metrics and the scoring procedure. As a consequence, I checked all the metrics of both approaches individually for correlations to the expert's judgment of the lakes. The aim was to identify the metrics, which can be used for the whole GIG. A first analysis showed that the value distributions make the successful development of a 5 class system unlikely. Therefore, I focused on the most essential requirement - separating the good from the bad lakes (i.e. identifying the M/P/B lakes). Preliminary results are promising, but for further investigations some important data needs to be included that was missing until termination of this milestone (DK assessment results, WISER results).

Others

- e) **WISER results:** The LakeFish team of the WISER project aims at developing a common European assessment system. The suitability of this system for the Central/Baltic subgroup has to be evaluated.

The future strategy of the CB intercalibration process is to continue the work without interruption beyond Phase II. Within 2011, more information can be gathered (DK and WISER). The WISER proposal will be tested for suitability on a GIG level. It will be tested, if approach d) can provide a common benchmark for the G/M-boundary.

The MS will be asked for their suggestions for the future IC process and a decision on the future of CB LakeFish intercalibration will be made.

DEVELOPMENT OF THE COMMON METRIC

General

- Decision on the common metrics in the CB Berlin meeting December 2010
 - WPUE [kg/m²]
 - Bream + Roach [% weight] or [% number] - *I am personally unhappy with this metric because Bream and Roach are competing species and Bream increases while Roach decreases at high trophic levels of POLY lakes in northern Germany. I kept it because it was a GIG decision and the assumption is, that total biomass increases and that Rutilus replaces Abramis in some geographical regions.*
 - Benthic species reacting on eutrophication: *Abramis brama, Blicca bjoerkna, Carassius gibelio, Carassius carassius, Coregonus 'Whitefish', Cyprinus carpio*
 - Littoral species: species reacting on shoreline degradation, macrophytes
- Realisation: I selected the following metrics, defined the class boundaries using the common CB database and drafted an Excel template. In the template, Species specific abundances and total WPUE have to be filled in. Then the metric-values are calculated and scored:
 - WPUE [kg/m²]
 - Bream + Roach [% weight]
 - Benthic species reacting on eutrophication [% weight], species listed above
 - Littoral species: Presence of species: *Esox lucius, Scardinius ery.* and *Tinca Tinca* (max. 3)
- Class boundaries were set using the following strategy
 - WPUE [kg/m²]: based on common database
 - Bream + Roach: based on common database
 - Benthic eutrophication species: based on common database
 - Littoral species: expert's definition! 3/3 species present score HIGH, 2/3 species score MODERATE, 1/3 species present score POOR, 0/3 species score BAD

Data handling for class boundary setting

Database was provided by the CEMAGREF (DATE)

Typology

- The natural lakes in the CB GIG are selected (N = 287)
- The PSD-Typology is applied
 - All lakes with max depth > 30 m were assigned to be DEEP type (38 lakes)
 - All lakes with mean depths > 7.1 and max depths <= 30 m were termed STRAT (37 lakes)
 - All lakes with max. depths < 11 m are POLY, lakes with mean depths < 4 m are POLY as well (182 lakes)
 - 30 lakes could not be assigned to types, expert judgment is needed for lakes with max. depths above 11 m AND mean depths between 4.0-7.1 m

Gear selection

The MS with **comparable gear** following CEN 14757 were selected

- Only benthic nets following CEN 14757 are used

- MS with fishing campaigns following CEN 14757 are: Denmark, Estonia, France, Germany
- Lithuania has to be excluded (uses gillnets different from the CEN standard, net height 3 instead of 1.5m and length 40m instead of 30m. Nets have 8 panels with 5m each instead of 12 panels with 2.5 m each. Mesh sizes 14/18/22/25/30/40/50/60 without small mesh sizes of CEN (5/6.25/8/10/12.5))
- Latvia has to be excluded. Uses multimesh-nets with 12 panels with 20/22/25/27/30/33/35/40/43/45/50/60 mm mesh size. Small mesh sizes of CEN are missing (5/6.25/8/10/12.5/15.5)

Fish data selection

The campaigns with **comparable data types** were selected. The following type of data is used in the database:

GSW: Total number and total weight by fish species (in the mesh size or in the net). Min and Max. sizes of the batch of fishes.

GSS: Same as the batch of fish GSW (Min. & max. size (mm) and Total weight (gr)). All individuals are assigned to the group GSS and a sub-sample SSF is realised at random to restore size distribution

INCDA: INComplete Data. At least number of fish species. Total weight and/or size are missing.

IST: individual size by species and mesh size BUT total number & weight (on the last row) by species in the net.

ISW: Individual size and weight: maximum size, total number (==1) and total weight fulfilled.

IW: Individual weight

NC: no Catch in the net

NI: not identified or empty net.

SSF: Sub-sample of fish belonging to the group GSS. (All ind. Belonging to the group GSS could be measured)

TNW: Total number and weight of fish species (all mesh incl.="88":(DK, ES, LT and AT); or by mesh size). No minimum and maximum sizes for the batch of fishes counted and weighted.

The following types of data was used for the analyses:

GSW: is included

GSS: included (but not subsamples SSF)

IST: included (number and weight are not repeated for ind. measures), data excluded using n = NOT NULL

ISW: is included

IW: included (number and weight are known)

NC: included to obtain correct total number of nets settled

TNW: included

SSF: a subsample of GSS repeating values already used for GSS, not included

CPUE: not occurring in natural CB lakes

NI: not occurring in natural CB lakes

INCDA: I checked for every lake with incomplete data, if the incomplete data could significantly change WPUE and/or species specific abundance. Lakes are included in further analyses, if only single/few individuals or small sized species not changing WPUE and %weight had INCDA.

--> INCDA has to be included for presence/absence of species

16 Danish lakes were excluded from further analyses because of incomplete data (all POLY type): TORMÅL, FØRBYSØ, NEDENSKOV SØ, THORSØ, ELLESØ MIDTJYLLAND, GULDAGER SØ, VOMME SØ, DALLUND SØ, NØRRESØ FYN, LYNGBY SØ, ØSTRUP GUNDSØMAGLE SØ, KORNERUP SØ, HØHOLMSØ VED E, UTTERSLEV MOSE, BIRKEDAM, MARIBO SØNDERSØ

Further data selection

- Two Estonian campaigns with only 2 nets settled are excluded from the analyses (Lake Pangodi, Lake Agali)
 - Lake Gjeller So excluded, only 3 nets/500 ha
 - Excluded all lakes with less than 0.8 Nets/km² (which is the minimum of CEN to detect differences - 16 Nets/20 km²): Grand-lieu, Müritz, Carcans-Hourtin, Großer Müggelsee
- > many German lakes were sampled twice with ½ of CEN effort - this data is included in the calculation

Number of campaigns (and corresponding number of lakes) in the MS used for class boundary setting

Lake Type	DE	DK	EE	FR	Total
POLY	65 (37 lakes)	70	15	10	160 (132 lakes)
STRAT	38 (22 lakes)	3	4 (3 lakes)	3	48 (31 lakes)
DEEP	29 (15 lakes)	1	-	4 (3 lakes)	34 (19 lakes)

Summary: database queries

- Lakes: Central/Baltic, natural, not excluded because of INCDA
- Gear: CEN_bent or NCEN_bent, benthic
- Lot: NOT SSF, total number not NULL

Class boundary setting

I use the same class boundary setting procedure used for the German system development.

Class boundaries are set for each type separately.

Classes are set using value distribution of metric

Clusters are identified with a cluster analysis (Ward-method, Euclid distance, SPSS program)

The ecological status classes are assigned by expert judgment with the help of a visual control of the plotted value distribution.

The expert's principles are

- Bad lakes are very rare!
- Poor lakes are rare.
- Moderate and Good are the most frequent status classes.
- High lakes are not rare (the fish community is resilient). However, High lakes are not as frequent as good or moderate ones.
- The distribution of lakes status classes should be similar for all metrics (type-specific)
- WPUE: High values of the metric indicate a low ecological status.
- Bream+Roach [% weight]: High values of the metric indicate a low ecological status.
- Benthic species reacting on eutrophication [% weight]: High values of the metric indicate a low ecological status.

- For POLY lakes, there were some lakes with very low values of WPUE. Extreme eutrophication and other pollution can lead to a decline of WPUE. Therefore, I don't know if these values indicate a very good or a very bad status. For very low values, the metric is not applicable.
- If the metrics 'Bream + Roach' or 'Benthic species reacting on eutrophication' equal 0, they are not applicable (n. a.) either. The complete absence of species most likely is not caused by eutrophication or other human pressures, but might be natural. If it is anthropogenic, it is not a good sign and should not be scored with HIGH.

Metric	Ecological status class	POLY	STRAT	DEEP
WPUE [kg/m ²]	n. a.	< 0.01	-	-
	5	≤ 0.09	≤ 0.03	≤ 0.022
	4	≤ 0.14	≤ 0.05	≤ 0.036
	3	≤ 0.19	≤ 0.07	≤ 0.050
	2	≤ 0.24	≤ 0.14	≤ 0.074
	1	> 0.24	> 0.14	> 0.074
Bream + Roach [% weight]	n. a.	0	0	0
	5	≤ 40	≤ 33	≤ 30
	4	≤ 65	≤ 53	≤ 40
	3	≤ 80	≤ 73	≤ 55
	2	≤ 95	≤ 93	≤ 65
	1	> 95*	> 93*	> 65
Benthic eutrophication species [% weight]	n. a.	0	0	0
	5	≤ 17	≤ 03	≤ 02
	4	≤ 37	≤ 10	≤ 07
	3	≤ 70	≤ 20	≤ 21
	2	≤ 90	≤ 30	≤ 35
	1	> 90	> 30*	> 35

* no bad lakes in the dataset for this metric and type, poor/bad boundary is an educated guess based on the previous boundaries

Scoring

- Each metric gets a score, ranging from 1-5
- The total sum of scores is calculated (X)
- An EQR value is calculated using $EQR = (X - X_{min}) / (X_{max} - X_{min})$. With 4 metrics, a minimum of 4 points and a maximum of 5 points it is $EQR = (X - 4) / 16$
- Total status was assigned to the EQR values:
 - HIGH ≥ 0.95
 - GOOD ≥ 0.75
 - MODERATE ≥ 0.50
 - POOR ≥ 0.25
 - BAD < 0.25

For the definition of the EQR boundaries for the status classes I used the German dataset. In this data, the ecological status of lakes was estimated by experts. I modified the class boundaries manually in order to reach a minimum sum of squares of the difference [expert status lake - common metric status fish].

SUITABILITY OF THE COMMON METRIC

Data compilation

- Development of a common metric (metric selected by expert judgment at the GIG meeting) - CB_LaFi
- Elaboration of an Excel-template which provides the corresponding assessment results for CB_LaFi and DE_LaFi
- Request for the assessment results of the corresponding national assessment system MS_LaFi and an expert judgement of the total ecological status of the lake (not of the fish fauna) MS_exp

Dataset

- CZ: Data for 16 reservoirs. CB_LaFi and DE_LaFi suitability is tested. No national method present and water bodies are anthropogenic therefore CZ is not included in intercalibration.
- DE: Data for 84 fishing campaigns at 78 lakes (6 lakes were sampled twice)
- EE: Data for 11 natural lakes (CEN 14757 but nets are not strictly set in random stratified procedure)
- LT: Data for 38 natural lakes. Data based on special fishing procedure without mesh sizes < 14 mm. CB_LaFi and DE_LaFi suitability is tested and correlations are checked with the provision that the mesh sizes will influence the results.
- NL: Data for 8 natural lakes in Netherland. Lakes were sampled with CEN 14757 multimesh-nets and electrofishing to be assessable with CEN-based methods. Simultaneously, lakes were sampled with the Dutch Trawl fishing procedure and assessed with the Dutch system to make the results comparable. Additionally, 2 Danish and 2 German lakes were sampled with both methods simultaneously.
- PL: Data for 10 natural lakes, suitability and correlations checked

System validation - methods

Suitability check

Is the method useful to assess the lakes of the specific MS?

- national methods: MS_LaFi ↔ MS_exp (status class),
- CB common metric: CB_LaFi ↔ MS_exp (status class),
- German method: DE_LaFi ↔ MS_exp (status class),

The suitability check is used to assess if the method provides reliable results for assessment of the lakes of a MS.

National method assessment results should correlate with national expert judgment of lake status. Otherwise the national method might be of limited suitability. This might eventually be the case for systems under development. Methods that do not fulfil this criterion should be improved.

For the CB common metric and the German system it is just a test, if they could be used for assessment. In terms of status classification, neither CB nor German systems need to be very useful for the direct assessment of lakes of different MS. But missing or contradictory correlations indicate a stronger level of disharmony which could be caused by biogeographical incomparability of metrics, differences of the system concept or other reasons.

The German system has to be intercalibrated like other MS systems. The direct comparison of national systems with the German system is kept, because the German approach was

originally developed for use as a common CB assessment system. The applicability of a common system was rejected at the CB meeting in Berlin. However, the comparisons are done to substantiate the decision.

Correlation check

Are the methods sufficiently correlated?

- between MS_LaFi and CB_LaFi (EQR)
- between MS_LaFi and DE_LaFi (EQR)

This is the most important prerequisite for intercalibration. A correlation indicates similarity of the concepts, metrics and systems.

Acceptance criteria

- for status class measures
 - Spearman $\rho \geq 0.5$, - roughly following IC Guidance, Annex V (CIS 2009)
 - mean absolute class difference < 1 - following IC Guidance, Annex V (CIS 2009)
 - critical misclassification (H/G \leftrightarrow M/P/B) $< 20\%$ of all classifications - my assignment
 - no systematic misclassification: 25 - 75 % of mis-classifications are false-positive - - my assignment
- for EQR measures
 - correlations with $\rho \geq 0.5$ - following IC Guidance, Annex V (CIS 2009)

System validation: results for MS

CZ

Notes: The Czech Republic has no natural lakes, but exclusively anthropogenic reservoirs. For this reason, no official participation in the intercalibration process takes place. No national method has been developed until now. The application of the common metric and the German LaFi is a test to ease the future national system development with the knowledge gained in the IC process. I decided to modify the DE_LaFi assessment a little bit (this is an allowed option of the German system). I excluded the size metrics for DEEP reservoirs (Flaje, Klicava, Nýrsko, Orlik, Rimov, Sec, Vranov, Zelivka). Czech and German experiences have shown, that individual growth is misleading in this kind of water body (median of fish sizes are usually bigger in deep reservoirs than in deep natural lakes). I deselected the Bream and Ruffe metrics where the species are naturally absent (Flaje, Nýrsko, Zlutice). A Czech national method is not present; the corresponding suitability check and the check of correlations are cancelled.

Suitability check: The CB_LaFi system would be a suitable method to assess the ecological status of Czech reservoirs. Only 2 of 16 water bodies are misclassified, both of them by one class, but critical. The deviation is not in the same direction (one is too good, the other one too bad). The suitability criteria shown in Tab. 2 have excellent values. The system concept, metrics and class boundaries can be useful in the future for the national system development. The German assessment method is not applicable to Czech reservoirs. There is no correlation of assessment result to expert's status estimation, the percentage of critical misclassifications is app. 1/5 and close to the acceptability limit, all misclassifications are false-positive (i.e. the German system systematically scores too badly).

Tab. 2: Comparison of ecological status classes assigned by CB_LaFi and De_LaFi results with expert's judgment of the ecological status/potential of Czech reservoirs (experts: Marie Prchalova, Jan Kubecka, number of lakes: 16). Check is ok, if acceptance criteria are fulfilled (p. 51).

criterion	CB_LaFi value	CB check	DE_LaFi value	DE check
rho ESC ↔ ESC	0.75, $p < 0.002$	ok	0.35, n. s.	X
macd	0.13	ok	0.5	ok
% critdev	13	ok	19	ok
% falsepos	50	ok	100	X

Conclusions: The results were sent to the Czech representative Marie Prchalova and it was taken note of. Conclusions will be drawn when a national system and an official participation of the Czech Republic in the IC process is given.

DE

Notes: Anthropogenic lakes were removed from the dataset for the system validation check. Special lakes were removed also (flushed lakes, saline lakes, heavily stocked lakes). For lakes which were fished in two different, non-consecutive years, I always took the first (earlier) campaign for the validation check. For the six lakes with two assessment results, all status classes were similar for the German system, but differed in three cases for the common metric.

Suitability check and correlation: The German assessment method is suitable to assess the German lakes, i. e. the status-classes correlate with the expert's expectations. The suitability criteria shown in Tab. 3 have good values except the low percentage of false-positive classifications. The German system scores too positive. *Favouring false-negative in comparison to false-positive assessments was intended in the German system development considering the one-out-all-out principle of four BQE proposed by the WFD.*

The CB common metric shows a correlation close to the limit of acceptance. The class deviation is in a good range, but approximately 1/3 of all lakes are critically misclassified. The CB system is not applicable to assess the German lakes without modifications.

Tab. 3: Ecological status classes: comparison of CB_LaFi and De_LaFi results with expert's judgment of the ecological status of German lakes (experts depend on federal country, number of lakes: 77). Check is ok, if acceptance criteria are fulfilled (p. 51).

critterion	DE_LaFi value	DE check	CB_LaFi value	CB check
rho ESC ↔ ESC	0.753, $p < 0.001$	ok	0.476, $p < 0.001$	X
macd	0.29	ok	0.55	ok
% critdev	14	ok	31	X
% falsepos	21	X	27	ok

The correlation between both systems is good; the common metric seems useful for calibration:

EQR DE_LaFi ↔ EQR CB common metric: $\rho = 0.710, p < 0.001$.

Conclusions: The German system reliably assesses the ecological status of German lakes. It can be included in the intercalibration process.

EE

Notes: For some lakes, I modified the German assessment system, which is an allowed option in the German approach. I excluded the Pikeperch metric, which is part of the German assessment of polymictic lakes. Pikeperch is naturally absent in some of the Estonian lakes; a low percentage is not indicating a good status, but directly the lack of stocking of that specific species.

Suitability check and correlations: The suitability checks of all three methods tested are not convincing (Tab. 4). The suitability values for Estonia are close to the acceptability limit of correlation between system and experts opinion. However, approximately 30 % of critical misclassifications do not seem reliable enough. Additionally, there is a strong tendency to assign too bad status classes. The application of the German system shows equal limitations. The common metric does not correlate with the expert's judgment of the lake ecology and thus is not suitable for Estonian lakes.

Tab. 4: Ecological status classes: comparison of EE_LaFi (current version 2), CB_LaFi and De_LaFi results with expert's judgment of the ecological status of Estonian lakes (experts: Anu Palm, Teet Krause, number of lakes: 21). A comparison for an outdated version of the Estonian system is added at the right in grey cells (EE_LaFi_old). Check is ok, if acceptance criteria are fulfilled (p. 51).

criterion	EE_LaFi_2 value	EE check	CB_LaFi value	CB check	DE_LaFi value	DE check
rho	0.478, $p < 0.05$	X	0.24, n. s.	X	0.57, $p < 0.01$	ok
macd	0.62	ok	0.52	ok	0.52	ok
% critdev	28	X	28	X	28	X
% falsepos	90	X	44	ok	82	X

The correlations between the Estonian system and the common metric EQR values are weak for both the outdated and the current version. The correlation between the current Estonian and the German system is above the acceptability limit.

New version of Estonian LaFi system

Est2 - EQR ↔ CB EQR: $\rho = 0.269$, n. s.

EQR: $\rho = 0.062$, n. s.

Est2 - EQR ↔ DE EQR: $\rho = 0.585$, $p < 0.01$

EQR: Spearman $\rho = 0.368$, n. s.

Old version of

Est1 - EQR ↔ CB

Est1 - EQR ↔ DE

Conclusion: The results show, that the continuous development of the Estonian LaFi system has led to better results in terms of identity of assessment of fish and expert's judgment of lake ecology. The common metric is unsuitable for intercalibration purposes. This is shown by low correlations with both lake status and the Estonian fish system. The German system shows a higher correlation and might be a basis for future comparisons.

LT

Notes: The Lithuanian fishing procedure is deviating from the CEN 14757 in the fact, that small mesh sizes < 14 mm are not included in the nets. In addition to bio-geographical constraints, methodological differences might prevent applicability of assessment systems developed for a CEN 14757 procedure to Lithuanian data. Anyhow, Tomas and I decided to perform the tests and include the differences in the interpretation of results.

The consequences of changed mesh sizes on the metrics are not obvious or inevitably, except for the German size metrics. Excluding small mesh sizes will increase the median of species-specific sizes and make the selected class boundaries obsolete. Therefore, I excluded the size metrics from the calculations of DE_LaFi. Additionally I excluded the Pikeperch metric, which is part of the German assessment of polymictic lakes. Pikeperch is naturally absent in the Lithuanian lakes; a low percentage is not indicating a good status, but directly the lack of stocking of that specific species.

Suitability check and correlations: The Lithuanian LaFi system is a suitable tool to assess the ecological status of Lithuanian lakes. The values shown in Tab. 5 indicate a good correlation of system result and expert judgment and a small class deviation. The Lithuanian system has a tendency to score too badly, although the values are within the acceptability limit set. The percentage of critical misclassification seems to be too high. Neither the status classes given by the common metric nor the German system significantly correlate with the expert's judgment of the lake. Main reason might be a displacement of the relative metrics (percentages) used in both systems due to the exclusion of small mesh sizes. Thus, wrong status classes might have been assigned.

Tab. 5: Ecological status classes: comparison of LT_LaFi, CB_LaFi and De_LaFi results with expert's judgment of the ecological status of Lithuanian lakes (expert: Tomas Virbickas, number of lakes: 38). Check ok, if acceptance criteria are fulfilled (p. 51).

criterion	LT_LaFi	LT check	CB_LaFi	CB check	DE_LaFi	DE check
rho	0.702, $p < 0.001$	ok	0.217, n. s.	X	0.304, n. s.	X
macd	0.37	ok	0.61	ok	0.68	ok
% critdev	24	X	33	X	33	X
% falsepos	71	ok	44	ok	73	ok

The presence/absence of small mesh sizes is a systematic difference in the methodology. Thus, a correlation between the EQR values can be given even if status classes deviate. The coefficients of correlation between the Lithuanian and the common metric / German method are significant, but rho is below the given threshold value:

LT EQR ↔ CB EQR: $\rho = 0.388$, $p < 0.05$,
0.408, $p < 0.05$

LT EQR ↔ DE EQR: $\rho =$

Conclusion: The correlations between the Lithuanian system and the common metric / German system are too weak to be used for an intercalibration based on direct comparison. One possibility for the future is to adjust the class boundaries for each metric of the calibration system (CB or DE). This could be done based on catch data with known mesh sizes by comparing the metric values with and without the small meshes. Alternatively, another IC option has to be found. The results are confirmed by the Lithuanian representative Tomas Virbickas.

NL

Notes: Eight shallow lakes in the Netherlands, two shallow lakes in Germany and two in Denmark have been sampled in 2009 with both CEN (gillnets) and trawling method. Data for eight Dutch lakes could be used for the following comparisons. Almost all lakes sampled are quite eutrophic and disturbed (POOR or MODERATE status of lakes according to expert judgment). The limited number of lakes investigated and the narrow range of expectable status classes has to be taken into account. Only the Dutch lakes are taken into account for the suitability check.

Suitability check and correlations: Due to the low number of lakes, the values given in Tab. 6 are not really representative. The Dutch system is a suitable tool to assess Dutch lakes. The assessment results highly correspond with the expert's expectations. In terms of status classes, only one out of eight lakes differs by one class within the poor/moderate range.

Both the CB common metric and the German system would be usable to assess the Dutch lakes, though with a lower performance. The correlation with the Dutch expertise on lake status is present, but the status classes assigned by both methods are too high.

Tab. 6: Ecological status classes: comparison of NL_LaFi, CB_LaFi and DE_LaFi results with expert's judgment of the ecological status of Dutch lakes (experts: Bob Bredervald, Nico Jaarsma, Eddy Lammens, number of lakes: 8). Check is ok, if acceptance criteria are fulfilled (p. 51).

critterion	NL_LaFi value	NL check	CB_LaFi value	CB check	DE_LaFi value	DE check
rho	0.774, $p < 0.05$	ok	0.715, $p < 0.05$	ok	0.677, n.s.	ok
macd	0.13	ok	0.75	ok	0.38	ok
% critdev	0	ok	25	-	12,5	-
% falsepos	-	-	0	X	0	X

The coefficients of correlation between both the common metric and the German system indicate comparability (although the significance is absent for the German LAFi due to the low number of lakes):

NL EQR ↔ CB EQR: $\rho = 0.756$, $p < 0.05$

NL EQR ↔ DE EQR: $\rho = 0.551$, n. s.

Conclusions: The comparisons show that the Dutch system can be intercalibrated thanks to the investigations with simultaneous application of different fishing methods. The intercalibration can be done by direct comparison with methods based on CEN 14757 (common metric, pseudo common metric, DE_LaFi, DK_LaFi and others). However, a low number of lakes covering only two degraded status classes might lead to problems with the analytical methodology.

PL

Notes: The results of the Polish assessment method are available for three lakes only. The suitability check of the national method and the EQR correlation checks are cancelled.

Suitability check: As can be seen from the results in Tab. 7, the suitability criteria indicate that both common metric and the German system are (completely) unsuitable to assess the ecological status of Polish lakes. The correlations of system results and expert judgments of the lake status are negative, indicating an assessment working in opposite directions. The other criteria are above the acceptability limits.

Tab. 7: Ecological status classes: comparison of CB_LaFi and De_LaFi results with expert's judgment of the ecological status of Polish lakes (expert: Witold Bialokoz, number of lakes: 10). Check is ok, if acceptance criteria are fulfilled (p. 51).

critterion	CB_LaFi value	CB check	DE_LaFi value	DE check
rho	- 0.72, $p < 0.05$	X	- 0.16, n. s.	X
macd	1.3	X	1.1	X
% critdev	80	X	50	X
% falsepos	44	ok	57	ok

Conclusions: It has to be kept in mind, that the most important suitability measure is the correlation between the Polish system and the common metric or German system. The comparisons should be repeated with a higher number of lakes and the Polish EQR values included. However, the values in Tab. 7 indicate, that common metric and German LaFi are unsuitable for direct comparison. Further decisions can be made when the Polish methods are established, what will be in 2012 for LFI+ and 2013 for a CEN 14757-based method. Rather than using direct comparison, it is likely that another solution for intercalibration will have to be found

Der nachfolgende Abschnitt beinhaltet die Arbeitsfassung zum Stand 07.12.11.

Draft submitted for comments and confirmation on November 18, 2011 by David
Sent to MS that provided data: DK (Torben), EE (Anu, Teet), LT (Tomas), NL (Nico, Eddy, Bob) and PL (Hannah, Witold)

Replies:

- EE: confirmed contents
- NL: Concept is clear, but realization is unclear, and therefore a final decision on acceptance is not possible. Could be sent to Sandra as conceptual proposal with the remark, that it still has to be discussed in the GIG (Eddy).

Aim and introduction

The WFD requires the assessment of water bodies with biological quality elements (BQE), i.e. phytoplankton, macrophytes, zoobenthos and fish. With these BQE, the ecological status has to be assessed and described with a 5-step classification: HIGH, GOOD, MODERATE, POOR and BAD. The biological assessment is the basis for evaluation of the current status. In case a MODERATE, POOR or BAD status is indicated, appropriate actions have to be initiated in order to set the water body in a good status.

With the process of intercalibration the status classes assigned on a national level are compared and eventually modified in order to guarantee comparable, harmonized and therefore equitable assessment on a larger geographical scale. Three basic options have been suggested to reach this goal (CIS 2011):

- Option 1 - Same data acquisition, same numerical evaluation: the intercalibration can concentrate on the harmonisation of reference conditions and class boundary comparison/setting
- Option 2 - Different data acquisition and numerical evaluation: the intercalibration can be done using common metrics which should cover all relevant parameters indicative of the BQE, respond to the pressures being intercalibrated, be ecologically meaningful and show no (or minor) bias due to biogeographical differences
- Option 3 - Similar data acquisition, but different numerical evaluation: the intercalibration is done by a pair-wise direct comparison at a sampling site/water body level and adjusted against a common metric using regressions.

The present paper deals with the intercalibration of ‘systems to assess the ecological status of lakes with fish’ (LaFi systems) in the Central/Baltic geographical intercalibration group (CB GIG). Fish are a well chosen BQE for the assessment of lakes, especially for an integrated assessment of the whole water body and for showing pressures taking effect for longer time ranges (see Annex, p. 80 ff.). On the other hand, fish community traits (metrics) or LaFi assessment results and individual pressures usually show only low correlations. The reasons are manifold: fish represent the highest trophic level with complex ecological cross-linking; i. e. they react on all pressures, not on individual ones. The correlation in between pressures parameters usually is much higher than between specific pressure parameters and fish metrics. For example eutrophication, shoreline degradation and lake use intensity always go hand in hand. On the other hand, fish are the most complex and resilient BQE; they do react on pressures - but they do not always react and they do not always react in the same way. Summarized, fish are suitable to identify degradation but they are not suitable as direct proxies of abiotic pressures parameters.

Furthermore, no agreement on relevant pressures or on parameter values equivalent to given ecological status classes could be found during the intercalibration process⁶. Besides the occurrence of reliable correlations to pressures, the identification of sites corresponding to a given status is an essential precondition of the above mentioned intercalibration options.

During the LaFi intercalibration process in the CB GIG, attempts were made to follow the current Intercalibration Guidance (CIS 2011) as close as possible. At the end of 2010, an intercalibration typology was developed and agreed on a GIG level; the selected lake types are polymictic lakes, stratified lakes and stratified lakes deeper than 30 m (abbreviated POLY, STRAT and DEEP)⁷. Roughly following Option 1, a common system was proposed by the GIG lead to assess the lake status based on 7 to 10 metrics, depending on lake type. The system was tested by the CB member states (MS), but the results were not satisfying in terms of correlation to either the expected status classes of the lakes or the national LaFi results. The fish communities in the CB GIG differ too much to allow assessment by a common system. As a consequence, a reduced list of promising metrics was collated by the CB experts⁸. Based on 4 metrics, a mini-system was developed for use as common metric in the GIG. However, the common metric also showed weak relationships to lake statuses and national assessment results.

This led to the situation that the intercalibration had to be done without reliable correlations of pressures and metric values, without benchmark sites, without common system and without correlated common metric. This made all options of the IC Guidance inadequate and an alternative solution was developed. This solution deals with the specific situation of Central Baltic LakeFish assessment, but it also takes into account some important principles of the guidance. It is based on the dataset obtained during the precedent process of intercalibration: a number of lakes with an expert judgment of the expected status class, corresponding values of some 10-15 fish metrics and the national assessment results in terms of status classes and EQR values.

The basic principles of the chosen method of intercalibration are explained in advance to make the following contents understandable. First of all, the IC process is limited to the harmonization of the good/moderate class boundary. This is justified by the fact that fish are the most resilient BQE and have a comparable low 'resolution' of indication. The good/moderate transition has the highest practical relevance as it's decisive for the need of mostly complex and expensive actions to obtain a GOOD status. In the IC guidance, the high/good transition is regarded to be equally important. However, lakes in HIGH status according to the national LaFi assessments systems are very rare (< 2 %). The dataset simply does not allow the comparison of high/good class boundaries for these few lakes.

The strategy of the CB LakeFish intercalibration is as follows:

- Use the available fish metrics as common metrics.
- Define common metric values that indicate a moderate/poor/bad lake with a 100 % certainty according to all experts.
- Use the common metrics to identify a subset of lakes in an unmistakable moderate/poor/bad status according to a harmonized expert's judgment (benchmark lakes). Combine metrics to increase the number of benchmark lakes.

⁶ see minutes of the Lake fish intercalibration available at CIRCA - Lakes - fish intercalibration

⁷ see Typology for CB LakeFish IC, available at CIRCA - Lakes - fish intercalibration/CB GIG

⁸ see minutes of the 1st CB Lake fish meeting in Berlin, available at CIRCA - Lakes - fish intercalibration/CB GIG

- Adjust the national good/moderate class boundaries to secure, that the benchmark lakes in a moderate/poor/bad status according to a harmonized expert's judgment are also in a moderate/poor/bad status according to the national assessment systems

With this procedure, the main aim of the intercalibration is achieved and the national assessment results are made comparable.

Data

The majority of CB member states have done fish sampling according to a standardized procedure with stratified multimesh-gillnet fishing (CEN 2005). Some MS have modified this procedure. Their results are included in the investigations to increase the number of intercalibrated MS. Tab. 8 gives an overview of the MS participating and the corresponding numbers of lakes investigated. For the further analyses, the lakes belonging to the STRAT and the DEEP type are combined to reach an adequate sampling size (termed STRAT+). The intercalibration dataset is limited to natural lakes - HMWB or AWB are not included.

Tab. 8: Member states, type-specific lake numbers and modification of fishing procedures included in the present intercalibration exercise.

MS	POLY	STRAT	DEEP	Method
DE	50	36	22	2*1/2 CEN and full CEN
DK	46			full CEN
EE	13	8		CEN mod (not randomly stratified)
LT	20	9	9	CEN mod (without small mesh sizes)
NL	8			full CEN
PL	5	3	2	full CEN
sum	142	56	33	

A special case is the Netherlands. The official national system is based on a combination of trawl fishing, electrofishing and fisheries statistics. The colleagues have done parallel fishing with CEN multimesh nets and trawl fishing anticipating the necessity for making their national system intercalibrateable. Therefore, the common metrics for Dutch lakes are evaluated using the CEN results.

The present investigation was preceded by different approaches to intercalibrate. The first attempt was a common system (end of 2010), the second a common metric consisting of 4 single metrics. Both common system and common metric were developed by the GIG lead. Excel templates were provided to the MS for calculation of metric values and scores. The MS were asked to insert national fishing data into these templates and put them to the GIG leads disposal for further analyses. This way, the fish metric values were gathered (see Annex p. 75 for description of the metrics).

For France, fishing results based on the CEN standardized procedure are present and a national assessment system was developed in October 2011⁹. However the requested metric values could not be provided due to limited resources for the time consuming completion of the templates.

⁹ see national system description (CAUSSE et al. 2011a), available at CIRCA - Lakes - fish intercalibration/CB GIG/CB method descriptions

An expert judgment of the ecological status in 5 classes according to the WFD is present for most lakes. It was provided by the persons in charge of LAFi assessment in the MS, thus the experts estimated the expected status for their 'own' lakes. This expert judgment estimates the combined intensity of all anthropogenic pressures affecting the lake, e.g. eutrophication, shoreline degradation, interruption of migration to in- and outflows, lake use intensity, and fisheries. It represents the estimated status of the lake and it was originally intended for checking the plausibility of the different fish assessment systems (national systems, common system, and common metric).

Expert's judgment is present for 118 POLY lakes originating from Germany, Denmark, Estonia, Lithuania, Netherlands and Poland. It was present for all 89 STRAT+ lakes (from Germany, Estonia, Lithuania and Poland).

Use of fish metrics to identify disturbed lakes

The dataset of lakes used for intercalibration includes an unknown number of lakes in a moderate/poor/bad condition. The exact number is unknown because expert's judgment might differ between member states, national fish assessment systems might be of different indicative value and/or the fish communities might react in different ways throughout the broad geographic range of the GIG. The intercalibration of CB LakeFish methods is based on a subset of benchmarking lakes that are unmistaken in moderate/poor/bad condition.

An estimation of the whole level of anthropogenic alteration including all pressures was used as pressure parameter for the intercalibration process because of the absence of reliable correlations between metrics and specific pressures throughout the whole GIG and because of the strong interactions of the pressures. This is the expert judgment of the total pressures intensity. The expert judgment combines estimations of the effects of eutrophication, other kinds of environmental pollution, shoreline degradation, lake use intensity, water level manipulation, fisheries intensity, effects of massive stocking and others. This procedure is not exactly in line with the precise requirements of the IC guideline (CIS 2011). The specific challenges of LakeFish assessments were delineated and the replacement of specific pressures by an expert's estimation of the overall lakes status was principally accepted for LakeFish intercalibration by ECOSTAT¹⁰.

For the identification of the good/moderate class boundary, the national expert's judgments of the CB lakes are reduced to two groups: GOOD (including high and good according to the WFD classification) and NOTGOOD (moderate, poor and bad lakes). Afterwards, the fish metric values belonging to these two groups are compared. All 12-15 metrics available are tested (4 metrics of the common metric and a type specific number of metrics of the common system). Additionally, the EQR values of both common metric and common system were tested for their suitability to discriminate between GOOD and NOTGOOD lakes.

The aim is to identify lakes that are unmistaken in NOTGOOD condition. As a basic principle, the 'worst' metric value of GOOD lakes represents a threshold value for identification of NOTGOOD lakes. Values below the threshold represent a characteristic situation of the fish community that describes a NOTGOOD lake according to all experts of all member states (a kind of least common denominator). The threshold values were identified using a figure. A few 'outliers' are excluded from the threshold setting in most cases. This is

¹⁰ See VAN DE BUND et al. (2011): Method compliance checking: which methods can be included in the intercalibration exercise? Document discussed at the ECOSTAT meeting of 30-31 March 2011. Document with ECOSTAT conclusions available at CIRCA: Lakes - fish intercalibration/CB GIG/official documents.

done if an appreciable number of additional lakes correctly identified as NOTGOOD is obtained (see Fig. 2). If outliers are excluded, the threshold values represent the situation in a NOTGOOD lake according to most experts and most cases. All fish metrics and the EQR value could be used to identify at least a few unmistakable disturbed lakes.

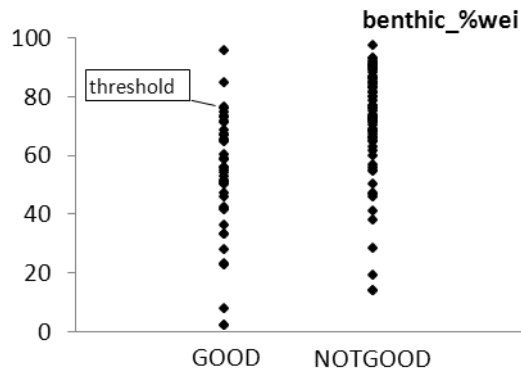


Fig. 2: Threshold setting for the common system metric benthic net species [%weight] for polymictic lakes. Two outliers are excluded to allow identification of 33 NOTGOOD lakes.

Results

Polymictic lakes

Step P1: thresholds and selection of useful metrics

The metrics tested for their suitability to identify NOTGOOD lakes are listed in Tab. 9. The threshold values for the identification of NOTGOOD lakes were determined as described before. An explanation of the metrics and the corresponding figures can be found in the Annex. According to expert judgment, 75 of 118 polymictic lakes them are in a NOTGOOD ecological status. Principally, all 15 metrics and the EQR values are useful to identify some NOTGOOD lakes. The exclusion of outliers in the threshold setting leads to a small number of lakes identified as NOTGOOD using fish metrics but GOOD according to expert judgment.

Tab. 9: Metrics, threshold values for identification of NOTGOOD lakes and performance in terms of number of NOTGOOD lakes correctly identified as NOTGOOD and GOOD lakes falsely classified as NOTGOOD (polymictic lakes, n= 118).

POLY metric	NOTGOOD if	NOTGOOD identified	GOOD false
Metrics of the common system			
Obligatory species [n species]	< 5 species	2	0
WPUE [kg/m ²]	> 0.116	16	1
Benthic DE [%weight]	> 79,0	29	2
Benthivor [%weight]	> 62,0	7	0
<i>Abramis brama</i> [%weight]	> 39,5	9	0
<i>Blicca bjoerkna</i> [% weight]	> 15,5	19	2
<i>Gymnocephalus cernuus</i> [%weight]	> 05,5	20	1
<i>Perca fluviatilis</i> [% weight]	< 10,3	35	3
<i>Sander lucioperca</i> [% weight]	> 13,5	10	0
Median <i>Perca fluviatilis</i> [g]	> 100	2	0
Median <i>Abramis brama</i> [g]	> 216	8	1
Median <i>Rutilus rutilus</i> [g]	> 160	2	0
Metrics of the common metric			
<i>Abramis</i> + <i>Rutilus</i> [% weight]	> 70,5	11	1
Benthic CB [% weight]	> 52,0	7	0
Littoral species missing	≥ 2	9	0

EQR values			
DE EQR	< 0,69	19	0
CB EQR	< 0,63	10	0

The suitability of the metrics to identify NOTGOOD lakes is different. Some of them identify only negligible number of NOTGOOD lakes, e.g. the German metric obligatory species. Others are conterminous or subsets of each others. The following metrics were discarded as unreliable or unnecessary:

- Obligatory species number: identifies only 2 lakes that are identified by at least 3 other metrics (species specific and WPUE)
- Benthivorous species (common system) and benthic species (common metric): both metrics identify the same lakes as NOTGOOD. These lakes are all identified by benthic net species (common system), too, but the latter one identifies much more lakes.
- Median *Perca fluviatilis* and median *Rutilus rutilus*: each metric identifies only 2 lakes that are identified by at least 2 other metrics (species specific and WPUE)
- Median *Abramis brama*: is the only NOTGOOD metric for 3 lakes, but one classification is correct, one incorrect and one is caused by cormorants, which is not an anthropogenic pressure
- Common system EQR and common metric EQR correctly identify a notable number of lakes as NOTGOOD. However, the lakes are also identified by at least one other metric (usually 3 or more). These values are complicated to calculate, not applicable in some MS and thus excluded.

Finally, nine metrics were selected for the identification of NOTGOOD lakes:

- All species specific metrics: *Abramis brama*, *Blicca bjoerkna*, *Perca fluviatilis*, *Sander lucioperca* and *Gymnocephalus cernuus* [% weight]
- WPUE [kg/m²]
- Benthic net species - common system definition [% weight]
- Abramis + Rutilus [% weight]
- Number of littoral species missing

Step P2: identification of the NOTGOOD benchmark lakes and performance

The subset of benchmarking NOTGOOD lakes is identified with the nine metrics selected in step P 1 (for **polymictic** lakes):

- *Abramis brama* [% weight]
- *Blicca bjoerkna* [% weight]
- *Gymnocephalus cernuus* [% weight]
- *Perca fluviatilis* [% weight]
- *Sander lucioperca* [% weight]
- *Abramis brama* + *Rutilus rutilus* [% weight]
- WPUE [kg/m²]
- Benthic net species - common system [% weight]
- Number of littoral species missing [n species]

The threshold values in Tab. 9 describe the ‘least common denominator’ for identification of a NOTGOOD lake. A lake identified as NOTGOOD with the fish metrics has a 100 % probability of being in a NOTGOOD ecological status. On the other hand, the number of lakes identified with each metric is low. Therefore, the benchmark set has to be based on a ‘one out - all out’-principle. If any of the fish metrics shows a NOTGOOD status, the ecological status of the lake is NOTGOOD.

The benchmark NOTGOOD dataset represent a subset of the unknown amount of lakes in a moderate/poor/bad status. The rest of the lakes can either be in a high/good or in a moderate/poor/bad status. The lack of identification of pressures with the metrics does not mean that there are no pressures. If no metric shows a NOTGOOD status, the ecological status of the lake is UNKNOWN.

A total number of 7 lakes are classified as NOTGOOD based on one or more outliers. All lakes classified as NOTGOOD be means of outliers are exclusively classified by outliers; in no case the indication of degradation is supported by a 'normal' indication of NOTGOOD. For the needs of a reliable benchmarking, the status of lakes classified as NOTGOOD by outliers is reset to UNKNOWN.

A high variability and outliers are a well known phenomenon in fish sampling. Besides a methodological uncertainty, the fish communities in the corresponding lakes might be influenced by factors that make it unsuitable for a generalized status assessment. Such factors can be uncommon natural or anthropogenic pressures: species specific predation leading to shifts in percentages (e. g. by cormorants), naturally low pH values, naturally high salinity, extreme influences of stocking/fishing or others.

With the procedure described, a set of benchmark lakes with a fish community that unmistakably describes a NOTGOOD ecological status is identified. The benchmark set can be related to the total dataset and the performance of its identification can be assessed with the following descriptors (CIS 2003):

correct positive: NOTGOOD lakes according to expert judgment that are classified as NOTGOOD by the fish benchmarking system

false positive: GOOD lakes according to expert judgment that are classified NOTGOOD with the fish system

correct negative: GOOD lakes according to expert judgment that are classified as UNKNOWN with the fish system

false negative: NOTGOOD lakes according to expert judgment that are classified as UNKNOWN with the fish system

Please note that 'positive' is the identification of NOTGOOD, comparable to medical tests. The values of the descriptors are shown in Tab. 10. They are expressed as percentages of the total number of polymictic lakes in the MS or GIG. The table shows a good suitability of the system: 86 % of all lakes are classified correctly, 14 % are unidentified degraded lakes and no GOOD lake is classified as NOTGOOD.

Tab. 10: Performance of the UNKNOWN/NOTGOOD fish benchmarking procedure for polymictic lakes. Values show percentage of total number of lakes for each member state and a summary. Further explanations can be found in the text.

POLY MS	N lakes	Exp. NOTGOOD Fish corr. pos. %	Exp. NOTGOOD Fish false neg. %	Exp. GOOD Fish corr. neg. %	Exp. GOOD Fish false pos. %
DE	49	50	12	38	0
DK	23	44	30	26	0
EE	13	31	8	61	0
LT	20	40	10	50	0
NL	8	88	12	0	0
PL	5	100	0	0	0
GIG	118	49	14	37	0

The confusing statistical terms and the use of percentages require a careful interpretation of Tab. 10. If the benchmarking procedure indicates a NOTGOOD status, the lake is unmistakably in a NOTGOOD status according to expert judgment. If the procedure does not indicate NOTGOOD, there is an app. 1/3 probability, that the lake nevertheless is NOTGOOD (and not GOOD). The benchmark procedure can show degradation, but it can not prove a good status.

Step P3: modification of the national class boundaries for the G/M transition

The goal of the intercalibration process is to assure, that the MS assign comparable status classes. The idea how to make the national class boundaries comparable is to use the NOTGOOD benchmark procedure to identify lakes degraded according to a harmonized expert judgment of the GIG. For the lakes rated as NOTGOOD, all involved experts agree that the fish community traits indicate a NOTGOOD status. The ecological status of these lakes is unanimously agreed, and thus can be used as common benchmark. The NOTGOOD lakes are used as a subset of the ‘real’ NOTGOOD lakes. Based on this subset, the national EQR class boundaries for the good/moderate transition are evaluated. For this purpose, the set of NOTGOOD lakes does not necessarily have to include all national NOTGOOD lakes. But to be useful it should fulfil three criteria

- 1) No GOOD_{exp} lake should be identified as NOTGOOD_{fish} with the benchmarking procedure
- 2) A significant proportion of NOTGOOD lakes should be correctly identified with the fish benchmark system
- 3) For each MS, the highest EQR value of the NOTGOOD lakes should not be much lower than the national good/moderate class boundary

The first criterion is needed to avoid an unjustified raise of the national G/M class boundary. Please explain. The goal of reaching 0 % of false-positive classification is achieved.

The second criterion guarantees, that the subset of NOTGOOD lakes is representative. Tab. 11 shows the corresponding values. The correct identification of NOTGOOD ranges from app. 60 % to 100 %. Therefore, a representative proportion of disturbed lakes is included in the set of NOTGOOD lakes for every member state. Please note the differences between Tab.

10 and Tab. 11. The percentages change with the basis of calculation, e.g. 78 % of the NOTGOOD lakes (expert judgment) are correctly identified by the benchmarking procedure compared to 50 % of all lakes that are both NOTGOOD and correctly identified in Tab. 10. The third criterion assures that the set of benchmark lakes is not restricted to the lowest national EQR values, i.e. the ‘worst’ of the NOTGOOD lakes (poor/bad).

Tab. 11: Numbers of the lakes with NOTGOOD ecological status class according to expert judgment and percentages of correct positive identification with the CB fish system (polymictic lakes).

POLY MS	Exp. NOTGOOD N lakes	Exp. NOTGOOD Fish corr. pos. %
DE	30	80 %
DK	17	59 %
EE	5	80 %
LT	10	80 %
NL	8	88 %
PL	5	100 %
Total	75	74 %

The high percentages of correct positive identification support the use of NOTGOOD benchmark lakes as representative subsample of the disturbed lakes. The NOTGOOD lakes represent the intercalibrated expert opinions on the ecological status of lakes indicated by certain values of fish metrics. Now, the national class boundaries are controlled for their compliance with this harmonized CB definition of NOTGOOD. To guarantee that all NOTGOOD lakes according to the harmonized definition are also moderate or worse in the national assessment, the following rule is applied:

The national G/M class boundary has to be higher than the highest EQR value of the national NOTGOOD lakes.

The lakes of the NOTGOOD benchmark set with the highest EQR values in the national assessment systems are shown in Tab. 12. The national good/moderate class boundaries are added. In all cases, the highest EQR values of NOTGOOD lakes are close to the national good/moderate class boundary. This supports the assumption, that the set of benchmark lakes is representative.

Tab. 12: Polymictic NOTGOOD benchmark lakes with the highest EQR values of the corresponding member state’s national system (for Poland, national EQR values were not available) and current national good/moderate class boundaries are added.

MS	NOTGOOD Lake name	NOTGOOD Lake EQR moderate if	MS EQR moderate if	comment
DE	Malchower See Seddiner See	≤ 0.80 ≤ 0.80	< 0.80	Accepted, raise of the national class good/moderate boundary
DK	Hanvejle	≤ 0.58	$< 0.60 ?$	Accepted, no change of the national class boundaries

EE	Korijärv	≤ 0.16	< 0.19	Accepted, no change of the national class boundaries
LT	Siesikai	≤ 0.67	< 0.62	Not accepted
	Kavalys	≤ 0.56		Accepted, no change of the national class boundaries
NL	Veluwe	≤ 0.54	$< 0.60 ?$	Accepted, no change of the national class boundaries

For Germany and Lithuania, the highest EQR values of the benchmark lakes are higher than the national good/moderate class boundary and thus modifications are necessary to achieve comparability to other MS.

For Germany, the good/moderate class boundary has to be raised from < 0.8 to ≤ 0.8 (is moderate). This might appear negligible, but the EQR values obtained by the German system are not continuous. Raising the class boundary changes the national assessment of 4 lakes, which is almost 8 % of all German lakes included in the intercalibration dataset.

A special case is Lake Siesikai in Lithuania. This lake has a moderate ecological status according to expert judgment and a NOTGOOD status according to the CB system for identification of degradation. However, the LT LaFi system does not indicate degradation. Lake Siesikai is impounded not only by eutrophication but also by artificial water level raising and high turbidity caused by erosion of soils. The latter two pressures are unusual in Lithuania and the national system is not suitable to indicate them. Lake Siesikai is in a degraded condition, but can not be assessed properly with the Lithuanian system. A raise of the national class boundaries based on this lake is not justified and would lead to an unacceptable amount of wrong indication of degradation (T. VIRBICKAS, pers. comm. Nov. 2011). Therefore, the second highest EQR value of the NOTGOOD lakes should be taken as relevant for setting the national class boundaries.

The national LakeFish systems of Denmark, Estonia and The Netherlands are precautionary; the national class boundaries are higher than the highest EQR values of degraded lakes. It should be noted that for The Netherlands, selection of benchmark lakes and national assessment results are based on different fishing methods that have been made comparable using parallel fishing.

Tab. 12 and the preceding results show:

- A harmonization of the opinions of the CB experts on which traits of the fish community indicate a degraded status of a lake can be achieved.
- The NOTGOOD lakes identified by the proposed benchmarking procedure cover a representative subset of all moderate/poor/bad lakes.
- The national class boundaries roughly reflect the ecological status as determined by the harmonized CB expert opinions.

Stratified lakes

The procedure of identifying threshold value for fish metrics indicating a NOTGOOD status, of selecting benchmark lakes based on the metrics and of the modification of the national class boundaries for STRAT+ lakes is similar to the one chosen for POLY lakes. The description is abbreviated, but some contents are repeated for better understanding.

Step S1: thresholds and selection of useful metrics

The metrics tested for their suitability for a common system to identify NOTGOOD stratified lakes are listed in Tab. 11. For all metrics, the threshold values for the identification of NOTGOOD lakes are determined as described before. An explanation of the metrics and the figures can be found in the Annex. Data for 89 polymictic lakes is present. According to the expert judgment, 48 of them are in a NOTGOOD ecological status. Principally, 11 metrics and the EQR values are useful to identify NOTGOOD lakes (Tab. 11). Two metrics are not suitable, because the range of values belonging to GOOD lakes encompasses the range for NOTGOOD lakes.

The following metrics were discarded for the identification of benchmarking NOTGOOD lakes:

- obligatory species number: the lakes are also identified by littoral species missing, WPUE and/or benthic
- Median *Abramis brama* and *Abramis*+*Rutilus* [%weight]: not suitable
- Median *Perca fluviatilis*: only one lake identified also by benthic and *Abramis brama* [%number]
- Median *Rutilus rutilus*: lakes are also identified by WPUE
- Benthivora common system and benthic common metric: very close to benthic common system, additional lakes identified with these two metrics are also identified by WPUE/*Abramis brama* [%number]
- EQR values: all lakes identified by WPUE, benthic or *Abramis brama* [%number]

Selected metrics:

- *Abramis brama* [%number]
- *Gymnocephalus cernuus* [%weight or %number]
- WPUE
- benthic common system
- species specific percentages of *Abramis brama* and
- littoral species missing

The exclusion of outliers in the threshold setting leads to an amount of 4 lakes identified as NOTGOOD using fish metrics but GOOD according to expert judgment.

Tab. 13: Metrics, threshold values for identification of NOTGOOD lakes and performance in terms of number of NOTGOOD lakes correctly identified as NOTGOOD and GOOD lakes falsely classified as NOTGOOD (stratified lakes, n= 89).

POLY metric	NOTGOOD if	NOTGOOD identified	GOOD false	comment
Metrics of the common system				
Oblspecnum [n species]	< 6 species	5	1	
WPUE [kg/m ²]	> 0.06	10	0	
Benthic DE [%weight]	> 68	13	1	
Benthivor [%weight]	> 37	6	0	
Abramis brama [%number]	> 4.4	9	2	
Gymnocephalus cernuus [%weight]	> 6.5	2	0	For stratified lakes (STRAT)
Gymnocephalus cernuus [%number]	> 21.0	3	0	For deep stratified lakes (DEEP)
Median Perflu [g]	> 62	1	0	
Median Abrabr [g]				Not suitable
Median Rutrut [g]	> 83	4	0	
Metrics of the common metric				
Abramis + Rutilus [% weight]				Not suitable
Benthic CB [% weight]	> 30	6	0	
Littoral species missing	2	2	0	
EQR values				
DE EQR	< 0.57	6	0	
CB EQR	< 0.63	10	0	

Step S2: identification of the NOTGOOD benchmark lakes and performance

The subset of benchmarking NOTGOOD lakes is identified with the five metrics selected in step S 1 (for **stratified** lakes):

- Abramis brama [% number]
- Gymnocephalus cernuus [% weight] for stratified and [% number] for deep stratified lakes
- WPUE [kg/m²]
- Benthic net species common system [% weight]
- Littoral species missing [n species]

Similar to the previous strategy, 3 lakes were reset to the status UNKNOWN that had been classified as NOTGOOD on the basis of outliers only. The performance of the NOTGOOD identification procedure is shown in Tab. 14. A good percentage of app. 80 % of lakes is assigned correctly; about 20 % are not identified degraded lakes.

Tab. 14: Performance of the UNKNOWN/NOTGOOD fish benchmarking procedure for stratified lakes. Values show percentage of total number of lakes for each member state and a summary. Further explanations can be found in the text.

STRAT+ MS	N lakes	Exp. NOTGOOD Fish corr. pos. %	Exp. NOTGOOD Fish false neg. %	Exp. GOOD Fish corr. neg. %	Exp. GOOD Fish false pos. %
DE	58	27	33	40	0
EE	8	38	0	62	0
LT	18	44	6	50	0
PL	5	20	0	80	0
GIG	89	32	22	46	0

Step S3: modification of the national class boundaries for the G/M transition

The preconditions to justify the modification on the basis of a harmonized subset of NOTGOOD lakes are:

- No GOOD lake should be identified as NOTGOOD with the benchmark system
- A significant proportion of NOTGOOD lakes should be correctly identified with the fish benchmark system

The values are shown in Tab. 15. The percentages of correctly identified NOTGOOD lakes range from 46 to 100 %; this is rated to be a reliable subsample of the total NOTGOOD lakes.

Tab. 15: Numbers of the lakes with NOTGOOD ecological status class according to expert judgment and percentages of correct positive identification with the CB fish system (stratified lakes).

STRAT+ MS	Exp. NOTGOOD N lakes	Exp. NOTGOOD Fish corr. pos. %
DE	35	46 %
EE	3	100 %
LT	9	89 %
PL	1	100 %
Total	48	58 %

The principle of the proposed way of intercalibration is, that the national good/moderate class boundary has to be higher than the highest EQR value of the subset of national NOTGOOD lakes. For the definition of the benchmark lakes, the CB lake types STRAT and DEEP were

combined. Germany and Lithuania use type specific metrics and/or type specific class boundaries for the good/moderate transition. Therefore, the combined set of STRAT+ benchmark lakes is re-divided into the original lake types STRAT and DEEP for adaptation of the national class boundaries. The corresponding values are shown in the Tab. 16.

For stratified lakes of the STRAT type, the national class boundaries cover all NOTGOOD lakes of the benchmarking set. There is a common agreement on the good/moderate transition and no modification of national class boundaries is necessary.

For deep stratified lakes (DEEP type), increases of the national class boundaries are necessary for the German and the Lithuanian assessment method.

Tab. 16: Stratified lakes with a NOTGOOD status according to the CB identification procedure and the highest EQR values of the corresponding member state's national system. National good/moderate class boundaries are added for direct comparison.

MS	Lake Type	NOTGOOD Lake name	NOTGOOD Lake EQR	MS EQR moderate if	comment
DE	STRAT	Wolziger See	0.75	< 0.80	accepted, no change of the national class boundaries
	DEEP	Schaalsee	0.82	< 0.80	accepted, raise of the national class good/moderate boundary
EE	STRAT/DEEP	Aheru	0.16	< 0.19	accepted, no change of the national class boundaries
LT	STRAT	J. Kauknoris	0.52	< 0.61	accepted, no change of the national class boundaries
	DEEP	Druksiai	0.75	< 0.60	accepted, raise of the national class good/moderate boundary

Similar to the situation for polymictic lakes, the expert's opinions on the diagnostic relevance of certain fish metrics are comparable in terms of the ecological status. There is a prevailing consensus on what is good and what is bad. The national class boundaries cover the EQR values of the benchmarking NOTGOOD lakes for stratified lakes, but two of three member states have to raise their national class boundaries for deep, stratified lakes.

Summary and Discussion

The intercalibration of systems to assess the ecological status of lakes with the fish fauna in the Central/Baltic group has to face specific issues: no agreement on significant pressures and relevant threshold values for abiotic characterization of ecological status of lakes, no reliable correlations of pressures and metric values, no benchmark sites, no working common system and no common metric with adequate correlations. The disadvantageous situation interdicts the application of one of the three options for intercalibration proposed by the Intercalibration Guidance (CIS 2011).

An alternative solution to compare the national class boundaries is developed. It focuses on the good/moderate class boundaries - assessing the high/good class boundary is not possible due to a negligible number of reference lakes in the dataset. Based on expert judgments, fish metrics were tested for their suitability to reliably identify lakes in a degraded ecological condition. It is possible to identify metric thresholds that separate unequivocally degraded lakes from lakes with an unknown ecological status. Each metric can identify a limited number of degraded lakes, but the combination of metrics provides an adequate set of lakes in degraded status according to the CB experts. The set of degraded lakes was used as benchmark to control the national class boundaries. The basic precondition is that all lakes in a degraded status according to all CB experts should also be identified as degraded with the national assessment system.

The results show, that the CB experts have a roughly comparable opinion on which fish metric values/fish communities have a diagnostic value for the difference between high/good status and significant degradation (moderate/poor/bad). The national class boundaries are always close to the EQR values determined with the benchmark set. Thus a general comparability of the levels of ecosystem alteration associated with the national transition from good to moderate assessment results is shown. In most cases, the national EQR values assigned to the good/moderate transition are above the values indicated with the benchmark, indicating a precautionary approach. The possibility to lower national class boundaries is not considered in the present investigation. However, the need to increase national class boundaries was identified for Germany and Lithuania. Tab. 17 shows the results of the intercalibration process. It has to be taken into account that the evaluation of the national class boundaries is specific for the intercalibration types while national systems might be independent on them.

Tab. 17: Summary of the intercalibration results and consequences for adjustment of national assessment systems.

Member state	Lake Type	Intercalibrated class boundary moderate if	National class boundary moderate if	effect
Denmark	POLY	≤ 0.58	$< 0.60 ?$	National boundary can be maintained
Estonia	POLY	≤ 0.16	< 0.19	National boundary can be maintained
	STRAT	≤ 0.16	< 0.19	National boundary can be maintained
	DEEP	≤ 0.16	< 0.19	National boundary can be maintained
Germany	POLY	≤ 0.80	< 0.80	National class boundary has to be raised
	STRAT	≤ 0.75	< 0.80	National boundary can be maintained
	DEEP	≤ 0.82	< 0.80	National class boundary has to be raised
Lithuania	POLY	≤ 0.56	< 0.62	National boundary can be maintained
	STRAT	≤ 0.52	< 0.61	National boundary can be maintained
	DEEP	≤ 0.75	< 0.60	National class boundary has to be raised
Netherlands	POLY	≤ 0.54	$< 0.60 ?$	National boundary can be maintained

A confirmation of the intercalibration results by the national experts and a remark on the implementation into the national assessment systems is to be added.

The lake fish assessment is an ongoing process and more lakes will be assessed with the national systems in the future. The intercalibration exercise is meant to reach a finalized state. However, it is suggested that the identification of disturbed lakes with the procedure presented is continuously applied by the member states besides the national assessment. This can warrant the enduring similarity of the interpretation of ecological alterations.

The performance of the UNKNOWN/NOTGOOD fish benchmarking procedure is surprisingly good (Tab. 10, Tab. 14). It is left to the member states to check if the metrics and corresponding threshold values can help increasing the reliability of the national systems. The procedure was developed to identify lakes in an unmistakable disturbed condition. This could be done with the help of expert judgments of the lake statuses and the exclusion of outliers. The applicability as an assessment system is limited. It would require an expert's judgment on if or if not a metric indicating a NOTGOOD status is likely to be an outlier.

The concept of using a subset of benchmark lakes in surely disturbed conditions for evaluation of class boundaries can easily be transferred to other BQE and thus used as an additional intercalibration option (if it becomes officially accepted). The distinction of GOOD and NOTGOOD has not necessarily to be based on expert judgment but could also be based on threshold values of certain abiotic measures (e.g. total phosphorous).

Annex for section ‘alternative option’

Description of metrics tested for suitability for the benchmark procedure

All metrics are based on randomized multimesh-net fishing procedure follow the European standard (CEN 2005). The metric values are calculated for benthic nets. All nets are included, there is no selection of depth strata and empty nets are kept. For species numbers, additional methods for the evaluation of presence/absence are necessary (electrofishing, fisheries statistics, literature).

The metric descriptions include an explanation, the expected reaction to pressures, literature sources, national assessment systems using the metric and the measure used in the present investigation. An abbreviation is given for metrics with figures in the annex.

Metrics of the common system

Obligatory species number: Obligatory species are species that should occur in all lakes in high/good condition. They are widespread and have comparably low demands on environmental conditions. Their absence indicates severe alteration of the fish community. Species are *Abramis brama*, *Esox lucius*, *Gymnocephalus cernuus*, *Perca fluviatilis*, *Rutilus rutilus*, *Scardinius erythrophthalmus* for POLY lakes also *Blicca bjoerkna*. The metric is used in the German LaFi system. Measure is [n species].

Abramis brama (Abrbra): *Abramis brama* is a widespread species with a usually high abundance in polymictic lakes. The species’ abundances increase with increasing eutrophication (BELPAIRE et al. 2000; GARCIA et al. 2006; JAARSMA 2007; JEPPESEN et al. 2000; MEHNER et al. 2004). Like all species-specific metrics, *Abramis brama* can only be assessed where it is naturally present. It is used in the national LaFi systems of Denmark (in combination with *Rutilus*), Germany and The Netherlands and Poland. Measure is [%weight] for POLY lakes and [%number] for STRAT+ lakes.

Blicca bjoerkna (Blibjo): The species is ecologically similar to *Abramis brama* and reacts similar to eutrophication. The metric is used in the German and the Lithuanian LaFi systems. Measure is [% weight].

Gymnocephalus cernuus (Gymcer): The species indicates degradation, although unspecific. Increased abundances are observed in eutrophied, constructional modified, intensively used or heavily fished lakes. The metric is used in the German and the Estonian LaFi systems. Measure is [%weight] for POLY/STRAT lakes and [%number] for DEEP lakes.

Perca fluviatilis (Perflu): An indicator of degradation, either eutrophication or structural degradation (the latter can be caused by eutrophication). Abundance decreases in the mesotrophic/eutrophic region and the metric is restricted to polymictic lakes. It is used in most national LaFi systems: Denmark (in combination with *Rutilus*), Germany, Lithuania, The Netherlands and Poland. Measure is [% weight].

Sander lucioperca (Sanluc): Pikeperch is a top-level predator with competitive advantages in turbid water (compared to Pike and Perch). The species’ abundance increases with increasing eutrophication. Occurrence is high enough to allow assessment in POLY lakes, only. The metric is used in the German and the Polish LaFi systems. Measure is [% weight].

Standardized catch of weight (WPUE): Increased nutrient supply leads to an increase in overall bioproduction. Therefore, high standardized catches indicate eutrophication. WPUE a widespread metric used in the national systems of France, and Germany. It is also proposed as metric by the WISER project (CAUSSE et al. 2011b; PEDRON et al. 2010).

Measure for WPUE is [kg/m²] in one night including dusk and dawn (no reference to the hours of settlement). Denmark and Estonia use the standardized catch of number (NPUE).

Benthic species - common system (benthic): Full term is ‘benthic net species’; the metric is a combination of functional traits of the fish community and their methodological evidence. It includes the following species that are best caught with benthic nets: *Abramis brama*, *Ballerus* (syn. *Abramis*) *ballerus*, *Blikka bjoerkna*, *Carassius gibelio*, *Carassius carassius*, *Coregonus* sp. (phenotype *lavaretus*, not *albula*), *Cyprinus carpio*, *Gymnocephalus cernuus*, *Rutilus rutilus*, *Sander lucioperca*. The metric is a measure of the abundance of littoral species (and to a small amount of pelagic species). Perturbation primarily affects the shoreline/littoral zone. Eutrophication also leads to structural degradation of the littoral regions by reducing submersed macrophytes. The pressures should lead to a decrease of more shoreline orientated species. Therefore a high percentage of benthic net species indicates unspecific degradation. Measure is [% weight]. Benthic and/or benthivore species are used in the assessment systems of Germany, Lithuania, The Netherlands and Poland

Benthivore species: Is a functional trait of the fish community describing the food preference. Is close to the metric benthic net species and has the same dependence on pressures. Main difference is the exclusion of the prevailingly abundant *Rutilus rutilus* in the metric. Species are: *Abramis brama*, *Blikka bjoerkna*, *Coregonus* sp. (phenotype *lavaretus*, not *albula*), *Cyprinus carpio*, *Gymnocephalus cernuus*, *Tinca tinca*. Measure is [% weight].

Median weight *Abramis brama*, *Perca fluviatilis*, *Rutilus rutilus*: The medians of weight are size descriptors for the most abundant species (the calculation of a median requires enough individuals). As catches with multimesh nets are not representative for lower size classes, the metric values are calculated with the fish above a lower length limit; i. e. *Abramis* ≥ 10 g, *Perca* ≥ 6 g and *Rutilus* ≥ 14 g. Eutrophication should lead to an increased individual growth while all pressures can reduce the recruitment of lower age classes. A higher value of median of weight indicates degradation. Size metrics are used in the LaFi systems of Denmark (mean individual biomass of all species), Germany (median weight of selected species - this metric), Lithuania (mean weight of *Rutilus rutilus*), and Poland (fraction of small Cyprinids).

Metrics of the common metric

WPUE: see above

Abramis brama + Rutilus rutilus (Abrbra + Rutrut): Combines the most abundant Cyprinids to a single measure. Both species show increasing abundances with increasing eutrophication. However, both species compete and their combination to a single metric should counterbalance the suppression of one of them. The metric is used in the Danish LaFi System. Measure is [% weight].

Benthic species - common metric: Is a functional fish community trait describing the habitat preference. It encompasses the combined abundance of the following species: *Abramis brama*, *Blicca bjoerkna*, *Carassius gibelio*, *Carassius carassius*, *Coregonus* 'Whitefish', and *Cyprinus carpio*. The abundance of benthic species increases with increasing eutrophication. Measure is [% weight].

Littoral species missing: The metric includes the presence of three species that are ecologically bound to littoral structures, especially macrophytes: *Esox lucius*, *Scardinius erythrophthalmus*, and *Tinca tinca*. Their absence indicates massive deterioration of shoreline integrity. For this metric, the number of absent species that would naturally be present is counted. Measure is [n species].

EQR values

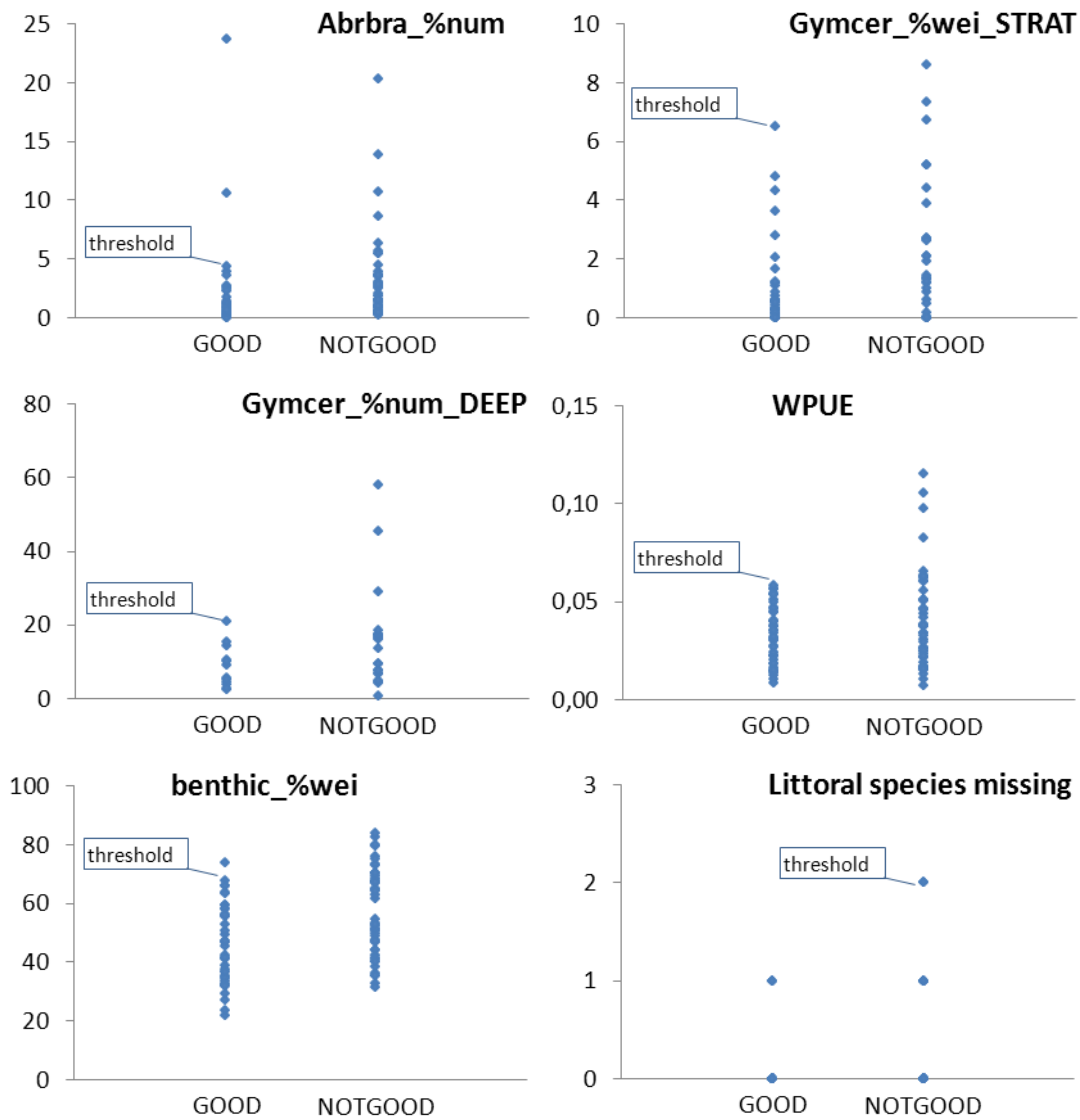
For both the common system and the common metric (4 metric mini-system) each individual metric is scored with one out of five classes according to the WFD nomenclature. The EQR values for both approaches are calculated in the 'traditional' way (LIT):

$$\text{EQR} = (\text{sum} - \text{min}) / (\text{max} - \text{min})$$

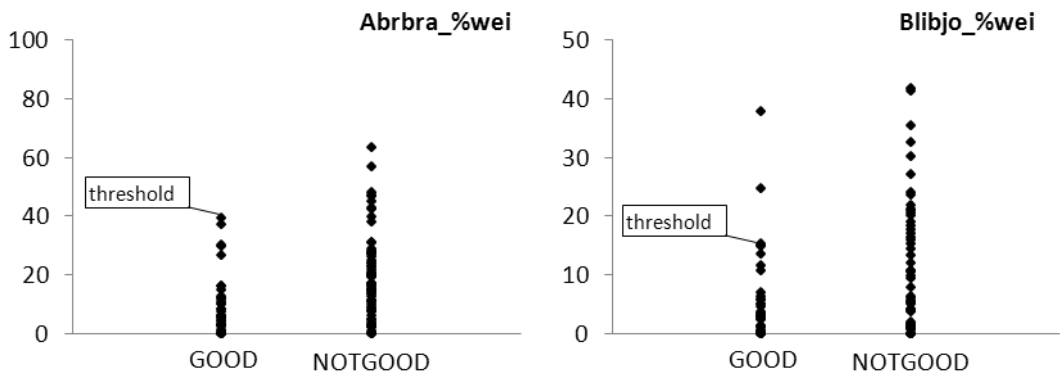
Sum: sum of scores, min: minimum sum of scores possible (= number of metrics), max: maximum sum of scores (5 times the number of metrics)

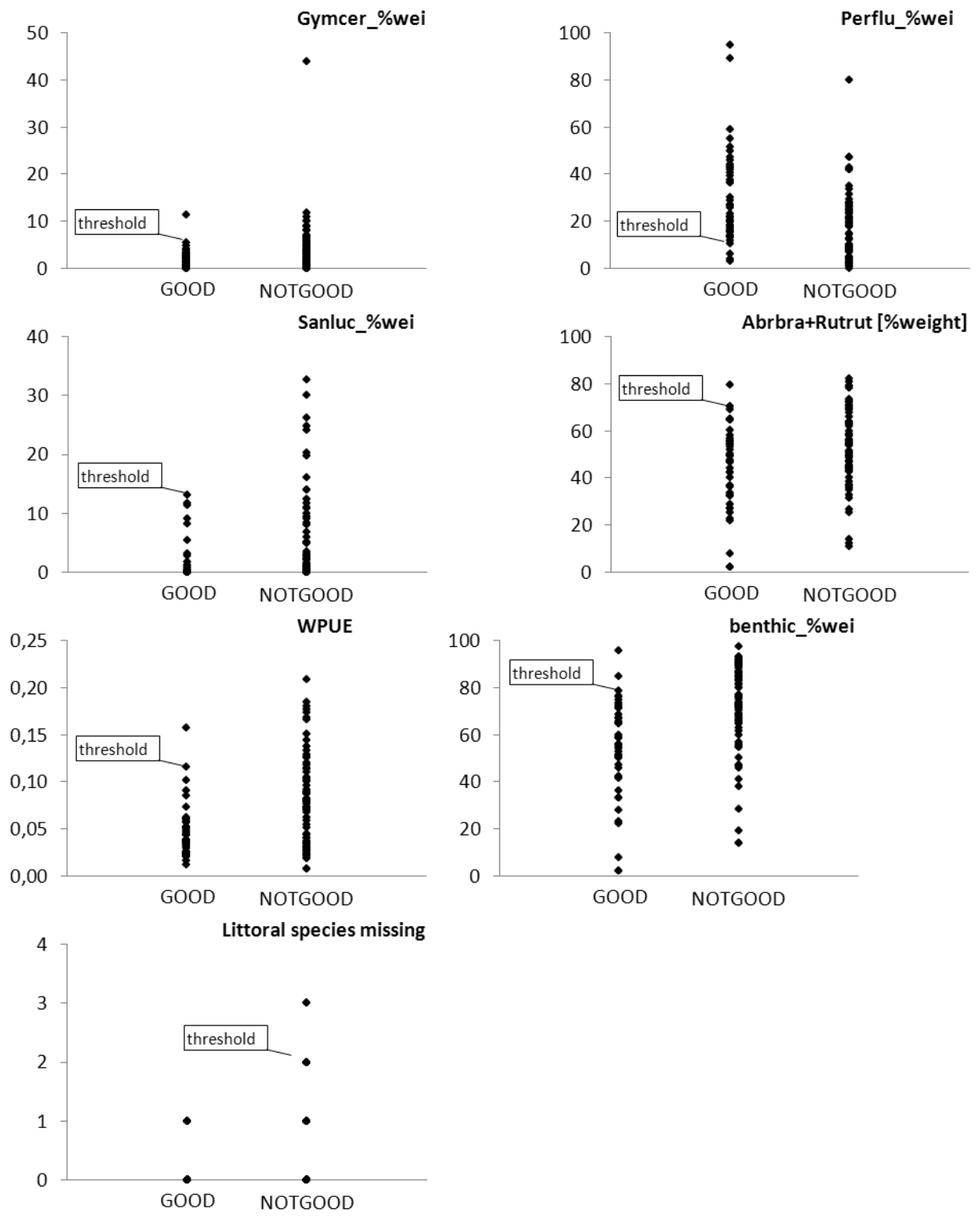
The values range from 0 to 1. Low EQR values show major alterations of fish community traits which should reflect major pressures.

Figures for stratified lakes (STRAT+)



Figures for polymictic lakes (POLY)





Comments on the suitability of the BQE fish to address specific pressures in Alpine and Central/Baltic lake assessment systems

The first version of this document (section) was drafted at the 2nd Alpine GIG meeting, March 10, 2011 in Scharfling, Austria.

It represents the point of view of the following MS:

Alpine: Austria (GIG lead), Germany, Italy, Slovenia

CB: Germany (GIG lead), MS were informed and asked for their opinion. General support: BE, critical view: EE, disagreement: FR.

Last update: November 2011

Fish as bioindicators in lakes

The WFD has included the fish fauna as an inevitable quality element to assess the ecological status of lakes. In addition to phytoplankton, macrophytes and makrozoobenthos the fish fauna completes an integrated assessment as the upmost trophic level in the ecosystem. Fish are long living and mobile and thus have an integrating value to indicate changes in the ecological status over a wider range of place and time. It is well known, that fish communities are suitable to assess the ecological status and/or the biological integrity of lakes. They show reactions on:

- eutrophication (BELPAIRE et al. 2000; GARCIA et al. 2006; HELMINEN et al. 2000; JEPPESEN et al. 2005b; LAUNOIS et al. 2011; MEHNER et al. 2004; PEDRON et al. 2010; PERSSON et al. 1991; SONDERGAARD et al. 2005; WHITTIER 1999)
- habitat destruction, shoreline degradation, lake use intensity (BELPAIRE et al. 2000; JENNINGS et al. 1999; WHITTIER 1999)
- HYMO degradation, connectivity (DEGERMAN et al. 2001)
- acidification (APPELBERG et al. 2000)
- combined degradation (WHITTIER 1999)

A summary can be found in (BECK & HATCH 2009).

Based on this knowledge, scientists in many European countries investigated the suitability of the fish communities in Lakes to indicate anthropogenic deterioration (MED and EC GIG not included):

- Austria (GASSNER & WANZENBÖCK 2005; GASSNER et al. 2003; GASSNER et al. 2005; ZICK et al. 2006)
- Belgium (BELPAIRE et al. 2000)
- Denmark (ANDERSON et al. 2005; JEPPESEN et al. 2005a; JEPPESEN et al. 2000; JEPPESEN et al. 2007; JEPPESEN et al. 2005b; SONDERGAARD et al. 2005)
- Finland and Sweden (APPELBERG 2000; APPELBERG et al. 2000; HOLMGREN & APPELBERG 2000; TAMMI et al. 2001; TAMMI et al. 2003)
- France (CAUSSE et al. 2011a; CAUSSE et al. 2011b; LAUNOIS et al. 2011; PEDRON et al. 2010)
- Germany (BRÄMICK et al. 2008; DIEKMANN et al. 2005; GARCIA et al. 2006; MEHNER et al. 2005a; MEHNER et al. 2004; MEHNER et al. 2005b)
- Netherlands (JAARSMA 2007)
- international (JEPPESEN et al. 2005b; MEHNER et al. 2007)

A huge additional amount of scientific knowledge gained in the process of WFD-compliant LakeFish assessment methods is unpublished or hidden in grey literature. In both the Alpine and the Central/Baltic GIG working LakeFish assessment systems have been developed in

Austria, Denmark, Estonia, France, Germany, Italy, Lithuania, The Netherlands, Poland. In summary, fish are a suitable and well chosen BQE to assess human impacts on lakes.

Problems using pressure-impact relationships for fish in lakes

For the development of assessment systems, the accompanying guidance demands well established relationships between human pressures and reaction of the metrics. This is regarded as essential for system development, class boundary setting and intercalibration. In spite of the well known impacts of certain pressures, such highly significant correlations are not present in most pressure-impact-analyses for fish in lakes. The reasons are manifold:

- A lake is an individual water body but is comparably huge and usually consists of sub-habitats, which occur several times. Fish are mobile and can avoid local pressures. Therefore, the whole fish community will not show reactions (a difference to the assessment of river sections).
- Lakes are complex ecosystems containing a diversified network of interacting trophic levels. Pressures affecting the ecological status of lakes are highly interdependent, e.g. lakes in urban regions are often influenced by lake use, shoreline degradation, eutrophication and water level regulation. Lakes influenced by single pressures are rare or absent¹¹. Many pressures impact the fish community in similar ways, e.g. via destruction of submersed macrophytes. Furthermore, natural variation can hide the effect of anthropogenic impacts in all trophic levels. Fish as the topmost level of the trophic cascade are indirectly integrating effects of pressures acting on any level below (bottom up-hypothesis). Thus linear relationships between individual pressures and fish-metrics may be lacking and it make it impossible to identify pressure specific effects (a difference to pressure specific BQE like phytoplankton for eutrophication)
- Fish represent the upmost trophic level in a lake and are a BQE integrating space and time. Fish are resilient to pressures. Fish should be used for an integrated assessment of pressure intensity while other BQE can be used to address specific pressures.
- Using highly correlated relationships leads to the result, that direct pressure measurements are replaced by imprecise and ineffective BQE metric measurements. The fish experts feel that this is not the intention of the WFD. The aim of the fish assessment should be to compare fish communities in reference condition with the current status. This can be done effectively with the present approaches.

The conclusion for fish in lakes is that highly correlated pressure-impact relationships can not be expected. Additionally, they are not necessary for effective assessment of ecological status. For fish in lakes, a reaction of the total assessment result (EQR or status class) to the combined intensity of all pressures is a suitable measure of assessment quality. This has successfully been tested, e.g. using expert's judgment.

Recommendations

We strongly recommend recognizing fish as a valuable BQE to assess the ecological status of lakes. Effective and WFD compliant lake assessment methods based on the fish communities are developed. However, some of the recommendations in the IC guidance are specific to certain BQE or water bodies and can not be transferred to the LakeFish assessment. We suggest leaving this freedom to the fish experts.

¹¹ This is not valid for acidification which occurs in lakes not impacted by other pressures. Acidification is not relevant for the Alpine GIG and of limited relevance for the CB GIG.

Consequences: The previously described point of view towards pressure-impact relationships was submitted as a scientific explanation of the specific challenges of the assessment of lakes with fish. The points were discussed in the ECOSTAT meeting of 30-31 March 2011. Conclusion was:

‘It is agreed to MS will demonstrate a reaction of the total assessment result to the combined intensity of all pressures (if necessary, using expert judgment) and clearly document it in the reports’.

Citation refers to VAN DE BUND et al. (2011): Method compliance checking: which methods can be included in the intercalibration exercise? Document discussed at the ECOSTAT meeting of 30-31 March 2011. Document with ECOSTAT conclusions available at CIRCA: Lakes - fish intercalibration/CB GIG/official documents.