Endbericht im Vorhaben

„Interkalibrierung Makrophyten und sehr große Fließgewässer“

im Auftrag der Länderarbeitsgemeinschaft Wasser (LAWA), Projekt-Nr. O 7.09 im Länderfinanzierungsprogramm "Wasser, Boden und Abfall".

Dr. Sebastian Birk

Essen, März 2011

Inhaltsverzeichnis

Erläuterung der Abkürzungen

BQE	Biologische Qualitätskomponente
$CB_{riv}GIG$ Makrophyten	Arbeitsgruppe „Interkalibrierung Makrophyten in Fließgewässern im GIG Mitteleuropa/Baltikum“
CIS	Gemeinsame Umsetzungsstrategie der EU
ECOSTAT	Arbeitsgruppe „Ökologischer Zustand“ der Gemeinsamen Umsetzungsstrategie für die WRRL (CIS)
EQR	Ökologischer Qualitätsquotient (Ecological Quality Ratio)
GIG	Geographische Interkalibrierungs-Gruppe
IK	Interkalibrierung
LAWA	Bund/Länderarbeitsgemeinschaft Wasser
mICM	Interkalibrierungsindex Makrophyten (macrophyte Intercalibration Common Metric)
WRRL	EG-Wasserrahmenrichtlinie

1 Einleitung

Die EG-Wasserrahmenrichtlinie 2000/60/EC fordert für die Oberflächengewässer der Mitgliedstaaten der Europäischen Union den guten ökologischen und chemischen Zustand. Im Prozess der Interkalibrierung werden die Ergebnisse der nationalen Verfahren zur Zustandsbewertung verglichen und harmonisiert. Der erste Teil des vorliegenden Berichts fasst die im Projektjahr 2010 geleisteten Arbeiten zur Interkalibrierung der Bewertungsverfahren für Makrophyten in Fließgewässern zusammen. Viele der vorgestellten Inhalte basieren auf Vorarbeiten zu diesem Themengebiet, die in Birk et al. (2007) und Birk (2008, 2009, 2010) dargestellt sind. Im zweiten Berichtsteil werden die Arbeiten zur Interkalibrierung sehr großer Fließgewässer präsentiert.

Der Arbeitsauftrag für das Projektjahr 2010 umfasste verschiedene Teilaspekte der Interkalibrierung. Inhaltlich zielten die Leistungen auf die Erfüllung der Vorgaben der neuen Interkalibrierungsleitlinie ab (European Communities 2010), welche die Aufgaben zur Umsetzung der Interkalibrierung im Jahr 2010 klar umgrenzt: Die Überprüfung der „Interkalibrierbarkeit“ der nationalen Bewertungsverfahren, die Schaffung von einheitlichen Grundlagen zum Vergleich der nationalen Klassengrenzen sowie die Definition von Ankerpunkten für die Interkalibrierung. Die im Projektzeitraum durchgeführten Arbeiten lassen sich folgendermaßen zusammenfassen:

Interkalibrierung Makrophyten in Fließgewässern

- Aktualisierung der Übersicht der nationalen Verfahren zur Bewertung von Makrophyten in Fließgewässern;
- Herleitung von alternativen Ankerpunkten für die Interkalibrierung der Makrophyten im GIG Mitteleuropa/Baltikum;
- Berichterstattung an die Europäische Kommission;
- Vorbereitung und Durchführung von Treffen der IK-Gruppe.

Interkalibrierung sehr großer Fließgewässer

- Zusammenstellung der nationalen Verfahren zur Bewertung sehr großer Fließgewässer in Europa;
- Vorstudien zur Ableitung einer Interkalibrierungs-Typologie für sehr große Fließgewässer in Europa und zur Herleitung alternativer Ankerpunkte für die Interkalibrierung sehr großer Fließgewässer;
- Berichterstattung an die Europäische Kommission;
- Vorbereitung und Durchführung von Treffen der IK-Gruppe.

2 Zur CIS-Leitlinie Interkalibrierung

Ein Schwerpunkt der Arbeiten bestand in der Kommunikation der Inhalte der neuen CIS-Leitlinie zur Interkalibrierung (European Communities 2010). Die dort vorgestellten Grundprinzipien sowie Informationen zu Durchführung und Zeitplan sind bereits in Birk (2010) zusammengefasst. Anhand von zwei Präsentationen wurden im relevanten Projektjahr sowohl der LAWA-Expertenkreis „Biologische Bewertung und Interkalibrierung“ als auch die Gruppe der Experten zur Makrophyten-Interkalibrierung unterrichtet (Anhänge 1 und 2).

Die Inhalte der Präsentation vor dem LAWA-Gremium zum Themengebiet „Referenzbedingungen und alternative Ankerpunkte“ (Annex III der IK-Leitlinie) werden an dieser Stelle nochmals wiedergegeben. Der Annex III definiert zwei Ziele:

- Schaffung eines allgemeinen Rahmens zur Herleitung von Referenzbedingungen für die nationalen Bewertungsverfahren, und
- Definition verbindlicher Methoden zur Herleitung von Referenzbedingungen/alternativen Ankerpunkten für die Interkalibrierung.

Die Ausweisung eines Netzwerks von Referenzstellen wird ausdrücklich empfohlen. Dabei können diese Referenzstellen BQE-spezifisch ausgewählt werden, wenn keine signifikante anthropogene Veränderung der jeweiligen Lebensgemeinschaft angezeigt ist. Der Grad an anthropogener Belastung ohne Auswirkung auf die Lebensgemeinschaft (=Referenzkriterien) sollte durch statistische Verfahren quantifiziert werden. Die Möglichkeiten von (Dosis-Wirkungs-) Modellierungen von Referenzbedingungen sowie die Nutzung von Expertenwissen sind explizit erwähnt.

Für die Interkalibrierung sind Referenzbedingungen/alternative Ankerpunkte einheitlich herzuleiten. Diese Herleitung orientiert sich an den Empfehlungen für die nationalen Bewertungsverfahren. Alternative Ankerpunkte sollen definiert werden, wenn (genügend) „echte“ Referenzstellen nicht vorhanden sind. Der Grad der Abweichung dieser alternativen Ankerpunkte von der „echten“ Referenz ist zu quantifizieren.

3 Interkalibrierung Makrophyten in Fließgewässern der GIG Mitteleuropa/Baltikum

Die nach IK-Leitlinie im Projektjahr 2010 durchzuführenden Arbeitsschritte lassen sich wie folgt zusammenfassen:

Datengrundlage und Vorbereitung der Interkalibrierung

- Zusammenstellung eines Datensatzes für die Interkalibrierung („Common Dataset"),
- Entwurf eines Arbeitskonzepts sowie der Auswahl der Interkalibrierungsoption,
- Evaluierung von möglichen Allgemeinen Metriks („Common Metrics").

Ankerpunkte und Klassengrenzvergleich

- Definition von Ankerpunkten für die Interkalibrierung (entweder durch naturnahe Referenzen oder alternative Ankerpunkte);
- Vergleich der nationalen Klassengrenzen untereinander.

Beide Themengebiete waren Inhalt der obligatorischen Berichterstattung („milestone-reporting") an die Europäische Kommission. Die Berichte sind als Anhang 3 und Anhang 4 diesem Dokument beigefügt.

Zusammenfassung der Milestone-Reports

16 Mitgliedstaaten nehmen aktiv an der Interkalibrierung der Bewertungsverfahren für Makrophyten in Fließgewässern teil. Allerdings verfügen derzeit nur zehn Länder über Bewertungsverfahren. Alle Verfahren sind WRRL-konform, unterscheiden sich jedoch hinsichtlich ihrer Ausrichtung (Bewertung unterschiedlicher Belastungen) und der Integration unterschiedlicher biologischer Gruppen (v.a. Phytobenthos). Gewässertypen, die interkalibriert werden, umfassen „Sandbäche des Tieflandes", „Silikatische Mittelgebirgsbäche" sowie „Flüsse des Tieflandes".

Bedingt durch die umfangreiche Datenakquise können wir auf eine relativ breite Datengrundlage zurückgreifen. Die mittlere Anzahl an Vegetationsaufnahmen pro Interkalibrierungstyp beträgt über 600; im Durchschnitt trugen 13 Länder zum gemeinsamen Datensatz bei. Die favorisierte IK-Option ist der direkte Vergleich (Option 3), aber in einem parallelen Ansatz wird ebenfalls der Common Metric mICM (siehe Birk 2009) berechnet. Das Verfahren zur Herleitung des mICM wurde im Projektjahr von einer wissenschaftlichen Zeitschrift zur Publikation angenommen. Die Veröffentlichung ist als Anhang 9 diesem Bericht beigefügt.

Intensiv wurde der in Birk (2010) skizzierte Ansatz zur Herleitung von alternativen Ankerpunkten ausgearbeitet. Hintergrund war die generelle Problematik, dass die europäische Fließgewässerlandschaft nur noch wenige echte (=unbelastete) Referenzstellen aufweist. Um dieser Situation gerecht zu werden, definiert die IK-Leitlinie „alternative Ankerpunkte“ als Gewässerzustand, dessen Grad an Degradation einheitlich stark ist. Die Intensität wird mit Blick auf die tatsächliche Situation bzw. die verfügbaren Daten bestimmt.

Um die Praxistauglichkeit dieses Konzeptes zu prüfen, unternahmen wir eine breit angelegte Studie zur Herleitung alternativer Ankerpunkte bei der Makrophyten-Interkalibrierung. Basierend auf Messstellen von möglichst hoher ökologischer Qualität und ausgewählten nicht-biologischen Daten wurden in einer multimetrischen Analyse sog. Benchmark-Stellen ermittelt, deren Biologie nur noch von (bio)geographischen Faktoren, nicht aber durch anthropogene Belastung geformt ist. Das Verfahren sowie seine Resultate sind im Anhang 5 dargestellt.

4 Interkalibrierung sehr großer Fließgewässer in Europa

Große Fließgewässer stellen komplexe Ökosysteme mit spezifischen Funktionen und einer hohen Habitatvielfalt dar. Funktionsfähigkeit und Strukturen dieser Systeme sind allerdings oftmals stark beeinträchtigt und modifiziert, da weltweit die ehemals weiträumigen Fluss- Auen-Komplexe in regulierte und eingedeichte Wasserstraßen umgewandelt wurden. Zwei Problemfelder erschweren die Umsetzung von ökologischer Zustandsbewertung an sehr großen Fließgewässern: Probenahme und Referenzbedingungen. Vor diesem Hintergrund wurde im Herbst 2009 die Arbeit zur Interkalibrierung sehr großer Fließgewässer unter der Leitung von Dr. Franz Schöll (Bundesanstalt für Gewässerkunde, Koblenz) aufgenommen mit dem Ziel, die ökologische Zustandsbewertungen der Mitgliedstaaten für diesen Gewässertyp zu harmonisieren. Im Folgenden sollen die wichtigsten Fragestellungen und Aspekte dieser Arbeit skizziert werden.

Was sind sehr große Fließgewässer?

Innerhalb der Interkalibrierungsgruppe wurde sich auf eine allgemeine Definition für sehr große Fließgewässer geeinigt. In Anlehnung an die Größenklassen des Annex II der WRRL sind sehr große Fließgewässer Ströme mit Einzugsgebieten größer 10.000 km^2.

Ziele der Interkalibrierung sehr großer Fließgewässer

Folgende Ziele wurden für die Arbeiten an der Interkalibrierung sehr großer Fließgewässer formuliert:

- Entwicklung eines Ansatzes zur Interkalibrierung sehr großer Fließgewässer, für welche generell keine Referenzstellen im naturnahen Zustand verfügbar sind;
- Definition alternativer Ankerpunkte oder Modellierung von Referenzbedingungen für die Interkalibrierung sehr großer Fließgewässer;
- Aufstellung eines einheitlichen Ansatzes, der für alle biologischen Qualitätskomponenten gangbar ist;
- Austausch mit der GIG-übergreifenden Arbeitsgruppe „Referenzen in der Interkalibrierung" zur Feinabstimmung der gewählten Ansätze.

Bewertungsverfahren für sehr große Fließgewässer in Europa

15 Länder nutzen insgesamt 47 Verfahren zur Bewertung von großen Fließgewässern anhand verschiedener biologischer Qualitätskomponenten. An dieser Stelle soll ein kurzer Überblick über die Verfahren gegeben werden. Eine ausführliche Zusammenfassung ist als Anhang 8 diesem Abschlussbericht beigelegt.

Makrozoobenthos: Die Probenahme erfolgt meistens durch Multi-Habitat-Verfahren im seichten Uferbereich. Alternativ werden Airlift-Sampling und Bodengreifer genutzt. Das taxonomische Bestimmungsniveau ist unterschiedlich: Mehr als die Hälfte der Mitgliedstaaten bestimmt die Organismen auf Artniveau, der Rest nutzt Gattungs- oder Familienniveau. Alle Bewertungsverfahren klassifizieren die Auswirkungen von allgemeiner Belastung und/oder organischer Verschmutzung. Bei den genutzten biologischen Metriks bestehen nur geringe Unterschiede zur Bewertung kleinerer Fließgewässer (z.B. Anzahl EPT, Sabrobienindices, ASPT).

Phytobenthos: Außer Deutschland, Österreich und der Slowakei werden in allen Ländern ausschließlich die Diatomeen zur Qualitätsbewertung herangezogen. Die Slowakei berücksichtigt zusätzlich das Vorkommen von filamentösen Bakterien. Allgemein werden Kratzproben noch 5 bis 10 Steinen in Ufernähe genommen. Das holländische Verfahren greift auf die Besammlung an Schilfbeständen zurück. Die Anzahl zu bestimmender Diatomeenschalen variiert zwischen 200 und 500. Alle Verfahren zeigen den Grad an Eutrophierung an, manche Methoden sind zusätzlich sensitiv gegenüber organischer Verschmutzung oder allgemeiner Belastung. Unter den verschieden biologischen Metriks wird der französische IPS (Indice de Polluosensibilité Spécifique) am häufigsten genutzt.

Fisch Fauna: Zur biologischen Datenerhebung wird meist eine Elektrobefischung entweder vom Boot oder im seichten Uferbereich durchgeführt. Schleppnetze, stationäre Netze oder hydroakustische Methoden werden nur selten verwendet. Der ökologische Zustand wird mit multimetrischen Verfahren ermittelt, die verschiedene Aspekte der Artengemeinschaften berücksichtigen (u.a. Lebensraumpräferenz, Ernährungstypen und Altersstruktur).

Makrophyten: Die Vegetation der sehr großen Fließgewässer wird in unterschiedlichem Umfang aufgenommen. Zumeist erfolgt eine Kartierung von 100-Meter-Abschnitten. In den Niederlanden werden pro Wasserkörper 6 bis 10 Abschnitte kartiert und zusammengerechnet. Das slowakische Verfahren berücksichtigt einen Abschnitt von 1000 Metern. Die aufgenommenen Pflanzengruppen sind ebenfalls verschieden: Alle Verfahren bestimmen Gefäßpflanzen, die meisten Verfahren berücksichtigen Moose und Armleuchteralgen, wenige kartieren die Großalgen. Die biologischen Metriks sind sehr verschieden und bewerten Effekte von Eutrophierung, hydromorphologischer Beeinträchtigung und allgemeiner Belastung. Eine Kombination von Phytobenthos und Makrophyten in der Bewertung erfolgt nur durch Deutschland und die Niederlande.

Phytoplankton: Die Bewertung des Potamoplanktons basiert auf Vegetationsmittelwerten von Proben, die ein- bis zweimal monatlich zwischen April und September/Oktober genommen werden. Die gesammelten Organismen werden auf Gattungs- oder Artniveau bestimmt. Alle Verfahren klassifizieren die Chlorophyll-a-Konzentration der Proben. Die taxonomische Zusammensetzung wird entweder durch Metriks berücksichtigt, welche die verschiedenen Anteile der Lebensgemeinschaften ins Verhältnis setzen (z.B. relative Abundanzen von Chlorophyta oder Cyanophyta), oder durch Indizes, die sensitiv gegenüber Belastung reagieren.

Zur Typologie sehr großer Fließgewässer in Europa

Auf Grundlage einer umfangreichen, aber späten Datenlieferung der teilnehmenden Länder konnten im Projektjahr nur Vorstudien zur Typologie sehr großer Fließgewässer in Europa durchgeführt werden. Hintergrund dieser Aktivitäten war die Fragestellung, inwiefern unterschiedliche Gewässertypen für den Vergleich der nationalen Bewertungsverfahren eine Rolle spielen. Wenn sich die sehr großen Fließgewässer in Europa typologisch stark voneinander abgrenzen, ist dies bei Interkalibrierung zu berücksichtigen.

Auf der Basis von Makrozoobenthos-Daten auf Familienniveau führten wir Clusteranalysen mit Datensätzen durch, welche aggregierte Taxalisten pro Gewässer und Land beinhalteten. Dabei machten wir keine Unterscheidungen zwischen gering und stark belasteten Gewässern. Die Ergebnisse zeigen vier klar abgrenzbare Gruppen von Gewässern (Abbildung 4.1), welche sich, basierend auf ihrer taxonomischen Zusammensetzung, den Typen „rhithral/epipotamal“, „potamal“, „mediterran“ und „degradiert“ zuordnen lassen. Eine Diskriminanzanalyse ergibt, dass die mittlere jährliche Wassertemperatur signifikante Unterschiede zwischen Rhithral- und Potamal-Typus aufweist.

Diese Ergebnisse vermitteln einen ersten, statistisch abgesicherten Hinweis auf die Existenz von zwei Haupttypen von sehr großen Fließgewässern in Europa: Auf der einen Seite stehen die rhithralen/epipotamalen Ströme mit geringer Wassertemperatur, höheren Fließgeschwindigkeiten und dem Vorkommen von Insektenfamilien wie Ephemerellidae, Perlidae und Philopotamidae. Typische Gewässer sind die Flüsse des mitteleuropäischen Berglandes bzw. Nordeuropas (z.B. Ober- und Mittelrhein, Inn, Drau, Obere Donau, Teno oder Torni). Potamale Gewässer weisen höhere Wassertemperaturen, geringere Fließgeschwindigkeiten und Invertebratenfamilien wie Valvatidae, Hydrospsychidae, Calopterygidae und Aeshnidae auf. Mittlere und Untere Donau, Garonne oder Ebro zählen zu diesem Typus. Wir planen, die Arbeiten zur Typisierung der sehr großen Fließgewässer in Europa fortzusetzen. Hierzu werden wir einen erweiterten Datensatz nutzen und verschiedene Optionen der Datenaggregation testen (u.a. mit Rückgriff auf frei-fließende bzw. gering belastete Probestellen).

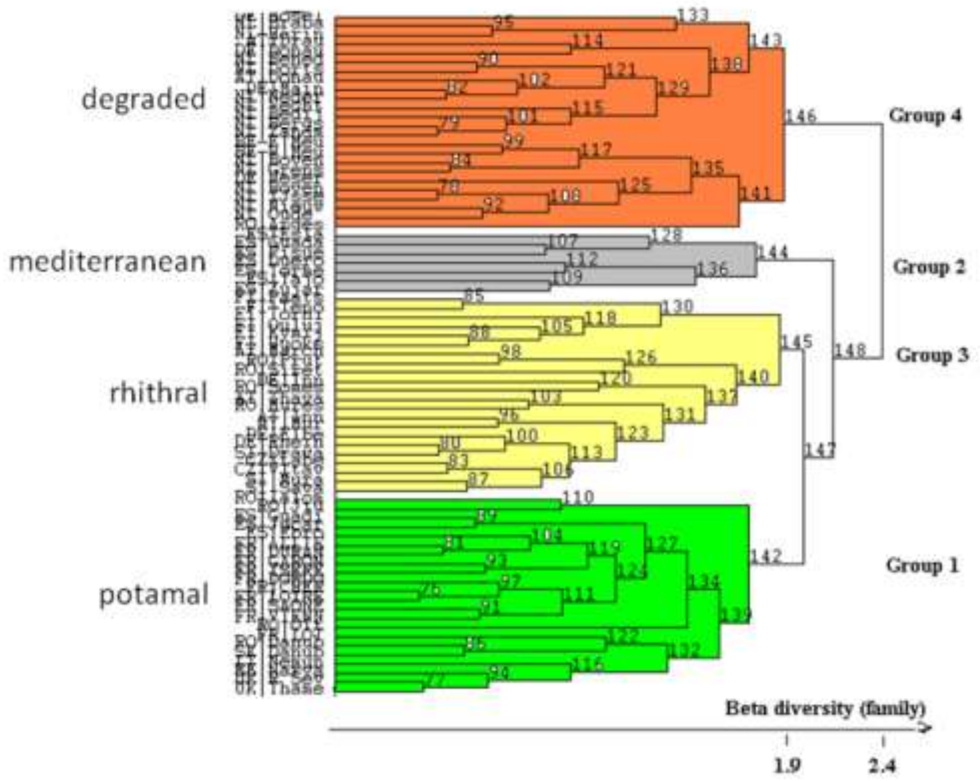


Abbildung 4.1: Ergebnisse der Clusteranalyse von Makrozoobenthos-Daten an sehr großen Fließgewässern in Europa

Alternative Ankerpunkte für die Interkalibrierung

Eine weitere Vorstudie wurde für die Herleitung alternativer Ankerpunkte durchgeführt. Basierend auf einem Datensatz von 80 Wasserkörpern (Diatomeen-Aufnahmen, Parameter zu chemischer und hydromorphologischer Belastung) wurde eine Faktoranalyse durchgeführt, um die Hauptkomponenten anthropogener Belastung im Datensatz zu bestimmen. Die ersten zwei Komponenten sind korreliert zu Parametern der chemischen und hydromorphologischen Qualität (Abbildung 4.2). Die Wasserkörper eines Mitgliedstaats zeigen jeweils eine klare Gruppierung, d.h. die Wasserkörper eines Landes/einer Region weisen oft einen ähnlichen Grad an anthropogener Belastung auf. Die erste Hauptkomponente (Wasserqualität) ist gut mit den möglichen Common Metrics für die Diatomeen-Interkalibrierung korreliert (IPS: r=-0.45, Trophieindex: r=0.37).

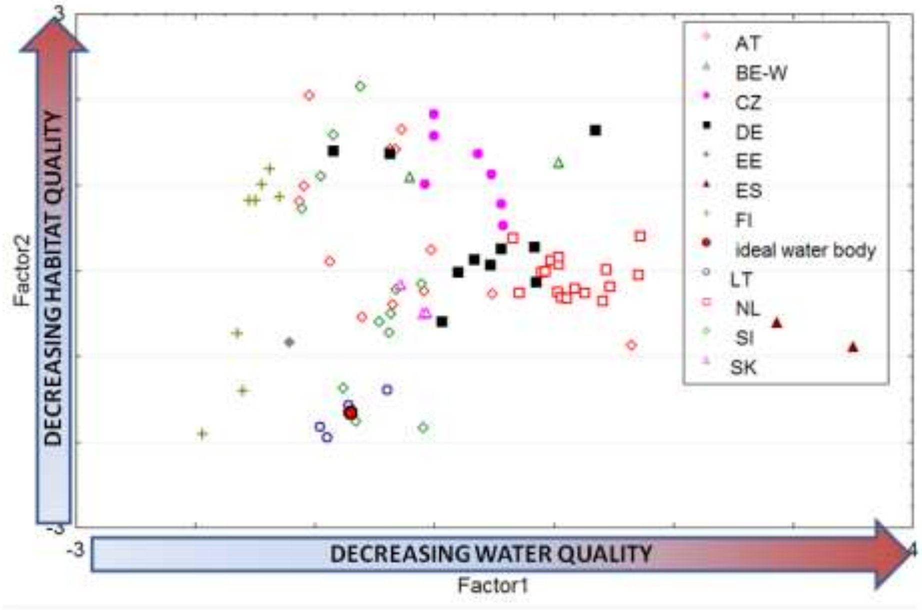


Abbildung 4.2: Position der nationalen Wasserkörper im „Belastungsraum“, definiert durch eine Faktorenanalyse der nicht-biologischen Parameter

Unter Nutzung der deutschen Richtwerte für den chemischen Zustand großer Flüsse verwendeten wir Grenzwerte zur Ausweisung von Gewässerstellen im guten physiko-chemischen Zustand. Das 75. Perzentil der numerischen Werte der ersten Hauptkomponente wurde genutzt, um sog. Benchmark-Stellen auszuweisen. Die meisten nicht-biologischen Parameter sowie die Diatomeen-Metriks zeigen signifikante Unterschiede zwischen Benchmark-Stellen und den anderen, stärker belasteten Wasserkörpern. Die Benchmark-Stellen sind vorwiegend an Wasserkörpern in Nord- und Osteuropa lokalisiert (Abbildung 4.3). Des weiteren besteht zwischen den ausgewiesenen Stellen ein gewisser Belastungsgradient. Vor diesem Hintergrund gilt es, in zukünftigen Arbeitsschritten ein erweitertes Konzept zur Findung von alternativen Ankerpunkten anzuwenden.

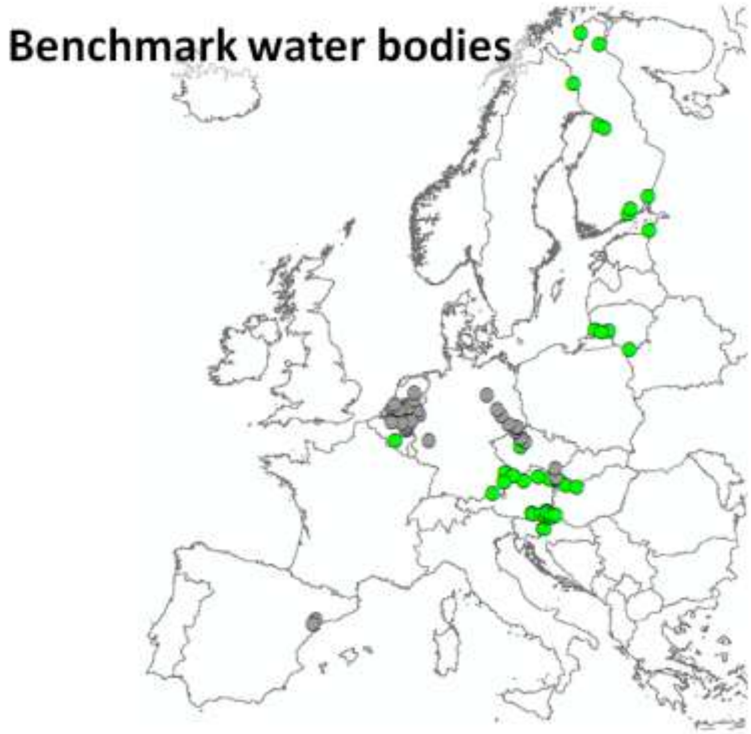


Abbildung 4.3: Lage der Wasserkörper (grün=Benchmark-Stellen)

5 Weitere Aktivitäten zur Umsetzung der Interkalibrierung

5.1 Wissenschaftliche Publikation zur Entwicklung eines Interkalibrierungsmetriks für Makrophyten

Unterschiedlichen Bewertungskonzepte erschweren die Harmonisierung der Qualitätsbewertung von Fließgewässern anhand von Makrophyten. Die im Projektjahr veröffentlichte Studie (Birk & Willby 2010; Anhang 9) beschreibt ein Verfahren, um Ähnlichkeiten zwischen den Bewertungsmethoden von Österreich, Belgien, Frankreich, Deutschland, Großbritannien und Polen zu erkennen. Auf der Basis eines internationalen Datensatzes von drei Europäischen Gewässertypen wiesen wir Vegetationsaufnahmen aus, die von der Mehrzahl der Methoden als „sehr gut" bewertet wurden. Der mittlere Bewertungsindex aller Methoden wurde anschließend mit den Abundanzen der einzelnen Pflanzenarten korreliert. Je nach Art und Stärke der Korrelationen definierten wir positive, negative oder neutrale Indikatoren. Dies ermöglichte die Beschreibung von Gewässertyp-spezifischen Artengemeinschaften im naturnahen und -fernen Zustand. Außerdem entwickelten wir auf dieser Grundlagen den Interkalibrierungsmetrik mICM, der mit allen Bewertungsindizes korreliert. Der geringere Zusammenhang mit der belgischen und deutschen Bewertung wurde durch geringfügige Anpassungen der Indikatorwerte ausgewählter Arten dieser Methoden verbessert. Die Analyse der Messstellen im allgemein sehr guten Zustand erlaubte die Definition von mICM-Referenzwerten. Die veröffentlichte Studie beschreibt einen allgemeinen Ansatz zur Harmonisierung von biologischen Bewertungsverfahren für unterschiedliche Gewässerkategorien.

5.2 Aufstellung aller im Projekt geleisteten Teilnahmen an nationalen und internationalen Aktivitäten zur Interkalibrierung

Zur Durchführung der Arbeitsaufgaben des Vorhabens bedurfte es der Teilnahme an verschiedenen Aktivitäten (Tabelle 5.1). Schwerpunkt bildete in diesem Projektjahr die Fortführung der Arbeiten zur Interkalibrierung Makrophyten sowie die Planung der Interkalibrierung sehr großer Fließgewässer. Innerhalb der Aktivität des CB_{riv}GIG Makrophyten wurden der Ansatz zur Herleitung alternativer Ankerpunkte im Rahmen eines mehrtägigen Arbeitstreffens in Stirling ausgearbeitet. Die Ergebnisse sind im Anhang 5 wiedergegeben. Die in Kapitel 2 erwähnte Vorstellung der Inhalte zur neuen IK-Leitlinie erfolgte auf Reisen nach Würzburg und Bordeaux. Das Protokoll des Treffens in Bordeaux ist als Anhang 6 beigefügt. Mehrere Treffen mit dem Koordinator der Interkalibrierung sehr großer Fließgewässer (Dr. Franz Schöll, Bundesanstalt für Gewässerkunde, Koblenz) bereiteten die Aktivitäten dieser Interkalibrierungsgruppe vor.

Tabelle 5.1: Aufstellung aller im Projekt geleisteten Teilnahmen an nationalen und internationalen Aktivitäten zur Interkalibrierung (I – international, N – national)

Nr.	Veranstaltung	Ausrichtung	Datum	Ort
1	LAWA Expertentreffen Fließgewässer	N	24. März 2010	Würzburg
2	Treffen der AG ECOSTAT	I	6.-7. April 2010	Brüssel (BE)
3	Expertenworkshop "IK-Große Fließgewässer"	I	19.-20. April 2010	Koblenz
4	CB_{riv}GIG Makrophyten Experten-Workshop	I	27.-28. Mai 2010	Bordeaux (FR)
5	Koordinationstreffen "IK-Makrophyten"	I	4.-7. August 2010	Stirling (UK)
6	Koordinationstreffen "IK-Große Fließgewässer"	I	3. September 2010	Koblenz
7	Expertenworkshop "IK-Große Fließgewässer"	I	22.-23. September 2010	Koblenz

6 Literatur

Birk S (2010) Endbericht im Vorhaben "Interkalibrierung – Teilprojekt: Vertretung in GIG und Makrophyten-Berechnungen" im Auftrag der Länderarbeitsgemeinschaft Wasser (LAWA), Projekt-Nr. O 7.09 im Länderfinanzierungsprogramm "Wasser, Boden und Abfall". Projektjahr 2009. Essen, Universität Duisburg-Essen.

Birk S (2009) Endbericht im Vorhaben "Interkalibrierung - Makrophyten in Fließgewässern im GIG Mitteleuropa/Baltikum im Jahr 2008" im Auftrag der Länderarbeitsgemeinschaft Wasser (LAWA), Projekt-Nr. O 6.08 im Länderfinanzierungsprogramm "Wasser, Boden und Abfall". Projektjahr 2008. Hamm (Sieg).

Birk S (2008) Endbericht im Teilvorhaben "Interkalibrierung Makrophyten in Fließgewässern im GIG Mitteleuropa/Baltikum" im Auftrag der Länderarbeitsgemeinschaft Wasser (LAWA), Projekt-Nr. O 1.07 im Länderfinanzierungsprogramm "Wasser, Boden und Abfall". Projektjahr 2007. Universität Duisburg-Essen, Essen.

Birk S, Willby NJ (2010) Towards harmonization of ecological quality classification: establishing common grounds in European macrophyte assessment for rivers. Hydrobiologia 652:149-163.

Birk S, Böhmer J, Meier C, Rolauffs P, Schaumburg J & Hering D (2007) EG-Wasserrahmenrichtlinie - Harmonisierung der Berichterstattung zur ökologischen Einstufung nach EG-Wasserrahmenrichtlinie (Interkalibrierung biologischer Untersuchungsverfahren in Deutschland). Universität Duisburg-Essen, Essen.

European Communities (2010) Guidance document on the intercalibration process 2008-2011. Guidance Document No. 14. Implementation Strategy for the Water

Framework Directive (2000/60/EC). Office for Official Publications of the European Communities, Luxembourg.

Anhang 1: Präsentation „Sachstand Interkalibrierung“ – 24. März 2010, Würzburg

Sachstand Interkalibrierung

22. Sitzung der LAWA-FG-Experten am 24.+25. März 2010 in Würzburg

Sebastian Birk

IK-Leitlinie – Anhang III
Leitlinie für die Herleitung von Referenzbedingungen und alternativen Ankerpunkten

Ziel der Leitlinie:

(A) Schaffung eines allgemeinen Rahmens zur Herleitung von Referenzbedingungen für die nationalen Bewertungsverfahren

(B) Definition verbindlicher Methoden zur Herleitung von Referenzbedingungen/alternativen Ankerpunkten für die Interkalibrierung

Netzwerk von Referenzstellen wird ausdrücklich empfohlen. Dabei können diese Referenzstellen BQE-spezifisch ausgewählt werden, wenn keine signifikante anthropogene Veränderung der Lebensgemeinschaft angezeigt ist.

Der Grad an anthropogener Belastung ohne Auswirkung auf die Lebensgemeinschaft (=Referenzkriterien) sollte durch statistische Verfahren quantifiziert werden.

2

IK-Leitlinie – Anhang III
Leitlinie für die Herleitung von Referenzbedingungen und alternativen Ankerpunkten

Die Möglichkeiten von
- (Dosis-Wirkungs-) Modellierungen von Referenzbedingungen sowie
- Expertenwissen

sind explizit erwähnt.

Für die **Interkalibrierung** sind Referenzbedingungen/alternative Ankerpunkte einheitlich herzuleiten.

Die einheitlich Herleitung von Referenzbedingungen orientiert sich an den Empfehlungen für die nationalen Bewertungsverfahren (erster Teil der Leitlinie).

Alternative Ankerpunkte sollen definiert werden, wenn (genügend) „echte" Referenzstellen nicht vorhanden sind.
Der Grad der Abweichung dieser alternativen Ankerpunkte von der „echten" Referenz ist zu quantifizieren.

3

IK-Leitlinie – Anhang V
Definition von Kriterien für die Vergleichbarkeit der Klassengrenzen

Vorgestelltes Verfahren schafft einheitliche Kriterien für alle IK-Optionen, Gewässerkategorien, BQEs
→ Möglichkeit einer übergreifenden Evaluierung der IK am Ende der zweiten Phase

Kernaspekte der Interkalibrierung:
(A) Wie ähnlich sind die Bewertungsverfahren (Variabilität der EQRs)?
(B) Wie ähnlich sind die nationalen Definitionen des GÖZ?

→ Zusammenführung der verschiedenen Ansätze von IK-Phase 1

4

IK-Leitlinie – Anhang V
Definition von Kriterien für die Vergleichbarkeit der Klassengrenzen

Implementierung des Ansatzes „einheitlicher Ankerpunkte" (Referenzbedingungen) bei allen IK-Optionen

Regression als analytische Grundlage für alle IK-Optionen

Einheitliches Maß für die Interkalibrierung:
Kappa-Koeffizient (Cohen 1960): Grad der Übereinstimmung des GÖZ

Das Verfahren vermeidet die Definition von konkreten Kappa-Werten, um der Vielfalt der möglichen IK-Situationen gerecht zu werden. Mindeststandards sind jedoch formuliert.

5

IK-Leitlinie – Anhang VI
Inhalte der Berichterstattung zur Interkalibrierung

- 1. Arbeitsschritt: Überprüfung der Vorbedingungen der Interkalibrierung[1].
  - „Sind die Bewertungsverfahren WRRL-konform?", z.B.: Werden die numerischen Ergebnisse in EQR ausgedrückt und in fünf Qualitätsklassen dargestellt?; Sind alle indikativen Parameter der normativen Begriffsbestimmungen des WRRL Annex V für die Qualitätskomponente durch das Bewertungsverfahren abgedeckt?
  - „Ist die Interkalibrierung durchführbar?", z.B.: Werden nur solche Bewertungsverfahren interkalibriert, die ähnliche Belastungen für ähnliche Gewässertypen durch ein ähnliches Bewertungskonzept einstufen?
  - Die erste Berichterstattung („Milestone 1") über einen Zwischenstand dieser Überprüfungen erfolgte im Oktober 2009 (siehe Anhang 5 für die Makrophyten in Fließgewässern im GIG Mitteleuropa/Baltikum).
- 2. Arbeitsschritt: Datengrundlage und Vorbereitung der Interkalibrierung.
  - Zusammenstellung eines Datensatzes für die Interkalibrierung („Common Dataset");
  - Entwurf eines Arbeitskonzepts;
  - Auswahl der Interkalibrierungsoption und Evaluierung von möglichen Allgemeinen Metriks („Common Metrics").
  - Die zweite Berichterstattung („Milestone 2") ist für März 2010 geplant.

6

IK-Leitlinie – Anhang VI

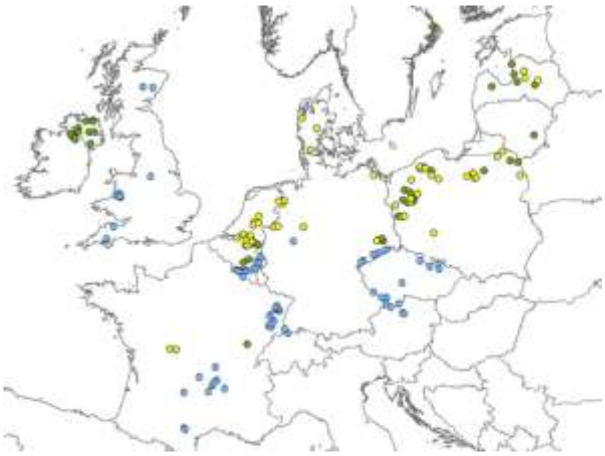
Inhalte der Berichterstattung zur Interkalibrierung

- 3. Arbeitsschritt: Ankerpunkte und Klassengrenzvergleich.
  - Definition von Ankerpunkten für die Interkalibrierung (entweder durch naturnahe Referenzen oder alternative Ankerpunkte);
  - Vergleich der nationalen Klassengrenzen untereinander.
  - *Dieser Arbeitsschritt wird durch den Inhalt des fehlenden Annex III (Leitlinien zur Bestimmung von Ankerpunkten für die Interkalibrierung) und Annex V (Definition von Vergleichbarkeitskriterien für die Interkalibrierung) konkretisiert werden.*
  - Die dritte Berichterstattung („Milestone 3") ist für September 2010 geplant.
- 4. Arbeitsschritt: Angleichung der nationalen Klassengrenzen.
  - Durchführung des Verfahrens zur Vereinheitlichung der nationalen Klassengrenzen.
  - *Dieser Arbeitsschritt wird durch den Inhalt des fehlenden Annex V (Definition von Vergleichbarkeitskriterien für die Interkalibrierung) konkretisiert werden.*
  - Die vierte Berichterstattung („Milestone 4") ist für März 2011 geplant.
- 5. Arbeitsschritt: Zusammenfassung.
  - Überprüfung und Diskussion der geleisteten Arbeit;
  - Zusammenstellung des Endberichts.
  - Die fünfte Berichterstattung („Milestone 5") ist für Juni 2011 geplant.

7

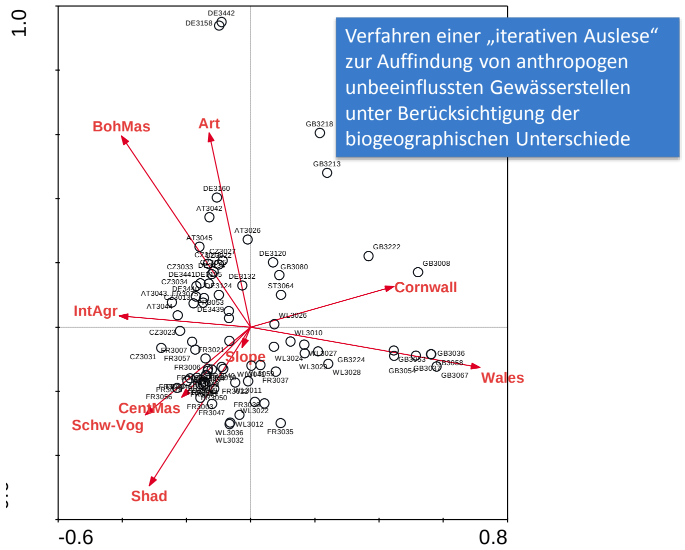
Sachstand Interkalibrierung Makrophyten

Definition von Ankerpunkten für die Interkalibrierung



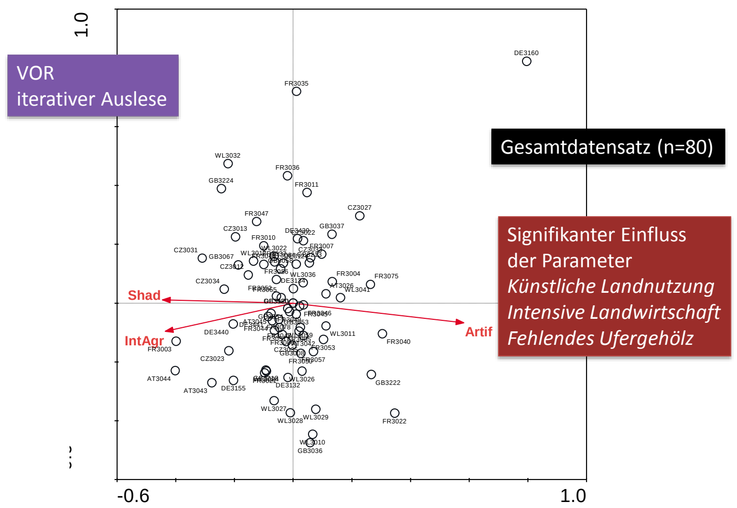
Grundlage: Referenzdatensatz (n=185) aus den Gewässertypen „Silikatische Mittelgebirgsbäche", „Sandbäche des Tieflandes", „Flüsse des Tieflandes"

8



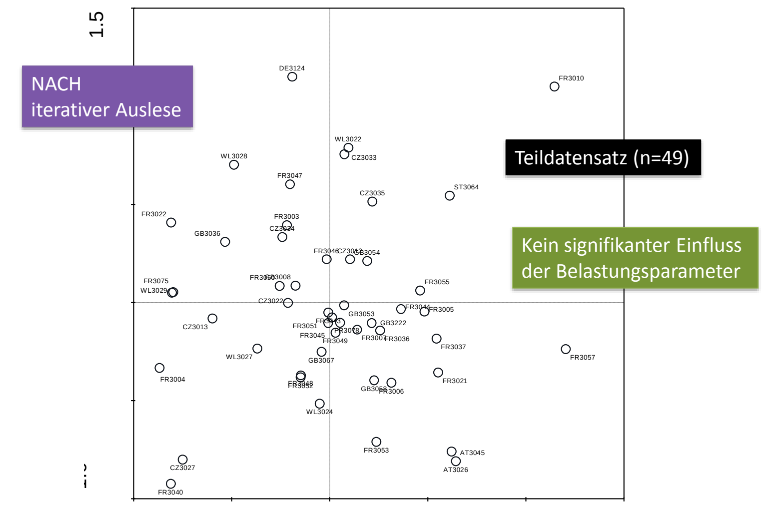
Kanonische Korrespondenzanalyse der Referenzstellen des IK-Typs „Silikatische Mittelgebirgsbäche“

9



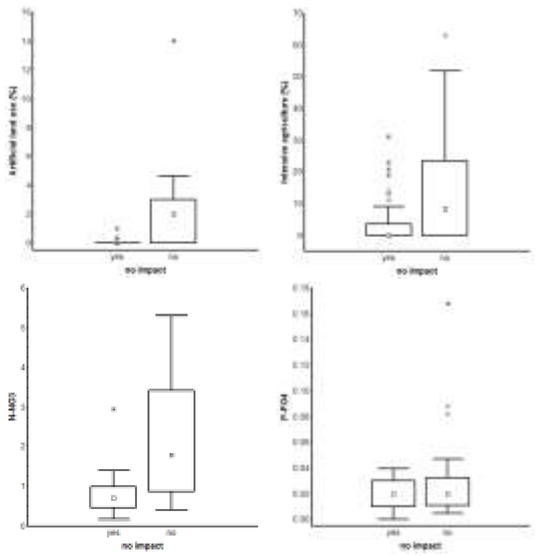
Partielle Kanonische Korrespondenzanalyse der Referenzstellen des IK-Typs „Silikatische Mittelgebirgsbäche“

10



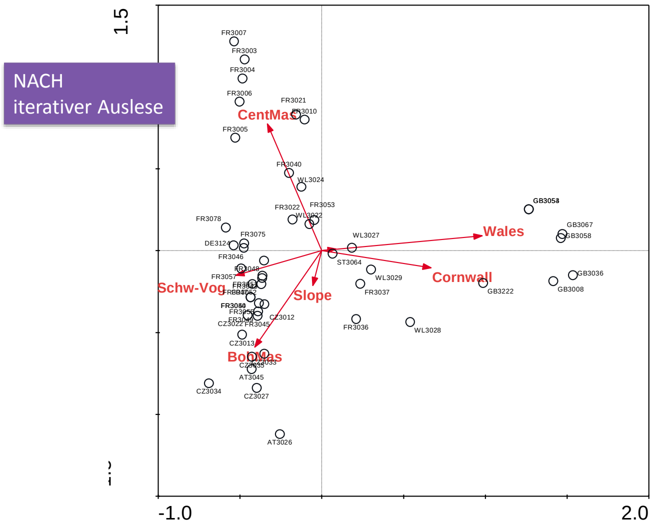
Partielle Kanonische Korrespondenzanalyse der Referenzstellen des IK-Typs „Silikatische Mittelgebirgsbäche“

11



Verteilung ausgewählter Belastungsparameter vor und nach iterativer Auslese

12



Kanonische Korrespondenzanalyse der Referenzstellen des IK-Typs „Silikatische Mittelgebirgsbäche“

13

Sachstand Interkalibrierung Makrophyten
Definition von Ankerpunkten für die Interkalibrierung

Arbeiten in 2010

Fortführung der Analysen zu Referenzen/alternativen Ankerpunkten für die drei Interkalibrierungstypen

Vorstellung der Ergebnisse auf Arbeitsgruppentreffen im April

Sammlung der Aktualisierungen der nationalen Bewertungsverfahren

Durchführung der Vergleichbarkeitsanalysen nach IK-Leitlinien-Anhang V

Vorstellung dieser Ergebnisse auf einem Arbeitsgruppentreffen im Herbst

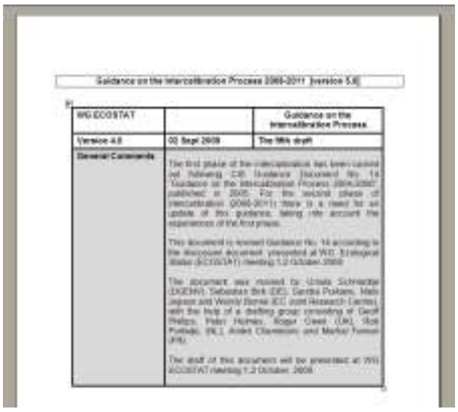
Allgemein wird dem Zeitplan der neuen IK-Leitlinie gefolgt.

14

Anhang 2: Präsentation „The new guidance on the intercalibration process“ – 27.-28. Mai 2010, Bordeaux (FR)

The New Guidance on the Intercalibration Process

Sebastian Birk



WG ECOSTAT		Guidance on the Intercalibration Process
Version 4.0	02 Sept 2009	The fifth draft
General Comments	[illegible]	

Content

- Key principles (Chapter 1)
- Steps (Chapter 2)
- Time-frame (Chapter 3)

Key principles
– Aims of intercalibration

1. Ensure the consistency of Member States classification results with the normative definitions of Annex V and to ensure the comparability of the classification results among Member States
2. Intercalibrate high-good and good-moderate boundary so that they correspond to comparable levels of ecosystem alteration
3. There is no specific role for the IC Register in Phase 2.
4. Close gaps of 1st phase of IC → intercalibrate full BQEs in all water categories
5. Review IC results for BQEs fully intercalibrated in Phase 1

Key principles
– WFD compliant assessment methods

6. Partial intercalibration is possible where full BQEs cannot be intercalibrated.
7. Where it is not possible to develop WFD compliant assessment methods based on scientific results address IC Steering Committee.
8. 'Macrophytes and phytobenthos' and 'macroalgae and angiosperms' – justification is needed if only one component is used.
9. Intercalibratability of assessment methods: If method is not intercalibratable, then MS needs to find an alternative intercalibration method → this needs to be approved by the GIG

Key principles
– Practical implementation

10. Intercalibration is carried out in GIGs for 'common intercalibration types'
11. Common intercalibration types should cover the main surface water types in the GIG, national types need to be linked with IC types → the existing common IC types need to be reviewed
12. All major combinations of IC types, BQEs and pressure(s) or combinations of pressures need to be covered
13. Comparability of reference conditions must be assured. If natural or near-natural reference conditions are not available or cannot be reliably derived for a certain type (for example for large rivers) intercalibration needs to be carried out against an alternative benchmark (e.g. good ecological status for that surface water type).
14. No intercalibration of ecological potential in Phase 2

Key principles
– Practical implementation (cont.)

15. Nonetheless, data from HMWBs and AWBs can be used
    - data of unimpacted BQEs (e.g. phytoplankton in reservoirs)
    - data of impacted BQEs to cover the whole gradient of the pressure (e.g. macroinvertebrates in diverted streams)
16. 3 different options can be used for intercalibration. Boundaries can be defined during the IC process.
17. Explanation of the 3 options
18. Results are expressed as ecological quality ratios (EQRs), but EQRs are not always directly comparable. Therefore, intercalibration is about demonstrating that the high-good and good-moderate boundaries represent a comparable level of anthropogenic alteration to the biological quality element.

Key principles
– Organisation and time-table

19. Time-table needs to followed closely.
20. Different roles for GIG leads, BQE leads, water category leads → together they form the IC Steering Group chaired by JRC and review issues that cannot be resolved. Ultimate decision lies with ECOSTAT.
21. The results of the intercalibration Phase 2 and the results of the review of intercalibration Phase 1 needs to be reported to ECOSTAT and Water Directors → COM Decision + Technical Report
22. Member States must translate the results of the intercalibration into their national assessment systems for all national types.

Steps of the intercalibration process

Preconditions

Q1. Do all national assessment methods meet the requirements of the Water Framework Directive?

WISER

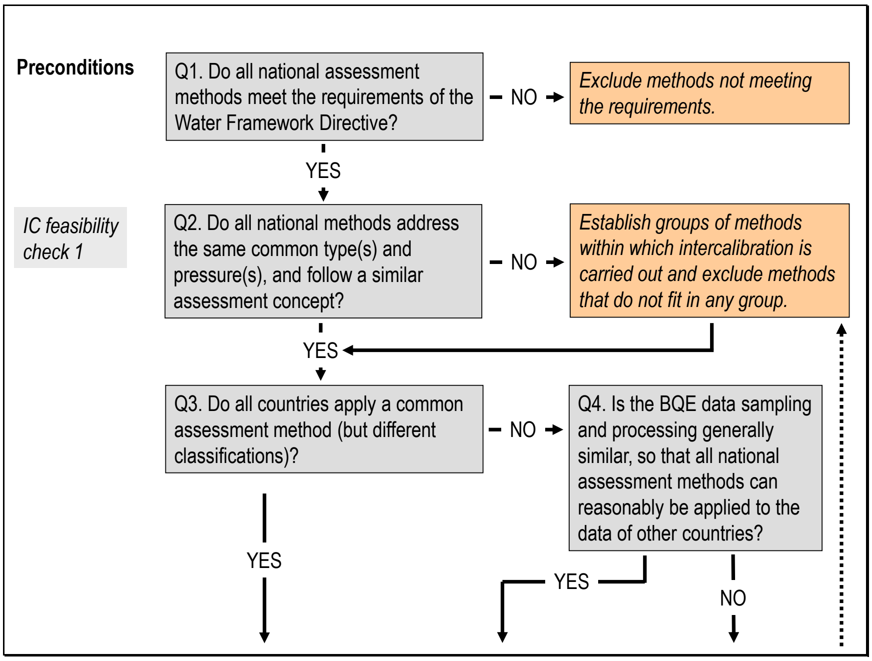
Questionnaire on biological assessment methods used in national WFD monitoring programmes

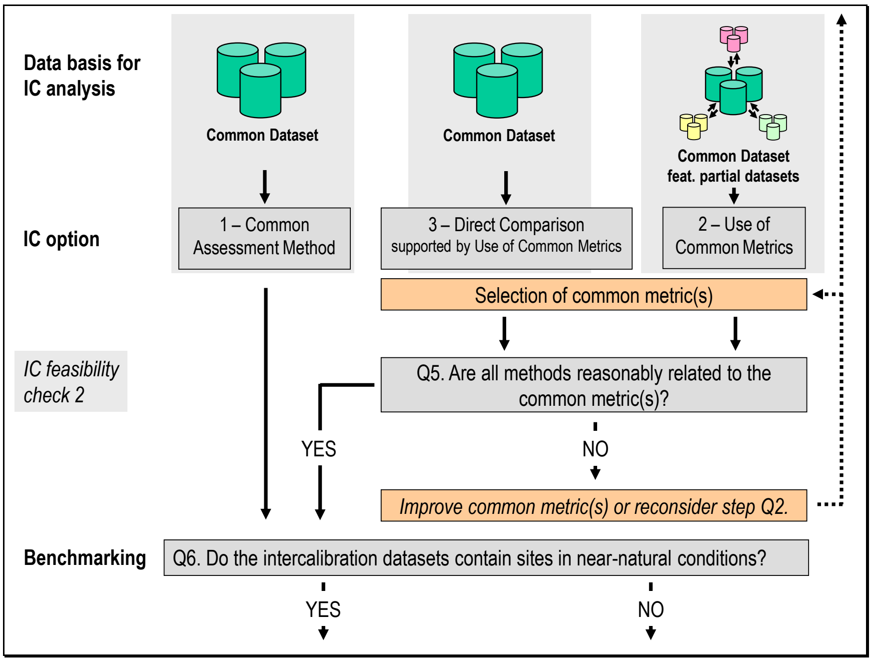
Introduction

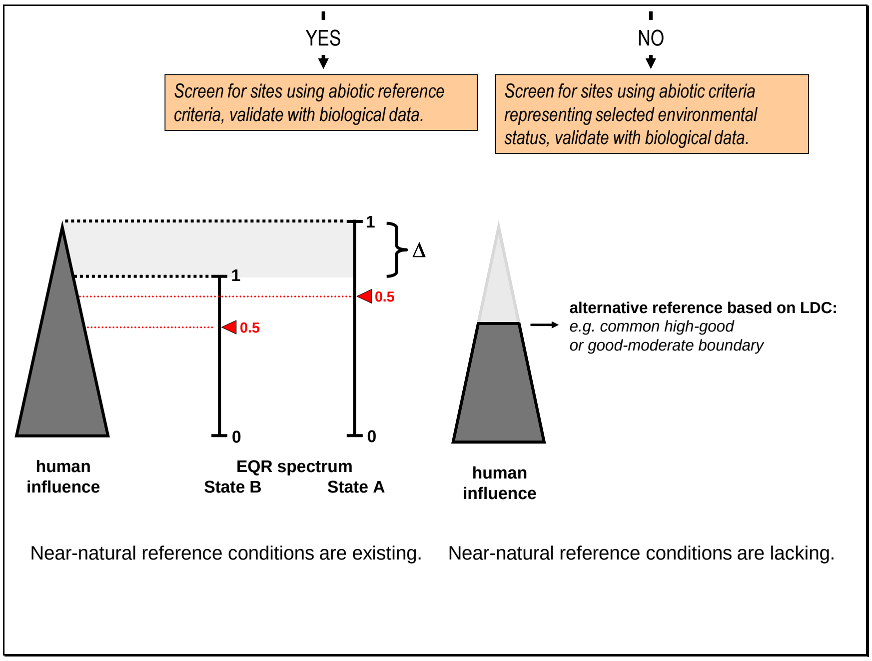
Purpose

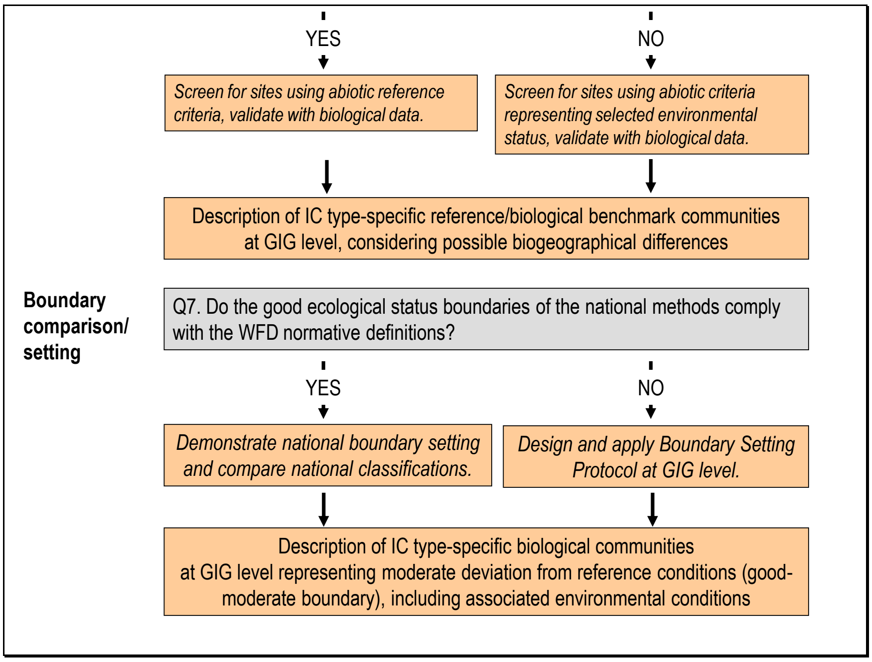
Questionnaire on biological assessment methods

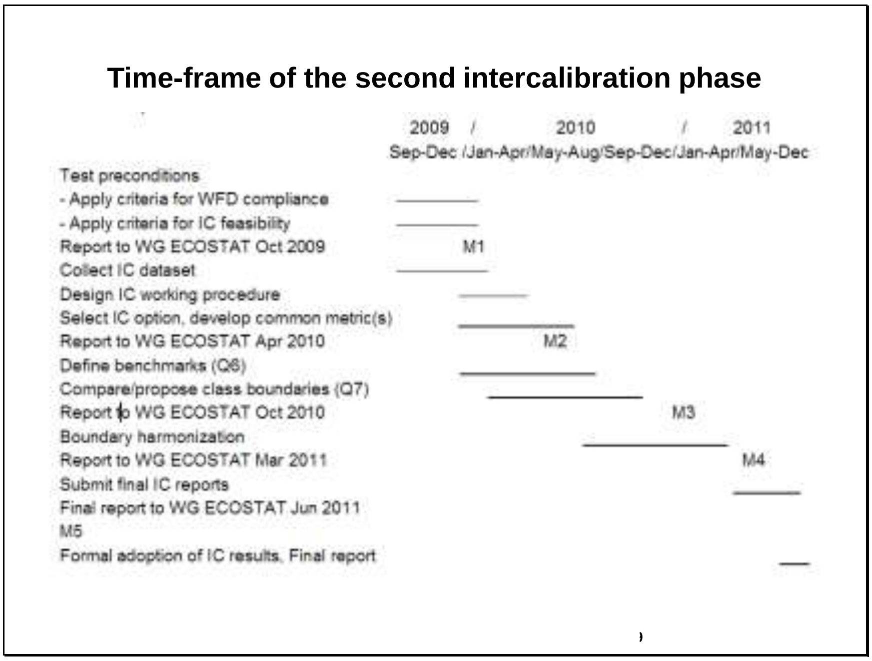
Joint activity of IC Steering Group and EU-Project WISER




Data basis for IC analysis
Common Dataset
Common Dataset
Common Dataset feat. partial datasets
IC option
1 – Common Assessment Method
3 – Direct Comparison supported by Use of Common Metrics
2 – Use of Common Metrics
Selection of common metric(s)
IC feasibility check 2
Q5. Are all methods reasonably related to the common metric(s)?
YES
NO
Improve common metric(s) or reconsider step Q2.
Benchmarking
Q6. Do the intercalibration datasets contain sites in near-natural conditions?
YES
NO


YES
NO
Screen for sites using abiotic reference criteria, validate with biological data.
Screen for sites using abiotic criteria representing selected environmental status, validate with biological data.
1
Δ
1
0.5
0.5
0
0
alternative reference based on LDC:
e.g. common high-good
or good-moderate boundary
human influence
EQR spectrum
State B
State A
human influence
Near-natural reference conditions are existing.
Near-natural reference conditions are lacking.


YES
NO
Screen for sites using abiotic reference criteria, validate with biological data.
Screen for sites using abiotic criteria representing selected environmental status, validate with biological data.
Description of IC type-specific reference/biological benchmark communities at GIG level, considering possible biogeographical differences
Boundary comparison/ setting
Q7. Do the good ecological status boundaries of the national methods comply with the WFD normative definitions?
YES
NO
Demonstrate national boundary setting and compare national classifications.
Design and apply Boundary Setting Protocol at GIG level.
Description of IC type-specific biological communities at GIG level representing moderate deviation from reference conditions (good-moderate boundary), including associated environmental conditions


Time-frame of the second intercalibration phase
2009 / 2010 / 2011
Sep-Dec /Jan-Apr/May-Aug/Sep-Dec/Jan-Apr/May-Dec
Test preconditions
- Apply criteria for WFD compliance
- Apply criteria for IC feasibility
Report to WG ECOSTAT Oct 2009
M1
Collect IC dataset
Design IC working procedure
Select IC option, develop common metric(s)
Report to WG ECOSTAT Apr 2010
M2
Define benchmarks (Q6)
Compare/propose class boundaries (Q7)
Report to WG ECOSTAT Oct 2010
M3
Boundary harmonization
Report to WG ECOSTAT Mar 2011
M4
Submit final IC reports
Final report to WG ECOSTAT Jun 2011
M5
Formal adoption of IC results, Final report

Anhang 3: Kopie der Milestone 2 Berichterstattung des CB$_{riv}$GIG Makrophyten

EUROPEAN COMMISSION
DIRECTORATE GENERAL JRC
JOINT RESEARCH CENTRE
Institute of Environment and Sustainability

WFD Intercalibration Phase 2 : Milestone 2 report (for ECOSTAT meeting 8-9 April 2009)

The new IC guidance foresees a 'Milestone 2' report for the spring of 2010, with the following key elements:

- Overview of the national assessment methods that will be intercalibrated;
- Results of WFD compliance and feasibility check;
- Data set collection;
- Design of IC procedure and development of IC common metric;
- Progress on Benchmarking and Boundary comparison/setting.

All GIGs are kindly requested to submit their progress reports for the relevant quality elements following this template as much as possible.
Also, you are requested to update the relevant sections of the intercalibration work plan (distributed as a separate document).

Depending on how the work has been organized, we expect one response for each quality element for each of the GIGs. In case of horizontal activities (e.g. large rivers) or where the work is carried out cross-GIG (e.g. fish in rivers), one coordinated response is expected.

The questions in the reporting template follow the new intercalibration guidance as closely as possible. Please contact the IC steering group if you need any further clarifications:

- Wouter van de Bund wouter.van-de-bund@jrc.ec.europa.eu - Rivers;
- Sandra Poikane sandra.poikane@jrc.eu.europa.eu - Lakes;
- Wendy Bonne wendy.bonne@jrc.ec.europa.eu - Coastal waters;
- Nikolaos Zampoukas nikolaos.zampoukas@jrc.ec.europa.eu - Transitional waters.

Please send your responses before 22nd March 2010 to eewai@jrc.ec.europa.eu

Water category/GIG/BQE/ horizontal activity:	**CBrivGIG Macrophytes**
Information provided by:	Roelf Pot, Sebastian Birk

1: Organisation

1.1. Responsibilities

Please indicate how the work is organised, indicating the lead country/person and **the list of involved experts of every country**

Lead, DE – Sebastan Birk
UK – Nigel Willby, Geoff Phillips
FR – Christian Chauvin
DE – Jochen Schaumburg, Christine Schranz
BE (WL) – Daniel Galoux
BE (FL) – An Leyssen, Kris van Looy
PL – Krzysztof Szoszkiewicz
NL – Roelf Pot
EE –Aive Kors
LT – Mindaugas Gudas
LV – Solvita Strazdina
AT – Karin Pall
IT – Maria Rita Minciardi
DK – Annette Baattrup-Pedersen, Lars K. Larsen
ES – Isabel Pardo
CZ – Libuse Opatrilova
NO – Marit Mjelde
SK – Lívia Tóthova
FI – Juha Riihimaki
PT – Francisca Aguiar

1.2. Participation

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

Participants in CB-GIG: AT, BE (WL, FL), CZ, DE, DK, EE, ES, FR, IT, LT, LV, NL, PL, UK

Participants bordering CB-GIG, exchanging information: FI, HR, NO, PT, SK

1.3. Meetings

Please list the meetings of the group (including planned future meetings) in 2010

Edinburgh (UK), March 13-14, 2008
Copenhagen (DK), April 29-30, 2009
Bordeaux (FR), April 22-23, 2010
Second meeting in 2010 t.b.c.

2. Overview of Methods to be intercalibrated

Identify for each MS:

- the national classification method that will be intercalibrated
- the status of the method (e.g., finalized agreed national method, method under development,)
- whether the WISER questionnaire has been completed for the method

Country	Method	Status	WISER questionnaire
		1 – finalized agreed national method 2 – method under evaluation for 2015 3 – method proposed 4 – method under development	1- Completed 2- Completed partially 3- Not submitted
Austria	Austrian Index Macrophytes for Rivers	1	1
Belgium (Flanders)	Flemish macrophyte assessment system	1	1
Belgium (Wallonia)	Macrophyte Biological Index for Rivers	2	1
France	Biological Macrophytes Index for Rivers	1	1
Germany	German Assessment system for Macrophytes & Phytobenthos according to the EU WFD	1	1
Netherlands	WFD-metrics for natural watertypes	2	1
Poland	Macrophyte Index for Rivers	3	1
United Kingdom	Ecological Classification of Rivers using Macrophytes	1	1
Italy	Macrophyte Biological Index for Rivers	2	1
Denmark		4	3

Do you have sufficient information of the national classification methods that will be intercalibrated to carry out the compliance and feasibility checks? Please, specify information gaps

Country	Information gaps
some	There is no English translation of the method yet.

3. Checking of compliance of national assessment methods with the WFD requirements

What is the progress with the compliance checking? (1 – completed; 2 – in progress; 3 – not started); in case of 'completed' or 'in progress, please fill in the detailed information below; if 'not started', indicate the reasons

List the WFD compliance criteria and describe the WFD compliance checking process and results (the table below lists the criteria from the IC guidance, please add more criteria if needed)

Compliance criteria	Compliance checking
Ecological status is classified by one of **five classes** (high, good, moderate, poor and bad).	Yes: all (AT, BE (FL), BE (WL), DE, UK, NL, FR, PL)
High, good and moderate ecological status are set in line with the WFD's **normative definitions** (**Boundary setting procedure**)	Class boundaries in all methods are set in line with the normative definitions. The way to achieve this differs. Some (UK, PL) explicitly use changes in ratio of tolerant/impact species, others use expert judgement for defining 'slightly different' and 'disturbed'.
All relevant parameters indicative of the biological quality element are covered (see Table 1 in the IC Guidance). A **combination rule** to combine parameter assessment into BQE assessment has to be defined. If parameters are missing, Member States need to demonstrate that the method is sufficiently indicative of the status of the QE as a whole	N.B. Macrophytes is only part of the BQE, phytobenthos is the other part; combination rules for these parts are unclear. All methods (AT, BE (FL), BE (WL), DE, UK, NL, FR, PL) use taxonomical composition and abundance of the taxa for weighing the indicator value. FR has extra metric for total abundance; BE (FL) and NL have extra metrics for total abundance for different growth forms. UK has two indicator species metrics related to the main pressures (eutrophication / hydromorphological degradation). All combination rules use weighted average, except UK combining both pressure-metrics: worst case.
Assessment is adapted to **intercalibration common types** that are defined in line with the typological requirements of the Annex II WFD and approved by WG ECOSTAT	Yes: all (AT, BE (FL), BE (WL), DE, UK, NL, FR, PL)
The water body is assessed against **type-specific near-natural reference conditions**	Yes: all (AT, BE (FL), BE (WL), DE, UK, NL, FR, PL)
Assessment results are expressed as **EQRs**	Yes: AT, BE (FL), BE (WL), DE, NL, UK No: FR, PL
Sampling procedure allows for **representative** information about water body quality/ecological status **in space and time**	all (AT, BE (FL), BE (WL), DE, UK, NL, FR, PL) sample once / 1-6 years on representative stretches of 100 m or more in (may) June-(August) September. Most sample 1 site/water body, UK 1-3 (recommended 3), BE(FL): 3, NL : 6

All data relevant for assessing the biological **parameters** specified in the WFD's normative definitions are covered by the **sampling procedure**	Yes: all (AT, BE (FL), BE (WL), DE, UK, NL, FR, PL): taxonomical composition, abundance of taxa and total abundance; in line with EN 14184
Selected taxonomic level achieves adequate **confidence and precision** in classification	Yes: all (AT, BE (FL), BE (WL), DE, UK, NL, FR, PL) basically species level; some groups of species within genus or whole genus with comparable indicator value.

Summarise the conclusions of the compliance checking

Design and concept of national assessment methods were extensively discussed and evaluated among experts at the intercalibration workshops. The WFD compliance criteria stated in the IC Guidance draft are met or will be met soon or can be dealt with through workarounds or conversions (EQR's).

4: Methods' intercalibration feasibility check

A feasibility check includes coverage of **intercalibration types, pressures and method concept**. The aim of the check is to address if all national methods address the same common type(s) and pressure(s), and follow a similar assessment concept.

4.1. Typology

What is the progress with the feasibility check on typology? (1 – completed; 2 – in progress; 3 – not started); in case of 'completed' or 'in progress, please fill in the detailed information below; if 'not started', indicate the reasons

The group has evaluated the common IC type delineation and defined sub-types based on alkalinity ranges. Furthermore, the analyses of the biological data showed distinct biological sub-types shared by two or more countries. Benchmarks will be defined accordingly.

How do you deal with common Intercalibration typology?

>> 1 – IC types from the 1st round

2 – New common IC types have been defined

3 – Not decided yet

Please, describe common intercalibration water body types and list the MS sharing each type

The tables below are from the intercalibration report of June 2007 (Birk et al.); the information has not been changed or updated since.

Common IC type	river characterisation	catchment area [km²]	altitude & geomorphology	alkalinity [meq/l]
R-C1	Small lowland siliceous sand	10-100	lowland, dominated by sandy substrate (small particle size), 3-8m width (bankfull size)	> 0,4
R-C2	Small lowland siliceous - rock	10-100	lowland, rock material, 3-8m width (bankfull size)	< 0,4
R-C3	Small mid-altitude siliceous	10-100	mid-altitude, rock (granite) - gravel substrate, 2-10m width (bankfull size)	< 0,4
R-C4	Medium lowland mixed	100-1000	lowland, sandy to gravel substrate, 8-25m width (bankfull size)	> 0,4
R-C5	Large lowland mixed	1000-10000	lowland, barbel zone, variation in velocity, max. altitude in catchment: 800m, >25m width (bankfull size)	> 0,4
R-C6	Small, lowland, calcareous	10-300	lowland, gravel substrate (limestone), width 3-10m (bankfull size)	> 2

Common Intercalibration SUB-Types (only relevant for R-C1, R-C4 and R-C6)

SUB TYPE	IC type	Altitude [m]	Alkalinity [meq/l]	Geology Specification
1.x.1	R-C1	0-200	<1	-
1.x.2	R-C1	0-200	>1	-
4.x.1	R-C4	0-200	<2	-

4.x.2	R-C4	0-200	>2	-
6.1	R-C6	0-200	>2	soft limestone
6.2	R-C6	0-200	>2	hard limestone (karst)

Sub-Type	country	Comment
R-C1.x.2	BE (FL), BE (WL), DE, *DK*, *IT*, *LT*, *LV*, NL, PL	French R-C1 sites are located in Aquitaine region: strong Atlantic influence put typological membership into question
R-C1.x.2		
R-C2	ES, *FR*, *UK*	-
R-C3	AT, BE (WL), *CZ*, DE, *ES*, FR, UK	-
R-C4.x.1	*ES*, FR, *LU*, *SE*, UK	-
R-C4.x.2	BE (FL), *BE (WL)*, DE, *DK*, FR, *LT*, *LU*, *LV*, NL, PL, UK	-
R-C6.1	*BE (WL)*, *DK*, FR, *LT*, *LU*, UK	-
R-C6.2	UK (Northern Ireland)	no common type (currently sub-type was only identified by Northern Ireland; also here: strong Atlantic influence)

What is the outcome of the feasibility evaluation in terms of typology? Are all assessment methods appropriate for the intercalibration water body types?

Method	Appropriate for IC types	Remarks
All	For all appropriate types, see table above	All methods deal with the typology within the method itself.

Conclusion

Intercalibration is feasible in terms of **typology:** as far as the countries share common types they are feasible in terms of typology

1 - feasible

~~2 - non feasible~~

~~3 - we need more information~~

4.2 . Pressures

What is the progress with the feasibility check on pressures? (1 – completed; 2 – in progress; 3 – not started); in case of 'completed' or 'in progress, please fill in the detailed information below; if 'not started', indicate the reasons

Most of the national assessment methods are primarily designed to assess eutrophication pressure. National studies show that the national indices are responsive to nutrient enrichment. However, several pressures are additionally addressed by all national methods. Most of these can be summarised as hydromorphological pressure (Flow modification, General degradation, Habitat destruction, Hydromorphological degradation, Riparian habitat alteration) or other pollutants (Catchment land use, Acidification, Pollution by organic matter, even Heavy metals and Organic compounds are indicated) or Impact of alien species.
The reflection of these differences on the national methods varies.
Some national methods are designed to assess any loss of quality, independent of pressure. This is especially reflected by the conceptually different assessments done by Flanders (using macrophyte growth forms) and the Netherlands (using taxonomic richness and macrophyte growth forms).

Do all national methods address the same pressure(s) ?

Method	Pressure	Remarks
Method A		
Method B		

This table is hard to fill correctly because the pressures are not equally important. Eventually eutrophication and hydromorphological degradation are the most important pressures in all methods and therefore the methods are comparable, thought there will be inevitable differences in assessment results for the same situations.
Because of the difference in combining and weighing the pressure-impact-relations which are both non-linear and due to interpretation differences concerning class boundary setting for the different pressures.

Conclusion
Intercalibration is feasible in terms of **pressures** addressed by the methods:
1 - feasible (likely)
~~2 - non feasible~~
~~3 - we need more information~~

4.3. Assessment concept

What is the progress with the feasibility check on assessment concept? (1 – completed; 2 – in progress; 3 – not started); in case of 'completed' or 'in progress, please fill in the detailed information below; if 'not started', indicate the reasons

The Flemish and Dutch assessment methods are conceptually different. They include growth form evaluation into quality classification. The intercalibration exercise is focussing only on the Flemish metrics related to the common metric. The Dutch method is different in terms of data acquisition (spot checks covering the entire water body) and metric calculation (sum of scores, not weighted average). Currently, this method is not taking part in intercalibration.

The Danish method (under development) is also conceptually different. It identifies syntaxa, i.e. assemblages of species that are typical for defines stages of degradation. Currently, this method is not taking part in intercalibration.

Do all national methods follow a similar assessment concept?

Examples of assessment concept:

- Assessment systems may focus on **different lake zones** - profundal, littoral or sublittoral - and subsequently may not be comparable.
- Additional important issues may be the **assessed habitat type** (soft-bottom sediments versus rocky sediments for benthic fauna assessment methods) or **life forms** (emergent macrophytes versus submersed macrophytes for lake aquatic flora assessment methods)
- **Different community characteristics** - structural, functional or physiological - can be used in assessment methods which can render their comparison problematic. For example, biodiversity indices vs species composition indices.

Method	Assessment concept	Remarks
Method A		
Method B		

Intercalibration might be feasible if parts of the metrics of national methods are excluded, but this has to be investigated further. We started to do so in the first round but were held by misunderstandings and further development of national methods.

Conclusion
Intercalibration is feasible in terms of **assessment concepts:**
~~1 - feasible~~
~~2 - non feasible~~
3 - we need more information

5. Collection of IC dataset

Please describe progress on data collection within the GIG

MS contribution	Data	
Data collected	Biological data	

	Physico-chemical data Pressure data	
Data amount	Number of sites/samples	Per country
IC type coverage	Type 1 Type 2	

List the data acceptance criteria and describe the data acceptance checking process and results

Data acceptance criteria	Data acceptance checking
Data requirements (obligatory and optional)	The common database holds sufficient numbers of data for the intercalibration types R-C1.x.2, R-C3 and R-C4.x.2. All data feasibility checks are completed. Intercalibration will focus on selected taxonomical groups only (exclusion of macroalgae).
The sampling and analytical methodology	
Level of taxonomic precision required and taxalists with codes	
The minimum number of sites / samples per intercalibration type	
Sufficient covering of all relevant quality classes per type	
Other aspects where applicable	

6. Progress on Design the work for IC procedure

Please describe progress of choice of the appropriate intercalibration option.

1 - Same data acquisition, same numerical evaluation → IC Option 1
2 - Different data acquisition and numerical evaluation → IC Option 2
3 - Similar data acquisition, but different numerical evaluation → IC Options 3 supported by the use of common metric(s)

In case of IC Option 2, please explain the differences in data acquisition

Both Option 2 (common metrics) and Option 3 (direct comparison) are evaluated in parallel.
In the NL method data are collected in multiple sites of the river and the metric copes with the differences between the sites, all others collect data in only one site per assessment.
Abundance data are collected differently; most MS collect abundance data per species on a scale that is more or less comparable after conversion into a harmonised scale, but some MS collect additional abundance data of growth forms and structures (BE (FL), NL).

7. Progress on development of IC common metrics

Which actions are ongoing/planned to develop/use common metrics for the IC exercise?
~~1 - Common metric is set~~
2 - Common metric is under development (to be optimalized)
~~3 - Common metric is planned in future~~
~~4 - Common metric is not planned (explain why)~~

Describe the development of IC Common metrics

The common metric mICM was developed based on the indication value for status in every national metric involved. The method for deriving the metric basically has been set, details will probably change with updates of national methods, indication value of the species will be updated with updates of the national methods.

8: Progress on Reference conditions/benchmarking

Have you set a harmonised set of reference criteria for screening of sites in near-natural conditions?

~~1 – Yes, screening of reference sites is done~~
2 – Yes, screening of reference sites will be done in future
~~3 – Not yet~~

Do the intercalibration dataset contain sites in near-natural conditions in a sufficient number to make a statistically reliable estimate. ?

1 – Yes, reference sites are available in a sufficient number (for some types only)
2 – No, we will use alternative benchmark (for other types)
~~3 - We don't know yet~~

From IC Guidance: If natural or near-natural reference conditions are not available or cannot be derived for a certain type (for example, for large rivers) intercalibration needs to be carried out against an alternative benchmark (e.g. good ecological status for that surface water type).

Have you described the biological communities at reference/benchmark conditions, considering potential biogeographical differences?

For the three relevant intercalibration types the communities have been described, but these descriptions need to be updated in the upcoming months.

9. Further comments

The time schedule of 3 years rather tight. The time period of the River Basin Management Plans is 6 years and the methods that will be applied will not change within this period. In the first round of intercalibration most of the methods have been evaluated, in the second round the methods that were not yet developed in 2007 are included in the intercalibration, but the earlier methods will not be updated before 2011. There is a third round of intercalibration needed to evaluate these updates.

Anhang 4: Kopie der Milestone 3 Berichterstattung des CB_{riv}GIG Makrophyten

EUROPEAN COMMISSION
DIRECTORATE GENERAL JRC
JOINT RESEARCH CENTRE
Institute of Environment and Sustainability

WFD Intercalibration Phase 2: Milestone 3 report

Water category/GIG/BQE/horizontal activity:	**CB and EC GIG**
Information provided by:	Sebastian Birk

1. Organisation

1.1. Responsibilities

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

The work is organised within GIGs. This document reports on the activities and achievements within the Central Baltic and Eastern Continental GIGs.

Lead, DE – Sebastan Birk

CB GIG
UK – Nigel Willby, Geoff Phillips
FR – Christian Chauvin
DE – Jochen Schaumburg, Christine Schranz
BE (WL) – Daniel Galoux
BE (FL) – An Leyssen, Kris van Looy
PL – Krzysztof Szoszkiewicz
NL – Roelf Pot
EE –Aive Kors
LT – Mindaugas Gudas
LV – Solvita Strazdina
AT – Karin Pall
IT – Maria Rita Minciardi
DK – Annette Baattrup-Pedersen, Lars K. Larsen
ES – Isabel Pardo
CZ – Libuse Opatrilova
IE – Martin McGarrigle
LU – Nora Welschbillig

EC GIG
SK – Lívia Tóthova
SI – Gorazd Urbanic
AT – Karin Pall, Franz Wagner
HR – Antun Allegro

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:

CB GIG
Austria, Belgium, Czech Republic, Denmark, Estonia, France, Germany, Ireland, Italy, Latvia, Lithuania, Luxembourg, Netherlands, Poland, Spain, United Kingdom

EC GIG
Austria, Croatia, Slovakia, Slovenia

1.3. Meetings

List the meetings of the group:

CB GIG
Edinburgh (UK), March 13-14, 2008
Copenhagen (DK), April 29-30, 2009
Bordeaux (FR), May 27-28, 2010
Turin (IT), spring 2011

EC GIG
Vienna (AT), January 20, 2011

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

1. finalized formally agreed national method,
2. intercalibratable finalized method,
3. method under development,
4. no method developed

Member State	Method	Status
CB GIG		
Austria	Austrian Index Macrophytes for Rivers	1
Belgium (Flanders)	Flemish macrophyte assessment system	1
Belgium (Wallonia)	Macrophyte Biological Index for Rivers	2
France	Biological Macrophytes Index for Rivers	1
Germany	German assessment system for Macrophytes & Phytobenthos according to the EU WFD	1
Italy	Macrophyte Biological Index for Rivers	2
Luxembourg	Biological Macrophytes Index for Rivers	3
Netherlands	WFD-metrics for natural water types	2
Poland	Macrophyte Index for Rivers	1
United Kingdom	Ecological Classification of Rivers using Macrophytes	1
EC GIG		
Austria	Austrian Index Macrophytes for Rivers	1
Slovakia	Slovak Assessment of Macrophytes in Rivers	2
Slovenia	River Macrophyte Index	1

Make sure that the **national method descriptions** meet the level of detail required to fill in the table 1 at the end of this document !

3. Checking of compliance of national assessment methods with the WFD requirements (April 2010 + update in October 2010)

Do all national assessment methods meet the requirements of the Water Framework Directive? (Question 1 in the IC guidance)
Do the good ecological status boundaries of the national methods comply with the WFD normative definitions? (Question 7 in the IC guidance)

List the WFD compliance criteria and describe the WFD compliance checking process and results (the table below lists the criteria from the IC guidance, please add more criteria if needed)

Compliance criteria	Compliance checking conclusions
1. Ecological status is classified by one of **five classes** (high, good, moderate, poor and bad).	Yes – all methods classify ecological status by one of five classes.
2. High, good and moderate ecological status are set in line with the WFD's **normative definitions (Boundary setting procedure)**	Yes – all methods set the high, good and moderate ecological status in line with the WFD's normative definitions.
3. **All relevant parameters** indicative of the biological quality element are covered (see Table 1 in the IC Guidance). A **combination rule** to combine parameter assessment into BQE assessment has to be defined. If parameters are missing, Member States need to demonstrate that the method is sufficiently indicative of the status of the QE as a whole.	"Macrophytes" is one of two components of the BQE "Macrophytes and Phytobenthos". Within the macrophyte methods all relevant parameters of the BQE (i.e. composition and abundance) are covered.
4. 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 – all methods adapt their assessment to intercalibration common types.
5. The water body is assessed against **type-specific near-natural reference conditions**	Yes – all methods assess the water body against type-specific near-natural reference conditions.
6. Assessment results are expressed as **EQRs**	Yes – all methods express the assessment results as EQRs.
7. Sampling procedure allows for **representative** information about water body quality/ ecological status **in space and time**	Yes – the sampling procedures allow for representative information about the ecological status.
8. All data relevant for assessing the biological **parameters** specified in the WFD's normative definitions are covered by the **sampling procedure**	Yes – all relevant data are covered by the sampling procedures.
9. Selected taxonomic level achieves adequate **confidence and precision** in classification	Yes – the selected taxonomic level achieves adequate confidence and precision in classification.

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

All national methods using river macrophytes are WFD compliant. The combination of sub-components to full BQE assessment is still (partly) unclear.

4. Methods' intercalibration feasibility check

Do all national methods address the same common type(s) and pressure(s), and follow a similar assessment concept? (Question 2 in the IC guidance)

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
CB GIG		
Sandy lowland brooks	Abbreviation: R-C1 Catchment area: 10 - 100 km^2 Altitude: <200 m Geology: siliceous Channel substrate: Sand Alkalinity: >1 meq/l	Belgium (Flanders), Denmark, Estonia, Germany, Italy, Latvia, Lithuania, Netherlands, Poland, United Kingdom
Siliceous mountain brooks	Abbreviation: R-C3 Catchment area: 10 - 100 km^2 Altitude: 200 - 800 m Geology: siliceous Channel substrate: Boulders, cobble and gravel Alkalinity: <0.4 meq/l	Austria, Belgium (Wallonia), Czech Republic, France, Germany, Luxembourg, Poland, Spain, United Kingdom
Medium-sized lowland streams	Abbreviation: R-C4 Catchment area: 100 - 1,000 km^2 Altitude: <200 m Geology: mixed Channel substrate: Gravel and sand Alkalinity: >2 meq/l	Belgium (Flanders), Czech Republic, Denmark, Estonia, France, Germany, Ireland, Latvia, Lithuania, Luxembourg, Netherlands, Poland, Spain, United Kingdom
EC GIG		
Plains: medium-sized, lowland	Abbreviation: R-E2 Catchment area: 100 - 1,000 km^2 Altitude: <200 m Geology: mixed Channel substrate: Sand and silt	Bulgaria, Croatia, Hungary, Moldova, Romania, Serbia, Slovakia, Slovenia, Ukraine
Plains: large to very large, lowland	Abbreviation: R-E3 Catchment area: 1,000 - >10,000 km^2 Altitude: <200 m Geology: mixed Channel substrate: Sand, silt and gravel	Bulgaria, Croatia, Hungary, Moldova, Romania, Serbia, Slovakia, Slovenia, Ukraine
Plains: medium-sized, mid-altitude	Abbreviation: R-E4 Catchment area: 100-1,000 km^2 Altitude: 200 - 500 m Geology: mixed Channel substrate: Sand and gravel	Austria, Bulgaria, Croatia, Hungary, Moldova, Romania, Serbia, Slovakia, Slovenia, Ukraine

What is the outcome of the feasibility evaluation in terms of typology? Are all assessment methods appropriate for the intercalibration water body types, or subtypes?

Method	Appropriate for IC types / subtypes	Remarks
CB GIG		
Austrian Index Macrophytes for Rivers (AT)	R-C3	none
Flemish macrophyte assessment system (BE-FL)	R-C1, R-C4	
Macrophyte Biological Index for Rivers (BE-WL)	R-C3	
Biological Macrophytes Index for Rivers (FR)	R-C3, R-C4	
German assessment system for Macrophytes & Phytobenthos according to the EU WFD (DE)	R-C1, R-C3, R-C4	
Macrophyte Biological Index for Rivers (IT)	R-C1	
Biological Macrophytes Index for Rivers (LU)	R-C3, R-C4	
WFD-metrics for natural water types (NL)	R-C1, R-C4	
Macrophyte Index for Rivers (PL)	R-C1, R-C3, R-C4	
Ecological Classification of Rivers using Macrophytes (UK)	R-C1, R-C3, R-C4	
EC GIG		
Austrian Index Macrophytes for Rivers (AT)	R-E4	none
Slovak Assessment of Macrophytes in Rivers (SK)	R-E2, R-E3, R-E4	
River Macrophyte Index (SI)	R-E4	
Conclusion Intercalibration is feasible in terms of typology		

4.2. Pressures

Describe the pressures addressed by the MS assessment methods

Method	Pressures	Remarks
CB GIG		
Austrian Index Macrophytes for Rivers (AT)	Eutrophication, General degradation	none
Flemish macrophyte assessment system (BE-FL)	Eutrophication, General degradation, Alien species	
Macrophyte Biological Index for Rivers (BE-WL)	Eutrophication	
Biological Macrophytes Index for Rivers (FR)	Eutrophication	
German assessment system for Macrophytes & Phytobenthos according to the EU WFD (DE)	Eutrophication, General degradation	
Macrophyte Biological Index for Rivers (IT)	Eutrophication	
Biological Macrophytes Index for Rivers (LU)	Eutrophication	
WFD-metrics for natural water types (NL)	General degradation	
Macrophyte Index for Rivers (PL)	Eutrophication	
Ecological Classification of Rivers using Macrophytes (UK)	Eutrophication, Hydromorphological degradation	
EC GIG		
Austrian Index Macrophytes for Rivers (AT)	Eutrophication, General degradation	none
Slovak Assessment of Macrophytes in Rivers (SK)	Eutrophication	
River Macrophyte Index (SI)	Eutrophication, General degradation	
Conclusion Intercalibration is generally feasible in terms of pressures. However, the Dutch and parts of the Flemish method cannot be intercalibrated due to different pressure focus.		

4.3. Assessment concept

Do all national methods follow a similar assessment concept?

Examples of assessment concept:

- **Different community characteristics** - structural, functional or physiological - can be used in assessment methods which can render their comparison problematic. For example, sensitive taxa proportion indices vs species composition indices.
- Assessment systems may focus on **different lake zones** - profundal, littoral or sublittoral - and subsequently may not be comparable.
- Additional important issues may be the **assessed habitat type** (soft-bottom sediments versus rocky sediments for benthic fauna assessment methods) or **life forms** (emergent macrophytes versus submersed macrophytes for lake aquatic flora assessment methods)

Method	Assessment concept	Remarks
CB GIG		
Austrian Index Macrophytes for Rivers (AT)	Focus on taxonomic structure	none
Flemish macrophyte assessment system (BE-FL)	Focus on taxonomic structure and growth forms	
Macrophyte Biological Index for Rivers (BE-WL)	Focus on taxonomic structure	
Biological Macrophytes Index for Rivers (FR)	Focus on taxonomic structure	
German assessment system for Macrophytes & Phytobenthos according to the EU WFD (DE)	Focus on taxonomic structure	
Macrophyte Biological Index for Rivers (IT)	Focus on taxonomic structure	
Biological Macrophytes Index for Rivers (LU)	Focus on taxonomic structure	
WFD-metrics for natural water types (NL)	Focus on taxonomic structure and growth forms	
Macrophyte Index for Rivers (PL)	Focus on taxonomic structure	
Ecological Classification of Rivers using Macrophytes (UK)	Focus on taxonomic structure	
EC GIG		
Austrian Index Macrophytes for Rivers (AT)	Focus on taxonomic structure	none
Slovak Assessment of Macrophytes in Rivers (SK)	Focus on taxonomic structure	
River Macrophyte Index (SI)	Focus on taxonomic structure	

Conclusion

Intercalibration is generally feasible in terms of assessment concept. However, the Dutch and parts of the Flemish method cannot be intercalibrated due to different assessment concepts.

5. Collection of IC dataset

Describe data collection within the GIG.
This description aims to safeguard that compiled data are generally similar, so that the IC options can reasonably be applied to the data of the Member States.

Make the following table for each IC common type

IC Type	Country	Number of surveys	Number of pressure data
CB GIG			
R-C1	Belgium (Flanders)	124	116
	Belgium (Wallonia)	9	9
	Germany	174	154
	Denmark	26	26
	Spain	30	7
	France	6	4
	Italy	38	22
	Lithuania	1	1
	Latvia	17	17
	Netherlands	33	33
	Poland	37	37
	Sweden	7	7
	United Kingdom	27	27
	Total	***529***	***460***
R-C3	Austria	33	33
	Belgium (Wallonia)	43	42
	Czech Republic	49	49
	Germany	165	99
	France	65	57
	Luxembourg	31	31
	Poland	19	19
	United Kingdom	33	32
	Spain	8	2
	Total	***446***	***364***
R-C4	Belgium (Flanders)	15	9
	Belgium (Wallonia)	26	26
	Czech Republic	10	10
	Germany	88	70
	Denmark	15	15
	Estonia	5	5
	Spain	110	23
	France	114	120
	Ireland	88	88
	Lithuania	19	19
	Luxembourg	27	27
	Latvia	38	38
	Netherlands	10	10
	Poland	45	44
	Sweden	20	20
	United Kingdom	263	257
	Total	***893***	***781***

IC Type	Country	Number of surveys	Number of pressure data
EC GIG			
R-E2	Croatia	2	2
	Czech Republic	4	4
	Romania	14	14
	Slovakia	17	17
	Total	***37***	***37***
R-E3	Croatia	2	2
	Czech Republic	24	24
	Hungary	19	19
	Romania	18	18
	Slovakia	15	15
	Total	***78***	***78***
R-E4	Austria	21	21
	Romania	6	6
	Slovenia	7	7
	Total	***34***	***34***

List the data acceptance criteria used for the data quality control and describe the data acceptance checking process and results

Data acceptance criteria	Data acceptance checking
Data requirements (obligatory and optional)	**Standardization of database** Some national methods include algae taxa in bioassessment. Others do not consider these taxa in their macrophyte method. The taxonomic groups surveyed by all national methods are (selected) mosses and vascular plants. The application of indices requiring algae information to the common database, that includes surveys with and without algae records, introduces bias caused by the different national survey protocols. Therefore, we excluded the records of algae taxa from the surveys used in the analysis.
The sampling and analytical methodology	
Level of taxonomic precision required and taxalists with codes	
The minimum number of sites / samples per intercalibration type	
Sufficient covering of all relevant quality classes per type	
Other aspects where applicable	

6. Benchmarking: Reference conditions or alternative benchmarking

In section 2 of the method description of the national methods above, an overview has to be included on the derivation of reference conditions for the national methods. In section 6 the checking procedure and derivation of reference conditions or the alternative benchmark at the scale of the common IC type has to be explained to ensure the comparability within the GIG.

Clarify if you have defined
- common reference conditions (N)
- or a common alternative benchmark for intercalibration (Y)

6.1. Reference conditions

Does the intercalibration dataset contain sites in near-natural conditions in a sufficient number to make a statistically reliable estimate? (Question 6 in the IC guidance)

- Summarize the common approach for setting reference conditions (true reference sites or indicative partial reference sites, see Annex III of the IC guidance):

not applicable

- Give a detailed description of **reference criteria** for screening of sites in near-natural conditions (abiotic characterisation, pressure indicators):

not applicable

- Identify the **reference sites** for each Member State in each common IC type. Is their number sufficient to make a statistically reliable estimate?

not applicable

- Explain how you have screened the biological data for impacts caused by pressures not regarded in the reference criteria to make sure that true reference sites are selected:

not applicable

- Give detailed description of **setting reference conditions** (summary statistics used)

not applicable

6.2. Alternative benchmarking (only if common dataset does not contain reference sites in a sufficient number)

- Summarize the common approach for setting **alternative benchmark** conditions (describe argumentation of expert judgment, inclusion of modelling)

Deriving alternative benchmarks in intercalibration: general procedure

According to the intercalibration guidance, reference or alternative benchmark conditions have to be established for the common intercalibration types in order to be able to compare the national class boundary settings. If natural reference conditions are not available or cannot be derived for a certain type, intercalibration needs to be carried out against an alternative benchmark. To enhance the transparency of the intercalibration process defining benchmark conditions shall be done using the common dataset. This requires finding benchmarks based on actual data sampled at existing sites. The availability of a comprehensive database that especially covers sites in alternative benchmark conditions (impacted by similar levels of impairment) is essential.

Selecting benchmark sites

First step in the delineation process of alternative benchmarks in intercalibration is the selection of suitable candidate benchmark sites from the common dataset. In general, benchmark sites shall be representative of the countries and tentative biotypes involved in the intercalibration exercise of a common type, balanced between regions regarding the level and type of pressure acting, and informative concerning the scope of supporting data. These data shall cover at least the pressures (and drivers) most relevant to the biological quality element of interest, and provide sufficient information about the environmental parameters characterising the water body or sampling site.

Alternative benchmarking relies on the homogeneous distribution of similarly disturbed sites among biotypes within a common type. If certain regions feature more pronounced anthropogenic disturbance, the common benchmark will be imbalanced. Balancing the selection of benchmark sites is not always trivial if, for instance, countries featuring contrasting population densities or land use practices like Poland and the Netherlands are involved in the exercise. This would then require additional adjustment procedures among biotypes or the exclusion of the least disturbed sites to equilibrate the differences. However, the causes for imbalances need to be evaluated carefully as different chemical value ranges indicative of certain pressures, for example, may simply reflect the natural characteristics of each biotype within the broader intercalibration type (e.g. in relation to geology or catchment area).

Benchmark sites are selected by a screening process similar to the delineation of reference sites. To identify candidate benchmark sites the common dataset is pre-screened using biological criteria. This optional step includes, for instance, the selection of common high status sites, i.e. sites classified in high ecological status by the majority of countries participating in the intercalibration exercise. Another option is to define thresholds on the common metric scale, and to exclude sites whose biological communities do not comply with these criteria.

The actual benchmark sites to be used in intercalibration result from pressure screening. This process can be carried out in various ways that differ in the scope of required data and the analytical efforts involved. In the following case study on river macrophyte benchmarking we describe an approach based on the analysis of pres-sure-impact-relationships. Benchmark sites were selected by sequentially removing the most impacted samples from the dataset until the pressure factors no longer influence the biological communities. This process required a broad range of environmental data and multivariate analysis. Less extensive approaches may comprise the definition of fixed pressure ranges from which sites are selected, or the reference to water bodies "not at risk" according to the national risk assessments following Article 5 of the WFD. These latter methods are suitable if the national data are incomparable (e.g. due to different data acquisition). In any case we advice to demonstrate the relevance of selected screening criteria for the intercalibrated quality element, e.g. by providing pressure-impact-relationships and rationales for the threshold selections.

Identifying biotypes

Biotypes are characterised by distinct biological communities that differ in taxonomic composition and/or dominance structure. These subtypological variants occur within a common intercalibration type due to divers patterns of species dispersal, climatological gradients or regional specificities of the common type (e.g. caused by lithology, topography). Biotypes can be identified by cluster analysis. In the case study described below we used Two-Way INdicator SPecies ANalysis. TWIN-SPAN is a divisive clustering method that first ordinates the samples using Correspondence Analysis (CA), then divides the samples into two groups using the CA centroid line. These steps are repeated for each subset of samples until a user-defined thresh-old is reached.

The level of detail in identifying biotypes is constraint by the definition of the national water body types to which the assessment method is adapted. The definition of biotypes must not be more specific than the delineation of national types. If, for instance, the cluster analysis reveals distinct biotypes among sites of a country's common intercalibration type, but the country's assessment system classified these biotypes using the same reference definition, a biotypological distinction for this country within the intercalibration exercise is inappropriate.

The biological and environmental features of the identified biotypes shall be de-scribed to evaluate their relevance in the intercalibration exercise. Significant differences in the common metric ranges among biotypes, demonstrated by statistical testing, are indicative of regional heterogeneities within an intercalibration type.

Adopting national benchmarks

The transnational benchmark dataset acts as a standard for comparing the biological communities at similarly disturbed conditions across large geographical regions. The aim of benchmarking is to adjust for the effects of regional differences among countries when comparing the national quality classifications. Provided that the benchmark sites selected for a common type are sufficiently representative, balanced and informative, the benchmark information needs to be transferred onto the level of the national classifications, i.e. the country being intercalibrated. The technical procedure including the different possible options is explained in Annex V (IC Guidance) in the step about "benchmark normalisation". Here, we specify how to select the appropriate national benchmark from the transnational biotypes identified. This step is not required if benchmarks have already been individually derived from separate national datasets (e.g. due to different data acquisition).

Figure 1 depicts the possible combinations of two significantly different biotypes occurring in two countries within a common intercalibration type. Constellations 1 and 2 show that the same biotype(s) occur(s) in both countries. In either case no benchmark normalisation is required. The comparison of ecological quality classifications can be performed between these countries without adjusting for biotypological differences. In constellations 3 and 4 the biotypes are differently distributed between countries. Both cases require benchmark normalisation, using the median (common metric or national EQR) value of all benchmark sites belonging to a biotype. While constellation 3 portrays a clear distinction between biotype occurrences in the two countries, country B features both biotypes in constellation 4. We suggest to benchmark this country by selecting the average of the median values of both biotypes.

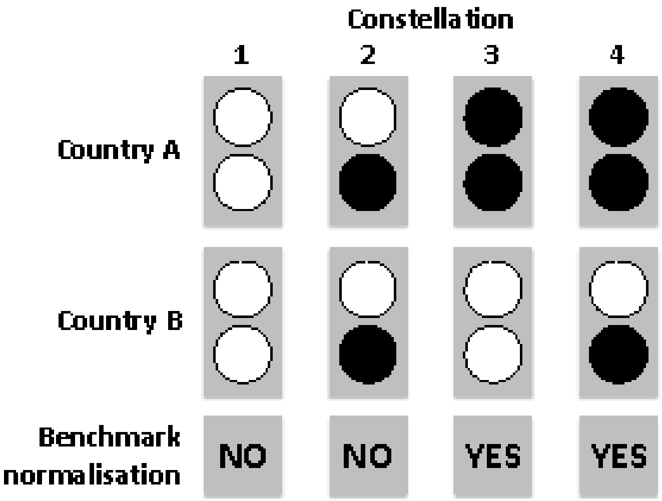


Figure 1: Potential distribution of two biotypes (black and white dots) within an intercalibration type between countries A and B (see text for explanation)

- Give a detailed description of **criteria** for screening of **alternative benchmark** sites (abiotic criteria/pressure indicators that represent a similar low level of impairment to screen for least disturbed conditions)

In plotting the ranges of environmental parameter values we could demonstrate the natural and anthropogenic gradients covered by the data. Both gradients were assumed to influence the biology at the survey sites, but we were only interested in the role of natural factors shaping the aquatic assemblages. Therefore, the aim of the subsequent analysis was to differentiate between the individual effects of each gradient on the macrophyte communities. We used partial Canonical Correspondence Analysis (pCCA) with pressure factors as explanatory variables and natural factors as co-variables. The

pCCA identified those factors that (significantly) explained the remaining variance in the biological data after the effects of the natural factors (and combined effects) were excluded.
To reduce the length of the pressure gradient in the next step, we iteratively removed sites from the data set until none of the pressure factors revealed significant impacts on the community ("benchmark screening"). In this removal process we were guided by the factors with the highest explanatory power and omitted the sites showing the worst condition concerning the relevant factor. If, for instance, orthophosphate was best in explaining community variance, we removed the site with the highest PO_4-P concentration from the data set. We sequentially removed the sites with the next lowest phosphate concentrations until the influence of the anthropogenic pressure was insignificant for the pool of remaining sites. These sites were, thus, referred to as "benchmark sites". To review the effect of the screening process we compared the spread of environmental parameter values between benchmark and removed sites.
The benchmark data showed significantly lower value ranges for most of the pres-sure factors, which, however, identified the benchmark sites as more or less disturbed (as opposed to pristine). Differences in the natural factors pointed at imbalances in the pressure gradient within the common intercalibration type (e.g. some countries feature generally less degraded sites, larger rivers are more disturbed).

- Identify the **alternative benchmark sites** for each Member State in each common IC type

still in progress

- Describe how you validated the selection of the alternative benchmark with biological data

To discover possible biotypes based on community variants within the common types we classified the macrophyte data at the benchmark sites using TWINSPAN. This analysis identified distinct groups of survey sites featuring similar macrophyte communities. By means of IndVal analysis, we determined the most indicative taxa of each group, i.e. taxa found mostly in a single group of the subtypology and present in the majority of the survey sites belonging to that group. Taxa with a relative abundance ≥ 2% were listed to provide information about the dominance structure of the community variants.
To characterise the individual biotypes we reviewed the environmental parameter values for each group and tested for significant differences between groups. Emphasis was placed on detecting country-specific stream type variants that would allow for clear allocation of national benchmarks for the common intercalibration type. We carefully checked for differences in pressure values between biotypes.

- Give detailed description how you identified the position of the alternative benchmark on the gradient of impact and how the deviation of the **alternative benchmark** from reference conditions has been derived

still in progress

Describe the **biological communities** at reference sites or at the alternative benchmark, considering potential biogeographical differences:

still in progress

7. Design and application of the IC procedure

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

Which IC option did you use?

- IC Option 1 - Same assessment method, same data acquisition, same numerical evalua-tion (N)
- IC Option 2 - Different data acquisition and numerical evaluation (N)
- IC Options 3 - Similar data acquisition, but different numerical evaluation (BQE sampling and data processing generally similar, so that all national assessment methods can reasonably be applied to the data of other countries) → supported by the use of common metric(s) (Y)
- Other (specify) (N)

Explanation for the choice of the IC option:

We will perform an IC Option 3 (direct comparison) using pseudo-common metrics supported by mICM (common metric).

In case of IC Option 2, please explain the differences in data acquisition

not applicable

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

Describe the IC Common metric:

See
Birk S, Willby NJ (2010) Towards harmonization of ecological quality classification: establishing common grounds in European macrophyte assessment for rivers. Hydrobiologia 652:149-163.

Are all methods reasonably related to the common metric(s)? (Question 5 in the IC guidance)

Please provide the correlation coefficient (r) and the probability (p) for the correlation of each method with the common metric (see Annex V of IC guidance).

Member State/Method	r	p
A	still in progress	
B		

Explain if any method had to be excluded due to its low correlation with the common metric:

still in progress

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

Clarify if
- boundaries were set only at national level (Y)
- or if a common boundary setting procedure was worked out at the scale of the common IC type (N)

In section 2 of the method description of the national methods above, an overview has to be included on the boundary setting procedure for the national methods to check compliance with the WFD. In section 8.1 the results of a common boundary setting procedure at the scale of the common IC type should be explained where applicable.

8.1. Description of boundary setting procedure set for the common IC type

Summarize how boundaries were set following the framework of the BSP:

- Provide a description how you applied the full procedure (use of discontinuities, paired metrics, equidistant division of continuum)

not applicable

- Provide pressure-response relationships (describe how the biological quality element changes as the impact of the pressure or pressures on supporting elements increases)

not applicable

- Provide a comparison with WFD Annex V, normative definitions for each QE/ metrics and type

not applicable

8.2. Description of IC type-specific biological communities representing the "borderline" conditions between good and moderate ecological status, considering possible biogeographical differences (as much as possible based on the common dataset and common metrics).

still in progress

8.3. Boundary comparison and harmonisation

Describe comparison of national boundaries, using comparability criteria (see Annex V of IC guidance).

still in progress

- Do all national methods comply with these criteria ? (Y/N)

– **If not, describe the adjustment process:**

still in progress

Anhang 5: Alternative benchmarking in the intercalibration of assessment methods using river macrophytes

Alternative benchmarking in the intercalibration of assessment methods using river macrophytes

First Draft

Sebastian Birk – University of Duisburg-Essen, DE
Nigel Willby – University of Stirling, UK

May 2010

Table of contents

1 Why use alternative benchmarking?

The intercalibration of national assessment methods requires the harmonised definition of common benchmarks to enable valid comparisons of good ecological status classifications (Schmedtje et al. 2009). The exercises for river invertebrates and diatoms in the Central Baltic GIG, for instance, referred to sites in (near-)natural reference state (Kelly et al. 2009, Bennett et al. 2010). On this basis country-specific benchmark values were derived on the common metric scale that represented the biological status under conditions undisturbed by man.

The (near-)natural reference state implies that the aquatic community shows no or insignificant impacts due to anthropogenic pressures. Given that the biological data have been acquired through comparable methods, remaining differences in community structure between reference sites are, thus, largely due to natural reasons, i.e. spatial and temporal variations. Since the broadly defined common stream types form the basic unit of intercalibration, biogeographical and sub-typological differences within a common type presumably play a significant role.

The identification of such differences is of major importance in intercalibration, since the ecological status is defined as the relative deviation from reference conditions. These reference conditions can significantly differ among (eco)regions sharing a common intercalibration type. Therefore, the good ecological status can only be harmonised if trans-national baselines are defined, in which the anthropogenic influences are controlled to uncover the natural spatial variation inherent within a common intercalibration type. Although the use of data from near-natural reference sites seems most straightforward, the scarcity of truly undisturbed sites precludes this approach for most regions and water body types in Europe (e.g. Birk & Hering, 2009).

Alternative benchmarking is meant to provide a practical solution to this problem (Schmedtje et al. 2009). However, this concept remains theoretical unless applied to concrete examples. First exercises focussed on least disturbed conditions (Stoddard et al. 2006) and derived common benchmarks from the aquatic communities occurring under these conditions (Birk & Hering 2009; Birk et al. [in prep.]).

2 The CB_{riv}GIG Macrophyte Approach

The study at hand describes the benchmarking approach conducted in the intercalibration exercise of river macrophyte methods in the Central Baltic GIG (Birk et al. 2007). The exercise covered three common intercalibration types: sandy lowland brooks (R-C1), siliceous mountain brooks (R-C3), and medium-sized lowland streams (R-C4). Among the available data only few sites complied with the CB_{riv}GIG reference criteria (van de Bund 2009). Therefore, we referred to sites in “best available” ecological status (i.e. “common high status” according to Birk & Willby 2010). These sites featured macrophyte communities of mainly high or good quality, but we assumed various degrees of anthropogenic pressure acting at the sites.

2.1 Screening for benchmark sites

In plotting the ranges of environmental parameter values we could demonstrate the natural and anthropogenic gradients covered by the data. Both gradients were assumed to influence the biology at the survey sites, but we were only interested in the role of natural factors shaping the aquatic assemblages. Therefore, the aim of the subsequent analysis was to differentiate between the individual effects of each gradient on the macrophyte communities. We used partial Canonical Correspondence Analysis (pCCA; ter Braack 1988) with pressure factors as explanatory variables and natural factors as covariables. The pCCA identified those factors that (significantly) explained the remaining variance in the biological data after the effects of the natural factors (and combined effects) were excluded.

To reduce the length of the pressure gradient in the next step, we iteratively removed sites from the data set until none of the pressure factors revealed significant impacts on the community (“benchmark screening”). In this removal process we were guided by the factors with the highest explanatory power and omitted the sites showing the worst condition concerning the relevant factor. If, for instance, orthophosphate was best in explaining community variance, we removed the site with the highest PO_4-P concentration from the data set. We sequentially removed the sites with the next lowest phosphate concentrations until the influence of the anthropogenic pressure was insignificant for the pool of remaining sites. These sites were, thus, referred to as “benchmark sites”. To review the effect of the screening process we compared the spread of environmental parameter values between benchmark and removed sites.

The benchmark data showed significantly lower value ranges for most of the pressure factors, which, however, identified the benchmark sites as more or less disturbed (depending on the stream type investigated). Differences in the natural factors pointed at imbalances in the pressure gradient within the common intercalibration type (e.g. some countries feature generally less degraded sites, larger rivers are more disturbed).

2.2 Identifying subtype variants

To assess possible subtypes based on community variants within the common types we classified the macrophyte data at the benchmark sites using a Two Way Indicator Species Analysis (TWINSPAN, Hill 1979). This analysis identified distinct groups of survey sites featuring similar macrophyte communities. Using IndVal analysis, we determined the most indicative taxa of each group, i.e. taxa found mostly in a single group of the subtypology and present in the majority of the survey sites belonging to that group (Dufrene & Legendre 1997). Taxa with a relative abundance ≥ 2% were listed to provide information about the dominance structure of the community variants.

To characterise the individual subtypes we reviewed the environmental parameter values for each group and tested for significant differences between groups. Emphasis was placed on detecting country-specific stream type variants that would allow for clear allocation of national benchmarks for the common intercalibration type. We carefully checked for differences in pressure values between subtype variants. Alternative benchmarking relies on the homogeneous distribution of similarly disturbed sites among subtype variants within an intercalibration type. If, for instance, certain regions feature more pronounced anthropogenic disturbance, the common benchmark would be imbalanced. This would require additional adjustment procedures among subtypes. However, different chemical value ranges may simply reflect the natural characteristics of each subtype within the broader intercalibration type (e.g. in relation to alkalinity or catchment area).

3 Data basis

Basis for the benchmark analyses were data comprising information about composition and abundance of river macrophytes at 236 sites. Data were surveyed at the three aforementioned stream types shared by 13 countries (Table 1; Figure 1). We selected these survey data because they complied with the criteria of “common good status sites” according to Birk & Willby 2010, i.e. the majority of national assessment methods[1] classified the surveys in good status and none of the methods classified them in poor or bad status. For the siliceous mountain brooks we modified the criteria to gain “common high status sites”, i.e. high status was assigned by most methods and no method classified these surveys in moderate or worse status.

Table 1: Number and origin of macrophyte surveys used in the analyses

IC type	Country	#Best-Available	#Benchmark
Sandy lowland brooks	Belgium (Flanders)	13	5
	Denmark	3	3
	Germany	7	4
	Latvia	3	3
	Lithuania	1	1
	Netherlands	4	3
	Poland	25	23
Siliceous mountain brooks	Austria	5	2
	Belgium (Wallonia)	13	5
	Czech Republic	9	7
	France	29	26
	Germany	12	2
	Great Britain	12	7
Medium-sized lowland streams	Belgium (Wallonia)	2	2
	France	48	33
	Germany	3	3
	Great Britain	26	14
	Latvia	5	4
	Lithuania	1	1
	Poland	14	13
	Sweden	1	1

The data originated from national monitoring programmes or scientific projects (e.g. Furse et al., 2006). Countries applied national macrophyte survey protocols that were in line with the requirements of the European Standard EN 14184:2003. Representative river stretches were visually inspected during the growing season (June to September) by wading, diving or boating, using rake, grapnel or aqua-scope where necessary (Birk et al., 2007). Representative sites spanned about 100 metres of river length.

Before commencing the benchmark analyses we harmonised the data to minimise the differences among national techniques of data acquisition. Firstly, we removed all macroalgal records (except for Charales) from the data since these were not consistently surveyed by all countries. Secondly, all riparian taxa were excluded (level of aquaticity ≥ 6; Birk et al. 2007) to focus the analysis only on primarily aquatic taxa. And

[1] Classification schemes of Austria, Belgium, Germany, France, Great Britain and Poland

finally, the national abundance scales were converted into an international 5-class scale (see Birk & Willby 2010).

We collected supportive environmental information for all sites that featured biological surveys data (Table 2). Most information was acquired within the national monitoring programmes and kindly provided by the national intercalibration delegates. The chemical measurements varied in level of aggregation and spatio-temporal match with the biological surveys (annual or growing season means: 75% of surveys, rest: spot measures; measured at exact site of macrophyte survey: 85%, rest: measured at same water body; measured in the same year: 50%, rest: measured in previous year(s)). We additionally derived data climatic conditions (Mitchell et al. 2004), stream gradient (Google Inc. 2010) and the land use at the survey sites (Bossard et al. 2000) and assigned sites to hydro-ecoregions (Wasson et al. 2003). Information was incomplete for some parameters (chemical data: approx. 15% of surveys missing; catchment land use data: 5%; Substrate, catchment size, channel width and depth: 1-3%). Before entering the data into the pCCA, missing values we replaced by global or country-specific averages, or inferred by regression analysis in case of significant relationships with other parameters. For the test of group differences only available data were used, i.e. we excluded sites with missing data.

Table 2: Supportive environmental information of the survey sites ([1] National monitoring, [2] Mitchell et al. 2004, [3] Wasson et al. 2003, [4] Google Inc. 2010, [5] Bossard et al. 2000

Group	Parameter	Unit	Source
Typology/(Bio)Geography	Longitude	decimal °	[1]
	Latitude	decimal °	[1]
	Country	-	[1]
	Mean annual precipitation	mm	[2]
	Mean annual air temperature	°C	[2]
	Hydro-ecoregion	-	[3]
	Slope	*relative scale (min-max)*	[4]
	Substrate score	*ranked scale*	[1]
	Altitude	M	[1]
	Catchment	km2	[1]
	Average channel depth	M	[1]
	Average channel width	M	[1]
Catchment land use	Artificial land use	%	[5]
	Intensive agriculture	%	[5]
	Low intensity agriculture	%	[5]
	(Semi-) natural land use	%	[5]
Site	Shading	*ranked scale*	[1]
	Land use at site	*ranked scale*	[5]
Chemistry	Conductivity	µS/cm	[1]
	Alkalinity	µeq/l	[1]
	$N\text{-}NH_4$	mg/l	[1]
	$N\text{-}NO_3$	mg/l	[1]
	$P\text{-}PO_4$	mg/l	[1]
	BOD_5	mg/l	[1]

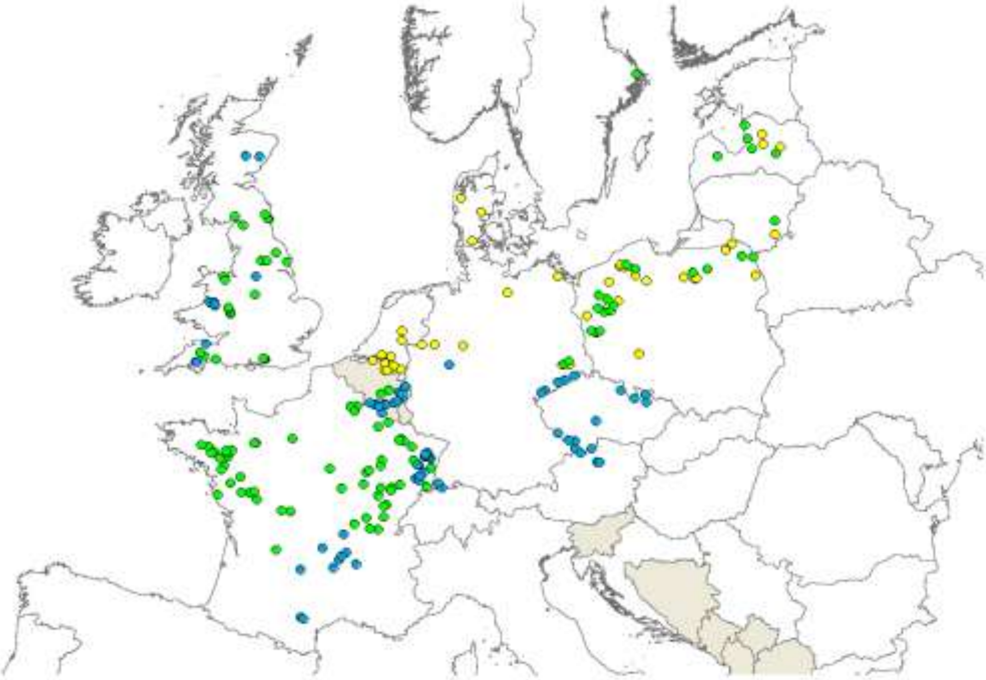


Figure 1: Location of the 236 "best available"[2] macrophyte survey sites (yellow: Sandy lowland brooks; blue: Siliceous mountain brooks; green: Medium-sized lowland streams)

[2] sites in common high/good status

4 Results

4.1 Screening for benchmark sites

The survey data cover a broad range of environmental conditions from torrential, base-poor cheeks exposed to a strong Atlantic influence, to sluggish lowland streams in late-Pleistocene, naturally enriched landscapes of continental climate. The geographical spread of site locations amounts to over 2000 km in latitude. 23 different hydroecoregions (Wasson et al. 2003) are included in the data set. The average annual precipitation, for instance, ranges from >2.1 litres for sites in Scotland to less than half a litre for sites in Eastern Poland. The maximum difference of the mean annual air temperature between sites in Southern France and Latvia or the Czech Republic is 8°C.

The level of pressure varies among the sites in the data set. The anthropogenic land use in the catchment as an indicator of the driving forces of human disturbance reveals generally higher shares for the lowland rivers (Figure 2). The effects of settlements and intensive agriculture are clearly observable in the nutrient concentrations. Some site feature dramatically high values especially for ammonium and orthophosphate (see Figures 3 to 5).

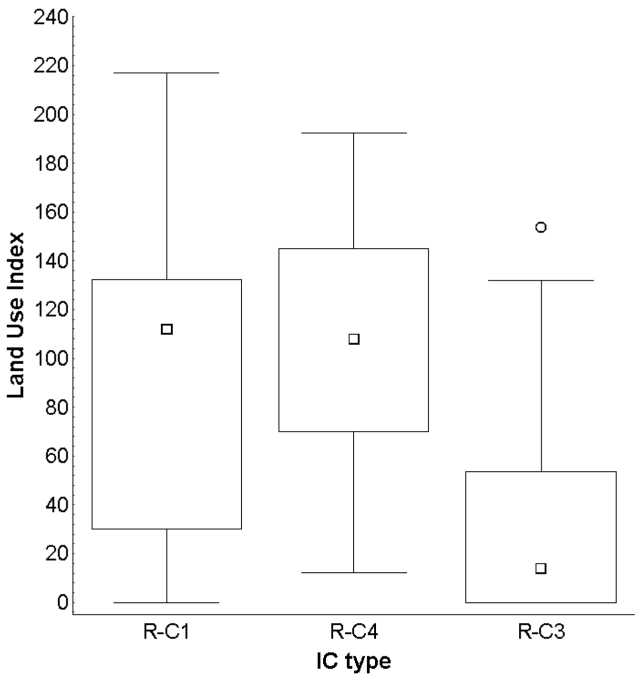


Figure 2: Value ranges of the Land Use Index[3] among common intercalibration types

[3] Calculation of the Catchment Land Use Index (LUI) according to Böhmer et al. 2004: LUI = 4 * Artificial land use + 2 * Intensive agricultural land use + 1 * Extensive agricultural land use

The benchmark screening resulted in the removal of a quarter to more than one-third of the sites, depending on common stream type. Most of the pressure variables revealed significant differences between benchmark and impacted sites (see Table A2 in the annex). Figures 3 to 5 display the value ranges of both groups for selected parameters.

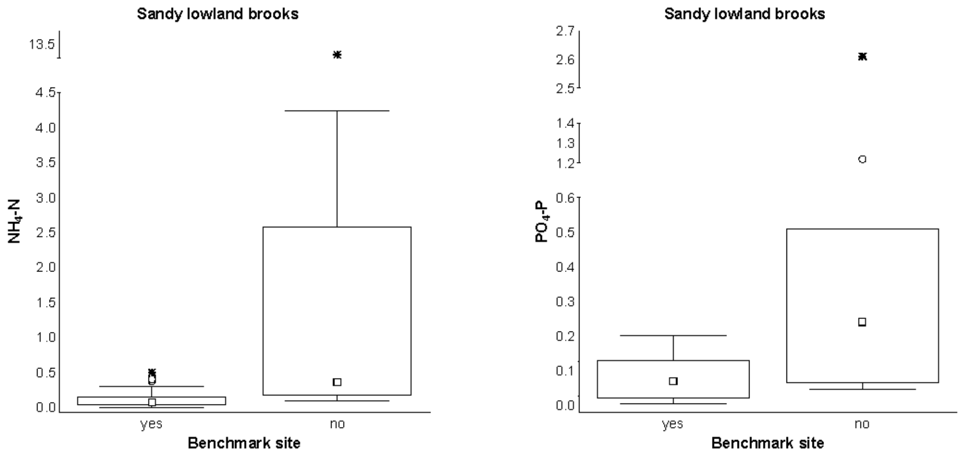


Figure 3: Spread of selected parameter values between benchmark and impacted sites for the sandy lowland brooks (R-C1)

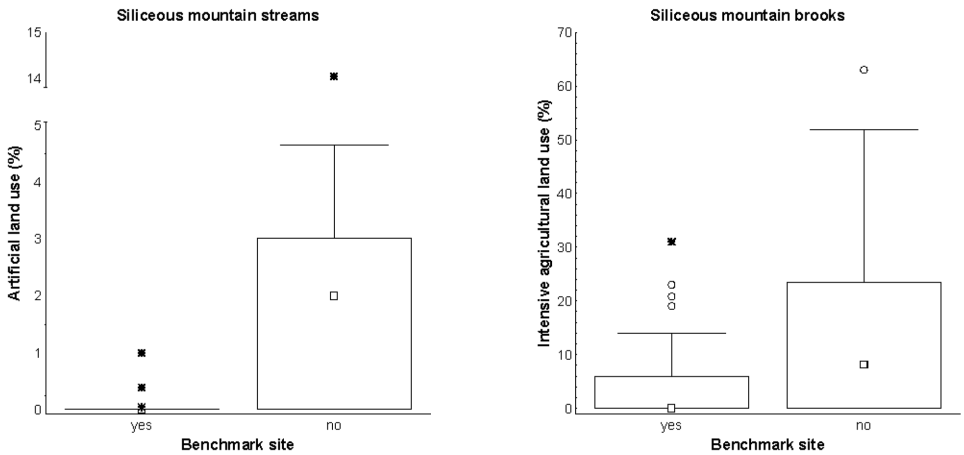


Figure 4: Spread of selected parameter values between benchmark and impacted sites for the siliceous mountain brooks (R-C3)

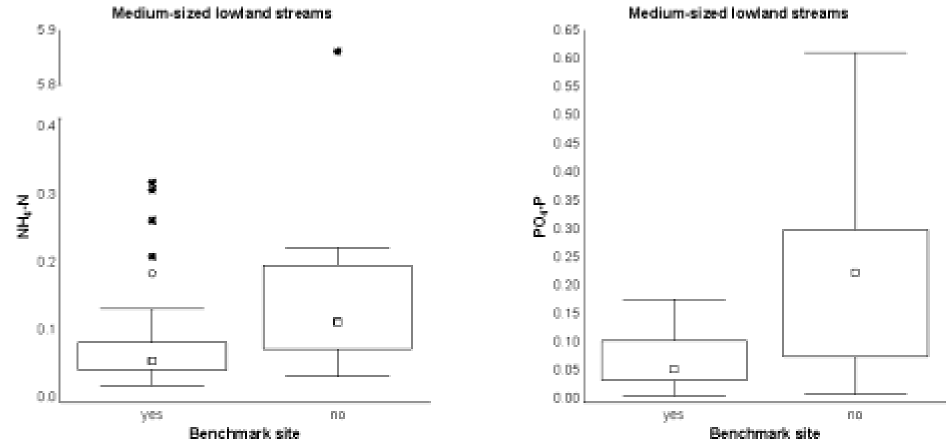


Figure 5: Spread of selected parameter values between benchmark and impacted sites for the medium-sized lowland streams (R-C4)

4.2 Identifying subtype variants

We identified three subtype variants within each common intercalibration type based on the TWINSPAN of biological survey data at benchmark sites. Using IndVal analysis and reviewing the dominance structure we distinguished between the community variants listed in Table 3. Details about the national affiliation of survey sites are given in the Annex (Table A1).

The variants reveal significantly different values for most natural and some pressure variables (Table 4).

Table 3: Subtype variants of the common intercalibration types as identified by TWINSPAN

IC type	TWINSPAN Group	Name of subtype variant
Sandy lowland brooks	1	*Potamogeton natans-Callitriche* variant
	2	*Berula erecta* variant
	3	*Berula-Glyceria*-Bryophyte variant
Siliceous mountain brooks	3	*Pellia-Racomitrium-Hyocomium* variant
	4	*Rhynchostegium-Fissidens-Amblystegium* variant
	5	*Brachythecium rivulare* variant
Medium-sized low-land streams	1	*Rhynchostegium-Amblystegium fluviatile* variant
	2	*Fontinalis-Sparganium emersum* variant
	3	*Nuphar-Phragmites-Lemna* variant

Table 4: Significant differences of parameter values among Twinspan group within the common intercalibration types (Kruskal-Wallis ANOVA by ranks; x=p>0.05)

Parameter	R-C1	R-C3	R-C4
mICM	x	x	x
Number of mICM species			x
Sum of mICM species abundances	x		x
Relative number of moss species	x	x	x
Relative abundance of moss species	x		x
Latitude	x	x	
Longitude	x	x	x
Average annual air temperature	x		
Average annual precipitation	x	x	x
Altitude	x	x	
Catchment		x	
Average channel width			
Average channel depth		x	x
Slope	x		
Dominant channel substrate			x
Shading	x		
Artificial land use	x		x
Intensive agricultural land use			
Extensive agricultural land use	x		x
Natural land use			
Land use at site	x		
NH_4-N	x	x	
NO_3-N	x		x
PO_4-P		x	
BOD_5		x	
Conductivity			

4.2.1 Subtype variants of the sandy lowland brooks (R-C1)

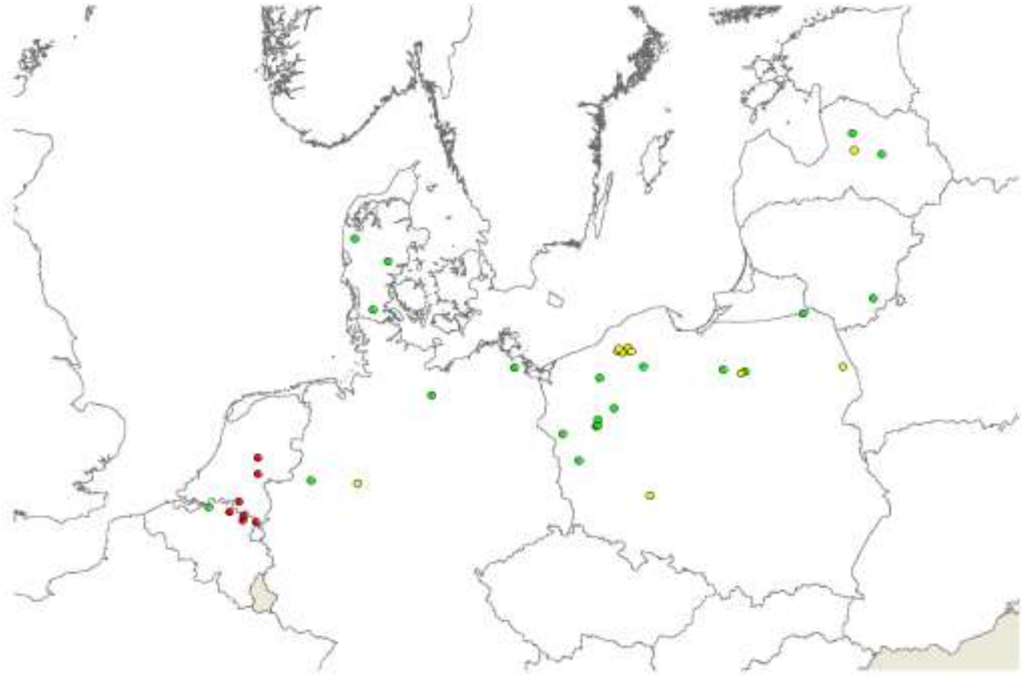


Figure 6: Location of benchmark sites for the sandy lowland brooks (red: *Potamogeton natans-Callitriche* variant, green: *Berula erecta* variant, yellow: *Berula-Glyceria*-Bryophyte variant)

Table 5: Significant indicator species of the TWINSPAN groups for the sandy lowland brooks (IndVal analysis)

TwSp-Group	Species	IndVal (%)	p-level
1	*Callitriche hamulata*	85.7	0.0002
	Potamogeton natans	57.1	0.0006
	Sparganium emersum	42.7	0.019
	Callitriche platycarpa	41.7	0.0024
	Polygonum hydropiper	35.7	0.039
	Potamogeton trichoides	28.6	0.0214
	Potamogeton polygonifolius	28.6	0.0222
	Lemna gibba	28.6	0.0242
2	*Berula erecta*	47	0.02
	Veronica anagallis-aquatica	46.4	0.011
	Elodea canadensis	36	0.0268
3	*Brachythecium rivulare*	63.6	0.0004
	Cratoneuron filicinum	54.5	0.0004
	Pellia endiviifolia	51.3	0.0018
	Conocephalum conicum	45.5	0.0012
	Glyceria notata	41	0.0048
	Amblystegium riparium	34.7	0.0154
	Rhynchostegium riparioides	34.5	0.0154
	Plagiomnium undulatum	27.3	0.0186

Table 6: Relative abundance of macrophyte species within the different TWINSPAN groups of the sandy lowland brooks (R-C1)
Only species occurring with >2% relative abundance in any of the subtypes are listed.

Species	TwSp-Group 1 (%)	TwSp-Group 2 (%)	TwSp-Group 3 (%)
Agrostis stolonifera	2.1	0.8	4.1
Amblystegium fluviatile	0.0	0.0	2.6
Amblystegium riparium	0.0	0.0	4.7
Berula erecta	0.6	18.4	11.9
Brachythecium rivulare	0.0	0.0	5.6
Callitriche hamulata	6.8	0.0	0.0
Callitriche platycarpa	10.7	0.0	0.0
Cardamine amara	0.0	1.3	4.3
Cratoneuron filicinum	0.0	0.0	4.9
Eleocharis acicularis	5.1	0.0	0.0
Equisetum fluviatile	0.1	2.1	0.1
Fontinalis antipyretica	0.0	1.1	6.9
Glyceria aquatica	4.3	3.9	0.7
Glyceria fluitans	6.4	4.0	6.0
Glyceria notata	0.0	0.5	8.3
Juncus bulbosus	2.1	0.0	0.0
Lemna gibba	2.8	0.0	0.0
Lemna minor	2.1	5.5	1.2
Luronium natans	2.1	0.0	0.0
Mentha aquatica	0.0	4.3	7.2
Myosotis scorpioides	3.4	1.0	2.4
Myriophyllum alterniflorum	2.1	0.2	0.0
Nasturtium officinale	0.0	0.1	7.4
Oenanthe aquatica	2.1	0.0	0.0
Phalaris arundinacea	2.8	3.6	0.7
Polygonum hydropiper	2.9	1.3	0.0
Potamogeton alpinus	2.1	5.7	0.0
Potamogeton natans	14.4	0.0	0.0
Potamogeton trichoides	2.8	0.0	0.0
Ranunculus peltatus	2.8	1.4	0.0
Rhynchostegium riparioides	0.0	0.0	4.1
Scirpus sylvaticus	0.0	2.5	0.8
Sparganium emersum	10.0	5.3	0.1
Sparganium erectum	2.2	3.1	0.0
Sphagnum denticulatum	2.1	0.0	0.0
Veronica anagallis-aquatica	0.0	4.4	0.3
Veronica beccabunga	0.0	2.3	6.1

In Figure 7 we see that TWINSPAN group 1 is clearly associated with lower lying, deeper and lower gradient rivers than subtype variants 2 and 3, the latter in particular. Accordingly we would expect a difference between biological types in relation to pressure variables. This is confirmed here in relation to ammonia, nitrate and artificial land cover. There is no clear difference between types in relation to the common metric, although, in this case, the relative number of moss taxa, may be a useful additional discriminator.

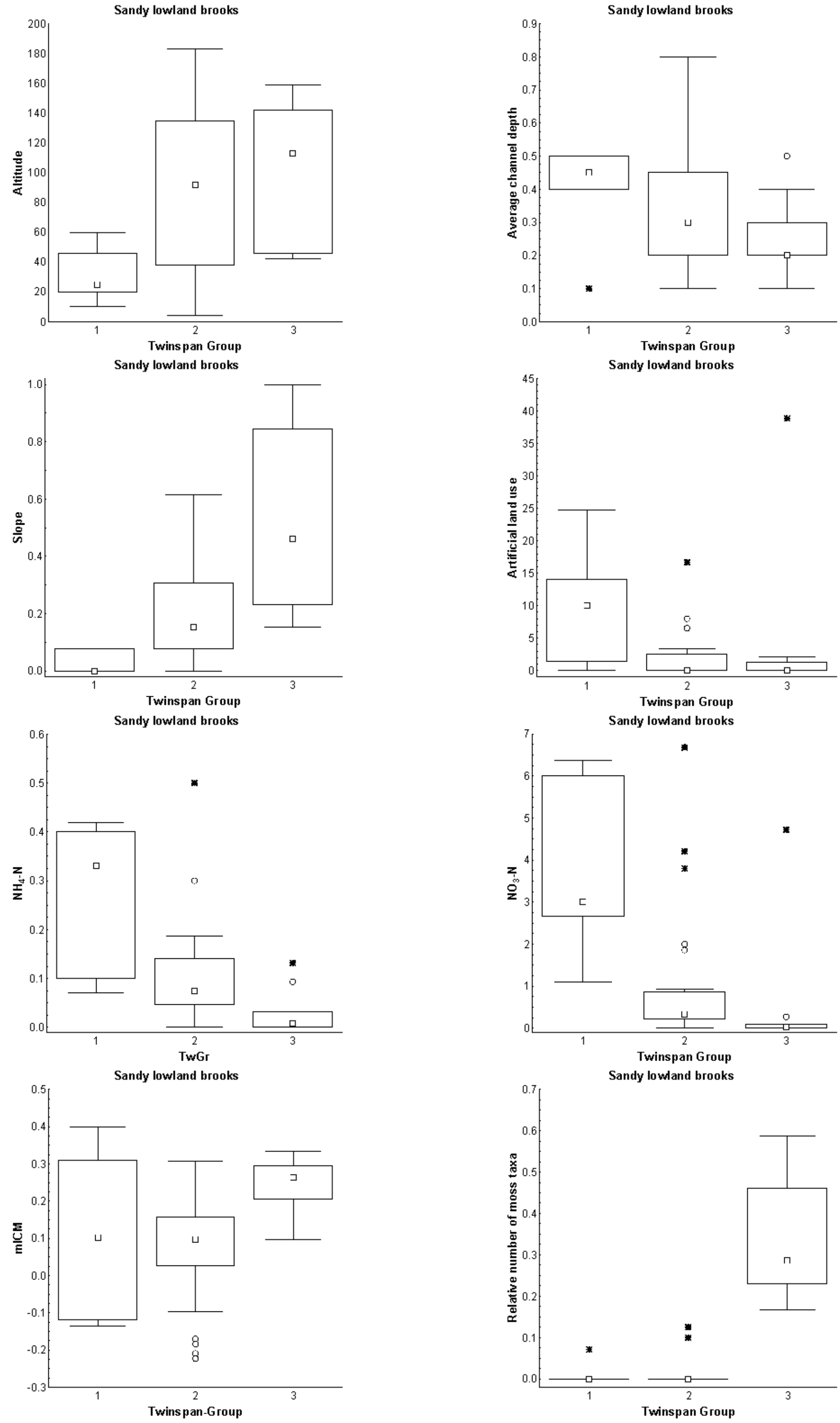


Figure 7: Selected environmental variables showing significant differences among subtype variants of the sandy lowland brooks (R-C1)

4.2.2 Subtype variants of the siliceous mountain brooks (R-C3)

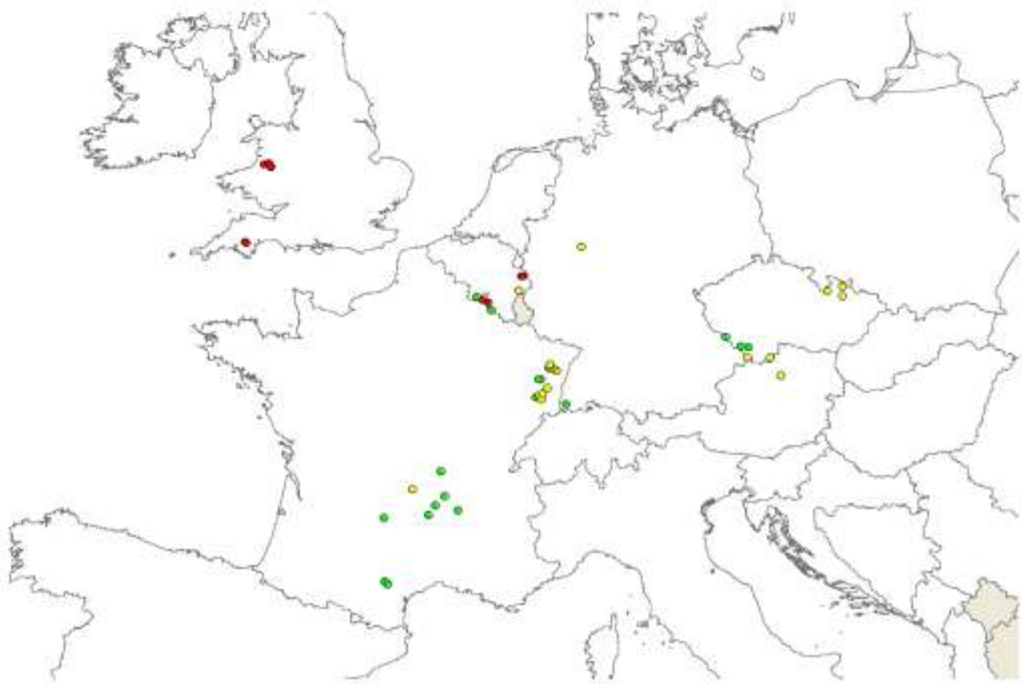


Figure 8: Location of benchmark sites for the siliceous mountain brooks (red: *Pellia-Racomitrium-Hyocomium* variant, green: *Rhynchostegium-Fissidens-Amblystegium* variant, yellow: *Brachythecium rivulare* variant)

Table 7: Significant indicator species of the TWINSPAN groups for the siliceous mountain brooks (IndVal analysis)

TwSp-Group	Species	IndVal (%)	p-level
3	*Pellia epiphylla*	70.3	0.0002
	Racomitrium aciculare	65.6	0.0002
	Hyocomium armoricum	63.6	0.0002
	Marsupella emarginata	54.5	0.0002
	Ranunculus flammula	45.5	0.0012
	Fontinalis squamosa	38.6	0.0102
	Hygrohypnum luridum	36.4	0.0012
	Oenanthe crocata	36.4	0.0016
	Thamnobryum alopecurum	33.7	0.007
	Jungermannia atrovirens	18.2	0.0498
	Sphagnum sp.	18.2	0.0462
4	*Rhynchostegium riparioides*	52.9	0.0006
	Fissidens crassipes	32.1	0.008
	Amblystegium fluviatile	26.4	0.0312
5	*Brachythecium rivulare*	59.2	0.0002
	Petasites hybridus	23.8	0.0244

Table 8: Relative abundance of macrophyte species within the different TWINSPAN groups of the siliceous mountain brooks (R-C3)
Only species occurring with >2% relative abundance in any of the subtypes are listed.

Species	TwSp-Group 3	TwSp-Group 4	TwSp-Group 5
Brachythecium plumosum	5.0	0.0	1.6
Callitriche hamulata	2.2	3.4	0.0
Fontinalis squamosa	24.2	7.9	0.0
Hygrohypnum luridum	4.7	0.0	0.0
Hygrohypnum ochraceum	5.4	3.7	4.8
Hyocomium armoricum	7.1	0.0	0.0
Marsupella emarginata	3.5	0.0	0.0
Myriophyllum alterniflorum	9.4	0.0	0.0
Pellia epiphylla	5.3	0.0	0.8
Petasites hybridus	0.0	0.0	5.1
Racomitrium aciculare	5.9	0.1	2.4
Racomitrium aquaticum	2.0	0.0	0.0
Scapania undulata	16.4	23.0	16.7
Thamnobryum alopecurum	3.4	0.0	0.8

In Figure 9 we see that TWINSPAN group 3, although found at lower elevations than subtypes variants 4 and 5, is associated with distinctly more Atlantic conditions. It has lower levels of ammonia and phosphate than types 4 and 5. This might be partly a function of the more base-rich catchment geology of the continental sub-types although we cannot establish this from the data available. Subtype 3 is generally more species-rich than the other two subtypes, and, being naturally rich in a number of high scoring bryophyte species, has a higher mICM score than the other two types.

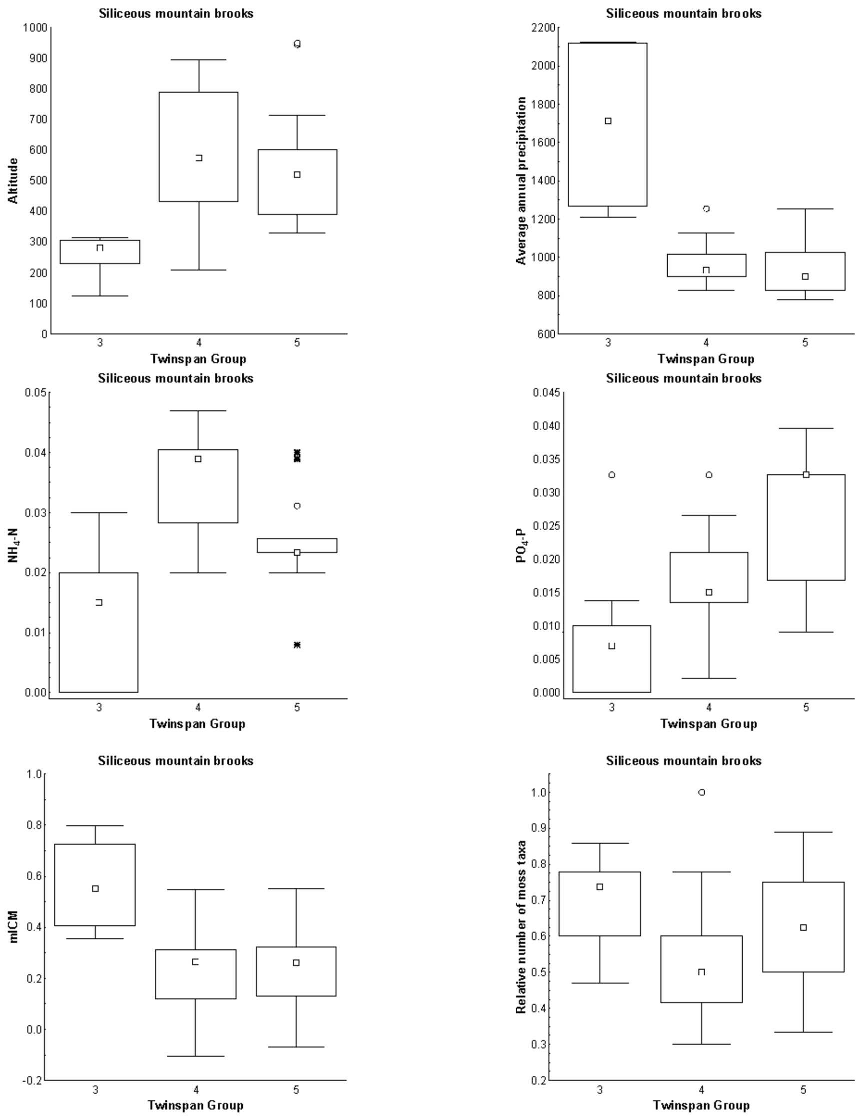


Figure 9: Selected environmental variables showing significant differences among subtype variants of the siliceous mountain brooks (R-C3)

4.2.3 Subtype variants of the medium-sized lowland streams (R-C4)

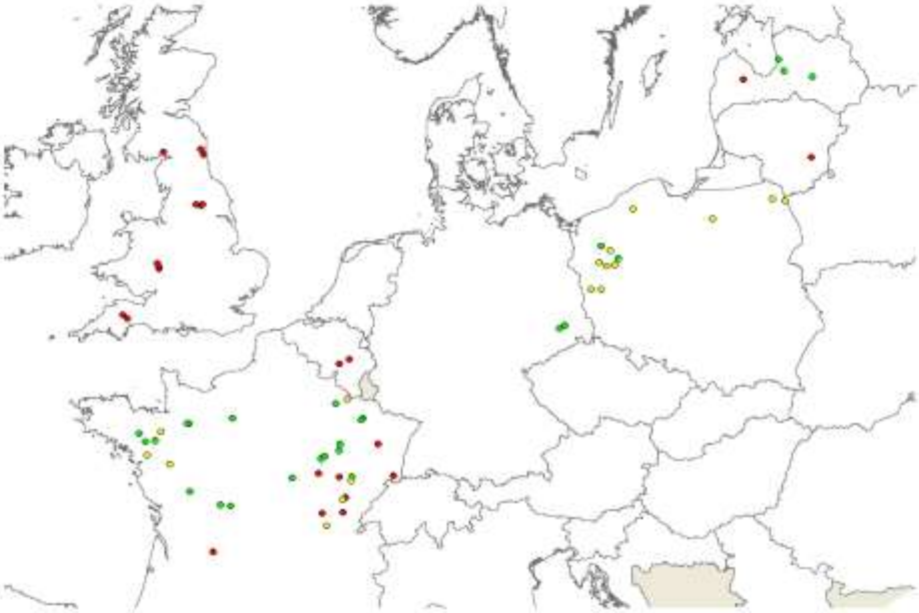


Figure 10: Location of benchmark sites for the medium-sized lowland streams (red: *Rhynchostegium-Amblystegium fluviatile* variant, green: *Fontinalis-Sparganium emersum* variant, yellow: *Nuphar-Phragmites-Lemna* variant)

Table 9: Significant indicator species of the TWINSPAN groups for the medium-sized lowland streams (IndVal analysis)

TwSp-Group	Species	IndVal (%)	p-level
1	*Rhynchostegium riparioides*	54.1	0.0002
	Amblystegium fluviatile	31.3	0.002
	Fissidens crassipes	23.7	0.0356
	Oenanthe crocata	19.1	0.0194
2	*Fontinalis antipyretica*	41.9	0.0064
	Amblystegium riparium	37.9	0.0036
	Sparganium emersum	29.6	0.0492
	Veronica anagallis-aquatica	28.1	0.0132
	Callitriche platycarpa	24.1	0.008
	Apium nodiflorum	22.7	0.018
	Octodiceras fontanum	20.8	0.0252
3	*Nuphar lutea*	52.4	0.0002
	Phragmites australis	43.1	0.0002
	Berula erecta	36.2	0.0004
	Lemna minor	33.4	0.0082
	Hydrocharis morsus-ranae	33.3	0.0002
	Alisma plantago-aquatica	33.2	0.0006
	Agrostis stolonifera	32.6	0.0048
	Solanum dulcamara	29.6	0.0218
	Iris pseudacorus	25.2	0.0442
	Glyceria aquatica	23.2	0.006
	Rumex hydrolapathum	22.6	0.0084
	Carex rostrata	22.2	0.0042
	Equisetum fluviatile	19	0.0138
	Ceratophyllum demersum	18.1	0.0126
	Potamogeton pectinatus	17.3	0.0322
	Epilobium hirsutum	16.7	0.0168

Table 10: Relative abundance of macrophyte species within the different TWINSPAN groups of the medium-sized lowland streams (R-C4)
Only species occurring with >2% relative abundance in any of the subtypes are listed.

Species	TwSp-Group 1	TwSp-Group 2	TwSp-Group 3
Agrostis stolonifera	0.1	0.6	2.3
Amblystegium riparium	4.3	6.2	0.7
Apium nodiflorum	0.5	2.8	0.0
Berula erecta	0.0	2.9	6.1
Ceratophyllum demersum	0.4	0.2	3.9
Cinclidotus riparius	3.5	1.7	0.0
Elodea canadensis	2.3	0.9	7.2
Fissidens crassipes	5.5	1.0	0.0
Fontinalis antipyretica	13.9	10.8	2.5
Glyceria aquatica	0.1	0.4	5.7
Lemna minor	0.2	2.1	1.3
Myriophyllum spicatum	6.5	2.5	5.9
Nuphar lutea	0.1	1.3	8.4
Octodiceras fontanum	1.5	7.7	0.0
Phalaris arundinacea	6.5	7.9	4.1
Phragmites australis	0.1	0.0	7.4
Potamogeton pectinatus	0.4	0.2	3.9
Potamogeton perfoliatus	4.3	1.0	1.1
Ranunculus fluitans	3.5	4.9	1.6
Rhynchostegium riparioides	16.1	2.0	0.0
Sparganium emersum	4.4	7.6	4.2
Sparganium erectum	4.1	1.1	1.2
Stratiotes aloides	0.0	0.0	2.3
Veronica anagallis-aquatica	0.4	2.3	0.7

There is little difference in channel depth between biological subtypes but in terms of substrate, finer material clearly dominates in TWINSPAN type 3 and there is an apparent gradient of declining oceanicity through types 1 to 3. In terms of pressures there are no consistent differences between the types with respect to land cover or nitrate, but for all parameters values are elevated above levels indicative of no-impact. Common metric values tend to be lower in TWINSPAN group 3 reflecting the lower incidence of high scoring moss taxa which are noticeably more prominent in types 1 and 2, in line with their coarser bed characteristics (Figure 11).

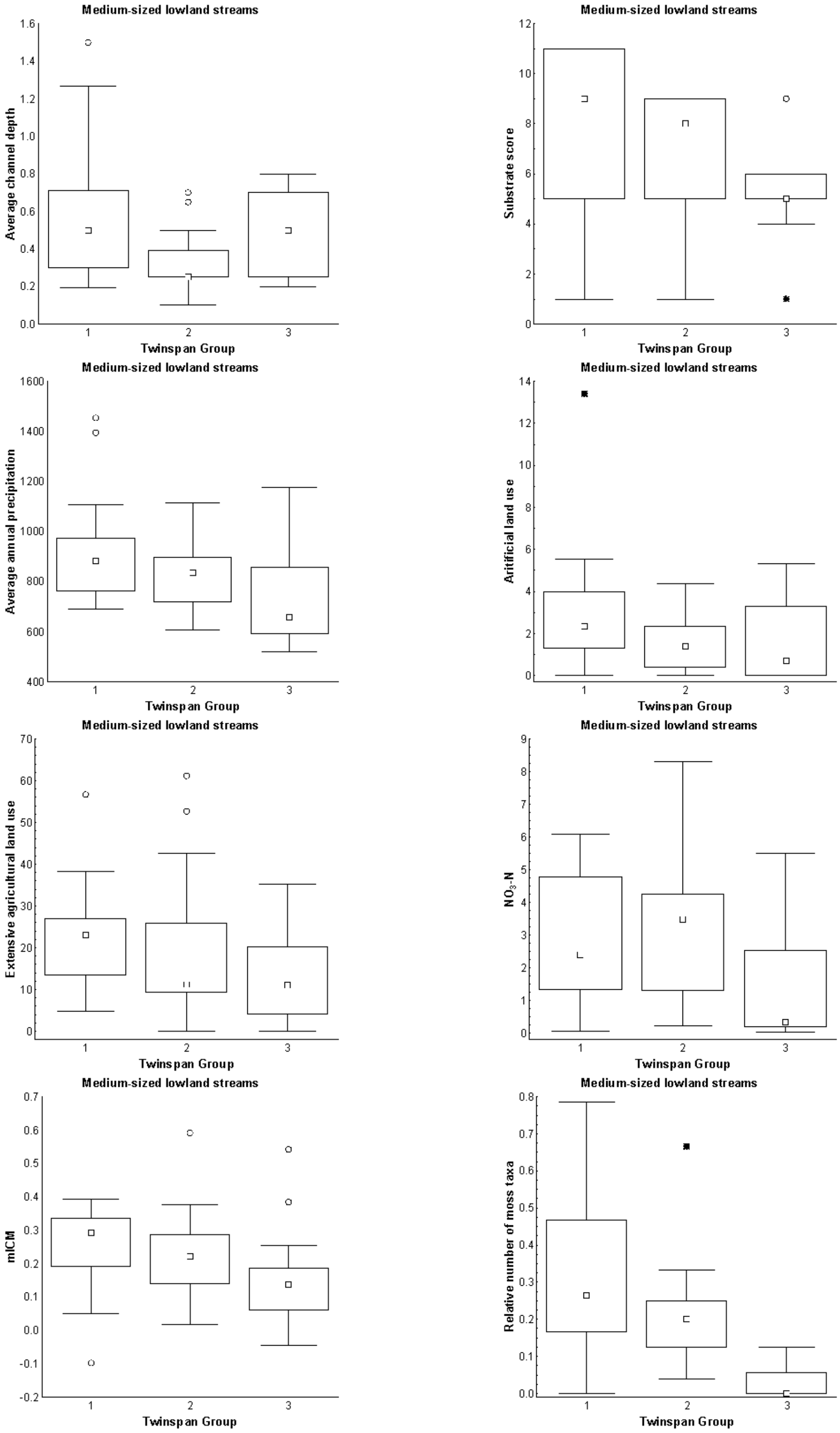


Figure 11: Selected environmental variables showing significant differences among subtype variants of the medium-sized lowland streams (R-C4)

5 References

Bennett, C., Owen, R., Birk, S., Buffagni, A., Erba, S., Mengin, N. et al. (2010). Bringing European river quality into line: an exercise to intercalibrate macro-invertebrate classification methods. Hydrobiologia (submitted).

Birk, S. & Hering, D. (2009). A new procedure for comparing class boundaries of biological assessment methods: A case study from the Danube Basin. Ecological Indicators, 9(3), 528-539.

Birk, S. & Willby, N. J. (2010). Towards harmonization of ecological quality classification: establishing common grounds in European macrophyte assessment for rivers. Hydrobiologia (accepted).

Birk, S., van Kouwen, L. & Willby, N. J. (in prep.). Harmonising large river bioassessment: How to define alternative benchmarks in the absence of pristine reference conditions? A case study from the Danube River.

Birk, S., Willby, N. J., Chauvin, C., Coops, H. C., Denys, L., Galoux, D. et al. (2007). Report on the Central Baltic River GIG Macrophyte Intercalibration Exercise, June 2007.

Böhmer, J., Rawer-Jost, C., Zenker, A., Meier, C., Feld, C. K., Biss, R., et al. (2004). Assessing streams in Germany with benthic invertebrates: Development of a multimetric invertebrate based assessment system, 34(4), 416-432.

Dufrene, M. & Legendre, P. (1997). Species assemblages and indicator species: The need for a flexible asymmetrical approach, 67(3), 345-366.

EN 14184:2003. Water quality - Guidance standard for the surveying of aquatic macrophytes in running waters.

Furse, M., Hering, D., Moog, O., Verdonschot, P., Johnson, R. K., Brabec, K. et al. (2006). The STAR project: context, objectives and approaches. Hydrobiologia, 566, 3-29.

Google Inc. (2010). Google earth (Version 5.1.3535.3218) [Software].

Hill, M. O. (1979). TWINSPAN: A Fortran Program for Arranging Multivariate Data in an Ordered Two-way Table by Classification of the Individuals and Attributes. Section of Ecology and Systematics, Cornell University. Ithaca, NY.

Kelly, M. G., Bennett, C., Coste, M., Delgado, C., Delmas, F., Denys, L. et al. (2009). A comparison of national approaches to setting ecological status boundaries in phytobenthos assessment for the European Water Framework Directive: results of an intercalibration exercise. Hydrobiologia, 621, 169-182.

Mitchell, T. D., Carter, T. R., Jones, P. D., Hulme, M. & New, M. (2004). A comprehensive set of high-resolution grids of monthly climate for Europe and the globe : the observed record (1901-2000) and 16 scenarios (2001-2100). Tyndall Centre for Climate Change Research and School of Environmental Sciences, University of East Anglia. Norwich.

Schmedtje, U., Birk, S., Poikane, S., van De Bund, W. & Bonne, W. (2010). Guidance document on the intercalibration process 2008-2011. Guidance Document No. 14. Implementation Strategy for the Water Framework Directive (2000/60/EC).

ter Braak, C. J. (1988). Partial canonical correspondence analysis. In H. H. Bock, Classification and related methods of data analysis (pp. 551-558). Amsterdam: North-Holland.

van de Bund, W. (2009). Water Framework Directive intercalibration technical report - Part 1: Rivers. Office for Official Publications of the European Communities. Luxembourg.

Wasson, J. G., Chandesris, A., Garcia-Bautista, A., Pella, H. & Villeneuve, B. (2003). European Hydro-Ecoregions. CEMAGREF. Lyon.

Annex

Table A1: Allocation of survey sites to common stream types and TWINSPAN groups (Group 0 = impacted sites)

IC type	TwSp-Group	AT	BE (FL)	BE (WL)	CZ	DE	DK	FR	LT	LV	NL	PL	SE	UK (GB)	Sum
	Sum		13			7	3		1	3	4	25			***56***
Sandy lowland	0		8			3					1	2			***14***
brooks	1		4								3				***7***
(R-C1)	2		1			3	3		1	2		14			***24***
	3					1				1		9			***11***
	Sum	5		13	9	12		29						12	***80***
Siliceous mountain	0	3		8	2	10		3						5	***31***
brooks	3			3				1						7	***11***
(R-C3)	4			1	3	1		12							***17***
	5	2		1	4	1		13							***21***
	Sum			2		3		48	1	5		14	1	26	***100***
Medium-sized	0							15		1		1		12	***29***
lowland streams	1			2				8	1	1				14	***26***
(R-C4)	2					3		18		3		3			***27***
	3							7				10	1		***18***
Total Sum		***5***	***13***	***15***	***9***	***21***	***3***	***77***	***2***	***8***	***4***	***39***	***1***	***38***	**236**

Table A2: Differences in environmental parameter values between benchmark and impacted sites (Mann-Whitney U Test) (n.s. = not significant)

IC Type	Parameter	#Benchmark	#Impact	p-level
Sandy lowland brooks (R-C1)	Altitude	42	14	n.s.
	Artificial land use	41	14	n.s.
	Average annual air temperature	42	14	0.001
	Average annual precipitation	42	14	0.027
	Average channel depth	41	13	n.s.
	Average channel width	42	14	0.007
	BOD_5	29	6	n.s.
	Catchment size	37	12	n.s.
	Conductivity	38	14	n.s.
	Dominant channel substrate	42	14	n.s.
	Extensive agricultural land use	41	14	0.004
	Intensive agricultural land use	41	14	n.s.
	Land use at site	42	14	0.009
	Latitude	42	14	0.000
	Longitude	42	14	0.001
	Natural land use	41	14	n.s.
	NH_4-N	37	13	0.000
	NO_3-N	40	12	0.001
	PO_4-P	39	12	0.002
	Shading	42	14	0.017
	Slope	42	14	0.018
Siliceous mountain brooks (R-C3)	Altitude	49	31	0.021
	Artificial land use	45	31	0.000
	Average annual air temperature	49	31	n.s.
	Average annual precipitation	49	31	n.s.
	Average channel depth	49	31	n.s.
	Average channel width	49	31	0.005
	BOD_5	42	25	n.s.
	Catchment size	49	31	0.000
	Conductivity	40	20	0.000
	Dominant channel substrate	48	29	n.s.
	Extensive agricultural land use	45	31	0.021
	Intensive agricultural land use	45	31	0.002
	Land use at site	49	31	n.s.
	Latitude	49	31	0.000
	Longitude	49	31	n.s.
	Natural land use	45	31	0.000
	NH_4-N	39	26	n.s.
	NO_3-N	42	22	0.000
	PO_4-P	42	26	n.s.
	Shading	49	31	0.026
	Slope	49	31	0.012
Medium-sized lowland streams (R-C4)	Altitude	71	29	0.004
	Artificial land use	67	27	0.008
	Average annual air temperature	71	29	n.s.
	Average annual precipitation	71	29	n.s.
	Average channel depth	70	26	n.s.
	Average channel width	68	26	n.s.
	BOD_5	62	27	0.000
	Catchment size	71	29	n.s.
	Conductivity	61	25	n.s.
	Dominant channel substrate	71	29	n.s.
	Extensive agricultural land use	67	27	n.s.
	Intensive agricultural land use	67	27	0.001
	Land use at site	71	29	n.s.
	Latitude	71	29	n.s.
	Longitude	71	29	0.003
	Natural land use	67	27	0.001
	NH_4-N	62	27	0.000
	NO_3-N	64	28	0.047
	PO_4-P	62	28	0.000
	Shading	71	29	n.s.
	Slope	71	29	n.s.

Figure A1: Benchmark sites of the sandy lowland brooks (R-C1), *Berula erecta* subtype variant. Upper: Omulew (PL) (photo: Google Earth); Lower: Furlbach (DE) (photo: S. Birk)

Figure A2: Benchmark sites of the medium-sized lowland streams (R-C4). Upper: Loue (FR), *Rhynchostegium-Amblystegium fluviatile* subtype variant; Middle: Veseta (LV), *Fontinalis-Sparganium emersum* subtype variant; Lower: Pasłęka (PL), *Nuphar-Phragmites-Lemna* subtype variant (all photos: Google Earth)

Figure A3: Benchmark sites of the siliceous mountain brooks (R-C3). Upper: Cherry Brook (UK) , *Pellia-Racomitrium-Hyocomium* subtype variant; Lower: Blanice (CZ), *Rhynchostegium-Fissidens-Amblystegium* subtype variant (all photos: Google Earth)

Anhang 6: Ergebnisprotokoll – CBrivGIG Workshop Bordeaux, 27. und 28. Mai 2010

Resolution minutes

7th Central-Baltic Rivers GIG Macrophyte Intercalibration Meeting

27-28 May 2010

Venue: Cemagref Bordeaux, 50 Avenue de Verdun, Cestas
Beginning: 27 May 2010, 09:30 h; Ending: 28 May 2010, 12:00 h

Participants:
Roelf Pot (NL), Daniel Galoux (WL), Francisca Aguiar (PT), Nigel Willby (GB), Christian Chauvin (FR), Sebastian Birk (DE), Aive Kõrs (EE), Libuse Opatrilova (CZ), Maria Rita Minciardi (IT), Nora Welschbillig (LU), Krzysztof Szoszkiewicz (PL), Karin Pall (AT), Vincent Bertrin (FR), Fany Roussel (FR)

A. New Member States to join the intercalibration exercise / additional intercalibration types

Italy has recently developed its national assessment methods for river macrophytes and will join the intercalibration exercise for the sandy lowland rivers (R-C1). Its participation in the other stream types will also be evaluated in the coming weeks.

Poland finalised the national assessment for the siliceous mountain brooks and will contribute to the final intercalibration for this type.

Luxembourg is currently evaluating its participation for the common types R-C3 (siliceous mountain brooks) and R-C4 (medium-sized lowland streams).

B. Common stream types to be intercalibrated in the second round

So far, the we concentrated our efforts on the intercalibration of the common stream types R-C1 (sandy lowland brooks), R-C3 (siliceous mountain brooks) and R-C4 (medium-sized lowland streams). There will be no intercalibration of the remaining common stream types in the CB_{riv}GIG within the second round of intercalibration.

C. Alternative benchmarking in the intercalibration of assessment methods using river macrophytes

The group agreed to use the approach of alternative benchmarking in intercalibration. The results of the recent benchmark analysis were discussed (see report). We concluded that the outcomes are relevant and useful. However, the level of detail resulting from the analysis is not always required. The final outcomes must relate to the actual number of national stream types to be included in the intercalibration of the common type (e.g., it is neither necessary nor desired to define more intercalibration benchmarks than national stream types used). Countries not represented in the benchmark data set are asked to assign their national types to the international benchmarks.

The final analysis on alternative benchmarks will be conducted with an amended data set including "best available" sites from Italy and Poland. Additional sites for the sandy lowland brooks (R-C1) will possibly be provided by Great Britain to supplement the data basis for the Western European subtype variant currently shared by Flanders and the Netherlands.

D. Technical procedure for boundary comparison

First results of applying the new comparability criteria on boundary comparison were presented for the siliceous mountain brooks. Nigel presented the different steps of the procedure. The group recognised the complex structure of the requirements for boundary comparison.

The effect of benchmark normalisation based on alternative benchmarking showed reasonable outcomes and allowed the British method to be in line with the reference expectations of the other countries. Although the outcomes provided first insights in the relative positions of national boundaries, we concluded to wait with clear recommendations for any harmonisation of boundaries until the analyses were carried out with the final methods' versions.

E. Differences in national assessment concepts

The difference between concepts of national methods was discussed as a result of the questions about the feasibility checks in the WISER questionnaire (IC Milestone 2 Report). Two main concepts were distinguished: weighted averaging of indicator species values on the one hand, and distance between species composition in reference situation and in actual situation on the other hand. It was argued that the first concept is suitable for measuring the impact of eutrophication, the second is appropriate for measuring the impact of hydromorphological deterioration. It was concluded that much development is still needed to cope with these differences and that at best two separate sub-metric are to be developed with diagnostic features. However, intercalibration is only about comparing the results.

F. Support in the reporting obligations

The final intercalibration report in the second round is due in June 2011. Member States were asked to support the drafting activity. Poland and the Netherlands indicated their willingness to contribute but need to ask for national funding to guarantee proper collaboration.

G. Next intercalibration meeting

The next meeting of the $CB_{riv}GIG$ macrophyte group will take place in spring 2011. Maria Rita Minciardi will evaluate the possibility to host this meeting in Turin (IT).

H. Actions

- All countries participating in the intercalibration exercise to deliver any updates or changes of their national method by **November 1st, 2010**. Later amendments cannot be considered for the second round of intercalibration.
- Maria Rita Minciardi and Krzysztof Szoszkiewicz to provide new data (macrophyte surveys and supporting environmental information) to include their methods in the intercalibration exercise. **Deadline: End of June 2010.**
- Nora Welschbillig to clarify the participation of Luxembourg in the intercalibration exercise. This includes evaluating the appropriate alternative benchmarks used for the rivers of Luxembourg since no benchmark sites are available. **Deadline: End of June 2010.**
- Roelf Pot to explore the relevance of the elevated nutrient concentrations on the Western European subtype variant of the sandy lowland brooks. The data basis will possible extended by sites delivered by Nigel Willby. **Deadline: End of July 2010.**
- Roelf Pot and Krzysztof Szoszkiewicz to evaluate the possibilities to support in the writing of the final intercalibration report 2011. **Deadline: Spring 2011** (next macrophyte meeting).

Anhang 7: Kopie der Milestone 3 Berichterstattung des X-GIG Große Fließgewässer

EUROPEAN COMMISSION
DIRECTORATE GENERAL JRC
JOINT RESEARCH CENTRE
Institute of Environment and Sustainability

WFD Intercalibration Phase 2: Milestone 3 report

Water category/GIG/BQE/ horizontal activity:	**XGIG Large Rivers**
Information provided by:	Franz Schöll, Sebastian Birk

1. Organisation

1.1. Responsibilities

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

Franz Schöll (DE, lead)
Sebastian Birk (DE)
Jürgen Böhmer (DE)
Wim Gabriels (BE-FL)
Kris van Looy (BE-FL)
Christine Keulen (BE-WL)
Jukka Aroviita (FI)
Stina Drakare (SE)
Joakim Lücke (SE)
Ana Lara Romero (ES)
Nuno Caiola (ES)
Carles Ibanez (ES)
Yves Souchon (FR)
Béla Csányi (HU)
Denisa Nemejcova (CZ)
Libuse Opatrilova (CZ)
Alena Slavikova (CZ)
Franz Wagner (AT)
Gisela Ofenböck (AT)
Richard Hemsworth (UK)
Rachel Benstead (UK)
Simone Ciadamidaro (IT)
Andrea Buffagni (IT)
Audrone Pumputyte (LT)
Gorazd Urbanic (SI)
Matus Haviar (SK)
Emilia Misikova Elexova (SK)
Roel Knoben (NL)
Eddy Lammens (NL)
Irja Truumaa (EE)
Steinar Sandøy (NO)
Elin M. Lien (NO)
Ann Kristin Schartau (NO)
Joao Ferreira (PT)
Samantha Jane Hughes (PT)
Maria Teresa Ferreira (PT)

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:

Germany, Austria, Belgium, Netherlands, Spain, United Kingdom, Estonia, Lithuania, Finland, Sweden, Italy, Slovenia, Slovakia, Czech Republic

The participation of France is highly appreciated, but no delegates were sent to the last workshops.

1.3. Meetings

List the meetings of the group:

First workshop – Koblenz 22-23 Sept. 2009
Second workshop – Koblenz 19-20 April 2010
Third workshop – Koblenz 22-23 Sept. 2010

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

1. finalized formally agreed national method,
2. intercalibratable finalized method,
3. method under development,
4. no method developed

BQE	Member State	Method	Status
Benthic Invertebrates	Austria	Assessment of the biological quality elements - part benthic invertebrates	1
	Belgium (Flanders)	Multimetric Macroinvertebrate Index Flanders	1
	Belgium (Wallonia)	Global biological index adapted to large watercourses and deep rivers	1
	Estonia	Estimation of freshwater quality using macroinvertebrates	1
	Finland	Finnish multimetric index	1
	Germany	Assessment method for rivers using benthic invertebrates	1
	Lithuania	Assessment system for rivers using macrozoobenthos indicators (Danish Stream Fauna Index)	1
	Netherlands	WFD-metrics for natural water types	1
	Slovakia	Slovak assessment of benthic invertebrates in large rivers	1
	Slovenia	Slovenian assessment system for rivers using benthic invertebrates	1
	Spain	Iberian Biological Monitoring Working Party	1
	United Kingdom	River Invertebrate Classification Tool	1
Fish Fauna	Austria	Fish Index Austria	1
	Belgium (Flanders)	Flemish Index of Biotic Integrity	1
	Belgium (Wallonia)	Biological Index for Fish integrity	1
	Finland	Finnish River Fish Index	1
	Germany	Fish-based Assessment System	1
	Lithuania	Assessment method of rivers using Lithuanian fish index	1
	Netherlands	WFD-metrics for natural water types	1
	Slovakia	Fish Index of Slovakia	1
	Spain	Index of Biotic Integrity using fish as indicators of the Ecological Status of Spanish Rivers	1
Macrophytes	Austria	Assessment of the biological quality elements - part macrophytes	1
	Belgium (Flanders)	Flemish macrophyte assessment system	1
	Germany	German Assessment system for Macrophytes and Phytobenthos according to the WFD	1
	Netherlands	WFD-metrics for natural water types	1
	Slovakia	Slovak assessment of macrophytes in rivers	1
	Slovenia	River Macrophyte Index	1
	United Kingdom	Ecological Classification of Rivers using Macrophytes	1

BQE	Member State	Method	Status
Phytobenthos	Austria	Assessment of the biological quality elements - part phytobenthos	1
	Belgium (Flanders)	Proportions of Impact-Sensitive and Impact-Associated Diatoms	1
	Belgium (Wallonia)	Pollution Sensitivity Index	1
	Estonia	Assessment system for rivers using phytobenthos in Estonia	1
	Finland	Indice de Polluosensibilité Spécifique (Specific Pollution Sensitivity Index SPI)	1
	Germany	German assessment system for Macrophytes and Phytobenthos according to the WFD	1
	Hungary	Improvement of the Hungarian ecological water qualification system - Phytobenthos in Rivers	1
	Netherlands	WFD-metrics for natural water types	1
	Slovakia	Slovak assessment of phytobenthos in rivers	1
	Slovenia	Ecological status assessment system for rivers using phytobenthos	1
	Spain	Pollution Sensitivity Index	1
	United Kingdom	WFD River Diatom method or Trophic Diatom Index version 3 Method	1
Phytoplankton	Belgium (Flanders)	German phytoplankton assessment method for rivers	1
	Germany	Index Phytoplankton PhytoFluss	1
	Hungary	Hungarian River Phytoplankton Index	1
	Slovakia	Slovak assessment of phytoplankton in rivers	1

Make sure that the **national method descriptions** meet the level of detail required to fill in the table 1 at the end of this document !

3. Checking of compliance of national assessment methods with the WFD requirements (April 2010 + update in October 2010)

Do all national assessment methods meet the requirements of the Water Framework Directive? (Question 1 in the IC guidance)
Do the good ecological status boundaries of the national methods comply with the WFD normative definitions? (Question 7 in the IC guidance)

List the WFD compliance criteria and describe the WFD compliance checking process and results (the table below lists the criteria from the IC guidance, please add more criteria if needed)

Compliance criteria	Compliance checking conclusions
1. Ecological status is classified by one of **five classes** (high, good, moderate, poor and bad).	All methods are compliant.
2. High, good and moderate ecological status are set in line with the WFD's **normative definitions (Boundary setting procedure)**	All methods are compliant.
3. **All relevant parameters** indicative of the biological quality element are covered (see Table 1 in the IC Guidance). A **combination rule** to combine parameter assessment into BQE assessment has to be defined. If parameters are missing, Member States need to demonstrate that the method is sufficiently indicative of the status of the QE as a whole.	All methods are compliant except for: ***Benthic Invertebrates*** - Austria: parameters *richness* and *diversity* are not covered. - Spain: parameter *abundance* is not covered. - United Kingdom: parameter *abundance* is not covered[1]. ***Macrophytes & Phytobenthos***

[1] However, the method used was intercalibrated in Phase 1 and is therefore considered as sufficiently indicative of the status of the QE as a whole.

	- Component *macrophytes* is not assessed by Belgium (Wallonia), Estonia, Finland, Hungary, Spain ***Fish fauna*** - Parameter *age structure* is not covered by Lithuania and the Netherlands ***Phytoplankton*** - Planktonic blooms are not considered by Belgium (Flanders), Germany, Hungary
4. 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	All methods are compliant.
5. The water body is assessed against **type-specific near-natural reference conditions**	All methods are compliant.
6. Assessment results are expressed as **EQRs**	All methods are compliant.
7. Sampling procedure allows for **representative** information about water body quality/ ecological status **in space and time**	All methods are compliant.
8. All data relevant for assessing the biological **parameters** specified in the WFD's normative definitions are covered by the **sampling procedure**	All methods are compliant.
9. Selected taxonomic level achieves adequate **confidence and precision** in classification	All methods are compliant.

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

The national assessment methods for very large rivers generally comply with the criteria stated above. For individual cases certain aspects need to be further analysed. This will be done during the upcoming months by requesting additional data from the Member States.

4. Methods' intercalibration feasibility check

Do all national methods address the same common type(s) and pressure(s), and follow a similar assessment concept? (Question 2 in the IC guidance)

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
Very large rivers (> 10,000 km^2 catchment area)	No specific type characteristics, sub-types will likely be defined within the benchmarking analyses currently carried out.	Austria - yes Belgium - yes Estonia - yes Finland - yes Germany - yes Hungary - yes Lithuania - yes Netherlands - yes Slovakia - yes Slovenia - yes Spain - yes United Kingdom – yes (under review)

What is the outcome of the feasibility evaluation in terms of typology? Are all assessment methods appropriate for the intercalibration water body types, or subtypes?

Method	Appropriate for IC types / subtypes	Remarks
No definitive answer possible - analyses are currently in progress.		
Conclusion No definitive answer possible - analyses are currently in progress.		

4.2. Pressures

Describe the pressures addressed by the MS assessment methods

BQE	Country	Pressure	Remarks
Benthic Invertebrates	Austria	Organic pollution, General degradation	none
	Belgium (Flanders)	Organic pollution, General degradation	
	Belgium (Wallonia)	General degradation	
	Estonia	Organic pollution, General degradation	
	Finland	General degradation	
	Germany	Organic pollution, General degradation	
	Lithuania	General degradation	
	Netherlands	General degradation	
	Slovakia	Organic pollution, General degradation	
	Slovenia	Organic pollution, Hydromorphological impairment	
	Spain	Organic pollution	
	United Kingdom	Organic pollution, General degradation	

<table>
<tr><th>BQE</th><th>Country</th><th>Pressure</th><th>Remarks</th></tr>
<tr><td rowspan="12">Phytobenthos</td><td>Austria</td><td>Eutrophication, Organic pollution, General degradation</td><td rowspan="34">none</td></tr>
<tr><td>Belgium (Flanders)</td><td>Eutrophication, Organic pollution, General degradation</td></tr>
<tr><td>Belgium (Wallonia)</td><td>Eutrophication, Organic pollution, General degradation</td></tr>
<tr><td>Estonia</td><td>Eutrophication</td></tr>
<tr><td>Finland</td><td>Eutrophication, Organic pollution, General degradation</td></tr>
<tr><td>Germany</td><td>Eutrophication, Organic pollution, General degradation</td></tr>
<tr><td>Hungary</td><td>Eutrophication, Organic pollution</td></tr>
<tr><td>Netherlands</td><td>Eutrophication, Organic pollution, General degradation</td></tr>
<tr><td>Slovakia</td><td>Eutrophication, Organic pollution, General degradation</td></tr>
<tr><td>Slovenia</td><td>Eutrophication, Organic pollution</td></tr>
<tr><td>Spain</td><td>Eutrophication</td></tr>
<tr><td>United Kingdom</td><td>Eutrophication, General degradation</td></tr>
<tr><td rowspan="9">Fish Fauna</td><td>Austria</td><td>Flow modification, General degradation, Hydromorphological degradation</td></tr>
<tr><td>Finland</td><td>Eutrophication, General degradation, Hydromorphological degradation</td></tr>
<tr><td>Belgium (Flanders)</td><td>Eutrophication, Flow modification, General degradation, Hydromorphological degradation</td></tr>
<tr><td>Belgium (Wallonia)</td><td>General degradation</td></tr>
<tr><td>Germany</td><td>Eutrophication, Flow modification, General degradation, Hydromorphological degradation, Pollution by organic matter</td></tr>
<tr><td>Lithuania</td><td>General degradation</td></tr>
<tr><td>Netherlands</td><td>Eutrophication, General degradation, Hydromorphological degradation</td></tr>
<tr><td>Slovakia</td><td>Flow modification, General degradation</td></tr>
<tr><td>Spain</td><td>Eutrophication, Flow modification, Hydromorphological degradation</td></tr>
<tr><td rowspan="7">Macrophytes</td><td>Austria</td><td>Eutrophication, Flow modification, General degradation, Hydromorphological degradation</td></tr>
<tr><td>Belgium (Flanders)</td><td>Eutrophication, Flow modification, General degradation, Hydromorphological degradation, Pollution by organic matter</td></tr>
<tr><td>Germany</td><td>Eutrophication, General degradation, Hydromorphological degradation</td></tr>
<tr><td>Netherlands</td><td>Eutrophication, Flow modification, General degradation, Hydromorphological degradation, Pollution by organic matter</td></tr>
<tr><td>Slovakia</td><td>Eutrophication, General degradation, Hydromorphological degradation, Pollution by organic matter</td></tr>
<tr><td>Slovenia</td><td>Eutrophication</td></tr>
<tr><td>United Kingdom</td><td>Eutrophication, Hydromorphological degradation</td></tr>
<tr><td rowspan="4">Phyto-plankton</td><td>Belgium (Flanders)</td><td>Eutrophication</td></tr>
<tr><td>Germany</td><td>Eutrophication, Hydromorphological degradation</td></tr>
<tr><td>Hungary</td><td>Eutrophication, Flow modification, Pollution by organic matter</td></tr>
<tr><td>Slovakia</td><td>Eutrophication, Flow modification, Pollution by organic matter</td></tr>
</table>

Conclusion

An a-priori analysis revealed that all methods address similar pressures. However, due to resource constraints the exercise is only doing in-depths evaluations of invertebrate and phytobenthos methods. A final statement of feasibility for these methods will be done on the basis of data analyses.

4.3. Assessment concept

Do all national methods follow a similar assessment concept?

Examples of assessment concept:

- **Different community characteristics** - structural, functional or physiological - can be used in assessment methods which can render their comparison problematic. For example, sensitive taxa proportion indices vs species composition indices.
- Assessment systems may focus on **different lake zones** - profundal, littoral or sublittoral - and subsequently may not be comparable.
- Additional important issues may be the **assessed habitat type** (soft-bottom sediments versus rocky sediments for benthic fauna assessment methods) or **life forms** (emergent macrophytes versus submersed macrophytes for lake aquatic flora assessment methods)

BQE	Country	Assessment concept	Remarks
Benthic Invertebrates	Austria	Similar assessment concept is followed (see conclusions).	none
	Belgium (Flanders)		
	Belgium (Wallonia)		
	Estonia		
	Finland		
	Germany		
	Lithuania		
	Netherlands		
	Slovakia		
	Slovenia		
	Spain		
	United Kingdom		
Phytobenthos	Austria		
	Belgium (Flanders)		
	Belgium (Wallonia)		
	Estonia		
	Finland		
	Germany		
	Hungary		
	Netherlands		
	Slovakia		
	Slovenia		
	Spain		
	United Kingdom		

Conclusion

The existing national assessment methods acquire their biological data from the main river channel and are based on concepts similar to the assessment of smaller rivers. Although the specific features of large rivers may require alternative, ecologically adapted classifications, the intercalibration exercise deals with the harmonization of the assessment methods that are currently used by the Member States.

5. Collection of IC dataset

Describe data collection within the GIG.
This description aims to safeguard that compiled data are generally similar, so that the IC options can reasonably be applied to the data of the Member States.

Make the following table for each IC common type

Country	BQE	Number of samples	Number of pressure data	Number of physico-chemical data
Austria	Benthic Invertebrates	16	13	13
	Diatoms	16		
Belgium (Flanders)	Benthic Invertebrates	3	6	3
	Diatoms	3		
Belgium (Wallonia)	Benthic Invertebrates	2	2	2
	Diatoms	4		
Czech Republic	Benthic Invertebrates	16	8	7
	Diatoms	24		
Estonia	Benthic Invertebrates	8	8	8
	Diatoms	5		
Finland	-	-	8	8
Germany	Benthic Invertebrates	84	50	37
	Diatoms	38		
Lithuania	Benthic Invertebrates	16	4	4
	Diatoms	4		
Netherlands	Benthic Invertebrates	67	17	40
	Diatoms	11		
Slovakia	Benthic Invertebrates	48	28	24
	Diatoms	40		
Slovenia	Benthic Invertebrates	49	12	12
	Diatoms	26		
Spain	Benthic Invertebrates	3	3	3
	Diatoms	3		
Sweden	-	-	18	18
United Kingdom	Benthic Invertebrates	64	3	9
	Diatoms	6		

List the data acceptance criteria used for the data quality control and describe the data acceptance checking process and results

Data acceptance criteria	Data acceptance checking
Data requirements (obligatory and optional)	Checking is still in progress
The sampling and analytical methodology	
Level of taxonomic precision required and taxalists with codes	
The minimum number of sites / samples per intercalibration type	
Sufficient covering of all relevant quality classes per type	
Other aspects where applicable	

6. Benchmarking: Reference conditions or alternative benchmarking

In section 2 of the method description of the national methods above, an overview has to be included on the derivation of reference conditions for the national methods. In section 6 the checking procedure and derivation of reference conditions or the alternative benchmark at the scale of the common IC type has to be explained to ensure the comparability within the GIG.

Clarify if you have defined

- common reference conditions (N)
- or a common alternative benchmark for intercalibration (Y)

6.1. Reference conditions

Does the intercalibration dataset contain sites in near-natural conditions in a sufficient number to make a statistically reliable estimate? (Question 6 in the IC guidance)

- Summarize the common approach for setting reference conditions (true reference sites or indicative partial reference sites, see Annex III of the IC guidance):

Not applicable

- Give a detailed description of **reference criteria** for screening of sites in near-natural conditions (abiotic characterisation, pressure indicators):

Not applicable

- Identify the **reference sites** for each Member State in each common IC type. Is their number sufficient to make a statistically reliable estimate?

Not applicable

- Explain how you have screened the biological data for impacts caused by pressures not regarded in the reference criteria to make sure that true reference sites are selected:

Not applicable

- Give detailed description of **setting reference conditions** (summary statistics used)

Not applicable

6.2. Alternative benchmarking (only if common dataset <u>does not contain reference sites</u> in a sufficient number)

- Summarize the common approach for setting **alternative benchmark** conditions (describe argumentation of expert judgment, inclusion of modelling)

Most national water bodies show a clumped distribution in "pressure space", i.e. the water bodies at large rivers of a region/country often show similar levels of disturbance instead of covering a broader gradient of degradation (see Figure 1). Preliminary analyses show that benchmarks are predominantly located in north-eastern and eastern Europe, i.e. their geographical distribution is highly skewed. Furthermore, the benchmark water bodies already cover a relevant pressure gradient. These findings disclose that we need to develop an improved concept of alternative benchmarking for the large river intercalibration exercise.

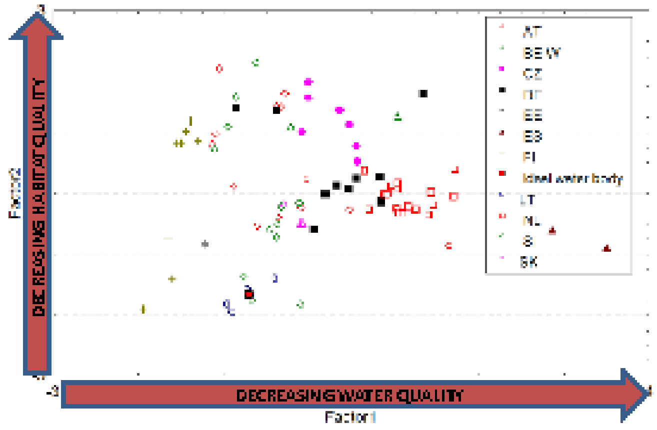


Figure 1: Position of national water bodies in "pressure space" defined by factor analysis of environmental data (Factor 1 related to concentrations of Orthophosphate, Nitrate, Chloride; Factor 2 related to degree of channelization and impoundment)

Improved concepts for alternative benchmarking are currently discussed. The common data basis and strong dose-response relationships for the diatoms, for instance, allow for an extrapolation of reference conditions (Figure 2). This technique will be further investigated for diatoms and benthic invertebrates in the next months.

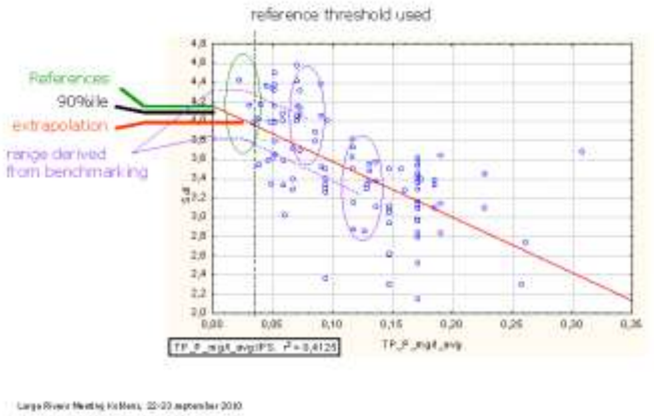


Figure 2: Comparison of different approaches to derive reference conditions using the dose-response relationship between TP and the diatom metric IPS (R2=0.41)

- Give a detailed description of **criteria** for screening of **alternative benchmark** sites (abiotic criteria/pressure indicators that represent a similar low level of impairment to screen for least disturbed conditions)

To be defined

- Identify the **alternative benchmark sites** for each Member State in each common IC type

To be defined

- Describe how you validated the selection of the alternative benchmark with biological data

To be defined

- Give detailed description how you identified the position of the alternative benchmark on the gradient of impact and how the deviation of the **alternative benchmark** from reference conditions has been derived

To be defined

Describe the **biological communities** at reference sites or at the alternative benchmark, considering potential biogeographical differences:

To be defined

7. Design and application of the IC procedure

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

Which IC option did you use?

- IC Option 1 - Same assessment method, same data acquisition, same numerical evaluation (N)
- IC Option 2 - Different data acquisition and numerical evaluation (Y)
- IC Options 3 - Similar data acquisition, but different numerical evaluation (BQE sampling and data processing generally similar, so that all national assessment methods can reasonably be applied to the data of other countries) → supported by the use of common metric(s) (Y)
- Other (specify) (N)

Explanation for the choice of the IC option:

We agreed to test intercalibration option 3 (direct comparison) in the diatom exercise, since national techniques of data acquisition for diatoms are sufficiently comparable. Thus, we will ask the Member States to deliver information about the national diatom typology of large rivers and specifications of the national diatom indices (indicator lists, metric algorithm and combination etc.). The calculation of the national diatom assessments will be done centrally using the common dataset. For invertebrates an intercalibration option 2 is envisaged. However, the national datasets used in intercalibration mostly cover only a narrow quality gradient and comprise few water bodies / monitoring sites. Therefore, we aim at merging those national datasets that were acquired by similar techniques (e.g. comparable sampling effort and level of identification). This should extent the data basis for the subsequent regression analyses. We finally plan to compare the national class boundaries against a common metric using these multi-national datasets of similar data characteristics.

In case of IC Option 2, please explain the differences in data acquisition

Invertebrate methods differ in sampling devices and techniques, sample processing and level of identification.

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

Describe the IC Common metric:

The diatom exercise can operate on the basis of pseudo-common metrics (i.e. average of national EQR excluding the country that is compared). The invertebrate exercise will refer to true (biological) common metrics that will be selected on the basis of the outcomes of further analyses (using the updated common database).

Are all methods reasonably related to the common metric(s)? (Question 5 in the IC guidance)

Please provide the correlation coefficient (r) and the probability (p) for the correlation of each method with the common metric (see Annex V of IC guidance).

Member State/Method	r	p
A	To be analysed	
B		

Explain if any method had to be excluded due to its low correlation with the common metric:

To be analysed

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

Clarify if

- boundaries were set only at national level (Y)
- or if a common boundary setting procedure was worked out at the scale of the common IC type (N)

In section 2 of the method description of the national methods above, an overview has to be included on the boundary setting procedure for the national methods to check compliance with the WFD. In section 8.1 the results of a common boundary setting procedure at the scale of the common IC type should be explained where applicable.

8.1. Description of boundary setting procedure set for the common IC type

Summarize how boundaries were set following the framework of the BSP:

– Provide a description how you applied the full procedure (use of discontinuities, paired metrics, equidistant division of continuum)

To be defined

– Provide pressure-response relationships (describe how the biological quality element changes as the impact of the pressure or pressures on supporting elements increases)

To be defined

– Provide a comparison with WFD Annex V, normative definitions for each QE/ metrics and type

To be defined

8.2. Description of IC type-specific biological communities representing the "borderline" conditions between good and moderate ecological status, considering possible biogeographical differences (as much as possible based on the common dataset and common metrics).

To be defined

8.3. Boundary comparison and harmonization

Describe comparison of national boundaries, using comparability criteria (see Annex V of IC guidance).

To be analysed

- Do all national methods comply with these criteria ? (Y/N)
- If not, describe the adjustment process:

Anhang 8: The WFD intercalibration exercise of assessment methods to classify the ecological status of very large rivers – Overview of national classification schemes

The WFD intercalibration exercise of assessment methods to classify the ecological status of very large rivers – *Overview of national classification schemes*

Sebastian Birk – University of Duisburg-Essen, DE

Version 3
October 2010
including comments of Ema Elexova (SK); Jukka Aroviita (FI), Wim Gabriels (BE-FL)

Table of contents

1 Introduction

The WFD intercalibration exercise

According to the European Water Framework Directive 2000/60/EC (WFD) all Member States of the European Union are obliged to achieve the good status of their surface waters. The assessment of ecological status is a key component of this obligation. To meet this requirement the Member States apply biological assessment methods to classify the status of so-called Biological Quality Elements (BQE). For rivers these are phytoplankton, macrophytes and phytobenthos, benthic invertebrates and fish fauna.

Member States are allowed to use their own methods for surface water monitoring and assessment, therefore a plethora of schemes has been implemented by the countries. To ensure a comparable classification of the good ecological status, the WFD introduces the concept of intercalibration (WFD, Annex V, Section 1.4.1 Comparability of biological monitoring results):

> (ii) In order to ensure comparability of … monitoring systems, the results of the systems operated by each Member State shall be expressed as ecological quality ratios for the purposes of classification of ecological status. These ratios shall represent the relationship between the values of the biological parameters observed for a given body of surface water and the values for these parameters in the reference conditions applicable to that body. The ratio shall be expressed as a numerical value between zero and one, with high ecological status represented by values close to one and bad ecological status by values close to zero.
>
> (iii) Each Member State shall divide the ecological quality ratio scale for their monitoring system for each surface water category into five classes ranging from high to bad ecological status … by assigning a numerical value to each of the boundaries between the classes. The value for the boundary between the classes of high and good status, and the value for the boundary between good and moderate status shall be established through the intercalibration exercise … .
>
> (iv) The [European] Commission shall facilitate this intercalibration exercise in order to ensure that these class boundaries are established consistent with the normative definitions … and are comparable between Member States.

Accomplishing intercalibration is prerequisite for a successful implementation of the WFD as it ensures an equal level of ambition between Member States to maintain and restore the good ecological status. Therefore, many scientists and water managers are currently involved in a Europe-wide intercalibration activity that is carried out for different water categories (i.e. rivers, lakes, coastal and transitional waters) and the various BQE.

Large river intercalibration

In autumn 2009 the Large River Intercalibration Exercise was launched with the aim to harmonise the ecological status classifications applied at large rivers across Europe. For the exercise large rivers were defined as running waters with catchment sizes above 10.000 km^2, i.e. “very large rivers” according to the size typology of the WFD Annex II. Large rivers represent complex ecosystems that feature specific functionality (Vannote et al. 1980), Junk et al. 1989) and a broad diversity of different habitats (Amoros and Bornette 2002). Yet, the multiple functions and structures of these systems have been severely modified, often turning the huge river-floodplain complexes into regulated and embanked waterways (Dynesius & Nilsson 1994).

Two main issues generally challenge the implementation of ecological status classifications for large rivers: the difficulties in data sampling and the lack of reference conditions (Yoder & Kulik 2003, Blocksom & Johnson 2009). The existing national assessment methods acquire their biological data from the main river channel and are based on concepts similar to the assessment of smaller rivers. Although the specific features of large rivers may require alternative, ecologically adapted classifications, the intercalibration exercise deals with the harmonisation of the assessment methods that are currently used by the Member States.

The report at hand provides a summary of the national assessment methods for large rivers and evaluates their WFD compliance regarding the definition of reference conditions, the setting of ecological class boundaries and the coverage of indicative parameters according to the WFD normative definitions.

2 National assessment methods for very large rivers

13 European Member States reported a total number of 43 national assessment methods for the ecological status classification of very large rivers (Table 1). The organism groups of benthic invertebrates, phytobenthos and fish fauna are most frequently assessed, while only few countries use macrophytes or phytoplankton in the quality classification of very large rivers (Figure 1).

Table 1: Overview of national assessment methods to classify the ecological status of very large rivers

BQE	Country	Method name
Benthic Invertebrates	Austria	Assessment of the biological quality elements - part benthic invertebrates
	Belgium (Flanders)	Multimetric Macroinvertebrate Index Flanders
	Belgium (Wallonia)	Global biological index adapted to large watercourses and deep rivers
	Estonia	Estimation of freshwater quality using macroinvertebrates
	Finland	Finnish multimetric index
	Germany	Assessment method for rivers using benthic invertebrates
	Italy	MacrOper, based on STAR_ICM index calculation
	Lithuania	Assessment system for rivers using macrozoobenthos indicators (Danish Stream Fauna Index)
	Netherlands	WFD-metrics for natural water types
	Slovakia	Slovak assessment of benthic invertebrates in large rivers
	Slovenia	Slovenian assessment system for rivers using benthic invertebrates
	Spain	Iberian Biological Monitoring Working Party
	United Kingdom	River Invertebrate Classification Tool
Fish Fauna	Austria	Fish Index Austria
	Belgium (Flanders)	Flemish Index of Biotic Integrity
	Belgium (Wallonia)	Biological Index for Fish integrity
	Finland	Finnish River Fish Index
	Germany	Fish-based Assessment System
	Italy	...
	Lithuania	Assessment method of rivers using Lithuanian fish index
	Netherlands	WFD-metrics for natural water types
	Slovakia	Fish Index of Slovakia
	Spain	Index of Biotic Integrity using fish as indicators of the Ecological Status of Spanish Rivers
Macrophytes	Austria	Assessment of the biological quality elements - part macrophytes
	Belgium (Flanders)	Flemish macrophyte assessment system
	Germany	German Assessment system for Macrophytes and Phytobenthos according to the WFD
	Netherlands	WFD-metrics for natural water types
	Slovakia	Slovak assessment of macrophytes in rivers
	Slovenia	River Macrophyte Index
	United Kingdom	Ecological Classification of Rivers using Macrophytes
Phytobenthos	Austria	Assessment of the biological quality elements - part phytobenthos
	Belgium (Flanders)	Proportions of Impact-Sensitive and Impact-Associated Diatoms
	Belgium (Wallonia)	Pollution Sensitivity Index
	Estonia	Assessment system for rivers using phytobenthos in Estonia
	Finland	Indice de Polluosensibilité Spécifique (Specific Pollution Sensitivity Index SPI)
	Germany	German assessment system for Macrophytes and Phytobenthos according to the WFD
	Hungary	Improvement of the Hungarian ecological water qualification system - Phytobenthos in Rivers
	Italy	Intercalibration Common Metrics Index
	Netherlands	WFD-metrics for natural water types
	Slovakia	Slovak assessment of phytobenthos in rivers
	Slovenia	Ecological status assessment system for rivers using phytobenthos
	Spain	Pollution Sensitivity Index
	United Kingdom	WFD River Diatom method or Trophic Diatom Index version 3 Method
Phytoplankton	Belgium (Flanders)	German phytoplankton assessment method for rivers
	Germany	Index Phytoplankton PhytoFluss
	Hungary	Hungarian River Phytoplankton Index
	Slovakia	Slovak assessment of phytoplankton in rivers

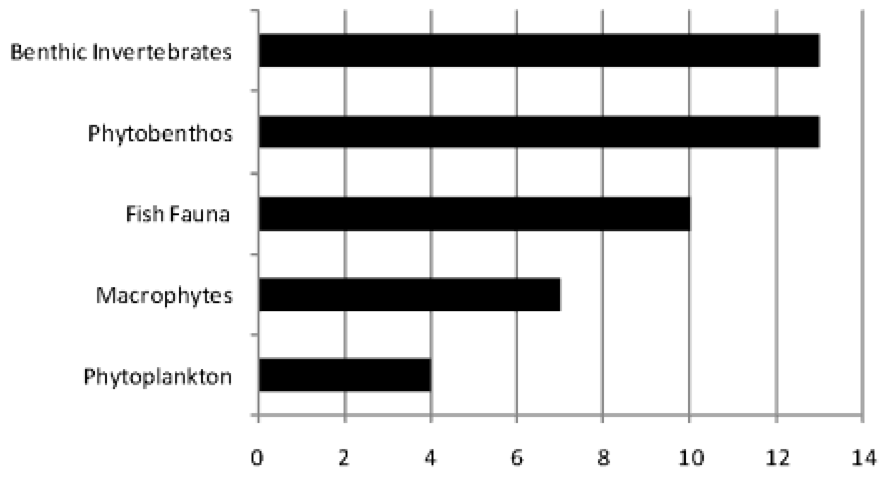


Figure 1: Number of national methods classifying the ecological status using various biological quality elements (47 methods)

Brief summary

Benthic invertebrates (Table A1 in the Annex): Data are acquired by multi-habitat-sampling at wadeable areas close to the rivers banks. Alternative techniques are airlift or grab sampling, or the additional hand-picking of animals from stones. The level of taxonomic identification differs among methods: More than half of the countries determines the sampled organisms to species-level, the rest identifies to the taxonomic level of genus or family. All assessment methods focus on the detection of general degradation and/or organic pollution. The selected assessment metrics are generally similar to those for the assessment of smaller rivers (e.g. number of EPT taxa, saprobic indices, Average Score Per Taxon).

Phytobenthos (Table A2): Except for Austria, Germany and Slovakia all countries only use diatoms in the assessment of the river phytobenthos. Slovakia additionally considers the occurrence of filamentous bacteria. Samples are generally acquired by scraping the surface of 5 to 10 stones at the river banks. In the Dutch method diatoms are sampled from reed stems. The number of valves identified for the assessment differs among countries and ranges from 200 to 500. All methods focus on the impacts of eutrophication, some also include the appraisal of organic pollution or general degradation. Amongst the various assessment metrics the Indice de Polluosensibilité Spécifique (IPS; Descy 1979) is most frequently used.

Fish Fauna (Table A3): Data on fish species are acquired by electrofishing along the banks by boat, or confined to wadeable stream areas. Beam trawls, nets or hydroacoustics are occasionally used. The fish status is assessed by multi-metric indices covering various community aspects such as habitat-, feeding- and age-guilds, sensitive taxa and migration indices.

Macrophytes (Table A4): Macrophyte taxa are surveyed by different techniques. Most countries record the macrophyte occurrences at 100 metre sections close to the river banks. The Dutch method pools the data obtained from 6 to 20 survey stretches within a water body. In Slovakia the length of the river surveyed for macrophytes amounts to 1 kilometre. Also plant groups included in the assessments differ among countries. All methods survey vascular plants, most methods additionally record mosses and stoneworts. Macroalgal data are acquired only by Slovakia and the United Kingdom. Each country uses a different set of biological metrics to assess eutrophication impacts, hydromorphological impairment or general degradation. Macrophyte growth forms are evaluated by the Flemish and British methods. Only Germany and the Netherlands combine the macrophyte results with the outcomes of the phytobenthos assessment.

Phytoplankton (Table A5): The phytoplankton assessments are based on seasonal means of samples taken once or twice per month from April to September/October. Sampled organisms are identified to genus or species-level. All methods evaluate the concentration of chlorophyll-a. Species composition is considered by assessing the proportions of various taxonomic or functional groups in the planktonic community (e.g. share of Chlorophytes or Cyanophytes), and applying specific sensitivity metrics (e.g. the Type-specific Index Potamoplankton TIP; Mischke & Behrendt 2007).

Definition of reference conditions

The ecological quality is assessed against reference conditions that resemble the biological communities at near-natural state. The WFD specifies three general approaches to derive biological reference conditions: (1) spatially based techniques using existing sites, (2) modelling, and (3) expert knowledge. Since large rivers are generally among the most degraded ecosystems worldwide (Dynesius and Nilsson 1994) and highly modified in the densely populated countries of Europe (e.g. Van Looy et al. 2008), pristine conditions no longer exist. Thus, most countries referred (at least partly) to expert knowledge in defining reference conditions for their assessment of very large rivers (Figure 2). Some methods are based on theoretical reference values gained from (re)constructing the biological communities that reflect potentially undisturbed conditions (e.g. Schaumburg et al. 2004). In other cases experts directly defined the reference values on the scale of the assessment metrics (e.g. Vanden Bossche & Usseglio-Polatera 2005, Gabriels et al. 2010). Schöll et al. (2005) established the reference state regarding the share of potamal invertebrate species occurring at the water body.

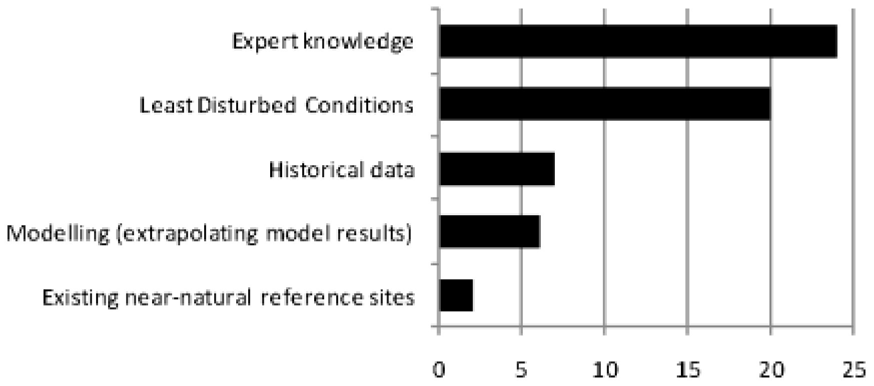


Figure 2: Number of methods referring to various techniques of defining reference conditions for large rivers (42 methods; multiple answers) (data taken from Birk et al. 2010)

Spatial reference approaches in large river assessment were mostly based on sites in least disturbed conditions (sensu Stoddard et al. 2006). Some countries defined abiotic criteria to screen for minimally degraded sites (e.g. Borics et al. 2007, Munne & Prat 2009, Vehanen et al. 2010). Others selected the best biological conditions of all available sites to fix the reference values ("best-available" approach; e.g. Rolauffs et al. 2003). Few methods referred to the use of historical data or modelling. Historical data were mostly consulted for the definition of fish reference communities (e.g. Kováč 2008, Schotzko et al. 2010). Modelling approaches included the extrapolation of biological reference values from pressure-impact-relationships (e.g. via nutrient emission models, Mischke & Behrendt 2007).

Several methods do not hold reference values specific to very large rivers. The invertebrate methods of Estonia and Finland, for instance, derived references from existing near-natural sites at rivers mainly below 10,000 km^2 catchment area. The site-specific models of the British methods are calibrated at smaller water bodies as only few very large rivers occur in the United Kingdom. River size is irrelevant especially for the definition of diatom reference conditions. Table A6 in the Annex provides an overview of the approaches to define references.

Ecological class boundary setting

All assessment methods classify the ecological status of water bodies by one of five classes ranging from high to bad quality. Class boundaries are expressed as Ecological Quality Ratios (EQR). In most national classifications boundaries were set by equidistant division of the EQR gradient, reflecting the (anticipated) continuous response of the biological elements to anthropogenic pressures (Figure 3). The high-good boundary was often defined by referring to the biological status at sites in least disturbed conditions. Other approaches comprised the calibration of the ecological classification against pre-classified sampling sites (e.g. Breine et al. 2006). The boundary setting of few methods was based on discontinuities in the relationship of human pressures and the biological response. Paired metrics that respond in different ways to the influence of pressures were also used (e.g. Willby et al. 2009). A more detailed description of the national boundary settings is given in the Annex (Table A7).

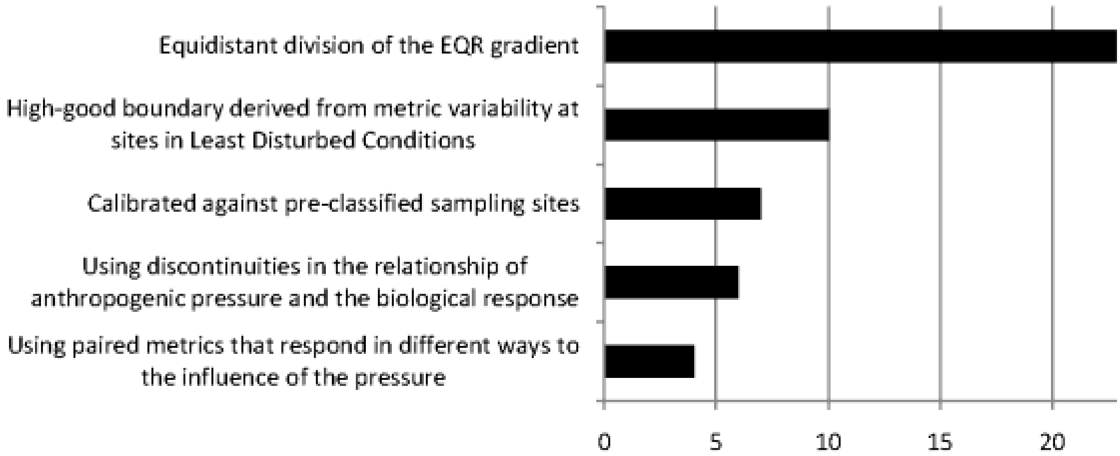


Figure 3: Number of methods using various approaches to define the ecological status classes (38 methods; multiple answers) (data taken from Birk et al. 2010)

Coverage of WFD indicative parameters

To assess the ecological status the WFD stipulates the coverage of certain parameters indicative of the biological quality element. Moreover, a combination rule to combine parameter assessment into full BQE assessment has to be defined (Schmedtje et al. 2009). Table 2 briefly presents the results of a preliminary checking of the national methods' design, pointing at possible issues of WFD compliancy. The presented outcomes require further discussions with the individual Member States.

Table 2: WFD compliancy issues of large river methods

BQE	Country	Problem
Benthic invertebrates	Austria	The rivers Danube, March and Thaya are only assessed by Saprobic Index, i.e. the parameters *richness* and *diversity* are not covered.
	Spain (Catalonia)	IBMWP index does not cover the parameter *abundance*. To be clarified.
	United Kingdom	Assessment does not cover the parameter *abundance*.
Macrophytes and Phytobenthos	~~Austria~~	Worst case ~~Component *macrophytes* is not assessed.~~
	Estonia	Component *macrophytes* is not assessed.
	Finland	Component *macrophytes* is not assessed.
	Hungary	Component *macrophytes* is not assessed.
	Slovakia	worst case
	Slovenia	average (eutrophication modules)
	Spain	Component *macrophytes* is not assessed.
	Belgium (Wallonia)	
	Germany	Component *macrophytes* is not assessed.
	Italy	
Fish Fauna	Lithuania	Assessment does not cover the parameter *age structure*.
	Netherlands	Assessment does not cover the parameter *age structure*.
Macrophytes	Belgium (Flanders)	worst case
	Slovakia	worst case
	United Kingdom	worst case
Phytoplankton	Belgium (Flanders)	Planktonic blooms are not considered.
	Germany	Planktonic blooms are not considered.
	Hungary	Planktonic blooms are not considered.
	Slovakia	Planktonic blooms are not considered. Generally problematic → Ute Mischke

3 References

Amoros C, Bornette G (2002) Connectivity and biocomplexity in waterbodies of riverine floodplains. Freshwater Biology 47:761-776.

Aroviita J, Koskenniemi E, Kotanen J, Hämäläinen H (2008) A priori typology-based prediction of benthic macroinvertebrate fauna for ecological classification of rivers. Environmental Management 42:894-906.

Belpaire C, Smolders R, van den Auweele I, Ercken D (2000) An Index of biotic integrity characterizing fish populations and the ecological quality of Flandrian water bodies. Hydrobiologia 434:17-33.

Berg MS, Coops HC, Pot R, Altenburg W, Nijboer R, Broek Tv, Fagel M, Arts G, Bijkerk R, Dam Hv, Ietswaart T, Molen Jv, Wolfstein K, Jong Dd, Hartholt H (2004) Achtergronddocument referenties en maatlatten waterflora. RIZA, Lelystad.

Birk S, Strackbein J, Hering D (2010) WISER methods database. Version: May 2010. Available at http://www.wiser.eu/results/methods-db/.

Blocksom KA, Johnson BR (2009) Development of a regional macroinvertebrate index for large river bioassessment. Ecological Indicators 9:313-328.

Borics G, Várbíró G, Grigorszky I, Krasznai E, Szabó S, Kiss KT (2007) A new evaluation technique of potamo-plankton for the assessment of the ecological status of rivers. Archiv für Hydrobiologie Suppl. Large Rivers 17:465-486.

Breine JJ, Maes J, Quataert P, Bergh E, Simoens I, Thuyne G, Belpaire C (2006) A fish-based assessment tool for the ecological quality of the brackish Schelde estuary in Flanders (Belgium). Hydrobiologia 575:141-159.

De Sostoa A, Caiola N, Casals F (2004) A new IBI (IBICAT) for local application of the E.U. Water Framework Directive. In: Jalón DG, Martínez PV Proceedings of the Fifth International Symposium on Ecohydraulics - Aquatic Habitats: Analysis and Restoration. IHAR, Madrid, pp. 187-191.

Descy J (1979) A new approach to water quality estimation using diatoms. Nova Hedwigia 64:305-323.

Dussling U, Berg R, Klinger H, Wolter C (2004) Assessing the ecological status of river systems using fish assemblages. Handbuch Angewandte Limnologie 20:1-84.

Dynesius M, Nilsson C (1994) Fragmentation and Flow Regulation of River Systems in the Northern Third of the World. Science (New York, N.Y.) 266:753-762.

Eloranta P, Soininen J (2002) Ecological status of some Finnish rivers evaluated using benthic diatom communities. Journal of Applied Phycology:1-7.

Expertteams (2007) Achtergronddocument referenties en maatlatten vissen. RIZA, Lelystat:110.

Gabriels W, Lock K, De Pauw N, Goethals PL (2010) Multimetric Macroinvertebrate Index Flanders (MMIF) for biological assessment of rivers and lakes in Flanders (Belgium). Limnologica - Ecology and Management of Inland Waters 40:199-207.

Hendrickx A, Denys L (2005) Toepassing van verschillende biologische beoordelingssystemen op Vlaamse potentiële interkalibratielocaties overeenkomstig de Europese Kaderrichtlijn Water - Partim "Fytobenthos". Rapport Instituut voor Natuurbehoud, Brussels.

Hlubikova D, Miettinen J, Hindakova A (2006) Testing of diatom water quality indices in monitoring and reference sites of Slovakian rivers (Central Europe). In: Ács É, Kiss KT, Padisák J, Szabó S 6th International Symposium on the Use of Algae for Monitoring Rivers. Magyar Algológiai Társaság, Göd, pp. 62-67.

Junk WJ, Bayley PB, Sparks RE (1989) The flood pulse concept in river-floodplain systems. In: Dodge D Proceedings of the international large river symposium. Can. Spec. Publ. Fish. Aquat. Sci. 106. pp. 110-127.

Kelly MG, Juggins S, Guthrie R, Pritchard S, Jamieson J, Rippey B, Hirst H, Yallop M (2008) Assessment of ecological status in U.K. rivers using diatoms. Freshwater Biology 53:403-422.

Kováč V (2008) National Method and the Fish Index of Slovakia (FIS). AQ-BIOS, Bratislava.

Leyssen A, Adriaens P, Denys L, Packet J, Schneiders A, Van Looy K, Vanhecke L (2005) Toepassing van verschillende biologische beoordelingssystemen op Vlaamse potentiële interkalibratielocaties overeenkomstig de Europese kaderrichtlijn water : partim 'Macrofyten'. Instituut voor Natuurbehoud, Brussels.

Leyssen A, Denys L, Schneiders A, Van Looy K, Packet J, Vanhecke L (2006) Afstemmen van referentiecondities en evaluatiesystemen voor de biologische kwaliteitselementen macrofyten en fytobenthos en uitwerken van een meetstrategie in functie van de Kaderrichtlijn Water. Rapport Instituut voor Natuurbehoud, Brussels.

Mischke U, Behrendt H (2007) Handbuch zum Bewertungsverfahren von Fließgewässern mittels Phytoplankton zur Umsetzung der EU-WRRL in Deutschland. Weißensee-Verlag, Berlin.

Munne A, Prat N (2009) Use of macroinvertebrate-based multimetric indices for water quality evaluation in Spanish Mediterranean rivers : an intercalibration approach with the IBMWP index. Hydrobiologia 10:203-225.

Novak, MA, Bode, RW (1992) Percent Model Affinity - A New Measure of Macroinvertebrate Community Composition. Journal of the North American Benthological Society 11:80-85.

Ofenböck T, Moog O, Stubauer I, Huber T, Leitner P (2005) Entwicklung eines flächendeckend anwendbaren Systems zur Beurteilung des ökologischen Zustandes auf Basis des Makrozoobenthos - Endbericht. Universität für Bodenkultur, Wien:91.

Pfister P, Pipp E (2010) Leitfaden zur Erhebung der biologischen Qualitätselemente - Teil A3 - Phytobenthos. Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft, Wien:96.

Rolauffs P, Hering D, Sommerhäuser M, Rödiger S, Jähnig S (2003) Entwicklung eines leitbildorientierten Saprobienindexes für die biologische Fließgewässerbewertung. Umweltbundesamt, Berlin.

Schaumburg J, Schranz C, Foerster J, Gutowski A, Hofmann G, Meilinger P, Schneider S, Schmedtje U (2004) Ecological classification of macrophytes and phytobenthos for rivers in Germany according to the Water Framework Directive. Limnologica 34:283-301.

Schmedtje U, Birk S, Poikane S, van De Bund W, Bonne W (2009) Guidance document on the intercalibration process 2008-2011. Guidance Document No. 14. Implementation Strategy for the Water Framework Directive (2000/60/EC).

Schotzko N, Haunschmid R, Petz-Glechner R, Honsing-Erlenburg W, Schmutz S, Unfer G, Wolfram G, Spindler T (2010) Leitfaden zur Erhebung der Biologischen Qualitätselemente Teil A1 - Fische. Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft, Wien.

Schöll F, Haybach A, König B (2005) Das erweiterte Potamontypieverfahren zur ökologischen Bewertung von Bundeswasserstraßen (Fließgewässertypen 10 und 20: kies- und sandgeprägte Ströme, Qualitätskomponente Makrozoobenthos) nach Maßgabe der EU-Wasserrahmenrichtlinie. Hydrologie & Wasserbewirtschaftung 49:234-247.

Stoddard JL, Larsen DP, Hawkins CP, Johnson RK, Norris RH (2006) Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications 16:1267-76.

Stubauer I (2002) Ausweisung saprobieller Grundzustände österreichischer Fließgewässer. Dissertation zur Erlangung des Doktorgrades an der Universität für Bodenkultur. Wien.

Vanden Bossche JV, Usseglio-Polatera P (2005) Characterization, Ecological Status and Type-specific Reference Conditions of Surface Water Bodies in Wallonia (Belgium) using Biocenotic Metrics based on Benthic Invertebrate Communities. Hydrobiologia 551:253-271.

Van Looy K, Meire P, Wasson JG (2008) Including riparian vegetation in the definition of morphologic reference conditions for large rivers: A case study for Europe's Western Plains. Environmental management 41:625-639.

Vannote RL, Minshall GW, Cummins KW, Sedell JR, Cushing CE (1980) The River Continuum Concept. Canadian Journal of Fisheries and Aquatic Sciences 37:130-137.

Vehanen T, Sutela T, Korhonen H (2010) Environmental assessment of boreal rivers using fish data - a contribution to the Water Framework Directive. Fisheries Management and Ecology 17:165-175.

Virbickas T, Kesminas V (2007) Development of fish-based assessment method for the ecological status of rivers in the Baltic region. Fisheries Management and Ecology 14:531-539.

Willby NJ, Pitt J, Phillips G (2009) The Ecological Classification of UK Rivers Using Aquatic Macrophytes. Environment Agency, Science report.

Yoder CO, Kulik BH (2003) The development and application of multimetric indices for the assessment of impacts to fish assemblages in large rivers: A review of current science and applications. Canadian Water Resources Journal 28:301-328.

van Dam H, Stenger-Kovács C, Ács E, Borics G, Buczkó K, Hajnal É, Soróczki-Pintér É, Várbíró G, Tóthmérész B, Padisák J (2007) Implementation of the European Water Framework Directive: Development of a system for water quality assessment of Hungarian running waters with diatoms. Archiv für Hydrobiologie, Suppl. Large Rivers 17:339-364.

Annex

Table A1: National assessment methods for very large rivers using **benthic invertebrates** (n.a. – data not available)

Country	Method name	db_ID[1]	Sampling	Determination level	Metrics	Pressure	Literature reference
Austria	Assessment of the biological quality elements - part benthic invertebrates	49	Multi-Habitat-Sampling at banks (occasionally Airlift)	Species	Saprobic Index (Danube, March, Thaya); Degradation Index, Number of EPT taxa, Proportion of EPT taxa, Lithal-profundal preferring taxa (Large Alpine rivers)	Organic pollution, General degradation	Stubauer 2002, Ofenböck et al. 2005
Belgium (Flanders)	Multimetric Macroinvertebrate Index Flanders	123	Kick-Sampling at banks combined with hand-picking of animals from stones	Genus and Family	Number of taxa, Number of EPT taxa, Number of sensitive taxa, Shannon-Diversity, Mean Tolerance Score	Organic pollution, General degradation	Gabriels et al. 2010
Belgium (Wallonia)	Global biological index adapted to large watercourses and deep rivers	290	Artificial substrates	Genus and Family	IBGN (incl. richness and sensitivity metrics)	General degradation	Vanden Bossche & Usseglio-Polatera 2005
Estonia	Estimation of freshwater quality using macroinvertebrates	46	Multi-Habitat-Sampling at banks	Species	Number of Taxa, Number of EPT taxa, Shannon-Diversity, Average Score Per Taxon (ASPT), Danish Stream Fauna Index	Organic pollution, General degradation	n.a.
Finland	Finnish multimetric index	146	Multi-Habitat-Sampling at banks within riffle zones	Species	Occurrence of type-specific taxa, Occurrence of type-specific EPT families, PMA (Percent Model Affinity)	General degradation	Aroviita et al. 2008, Novak & Bode 1992
Germany	Assessment method for rivers using benthic invertebrates	275	Multi-Habitat-Sampling at banks, Grab	Species	Saprobic Index, Potamon-Typie-Index	Organic pollution, General degradation	Rolauffs et al. 2003, Schöll et al. 2005
Italy	MacrOper, based on STAR_ICM index calculation	215	Artificial substrates (Multiple-plate sampler)	Genus and Family	n.a.	General degradation	n.a.
Lithuania	Assessment system for rivers using macrozoobenthos indicators (Danish Stream Fauna Index)	58	Multi-Habitat-Sampling at banks	Genus and Family	Danish Stream Fauna Index	General degradation	n.a.
Netherlands	WFD-metrics for natural water types	288	Multi-Habitat-Sampling at banks, Sediment corer, Brushed stone samples	Species	Number of typical species, Proportion of negative species, Proportion of positive species	General degradation	n.a.
Slovakia	Slovak assessment of benthic invertebrates in large rivers	165	Multi-Habitat-Sampling at banks	Species	Saprobic Index, Proportion of oligosaprobic, Rhithron-Type-Index, Index of biocoenotic region, Proportion of akal-lithal-psammal preferring taxa, BMWP	Organic pollution, General degradation	n.a.
Slovenia	Slovenian assessment system for rivers using benthic invertebrates	289	Multi-Habitat-Sampling at banks	Species	Saprobic Index, River Fauna Index, Proportion of akal-lithal-psammal preferring taxa	Organic pollution, Hydromorphological impairment	n.a.
Spain	Iberian Biological Monitoring Working Party	9	Multi-Habitat-Sampling (Surber + Drowning Box)	Family	Iberian BMWP	Organic pollution	Munne & Prat 2009
United Kingdom	River Invertebrate Classification Tool	20	Multi-Habitat-Sampling at banks, Airlift	Family	ASPT, Number of taxa	Organic pollution, General degradation	n.a.

[1] ID of the WISER methods' database; available online at *"http://www.wiser.eu/programme-and-results/data-and-guidelines/method-database/detail.php?id="* plus *"db_ID"*

Table A2: National assessment methods for very large rivers using **phytobenthos** (n.a. – data not available)

Country	Method name	db_ID	Diatoms	Other PB[2]	Sampling	#Valves	Metrics	Pressure	MA[3] combination	Literature reference
Austria	Assessment of the biological quality elements - part phytobenthos	168	yes	yes	Scraped from stones at river banks (5-10 replicates); full survey for other phytobenthos	500	Trophic Index, Saprobic Index, Reference-Taxa-Indices	Eutrophication, Organic pollution, General degradation	Worst-case	Pfister & Pipp 2010
Belgium (Flanders)	Proportions of Impact-Sensitive and Impact-Associated Diatoms	127	yes	no	Scraped from stones and plants at river banks (5 replicates)	500	IAD, ISD	Eutrophication, Organic pollution, General degradation	Worst-case	Hendrickx & Denys 2005, Leyssen et al. 2006
Belgium (Wallonia)	Pollution Sensitivity Index	14	yes	no	Scraped from stones at river banks	n.a.	IPS	Eutrophication, Organic pollution, General degradation	n.a.	n.a.
Estonia	Assessment system for rivers using phytobenthos in Estonia	111	yes	no	Scraped from stones at river banks (5 replicates)	n.a.	IPS, Trophic Diatom Index, Watanabe Index	Eutrophication	n.a.	n.a.
Finland	Pollution Sensitivity Index	247	yes	no	Scraped from stones at river banks (5-10 replicates)	400	IPS	Eutrophication, Organic pollution, General degradation	n.a.	Eloranta & Soininen 2002
Germany	German Assessment system for Macrophytes and Phytobenthos according to the EU WFD	218	yes	yes	Scraped from stones at river banks (5-10 replicates); full survey for other phytobenthos	400	Trophic Index, Saprobic Index, Reference-Index	Eutrophication, Organic pollution, General degradation	Average	Schaumburg et al. 2004
Hungary	Improvement of the Hungarian ecological water qualification system - Phytobenthos in Rivers	221	yes	no	Scraped from stones at river banks (5-10 replicates)	400	IPS	Eutrophication, Organic pollution	n.a.	van Dam et al. 2007
Italy	Intercalibration Common Metrics Index	248	yes	no	Scraped from stones at river banks (5-10 replicates)	n.a.	IPS, Trophic Index	Eutrophication, Organic pollution	n.a.	
Netherlands	WFD-metrics for natural water types	298	yes	No	Removal of 4-8 reed stems (several replicates)	200	IPS	Eutrophication, Organic pollution, General degradation	Average	Berg et al. 2004
Slovakia	Slovak assessment of benthic diatoms in rivers	169	yes	yes	Scraped from stones at river banks (5-10 replicates)	300-500	IPS, CEE, EPI-D, occurrence of filamentous bacteria	Eutrophication, Organic pollution, General degradation	Worst-case	Hlubikova et al. 2006
Slovenia	Ecological status assessment system for rivers using phytobenthos	37	yes	no	All available habitats	500	Saprobic Index, Trophic Index	Eutrophication, Organic pollution	Average (eutrophication module)	n.a.
Spain	Pollution Sensitivity Index	10	yes	no	Scraped from stones at river banks (5-10 replicates)	400	IPS	Eutrophication	n.a.	n.a.
United Kingdom	WFD River Diatom method or Trophic Diatom Index version 3 Method	54	yes	no	Scraped from stones at river banks (5 replicates)	500	Trophic Diatom Index	Eutrophication, General degradation	Worst-case	Kelly et al. 2008

[2] Phytobenthos
[3] Macrophytes

Table A3: National assessment methods for very large rivers using **fish fauna** (n.a. – data not available)

Country	Method name	db_ID	Sampling	Gear	Metrics	Literature reference
Austria	Fish Index Austria	42	"Strip fishing" (25 strips)	Electrofishing and nets, hydroacoustics	Proportion of dominant taxa, Proportion of subdominant taxa, Proportion of rare taxa, Habitat guilds, Fish Region Index, Reproduction guilds, Length-Frequency-Distribution	Schotzko et al. 2010
Finland	Finnish River Fish Index	120	Wadeable stream areas	Electrofishing	Proportion of tolerant/intolerant taxa, Number of age-0+ salmonid juveniles, Number of cyprinid individuals, Number of fish species	Vehanen et al. 2010
Belgium (Flanders)	Flemish Index of Biotic Integrity	52	Partial random sampling (shallow zones)	Electrofishing	Total number of species, Mean tolerance, Mean typical species value, Type species, Total biomass, Weight % of non-native species, Trophic composition, Natural recruitment	Belpaire et al. 2000
Germany	Fish-based Assessment System	246	n.a.	Electrofishing	Number of type-specific taxa, Number of accompanying taxa, Number of anadromous and potamodromous taxa, Number of habitat guilds, Number of reproductive guilds, Number of trophic guilds, Guild abundance, Abundance of perch and roach, Proportion of age-guilds, Fish Region Index, Migration Index, Guilding Species Index, Community Dominance Index	Dussling et al. 2004
Italy	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Lithuania	Assessment method of rivers using Lithuanian fish index	61	Partial random sampling (shallow zones)	Electrofishing	Proportion of tolerant/intolerant taxa, Proportion of lithophilous taxa, Proportion of rheophilic taxa, Proportion of omnivorous taxa, Relative number of tolerant taxa, Relative number of lithophilic taxa, Number of intolerant taxa	Virbickas & Kesminas 2007
Netherlands	Netherlands References and Metrics for Fish in Large Rivers	293	Along the banks by boat	Electrofishing, beam trawl	Number of rheophilic taxa, Number of diadromous taxa, Number of limnophilic taxa, Abundance of rheophilic taxa, Abundance of limnophilic taxa	Expertteams 2007
Slovakia	Fish Index of Slovakia	166	Along the banks by boat	Electrofishing	Relative abundance of insectivorous taxa, Relative abundance of phytophilous taxa, Relative abundance of lithophilous taxa, Relative abundance of benthic taxa, Relative abundance of rheophilous taxa, Relative abundance of potamodromous taxa, Relative abundance of piscivorous taxa, Relative abundance of salmonid taxa, Relative abundance of invasive taxa, Equitability Index	Kováč 2008
Spain	Index of Biotic Integrity using fish as indicators of the Ecological Status of Spanish Rivers	292	Along the banks by boat	Electrofishing	Number of introduced individuals, Number of native marine migratory species, Number of native piscivorous species, Number of intolerant species less than 15 cm, Number of invertivorous species, Number of omnivorous species, Number of rheophilic species, Number of tolerant introduced species, Number of intolerant native species, Number of lithophilic native species, Proportion of biomass of benthic natives, Proportion of individuals with injuries / diseases, Proportion of intolerant species, Proportion of omnivorous species, Proportion of introduced invertivorous species, Proportion of native lithophilic species, Proportion of native tolerant species	De Sostoa et al. 2004
Belgium (Wallonia)	Biological Index for Fish integrity	294	Along the banks by boat	Electrofishing	Number of native taxa, Number of benthic taxa, Proportion of intolerant taxa, Number of bullheads, Number of spawners, Age0-Presence	n.a.

Table A4: National assessment methods for very large rivers using **macrophytes** (n.a. – data not available)

Country	Method name	db_ID	Sampling	Taxonomical groups	Metrics	PB[4] combination	Literature reference
Austria	Austrian Index Macrophytes for Rivers	69	500 m survey close to river banks	Vascular plants, mosses, stone-worts	Austrian Macrophyte Index	Worst-case	Pall & Mayer-hofer 2008
Belgium (Flanders)	Flemish macrophyte assessment system	125	100 m survey close to river banks	Vascular plants	Percentage of type-specific species, Percentage of disturbance species, Number of growth forms	Worst-case	Leyssen et al. 2005
Germany	German Assessment system for Macro-phytes and Phytobenthos according to the EU WFD	218	100 m survey close to river banks	Vascular plants, mosses, stone-worts	Reference Index, additional criteria	Average	Schaumburg et al. 2004
Netherlands	WFD-metrics for natural water types	297	Pooling of 6-20 survey stretches (100 m close to river banks)	Vascular plants, mosses, stone-worts	Type-specific species, Growth form cover	Average	Berg et al. 2004
Slovakia	Slovak assessment of macrophytes in rivers	167	1 km close to river banks	Vascular plants, mosses, macro-algae	IBMR	Worst-case	n.a.
Slovenia	River Macrophyte Index	81	n.a.	Vascular plants, mosses, macro-algae	River Macrophyte Index	Average (eutrophi-cation module)	n.a.
United Kingdom	Ecological Classification of Rivers using Macrophytes	19	100 m survey close to river banks	Vascular plants, mosses, macro-algae	River Macrophyte Nutrient Index, River Macrophyte Hydraulic Index, Functional Group Diversity, Number of Taxa, Filamentous Algal Cover	Worst-case	Willby et al. 2009

Table A5: National assessment methods for very large rivers using **phytoplankton** (n.a. – data not available)

Country	Method name	db-ID	Sample size and frequency	Identification level	Metrics	Literature reference
Belgium (Flanders)	German phytoplankton assessment method for rivers	64	Seasonal mean of monthly samples (April to Sept.)	Species and genus	Chlorophyll-a, Proportion of Pennate taxa, Proportion of Chlorophytes, Proportion of Cyanophytes, Type-specific Index Potamoplankton (TIP)	Mischke & Behrendt 2007
Germany	Index Phytoplankton PhytoFluss	229	Seasonal mean of monthly samples (April to Oct.)	Species and genus	Chlorophyll-a, Proportion of Pennate taxa, Proportion of Chlorophytes, Proportion of Cyanophytes, Type-specific Index Potamoplankton (TIP)	Mischke & Behrendt 2007
Hungary	Hungarian River Phytoplankton Index	160	Seasonal mean of monthly samples (April to Oct.)	Species	Chlorophyll-a, Share of functional groups (defined by trophic state, turbulence character, residence time, indication of pollution)	Borics et al. 2007
Slovakia	Slovak assessment of phytoplankton in rivers	170	Seasonal mean of two monthly samples (April to Oct.)	Species and genus	Chlorophyll-a, Total abundance, Proportion of Cyanophytes (incl. dominance of *Microcystis*), Proportion of Chromophytes, Proportion of Chlorophytes, Proportion of Euglenophytes	n.a.

[4] Phytobenthos

Table A6: Definition of reference conditions (n.a. – data not available)

BQE	Country	RefCond	Comments
Benthic Invertebrates	Austria	Expert knowledge, Least Disturbed Conditions	75th percentile of least-disturbed sites
	Belgium (Flanders)	Expert knowledge	-
	Belgium (Wallonia)	Expert knowledge	-
	Estonia	Existing near-natural reference sites	Reference values adopted from smaller rivers
	Finland	Existing near-natural reference sites	Reference values adopted from smaller rivers
	Germany	Expert knowledge, Least Disturbed Conditions	Module PTI: theoretical reference value (dominance of potamal species). Module "Saprobic Index": Best available sites (10th percentile of available large river data).
	Italy	…	…
	Netherlands	Expert knowledge, Historical data	-
	Slovakia	Expert knowledge, Least Disturbed Conditions	95th percentile value of metric ranges at all sites ("best available")
	Slovenia	Expert knowledge, Least Disturbed Conditions, Modelling (extrapolating model results)	Module "Saprobic Index": Best available sites; Module "Hydromorphology Index": Maximum score of biological index plus 5%
	Spain	Least Disturbed Conditions	Chemistry - ammonium: < 0.2 mg/l (mean), < 1 mg/l (max); nitrate-N: < 10 mg/l (mean), < 20 mg/l (max); phosphate-P: < 0.1 mg/l (mean), < 1mg/l (max); Hydrology - minimum flow: > 20 % of natural flow; near-natural flow regime variation; Morphology - good riparian conditions (QBR index > 75)
	United Kingdom	Least Disturbed Conditions	-
Fish Fauna	Austria	Expert knowledge, Historical data	-
	Belgium (Flanders)	Expert knowledge, Least Disturbed Conditions	Least-disturbed sites at the Meuse river (best class of habitat quality)
	Belgium (Wallonia)	Least Disturbed Conditions	Reference values adopted from smaller rivers
	Finland	Least disturbed conditions	Best 20% fractal of the sum of morphology scores, human disturbance effects and water quality PCA scores
	Germany	Expert knowledge, Historical data	-
	Italy	…	…
	Lithuania	Least Disturbed Conditions	Reference values adopted from smaller rivers
	Netherlands	Expert knowledge, Historical data	Dutch situation of around 1900
	Slovakia	Historical data	-
	Spain (Catalonia)	Least Disturbed Conditions	Reference values adopted from smaller rivers
Macrophytes	Austria	Expert knowledge, Historical data	-
	Belgium (Flanders)	Expert knowledge, Historical data	-
	Germany	Expert knowledge	Theoretical reference value (ideal type-specific species composition)
	Netherlands	Expert knowledge, Historical data	Reconstruction of good species composition and physico-chemical conditions from fragmented historical data, derivation from conditions in tributaries and comparisons with selected rivers in other countries (e.g., Allier, Loire, Prypyat, Elbe, Danube)
	Slovakia	Expert knowledge	-
	Slovenia	Expert knowledge	Reference values adopted from smaller rivers
	United Kingdom	Least Disturbed Conditions	-

Table A6 (cont.): Definition of reference conditions

BQE	Country	RefCond	Comments
Phytobenthos	Austria	Expert knowledge, Least Disturbed Conditions	Bioregion-specific reference definition (allocation of bioregional reference to large river stretch located in bioregion)
	Belgium (Flanders)	Expert knowledge	-
	Belgium (Wallonia)	Least Disturbed Conditions	No river size-specific reference values
	Estonia	Least Disturbed Conditions	No river size-specific reference values
	Finland	Least Disturbed Conditions	-
	Germany	Expert knowledge	Theoretical reference value (ideal type-specific species composition)
	Hungary	Expert knowledge, Least Disturbed Conditions	-
	Italy	Expert knowledge	-
	Netherlands	Expert knowledge	-
	Slovakia	Modelling (extrapolating model results)	The regression model used for the index values: IPS = 15.821 + 0.005(altitude) - 0.491(HMQS) - 0.025(agricult.+urban) (R^2=0.36); CEE = 13.187 + 0.005(altitude) - 0.043(agricult.+urban) (R^2=0.412); EPI-D = 12.924 + 0.006(altitude) - 0.031(agricult.+urban) - 0.464(HMQS) (R^2=0.444). Catchment size is not statistically significant factor for either of the indices so it was not included in the models. HMQS was statistically significant only for IPS. Mid-points of the altitude and size categories were used for the predictions, except 150 m for altitude category < 200 m, and 950 m for altitude category > 800 m. Hydromorphological impacts were set to 1 and land use parameters were set to zero impact levels, to derive the type specific reference values (HMQS = 1, agricultural and urban land combined = 'agricult.+urban'=0).
	Slovenia	Expert knowledge, Modelling (extrapolating model results)	-
	Spain	Expert knowledge, Modelling (extrapolating model results)	-
	United Kingdom	Least Disturbed Conditions	-
Phytoplankton	Belgium (Flanders)	Expert knowledge	-
	Germany	Least Disturbed Conditions, Modelling (extrapolating model results)	Chl-a: extrapolation of reference value from nutrient emission model; taxonomic metrics: sites in least-disturbed conditions (multiplication by 1.5)
	Hungary	Least Disturbed Conditions	Sites in least-disturbed conditions (LDC) are not impounded and outside reservoirs. Chl-a reference: 50^{th} percentile of LDC sites corresponds to high-good boundary; Danube (no LDC sites): 33^{th} percentile of all monitoring data.
	Slovakia	Expert knowledge	-

Table A7: Ecological class boundary setting (general schemes) (n.a. – data not available)

BQE	Country	Boundary setting	Boundary description	Boundary communities
Benthic Invertebrates	Austria	Calibrated against pre-classified sampling sites Equidistant division of the EQR gradient	n.a.	Expressed by metric for each river types, no verbal description.
	Belgium (Flanders)	Equidistant division of the EQR gradient	EQR gradient is assumed to represent a continuous correlation with general degradation.	The EQR values at good status reflect metric values that are only slightly lower than at (expert-based) reference state, hence the community can be characterised as only slightly different from reference in terms of taxa richness, sensitivity and diversity.
	Belgium (Wallonia)	Equidistant division of the EQR gradient	n.a.	n.a.
	Estonia	Calibrated against pre-classified sampling sites	n.a.	Each single index yields either five balls (high quality), four balls (good quality), two balls (moderate quality), or zero balls (poor or bad quality). Good status is defined as the sum of balls ranging 18-22 (derived from five indices), or 14-17 (derived from four indices).
	Finland	Equidistant division of the EQR gradient High-good boundary derived from metric variability at near-natural reference sites	n.a.	Only EQRs are used to define ecological quality classes. Good status is defined by the 25 percentage point of reference site EQR in each type.
	Germany	Equidistant division of the EQR gradient	The national class boundary setting follows Option B of the REFCOND Guidance: The scale of EQR values was established by expert judgement proposing appropriate intervals from high to bad ecological status. The application of the scale to real datasets confirmed the proposed boundary setting. High status of the Module "General Degradation" corresponds to very low level of anthropogenic impact concerning land use und hydromorphological pressure.	n.a.
	Italy	...	...	...
	Netherlands	Calibrated against pre-classified sampling sites Equidistant division of the EQR gradient	The boundaries for the different EQR-classes (bad, poor, moderate, good and high) are set, based on expert judgement and follow a more or less equal division of quality. The WFDi and its class-boundaries were validated by experts judging species lists from anonymous sites, using normative definitions. In the validation of the method the response of the WFD-classes to pressures was tested. WFD-classes responded negatively to hydromorphological pressure. Of the chemical pressures studied, EQR is most related to oxygen content. EQR and oxygen availability are positively correlated. Influences of other chemical pressures considered (phosphate and nitrogen content) were less clear. Water bodies in the Netherlands are hydromorphologically altered, making physical pressure an important factor in assessment of Dutch water bodies.	Good status is characterized by a high diversity and abundance of typical species and an increasing abundance of dominant positive species. The abundance of dominant negative species is low.
	Slovakia	Equidistant division of the EQR gradient	n.a.	In Slovakia background taxa lists are not prescribed and especially created for good status conditions as well as for any other ES classes.
	Slovenia	Equidistant division of the EQR gradient Using paired metrics that respond in different ways to the influence of the pressure	Equidistant division was applied for Module Organic pollution, whereas for module Hydromorphological degradation paired metrics were used.	n.a.
	Spain	Equidistant division of the EQR gradient High-good boundary derived from metric variability at near-natural reference sites	The 25th percentile of the reference values is considered to be a slight deviation from the reference conditions and, therefore, is used as a class boundary value between high and good biological quality. Thus, the variability in values must be analysed at the reference sites, as this may affect the establishment of quality class boundaries. Beyond the 25th percentile of the reference values, the remaining quality class boundaries are defined in equal bands.	High number of families (more than 10-20), High number of EPT families (more than 6), High number of families of selected EPTCD (more than 4).
	Spain	Equidistant division of the EQR gradient Using discontinuities in the relationship of anthropogenic pressure and the biological response.	n.a.	n.a.

Table A7 (cont.): Ecological class boundary setting (general schemes)

BQE	Country	Boundary setting	Boundary description	Boundary communities
Benthic invertebrates	United Kingdom	High-good boundary derived from metric variability at sites in Least Disturbed Conditions Using paired metrics that respond in different ways to the influence of the pressure	Boundaries were checked against pressures.	High/Good boundary: Pressures may just be picked up in the biology. Middle of Good: Most Expected taxa present, with noticeable impact from pressures. Good / Moderate boundary: Major Taxonomic groups (dependent on typology) begin to be lost. NIEA response: There are slight changes in the composition and abundance of inverte-brate taxa from the type-specific communities. The ratio of disturbance-sensitive taxa to insensitive taxa shows slight alteration from type-specific levels. The level of diversity of invertebrate taxa shows slight signs of alteration from type-specific levels.
Fish Fauna	Austria	Calibrated against pre-classified sampling sites	Pre-classification of sites (status 1-5) by expert judgement - statistical difference testing among the pre-classification values for each metric - weighting of the metrics.	Minor deviation of the different metrics from the reference condition. More than 50% and <99% of the origin dominant fish species occur (>50<75% subdominant fish species; >20<50% rare species); <1 guild missing compared to reference condition, all age classes occur, juveniles are minor performed, higher amount of adult individuals.
	Belgium (Flan-ders)	Equidistant division of the EQR gradient Using discontinuities in the relationship of anthropogenic pressure and the biological response.	Class boundary setting was done using the metric values of the best available sites. These values were equally divided into 5 equal segments. The sum of the different metric scores gives an IBI. Initially, integrity classes were defined by dividing the maximum possible score into 8 equal distant intervals resulting into 9 classes. Later, to meet the criteria imposed by the WFD, these intervals were grouped into 5 classes and rescaled to an EQR ranging between 0 and 1. We rescaled applying the following formula: score = NUNS+ (S-UNS)/(UPS-UNS)*0.2 where previous score (not transformed)=S previous interval value: underscore (UNS) and upper score (UPS) will be transformed to an equal distant interval: new underscore (NUNS) and upper score (NUPS)	The EQR values at good status reflect metric values that are only slightly lower than at (expert-based) reference state, hence the community can be characterised as only slightly different from reference in terms of taxa richness, sensitivity and diversity.
	Belgium (Wal-lonia)	High-good boundary derived from metric variability at near-natural reference sites	n.a.	A good fish community is composed from the native species of the fish zone and is not disturbed by anthropogenic pressures or exotic species.
	Finland	High-good boundary derived from metric variability at sites in Least Disturbed Conditions	n.a.	n.a.
	Germany	Equidistant division of the EQR gradient	n.a.	n.a.
	Italy	...	...	...
	Lithuania	Using discontinuities in the relationship of anthropogenic pressure and the biological response.	Class boundary for each individual metric was set calculating averages between 25% (higher status class) and 75% (lower status class) (for increasing metrics, e.g. abundance of individu-als of tolerant species the opposite scheme was applied).	Nearly all type-specific intolerant species are present, however in lower abundances. Community is dominated by lithophilic and rheophilic fish individuals, diversity and abundance of less sensitive omnivorous fish is higher than at reference conditions.
	Netherlands	Equidistant division of the EQR gradient	Expert judgement. The method is considered to be sensitive to pressures.	Based on expert judgement. Rivers with rheophilic species (15 -19 species; 25 - 40% relative abundance), limnophilic species (4 species; 3 - 10% relative abundance) and diadromous species (6-10 species). The level of dominance varies per river type; there are three types of large rivers.
	Slovakia	Calibrated against pre-classified sampling sites	n.a.	This is expressed directly by FIS (which closely corresponds to EQR) that must be higher than 0.57.
	Spain (Catalonia)	Using discontinuities in the relationship of anthropogenic pressure and the biological response.	Intersection of Percentiles in a box-plot analysis.	Same as reference community description.
Macrophytes	Austria	Equidistant division of the EQR gradient	The class boundaries for each metric were defined according to the normative definitions and interpretations of the WFD as given in the REFCOND Guidance.	-
	Belgium (Flan-ders)	Equidistant division of the EQR gradient	EQR gradient is assumed to represent a continuous trend with general degradation.	The EQR values at good status reflect metric values that are only slightly lower than at (expert-based) reference state, hence the community can be characterised as only slightly different from reference in terms of taxa richness, sensitivity and diversity.
	Germany	Calibrated against pre-classified sampling sites	The boundaries were set at the zones of distinct changes of the biocoenosis (macrophytes and phytobenthos), and depending on indicator species lists derived from nutrient dependent Trophic Index (diatoms).	Type-specific reference species and tolerant species are still dominant, pressure indicators are rare, i.e. slight deviation from high status (WFD normative definitions).
	Netherlands	Equidistant division of the EQR gradient	The reference score for the sum of the scores of the species is derived from frequency data in the national vegetation database on well developed plant communities in the Netherlands,	Good status of very large rivers is characterized by macrophyte growth in areas with lower current and water level dynamics (e.g. non-flowing side channels), but moderate in

			which is considered a good estimate for the probability of finding the species in a fixed amount of samples. The fraction of species at G/M and H/G are estimated with expert judgment, and adjustment may be needed because of too low number of reference sites. Final adjustment of the reference scores are based on intercalibration results.	surface covered compared to reference conditions. Pressure tolerant species are present but only in low abundance; total cover of vegetation is moderate and type-specific.

Table A7 (cont.): Ecological class boundary setting (general schemes)

BQE	Country	Boundary setting	Boundary description	Boundary communities
Macrophytes	Slovakia	Equidistant division of the EQR gradient	Extrapolation of classification scheme - based on exact data of reference sites - based on expert judgment or historical data 1. establishing theoretical reference value of particular indices 2. extrapolation of reference value and boundary values between 5 classes of ecological status 3. classification of concrete locality for all indices 4. establishing ecological status = averaging of all particular EQR values	In Slovakia background taxa lists are not prescribed and especially created for good status conditions as well as for any other ecological status classes.

Table A7 (cont.): Ecological class boundary setting (general schemes)

BQE	Country	Boundary setting	Boundary description	Boundary communities
Macrophytes	Slovenia	Using paired metrics that respond in different ways to the influence of the pressure	Boundary values for the five classes of the ecological status were determined on the basis of the changing of portion of the frequency of so called "good" and "bad" RMI taxa. Portions were calculated on the basis of the frequency of the taxa. Taxa from the group A and AB were taken as "good" and taxa from the group C and BC as "bad". Boundary value between high and good ecological status was determined where so called "bad" taxa started to appear. Boundary value between good and moderate status was determined where there was a cross-point of curves of portion of "good" and "bad" taxa. Boundary value between moderate and poor ecological status was determined where the portion of frequency of "good" taxa drops below 10 %, and boundary value between poor and bad status where "good" taxa do not appear anymore.	-
	United Kingdom	High-good boundary derived from metric variability at sites in Least Disturbed Conditions Using paired metrics that respond in different ways to the influence of the pressure	The relative positions of High-Good and Poor-Bad boundaries are effectively symmetrical with sensitive species overwhelmingly dominant at one and tolerant species overwhelmingly dominant at the other. Using the same standard error from logistic regressions, a ratio of sensitive:tolerant species of 85:15 is used as the High-Good boundary, since this represents the upper error when tolerant species are predicted to be absent. These ratios are reversed at the Poor-Bad boundary with 15% sensitive species representing the lower error when sensitive species are predicted to be absent.	Sensitive taxa dominate, highly sensitive taxa are scarcer and account for about half the contribution of sensitive taxa. Tolerant taxa are present, but remain subordinate. Highly tolerant taxa, if present are rare. Macrophyte functions at high status all remain intact, undesirable disturbances are rare and macrophyte cover is stable over time.
Phytobenthos	Austria	High-good boundary derived from metric variability at near-natural reference sites	H/G: 10th percentile of high class Trophic Index / Saprobic Index values (all values lying within the defined type specific trophic / saprobic reference class based on Trophic Index / Saprobic Index classes according to ROTT's trophic / saprobic indication system) G/M: Upper Trophic Index boundary of next worse trophic class (following the type specific trophic reference class)	For good status defined common reference species and/or river type-specific species must obtain a certain percentage of all occurring algae (percentage varying in different bioregions).
	Belgium (Flanders)	Equidistant division of the EQR gradient	EQR gradient is assumed to represent a continuous trend with general degradation.	The EQR values at good status are characterised by a relatively low IAD and a ISD that is slightly reduced in comparison to reference.
	Belgium (Wallonia)	Equidistant division of the EQR gradient	The presence of very sensitive organisms has been taken in account to define high status.	A good community is relevant of the local typology and the presence of sensitive organisms or families is taken in account.
	Estonia	Equidistant division of the EQR gradient	n.a.	n.a.
	Finland	Using discontinuities in the relationship of anthropogenic pressure and the biological response.	The rivers studied were classified to five classes according to the degree of human impacts in the drainage basin in general or near the sampling station. Rivers with more or less natural state of very low degree of human impacts showed IPS values > 16, whereas those with slight human impact had the IPS from 14 to 16. The index values decreased markedly with increasing strength of human impact. Based on the results, the following limit values for IPS for evaluation of ecological water quality classes were proposed. High quality IPS>17; Good quality IPS 15-17; Moderate quality 12-15; Poor quality 9-12; Bad quality <9	n.a.
	Germany	Calibrated against pre-classified sampling sites	The boundaries were set at the zones of distinct changes of the biocoenosis (macrophytes and phytobenthos), and depending on indicator species lists derived from nutrient dependent Trophic Index (diatoms).	Type-specific reference species and tolerant species are still dominant, pressure indicators are rare, i.e. slight deviation from high status (WFD normative definitions).
	Hungary	High-good boundary derived from metric variability at sites in Least Disturbed Conditions Equidistant division of the EQR gradient	Reference conditions which could be applied across rivers in Hungary have not been established yet. Nevertheless unimpacted stretches or sites with low pollution and with smaller hydromorphological alterations can be found in almost every river type (i.e. sites in Least Disturbed Conditions - LDC). On basis of the pressure data (TP, BOD, CODCr, Electrical Conductivity) the LDS were selected.10th percentiles of the index values of the selected LDS sites were considered as high/good (H/G) class boundaries and 75th percentiles as good/moderate (G/M) boundaries in every type. The rest of data was divided into 3 equal parts between the minimum value of the index in a given river groups and the G/M value in order to set the further boundaries. Theoretical EQR values (H/G= 0.8; G/M= 0.6; M/P= 0.4; P/B= 0.2) were plotted against the index boundaries for all types. By equation of the actual line of best fit the EQR values can be calculated.	At good status stands of the sensitive taxa are well developed. They are dominant, but significantly decreasing at good-moderate boundary and replaced by tolerant taxa.The 10th percentiles of the index values of the selected LDS sites were considered as high/good (H/G) class boundaries.
	Italy	…	…	At the good status, sensitive taxa, such as *Achnanthidium minutissimum, Cymbella prostrata, Achanthidium biasolttianum* discrease and tolerant species grow up

Table A7 (cont.): Ecological class boundary setting (general schemes)

BQE	Country	Boundary setting	Boundary description	Boundary communities
Phytobenthos	Netherlands	Boundaries taken over from the intercalibration exercise	By using similarities in the geographic conditions the score on the IPS-scale in accordance with the reference condition is deduced for the Dutch situation. Next the scores of ten variants of an IPS-based metric were calculated for samples of CB-GIG type R-C1 and R-C4. For each of the ten variants the boundary values H/G and G/M at the intercalibration metric and several other performance characteristics were calculated, including the 95% confidence intervals of the boundary values. Finally a metric with a reference value has been chosen with boundary values which deviates less than the required 0.05 units from the mean values of all Member States.	The Good-Moderate boundary is based on the Intercalibration Metric.
	Slovakia	Equidistant division of the EQR gradient High-good boundary derived from metric variability at sites in Least Disturbed Conditions	Benthic diatoms: 2 modules - benthic diatoms and filamentous bacteria. a) benthic diatoms modul-4 altitude categories, based on reference sites within 2004 . For 200-500, 500-800 and above 800 - boundary between H/G = 25. Percentile of average based on reference sites in 2004. For altitude below 200 linear model used - derived from type of altitude 200-500 by means of modelling (this procedure - applied for all 3 metrics). The other boundaries calculated using the range of metrics values within high status (best value) and minimal calculated value of metric from the data set. The whole range was equally subdivided and boundaries were stated accordingly. b) filamentous bacteria module-percentage of bacteria in phytobenthos in vivo (5-class classification). Each class is classified by Score (below 1 %-5, 1-10 %-4, 11-25%-3, 25-40%-2, above 40%-1.) Result of both modules= the worse value classifies.	In Slovakia background taxa lists are not prescribed and especially created for good status conditions as well as for any other ecological status classes.
	Slovenia	Equidistant division of the EQR gradient	n.a.	n.a.
	Spain	Equidistant division of the EQR gradient	n.a.	n.a.
	United Kingdom	High-good boundary derived from metric variability at near-natural reference sites Using paired metrics that respond in different ways to the influence of the pressure	The high/good boundary is set at the 75th percentile of EQR values for reference sites within a particular type. 'Crossover' between nutrient-sensitive and nutrient-tolerant species. Biological metrics tend to show gradual change as the level of nutrient/organic pressure increases, with no distinct discontinuities that could act as criteria for setting class boundaries. An alternative approach – based on the proportions of nutrient-tolerant, nutrient-sensitive and indifferent taxa within samples – was used to define status class boundaries in the UK, with the good/moderate boundary set at the point where the proportion of sensitive taxa falls below that of tolerant taxa. In ecological terms, the diatom flora at high and good status is characterised by a number of taxa, often with relatively broad niches (e.g. *Achnanthidium minutissimum*, *Fragilaria capucina*) which occur at different phases of a micro-succession from colonisation of bare rock up to a mature biofilm. At high status, these are accompanied by other nutrient-sensitive taxa but as nutrient concentrations increase, the most sensitive of these taxa disappear whilst the numbers of nutrient tolerant taxa increases. Therefore, 'crossover' is the point at which the taxa which form the 'association' characteristic of a site in the absence of pressure become subordinate to taxa which are favoured by a pressure (nutrients, in this case). The EQR gradient below the good/moderate boundary is then divided into three equally-spaced portions from which the moderate/poor and poor/bad boundaries are derived.	*A. minutissimum*, *F. capucina*, *F. vaucheriae* and *N. dissipata* were present in a majority of sites at 99.5%, 69.2%, 70.6% and 70.6% respectively, but the maximum relative abundance recorded was lower than in samples at high status (62.5%, 25.7%, 26.0% and 41.6% relative abundance respectively). Other species including *G. parvulum*, *A. pediculus*, *Planothidium lanceolatum*, *Reimeria sinuata* and motile species including *N. gregaria*, *N. lanceolata*, *Navicula minima* and *N. dissipata* were present in over 70% of all samples in this status class. Concerning these species the highest maximum of relative abundance was recorded for *G. parvulum* (61.4%).
Phytoplankton	Belgium (Flanders)	Equidistant division of the EQR gradient	EQR gradient is assumed to represent a continuous trend with general degradation.	The EQR values at good status are characterised by metric values that are only slightly lower than at (expert-based) reference state, hence a slightly increased biomass per volume, a slightly disturbed distribution of proportions of different phytoplankton groups such as diatoms and green algae, and a slight increase of cyanobacteria are possible.

Table A7 (cont.): Ecological class boundary setting (general schemes)

BQE	Country	Boundary setting	Boundary description	Boundary communities
Phytoplankton	Germany	Calibrated against pre-classified sampling sites Using discontinuities in the relationship of anthropogenic pressure and the biological response.	No or few reference sites. Background limiting factor TP was reconstructed by modelling (MONERIS) for all types with a value of smaller than 50µg/l. HG boundary: On the base of background TP, the maximal potential of the phytoplankton biomass was reconstructed for each relevant river types by exponential regression line from existing data PB boundary: Point with no further effect of pressure increase (about 250µg/l TP). GM and MP boundaries by equidistant division of range between HG and PB boundary. GM boundary was checked and corrected by three methods: 1) Comparison with values in the few existing near natural sites (the 90percentile) 2) Identification of discontinuity in Chl-a / TP ratio (GM boundary was set before a sudden increase of Chl-a can be observed in a high portion (50%) of the cases in some river types. 3) Habitat reconstruction including grazing pressure (macrozoobenthos) and habitat structure (shading, charge range, flow velocity, mean water retention time etc.) and checking by historical data (River Rhine: historical museum filters from 1870 with diatom valves, River Havel: paleolimnological studies).	Above 90 µg/l TP (pressure) a sudden increase of total biomass of phytoplankton (chlorophyll a and phaeophytin) can be observed in a relevant portion of cases for the trophic sensitive national river types. The river types with high run-off or with small catchment areas 1000 - 5000km² are less sensitive to pressure, so TP concentrations less than 135 µg/l allow for good quality status.
	Hungary	High-good boundary derived from metric variability at sites in Least Disturbed Conditions	The functional groups of algae were evaluated on basis of their ecological characteristics. Nutrient status, tolerance of turbulent conditions, time sufficient for development of the given assemblage and general risk. All the groups were given a factor number (1-5). All the boundaries were set by the relative abundance of the reference (F=5) and good (F=4) taxa. These ratios were different in every river type.	Dominance of A, B, C, D, TIB (Borics et al. 2007) functional groups, in different ratio, depending on river type. Other groups can also be abundant.
	Slovakia	Equidistant division of the EQR gradient	Phytoplankton: no reference sites for large lowland rivers, derivation of reference values based on expert judgment; confirmation by calculations/statistics. Advance setting of boundaries (data from period 2001-2005): statistical values (mean of 6 measured values within vegetation period for each metric in monitored sites) were calculated for setting of boundaries. These were verified after calculations by expert judgment and compared (correlated) to chemical quality class boundaries.	In Slovakia background taxa lists are not prescribed and especially created for good status conditions as well as for any other ecological status classes.

Anhang 9: Veröffentlichung Birk & Willby (2010), Hydrobiologia

Hydrobiologia (2010) 652:149–163
DOI 10.1007/s10750-010-0327-3

PRIMARY RESEARCH PAPER

Towards harmonization of ecological quality classification: establishing common grounds in European macrophyte assessment for rivers

Sebastian Birk · Nigel Willby

Received: 14 January 2010 / Revised: 4 June 2010 / Accepted: 12 June 2010 / Published online: 26 June 2010

Abstract Different national assessment concepts impede the harmonization of river quality classifications using macrophytes in Europe. This study describes a procedure to identify similarities between the national methods for ecological quality assessment of Austria, Belgium (Flanders and Wallonia), France, Germany, Great Britain and Poland. Based on an international data set covering three European stream types we identified sites commonly assessed as high status by most methods. A mean index derived from averaging the national assessment results per stream site was then correlated with the abundance of each macrophyte taxon. We defined common macrophyte indicator scores using these correlation coefficients. This enabled the description of type-specific macrophyte communities under near-natural and degraded conditions, and the development of a common metric (mICM) that was correlated with all national methods. The weaker relations of the Flemish and German methods were improved by adjusting national indicator scores of selected macrophyte taxa that deviated from the common indicator scores. The analysis of common high status sites provided mICM reference values. This study offers a general approach to harmonize the national assessment methods for biological elements of any water category.

Electronic supplementary material The online version of this article (doi:10.1007/s10750-010-0327-3) contains supplementary material, which is available to authorized users.

Handling editor: P. Nõges

S. Birk (✉)
Department of Applied Zoology/Hydrobiology, University of Duisburg-Essen, Essen, Germany
e-mail: sebastian.birk@uni-due.de

N. Willby
School of Biological and Environmental Science, University of Stirling, Stirling, Scotland, UK

Keywords Intercalibration · Common metric · European Water Framework Directive · Common stream type · Reference community

Introduction

River macrophyte communities are determined by the characteristics of the local habitat in which they occur, namely light availability, current velocity, sediment patterns and nutrient supply. Biogeographical zone, catchment geology and stream hydrology establish the large-scale framework influencing occurrence and abundance of macrophytes (Lacoul & Freedman, 2006). Since most of these local factors are susceptible to anthropogenic alteration, macrophytes are potentially effective bioindicators that respond to various human pressures such as physical stream disturbance and alteration (Baattrup-Pedersen & Riis, 1999, 2004, O'Hare et al., 2006), river pollution (Schneider & Melzer, 2003; Hilton et al., 2006; Szoszkiewicz et al., 2006a), disruption of the riparian integrity (Ferreira et al., 2005) or general deterioration (Sand-Jensen et al., 2000). In summary, these studies reveal some general patterns of community change under human influence: Anthropogenic disturbance fosters a decline in species richness especially of submerged macrophytes, the reduction of oligotrophic and large, slow growing species, leading to fragmentation of populations and rare species threatened by extinction, and favours macrophyte taxa with high growth capacity and efficient dispersal.

Aquatic plants are an important influence on the ecology of rivers (Sand-Jensen, 1998). As a major source of primary production they interact directly with higher trophic levels in a variety of ways; they directly or indirectly determine the level of physical habitat support for macroinvertebrates and fish, and exert a significant influence on the efficiency and spatial variation in transport of both water and sediment (Dawson, 1978), as well as regulating fluxes of key nutrients (Kleeberg & Heidenreich, 2004). Combined with benthic microalgae, macrophytes thus form an obligatory element in the monitoring of ecological river quality as stipulated by the EU Water Framework Directive (WFD) 2000/60/EC (European Union, 2000). For discrete stream types the taxonomical composition and abundance of macrophytes are appraised by biological assessment methods. The status observed at the monitored river stretch is compared to the status expected under near-natural conditions. The resulting Ecological Quality Ratio (EQR) evaluates the river quality in a score ranging from 0 (worst status) to 1 (reference status). This range is divided into five classes of ecological quality: high, good, moderate, poor and bad.

The WFD requires that all water bodies should attain good ecological status within the near future. However, countries obliged to fulfil these requirements are applying different assessment methods. To achieve a common level of ambition in reaching the WFD's objective good ecological status is harmonized through the so-called "intercalibration exercise" (Heiskanen et al., 2004). The specific challenge of this exercise is to calibrate the national interpretations of good ecological status. Although the WFD provides general guidelines for the high, good and moderate quality status in the form of 'normative definitions', the practical implementation of these definitions has to be compared between the various countries.

For benthic diatoms (Kelly et al., 2009) and benthic invertebrates (Bennett et al., 2010) the national assessment methods were intercalibrated by the use of common metrics (Buffagni et al., 2007). These metrics allow the national definitions of good ecological status to be compared across different countries and stream types. Common metrics take advantage of similar assessment principles that all national methods share. For instance, anthropogenic pressure generally causes a decrease of taxa richness in invertebrate communities, making the total number of taxa a suitable common metric (Buffagni et al., 2005). However, Birk et al. (2006) reported on difficulties in finding common metrics for the intercalibration of river macrophyte methods. They identified differing national method designs: The German and Dutch classifications integrated the concept of stream type-specific macrophyte communities into assessment, the latter with a strong focus on taxa richness to evaluate community completeness (Schaumburg et al., 2004, Van der Molen & Pot, 2007). On the contrary, the British and French appraisals were derived from a global list of indicator species and related to stream types by defining near-natural reference values (Holmes et al., 1999; NF T90-395:2003). In the analysis of Birk et al. (2006) all methods except the Dutch were significantly related to the gradient of nutrient enrichment. The Dutch method showed best response to a combined gradient of physico-chemical, hydromorphological and land use parameters.

These fundamental differences raise the question, of what, if any conceptual similarities exist between the river macrophyte classifications of different countries. Expert discussions confirmed a common notion of type-specific macrophyte assemblages at high ecological status (Birk et al., 2007). Motivated by this finding, this study investigates whether this common notion can be empirically defined and captured through a bespoke common metric. In particular, our study is based on the following hypotheses:

- Certain macrophyte assemblages that occur in a common stream type are classified in high quality status by the majority of national assessment methods, regardless of the country of origin of the site.

- These assemblages include species that are regarded as indicators of near-natural conditions across national methods.
- Following this concept, indicators of general disturbance can also be identified based on data from sites that are commonly classified in poor or bad status.

These hypotheses were tested by applying seven national macrophyte assessment methods to an international data set that covered three European stream types. We correlated the abundance of individual macrophyte taxa with the average national EQR per survey and thus extracted indicators of a common high or low quality status. Based on these outcomes we

(1) described the macrophyte assemblages of each stream type under near-natural and degraded conditions,
(2) developed a common macrophyte metric expressed as a continuous variable and related it to the national methods,
(3) proposed amendments to certain national methods in order to improve their relationship with the common metric and
(4) identified type-specific reference values that allow the common metric scores to be expressed in the format of an EQR.

Methods

Data basis

In this study data on taxonomic composition and abundance of river macrophytes were used. Sampling sites were located on rivers belonging to three common stream types (ECOSTAT, 2004) that were shared by 13 countries in Central and Western Europe and the Baltic region (Table 1). The common stream types were delineated by their altitude, catchment size, geology, substrate composition and alkalinity. The types covered (i) small, sandy or (ii) medium-sized streams in the lowlands and (iii) small streams in the mountains. These delineations matched with most of the national macrophyte typologies used in quality assessment. Only Flanders and the Netherlands subdivided their lowland streams mainly into regional variants.

In total, 609 macrophyte surveys were provided by the countries listed in Table 2. The data originated from national monitoring programmes or scientific projects (e.g. Furse et al., 2006). Countries applied national macrophyte survey protocols that were in line with the requirements of the European Standard EN 14184:2003. Representative river stretches were visually inspected during the growing season (June to September) by wading, diving or boating, using rake, grapnel or aqua-scope where necessary (Birk et al., 2007). Representative sites spanned about 100 m of river length.

The macrophyte abundance was recorded in different scales (Table 3). Most countries specified the abundance as relative coverage of the surveyed area. Per cent values were graded into five, seven or nine classes. The Austrian method combined the number of single plant records per surveyed section and the plant quantity per habitat following Kohler (1978) in a five-class scheme. Germany estimated the plant quantity (Melzer et al., 1986) in one of five different classes. The Dutch abundance data were given in various scales (Braun-Blanquet, 1928, Tansley, 1946).

National assessment methods

Seven countries participated in this exercise with their national assessment methods (Table 4). Most methods

Table 1 Characterisation of the common stream types

Common stream type	Type abbreviation	Catchment area (km^2)	Altitude (m)	Geology	Channel substrate	Alkalinity (meq/l)
Sandy lowland brooks	R-C1	10–100	<200	Siliceous	Sand	>1
Siliceous mountain brooks	R-C3	10–100	200–800	Siliceous	Boulders, cobbles and gravel	<0.4
Medium-sized lowland streams	R-C4	100–1,000	<200	Mixed	Gravel and sand	>2

Table 2 Number of macrophyte surveys used in the analysis, listed per country and common stream type

Stream type	Country	Number of surveys
Sandy lowland brooks	Belgium (Flanders)	105
	Belgium (Wallonia)	1
	Denmark	15
	Germany	38
	Latvia	15
	Lithuania	1
	Netherlands	14
	Poland	11
	Total number	200
Siliceous mountain brooks	Austria	31
	Belgium (Wallonia)	43
	Czech Republic	13
	France	78
	Germany	81
	Great Britain	33
	Total number	279
Medium-sized lowland streams	Belgium (Flanders)	15
	Denmark	4
	Germany	32
	Great Britain	3
	Latvia	29
	Lithuania	9
	Luxemburg	3
	Netherlands	8
	Poland	27
	Total number	130

focussed on the assessment of specific human pressure (Austria, France, Great Britain, Poland, Wallonia). The principal component of this approach was formed by a list of indicator taxa graded by their response, mainly to nutrient enrichment. Numerical assessment results were obtained by computing a sensitivity metric, i.e. the average score of indicative species weighted by their abundance. In the case of the Austrian, French and Wallonian metrics, this also included a factor considering the taxon's ecological amplitude.

The Flemish and German methods were oriented towards the indication of non-specific anthropogenic disturbance. Besides sensitivity measures, these methods considered additional metrics, such as richness of macrophyte growth forms, or taxa richness and dominance. The basic element of the German Reference Index (RI) was the type-specific definition of reference and non-specific disturbance indicating taxa. The RI was a numerical expression of the relation of both response groups at a river site. The supplementing assessment criteria directly contributed to the score of the RI. The Flemish method integrated three metrics in the appraisal of ecological status by the "one out–all out" principle: The type-specific index for water vegetation, the perturbation index for water vegetation and the richness of various growth forms. Based on the experiences gained in earlier intercalibration studies (Birk et al., 2007) the assessment of macrophyte growth form was not considered in the main analysis. However, we additionally tested the performance of the Flemish method including the growth form metric.

Intercalibration analysis

Preparatory steps of the intercalibration analysis comprised the harmonization of the macrophyte taxonomy, especially resolving synonyms arising from national differences in taxonomic nomenclature. Furthermore, the abundance data were converted from the national into a standard international abundance scale (Table 3). The national metrics were applied to the macrophyte survey data. Using the national stream type-specific reference values all metric results were transformed into EQRs. The ecological quality of each survey was classified according to the national methods. Those surveys were identified that the majority of methods classified in high status and which none of the methods classified in moderate or worse status. These are henceforth referred to as "common high status sites".

For each common stream type the values of the national EQRs were normalized to a scale ranging from 0 to 1. These normalized values were then averaged for each survey. In the case of the two Flemish sensitivity metrics the lowest value per survey was taken ("worst case") according to the national protocol (Leyssen et al., 2005). As a result, a mean index score was assigned to each survey that was composed of the average of the normalized national metric values. Each national metric therefore

Table 3 Conversion table of national macrophyte abundance classes into the international abundance scale

Country	International abundance scale				
	1st class (rare)	2nd class (occasional)	3rd class (frequent)	4th class (abundant)	5th class (dominant)
Austria	1	2	3	4	5
Belgium (Flanders)	1	2	3	4 5	6 7
Belgium (Wallonia)	1	2	3	4	5
Czech Republic	1	2	3	4	5
Denmark	1	2	3	4	5
France	1	2	3	4	5
Germany	1	2	3	4	5
Great Britain Poland	1	2	3 4	5 6	7 8 9
Latvia	1	2	3	4	5
Lithuania	1	2	3	4	5
Luxembourg	1	2	3	4	5
The Netherlands (Braun-Blanquet, 1928)	1 2	3 4	5	6 7	8 9
The Netherlands (Tansley, 1946)	1	2	3 4	5 6	7 8

Table 4 National assessment methods using macrophytes in rivers

Country	Name of method	Intercalibrated assessment metric(s)	Relevant stream type(s)	Literature reference
Austria	Austrian Index for Macrophytes in Rivers (AIM Rivers)	Single metric combining ecological preference and abundance	Siliceous mountain brooks	BMLFUW (2006)
Belgium (Flanders)	MAFWAT (Makrophyten Waterlopen)	(1) Type-specific index for water vegetation (TSw) (2) Perturbation index (organic pollution, eutrophication) for water vegetation (Vw)	Sandy lowland brooks, Medium-sized lowland streams	Leyssen et al. (2005)
Belgium (Wallonia), France	Indice Biologique Macrophytique en Rivière (IBMR)	Single metric combining occurrence (indicator value per taxon), ecological amplitude and abundance	Siliceous mountain brooks, Medium-sized lowland streams (only France)	NF T90-395:2003
Germany	Deutsches Bewertungsverfahren für Makrophyten und Phytobenthos (PHYLIB)	Index relating Species Response Groups (Reference, Disturbance, Indifferent) plus additional criteria for Siliceous mountain brooks: acidification module; medium-sized lowland streams: evenness, number of submerged taxa, ratio of *Myriophyllum spicatum* and *Ranunculus* sp.	Sandy lowland brooks, Siliceous mountain brooks, Medium-sized lowland streams	Schaumburg et al. (2006)
Great Britain	River Nutrient Macrophyte Index (RNMI)	Single metric combining occurrence (indicator value per taxon) and abundance	Siliceous mountain brooks, Medium-sized lowland streams	Willby et al. (2009)
Poland	Macrophyte Index for Rivers (MIR)	Single metric combining occurrence (indicator value per taxon) and abundance	Sandy lowland brooks, Medium-sized lowland streams	Szoszkiewicz et al. (2006b)

had an equal contribution to the mean index score, the latter representing a global gradient of ecological quality.

In the following step this mean index was correlated with the abundance of each macrophyte taxon recorded in the surveys using the international abundance scale and including zero abundance (i.e. absences). The linear relation of taxa abundance to the mean index was quantified by Spearman's coefficient of correlation. The analysis yielded a coefficient for each taxon and comprised a spectrum of values identifying taxa correlated—either positively or negatively—or not correlated to the average national assessment results. Positive correlations indicated an increase in abundance with increasing mean index values (i.e. increasing quality), whilst negative correlations indicated an increase in abundance with decreasing mean index values (i.e. worsening quality).

We used the correlation coefficients to define taxon-specific indicator scores. These scores were assigned only to taxa records at species level except for selected algae and mosses. Taxa with only one record in the database or typically riparian taxa, such as *Filipendula ulmaria*, *Ranunculus repens* or *Urtica dioica*, were given no indicator score. We rescaled the coefficients of scoring taxa based on the maximum or minimum correlation separately for each common stream type. For instance, if the range of correlation coefficients was from −0.3 to +0.5 we rescaled from −0.5 to +0.5 which produced an actual range running from −0.6 to +1.0 with a zero score coinciding with a zero correlation. The indicator scores were used in a common type-specific, weighted average metric, the "macrophyte Intercalibration Common Metric" (mICM) following the terminology of Buffagni et al. (2005):

$$\mathrm{mICM}_x = \frac{\sum (s_i * \mathrm{abd}_i)}{\sum \mathrm{abd}_i}$$

where mICM_x was the macrophyte Intercalibration Common Metric value of a survey at the common stream type x, s_i was the taxon-specific correlation value of the ith taxon and abd_i was the international abundance class of the ith taxon.

The mICM was plotted against each national metric per common stream type. Linear regression models were applied, and the resulting coefficients of determination (R^2) were checked. In case of R^2 values <0.5 we compared the mICM scoring taxa list and the national indicator list. Obvious discrepancies between both lists were adjusted by proposing small amendments to the scoring of species within the national list. However, we focussed only on those amendments that allowed for an increase of the R^2 value $\geq$0.5 in the regression analysis. Furthermore, to demonstrate the performance of the Flemish method including the growth form metric the mICM was also correlated with the worst case of the three Flemish metrics.

In using the average of national assessments as a global quality gradient it is of course important that the variability in EQRs between national methods at a given site is as low as possible since this source of uncertainty is ignored when subsequently determining the correlations between each species and the global mean. Thus, a high correlation between a species and the global mean may not translate to a high correlation between the mICM and a national method if different countries diverge widely in their assessment of individual sites. In our study the coefficient of variation (standard error expressed as a percentage of the site mean) averaged from 9 to 29% in the different river types, with a global mean of 20%. The latter value is usually considered to represent an acceptable level of precision in ecological studies. High variability in national EQRs will be reflected in low levels of relatedness of national methods. This issue must therefore be addressed before any harmonization of national classifications can proceed.

For each stream type we determined the median mICM value from the pool of common high status sites. This value served as the common stream type-specific reference by which the mICM was transformed into an EQR. To characterize the distribution of mICM EQR values amongst the common high status sites we calculated the 5th and 10th percentile values. Lower percentiles, such as these, have been widely used in invertebrate classification as a statistical basis for the lower limit of high status sites (e.g. Clarke et al., 1996).

For the common high status sites in each type we then prepared a narrative describing the biological assemblage, drawing on available pressure data, published accounts of the autecology of individual taxa or the general response of aquatic vegetation to pressure gradients (e.g. Butcher, 1933; Haslam, 1978; Riis et al., 2000; Demars & Harper, 2002; O'Hare et al., 2006), and contrasting the assemblage with that found in sites that were classified by the majority of countries as being at poor or bad status. The narrative description of the biology at common high status sites should be seen as an important step in providing a guiding image of unimpacted biology.

Results

The common metric for intercalibration

For each common stream type Table 5 displays the list of macrophyte species significantly correlated to the mean index gradient. In total, 102 (sandy lowland brooks), 140 (siliceous mountain brooks), and 110 (medium-sized lowland streams) indicator taxa were defined. All indicator taxa and their rescaled indicator scores are shown in the Supplementary material. The relationships of the mean index score and the abundance of three selected species for the mountain brooks' data set is illustrated in Fig. 1.

Coefficients of determination (R^2) obtained in the regression analyses of the national metrics against the mICM range from 0.28 to 0.78. On average, the British metric is best correlated with the mICM whilst the German method shows the weakest overall relationship. Highest R^2 values are obtained in the analysis of the siliceous mountain brooks. The extent of correlation between each national method and the mICM is a reflection of relatively low uncertainty in the assessment of individual sites as indicated by the variation in national EQRs.

To improve weak relationships of the mICM with the national metrics we adjusted the national indicator lists of Flanders (including additional disturbance taxa in metric Vw for the sandy lowland brooks) and Germany (re-scoring of indicator taxa for the medium-sized lowland streams). Both adjustments lead to coefficients of determination ≥ 0.5 in the regression analysis. Adjustments and results of the regression analyses are specified in Table 6.

Stream type-specific macrophyte communities at high and low quality

In total, 111 common high status sites are identified, with most of these relating to the mountain brooks (Table 7). The mICM median values of these sites show a clear difference, distinguishing between the two lowland types on one hand and the mountain brooks on the other. However, the percentiles of the mICM EQR value distributions are rather similar, ranging from 0.79 to 0.84. This indicates that the spread of values is consistently narrow, and that the unit of turnover of the reference population (~ 0.2) would be appropriate for establishing a series of lower class boundaries. In so far as high status is defined as representing minimal departure in composition from unimpacted reference conditions this also implies that a sufficient degree of convergence exists between countries in their view of reference biological assemblages, despite the limited number of contemporary examples to draw upon. A narrative describing each stream type at high and low status is given below.

Sandy lowland brooks

The highest quality sites feature a combination of submerged rooted aquatic species of which *Callitriche hamulata*, *Potamogeton natans* and *Sparganium emersum* are by far the commonest. Scarcer associated species, which characterize high status sites include *Potamogeton alpinus*, *Myriophyllum alterniflorum*, *Elodea canadensis*, various other *Callitriche* spp. and *Ranunculus peltatus*. Emergent vegetation is dominated by *Phalaris arundinacea*, *Sparganium erectum* and *Phragmites australis*, the latter characteristic of high status sites. There is a range of moderate- to small-sized emergent species, of which *Persicaria hydropiper*, *Myosotis scorpioides*, *Glyceria fluitans*, *Berula erecta*, *Mentha aquatica* and *Veronica anagallis-aquatica* are the most abundant, but it is the less frequent elements, such as *Cardamine amara* and *Caltha palustris* that are most characteristic of high status sites. This assemblage is most likely to be associated with small, active, mesotrophic, shallow, sand-dominated, clear water, moderately fast-flowing, partially shaded streams. The very limited number of characteristic species suggests that this is a type with several geographically distinct variants under high status conditions.

With declining quality there is a shift to a community dominated by *Sparganium emersum*, alongside a range of species that are absent from or much scarcer in the highest status sites, including *Potamogeton pectinatus*, *P. trichoides*, *P. perfoliatus* and *P. crispus*, and the duckweeds *Lemna minor* and *L. minuta*. Amongst the emergent species that overlap with high status sites *Sparganium erectum*, *Persicaria hydropiper*, *Phalaris arundinacea*, *Phragmites australis*,

Table 5 List of macrophyte species significantly correlated to the mean index gradient ($P < 0.001$) (C_{Sp} = Spearman rank correlation coefficient, n = number of taxon records in the data set)

Stream type	Taxon name	C_{Sp}	n
Sandy lowland brooks	*Callitriche hamulata*	0.457	58
	Potamogeton trichoides	−0.235	38
	Lemna minuta	−0.238	33
	Rorippa amphibia	−0.240	40
	Potamogeton perfoliatus	−0.248	12
	Lemna minor	−0.336	91
	Potamogeton pectinatus	−0.338	22
Siliceous mountain brooks	*Scapania undulata*	0.611	78
	Racomitrium aciculare	0.482	37
	Pellia epiphylla	0.366	27
	Chiloscyphus polyanthos	0.363	112
	Brachythecium plumosum	0.348	17
	Hyocomium armoricum	0.342	14
	Fontinalis squamosa	0.321	62
	Hygrohypnum ochraceum	0.306	34
	Marsupella emarginata	0.260	8
	Brachythecium rivulare	0.237	69
	Ranunculus flammula	0.208	11
	Sphagnum sp.	0.203	5
	Melosira sp.	−0.179	38
	Myriophyllum spicatum	−0.196	6
	Rhizoclonium sp.	−0.202	7
	Oedogonium sp.	−0.243	28
	Lemna minor	−0.250	10
	Fontinalis antipyretica	−0.252	144
	Agrostis stolonifera	−0.257	63
	Cladophora sp.	−0.262	21
	Phalaris arundinacea	−0.330	113
	Amblystegium riparium	−0.529	84
Medium-sized lowland streams	*Fontinalis antipyretica*	0.488	45
	Hildenbrandia sp.	0.416	18
	Potamogeton alpinus	0.296	12
	Mentha aquatica	0.292	37
	Oedogonium sp.	−0.286	6
	Sagittaria sagittifolia	−0.299	39
	Lemna minor	−0.304	65
	Potamogeton pectinatus	−0.523	26

Myosotis scorpioides and *Berula erecta* are all much reduced, and are typically replaced by *Rorippa amphibia*, *Glyceria aquatica*, *Sagittaria sagittifolia* and *Alisma plantago-aquatica*. This change in structure suggests a shift to silty, stable, eutrophic, slow flowing, turbid conditions in managed or regulated channels with degraded riparian habitat.

Siliceous mountain brooks

High quality sites feature a combination of leafy liverworts (*Scapania undulata* or *Chiloscyphus polyanthus*, and less frequently, *Marsupella emarginata* or *Jungermannia atrovirens*), acrocarpous mosses (most notably *Racomitrium aciculare*, plus smaller

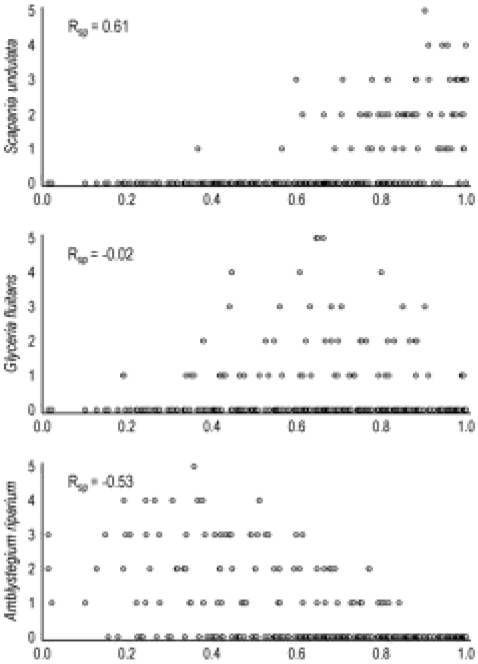


Fig. 1 Relationships between the mean index scores and the abundance of three macrophyte taxa for the mountain brooks' data set (R_{sp} = Spearman's coefficient of correlation)

quantities of marginal species, such as *Philonotis fontana* and *Dicranella palustris*, *Fissidens crassipes* and *F. rufulus*), thallose liverworts (*Pellia epiphylla*), and small macroalgae, including *Lemanea*, *Oscillatoria* and *Mougeotia* spp. These taxa occur against a backdrop of extensive growths of a range of pleurocarpous mosses, including *Rhynchostegium riparioides*, *Fontinalis squamosa*, *F. antipyretica*, *Hygrophypnum ochraceum* (and occasionally *H. luridum*), *Brachythecium rivulare*, *B. plumosum*, *Hyocomium armoricum*, *Thamnobryum alopecurum* and *Amblystegium fluviatile*. Vascular plants are likely to be restricted to *Callitriche hamulata*, plus occasional marginal growth of species such *Glyceria fluitans*, *Phalaris arundinacea* and *Ranunculus flammula*. The latter often occurs alongside a range of mire-forming mosses including *Sphagnum*, *Philonotis caespitosa* and *Plagiomnium undulatum*. This is an assemblage of small, shallow, turbulent, flashy, neutral to base-poor, oligotrophic, upland rivers, with a cobble and boulder substrate, often with extensive shading by deciduous trees.

Several of these species persist in the lowest quality sites, most notably *Fontinalis antipyretica* and *Rhynchostegium riparioides*, but most bryophytes are replaced by *Amblystegium riparium or attached growths of large green filamentous algae, such as Cladophora glomerata* and *Vaucheria* spp. Channel margins are likely to feature more extensive growth of *Phalaris arundinacea*, plus *Sparganium erectum* and a range of smaller amphibious species, such as *Agrostis stolonifera*, *Glyceria fluitans*, *Veronica beccabunga* and *Myosotis scorpioides*. The cover of instream vascular species is generally small, but may include *Callitriche hamulata*, *Ranunculus peltatus*, *Elodea nuttallii*, *Potamogeton crispus*, *Sparganium emersum* and *Ceratophyllum demersum*. This assemblage reflects a shift to more stable, moderate to slow flowing, fertile conditions, with reduced shading of the margins and mixed sand–gravel substrates. This change is therefore most likely to be associated with a combination of pollution and siltation from diffuse sources, flow regulation, channel realignment and overgrazing.

Medium-sized lowland streams

High quality sites are dominated by two species *Fontinalis antipyretica* and *Sparganium emersum*, each of which account for 10% of the total plant cover. Other common and widely distributed instream aquatics include *Nuphar lutea*, *Elodea canadensis* and *Amblystegium riparium*, plus the red encrusting alga *Hildenbrandia*, which is highly characteristic of these high status lowland streams. A diverse range of pondweed species (especially *Potamogeton alpinus*, *P. perfoliatus* and *P. natans*, but occasionally *P. praelongus* or *P. gramineus*) occur alongside batrachids, such as *Ranunculus fluitans* and *R. aquatilis*, plus *Myriophyllum spicatum* and *Callitriche hamulata*. Marginal vegetation is dominated by *Phalaris arundinacea* and *Sparganium erectum*, but other stand forming species are also frequently present and will include *Scirpus lacustris*, *Iris pseudacorus*, *Glyceria aquatica* and *Equisetum fluviatile*. Of the smaller marginal species *Mentha aquatica*, *Veronica anagallis-aquatica*, *Alisma plantago-aquatica*, *Glyceria fluitans*, *Berula erecta* and *Myosotis scorpioides* are especially well represented,

Table 6 Results of the linear regression analysis of mICM against the national metrics

Stream type	Country	R^2 (orig.)	R^2 (amend.)	Specification of amendment
Sandy lowland brooks	Belgium (Flanders)	0.28 (0.21)	0.50 (0.24)	Additional disturbance indicators in metric Vw: *Nymphoides peltata*, *Potamogeton berchtoldii*, *Potamogeton crispus*, *Potamogeton perfoliatus*, *Rorippa amphibia*, *Sagittaria sagittifolia*, *Sparganium emersum*
	Germany	0.55	–	–
	Poland	0.62	–	–
Siliceous mountain brooks	Austria	0.74	–	–
	Belgium (Wallonia)	0.77	–	–
	France	0.76	–	–
	Germany	0.54	–	–
	Great Britain	0.78	–	–
Medium-sized lowland streams	Belgium (Flanders)	0.59 (0.02)	–	–
	Germany	0.17	0.59	Rescoring of indicator taxa: *Fontinalis antipyretica* (Reference), *Lemna minor* (Disturbance)
	France	0.58	–	–
	Great Britain	0.62	–	–
	Poland	0.58	–	–

R^2 (orig.): coefficient of determination using national index with original indicator taxa list; R^2 (amend.): coefficient of determination using national index with amended indicator taxa list; R^2 values in brackets: coefficient of determination including Flemish growth form assessment

Table 7 Number of common high status sites (*N*), mICM reference values (REF), and 5th and 10th percentile values of the mICM EQR distributions

Stream type	*N*	REF	5th percentile	10th percentile
Sandy lowland streams	27	0.15	0.83	0.84
Siliceous mountain brooks	63	0.36	0.79	0.83
Medium-sized lowland streams	21	0.13	0.80	0.84
All data combined	111	–	0.79	0.83

and may harbour small patches of various lemnids. Within the marginal zone *Carex rostrata* and *Lysimachia thyrsiflora* are uncommon but are unique to high status sites. This is an assemblage of medium-sized, active, moderate to fast-flowing, shallow lowland rivers on neutral to base-rich geology with clear, mesotrophic to eutrophic water. A diversity of substrates occurs, and will include a mix of sand, gravel and unsilted coarser material. The vegetation itself is a major architect of hydromorphological diversity. Remnants of this vegetation occur in rivers in central and north west Europe (e.g. Wiegleb, 1984; Holmes et al., 1999), but it is only in the less densely populated countries of north east Europe that this vegetation can still be found with any regularity (e.g. Paal & Trei, 2004; Baattrup-Pedersen et al., 2008).

The most degraded sites are strongly characterized by *Potamogeton pectinatus*, which accounts for 21% of the total plant cover in common poor or bad status sites. Of the commoner instream associates *Fontinalis antipyretica*, *Elodea canadensis* and *Sparganium emersum* all persist but at greatly reduced cover, and are likely to be joined or replaced by *Potamogeton crispus*, *Elodea nuttallii*, *Lemna minor*, *Ranunculus penicillatus*, *Ceratophyllum demersum*, *Persicaria amphibia* or *Zannichellia palustris*, plus various large green filamentous algae, including *Cladophora*, *Rhizoclonium*, *Vaucheria* and *Oedogonium*. The status of

Nuphar lutea and *Amblystegium riparium* is little changed in comparison to the highest status sites. The margins remain dominated by *Phalaris arundinacea* and *Sparganium erectum* with *Glyceria aquatica* and *Scirpus lacustris* as common associates. However, in place of a range of smaller herbaceous species *Solanum dulcamara*, *Rorippa amphibia*, *Sagittaria sagittifolia* or *Typha latifolia* normally occur. Collectively this assemblage indicates a highly enriched, stable, sluggish, well-lit environment dominated by fine sediment. Such vegetation is often associated with streams in urban or intensely agricultural catchments, where management and physical modification of channels and their margins are the norm.

Discussion

Development of a common metric for intercalibration

Despite superficial differences in national methods our study confirms a strong underlying view of the highest and lowest quality sites which can be extracted as a quality gradient within different stream types. The mICM thus proves to be a suitable vehicle via which to compare national classifications of river macrophytes. Several previous studies have attempted to compare indicator scores of individual plant species across different river macrophyte assessment methods used within Europe (e.g. Szoszkiewicz et al., 2006a; Schneider, 2007), but such studies have focussed exclusively on trophic indices, have generally operated across types, or within very broad types (e.g. upland versus lowland) and have therefore avoided issues of benchmarking indices against concepts of minimal impact. Of course in creating an mICM to harmonize different national methods it is important to understand that this does not, in itself, legitimize the use of macrophytes as disturbance indicators, it merely successfully captures the global trend inherent in the different national classifications.

In terms of ability to capture the essence of national methods, except for two out of 13 cases, all regressions between the mICM and national EQR are characterized by a coefficient of determination ≥ 0.5, thus meeting an important intercalibration criterion given by Bennett et al. (2010) in the comparison of benthic invertebrate methods. On average, the mICM is related more strongly to the national methods than the common metric proposed by Birk et al. (2006). However, compared to the outcomes of a diatom intercalibration exercise its performance is poorer (Kelly et al., 2009). This is mainly attributable to the low average relation of the mICM to the Flemish and German methods, underlining their conceptual difference from the other countries.

Our approach to developing a common metric for intercalibration allows differences between methods to be detected at the level of national indicator lists. The mICM taxa scores actually represent a linear correlation with averaged national indicator ratings. Positive scoring taxa are common indicators of near-natural conditions, distinguished by an increase in their absolute abundance with increasing quality, whilst negative scoring taxa tend to increase in cover with declining quality. Taxa with a low correlation to the mean index are either ubiquitous, indicative of moderate conditions or are regarded inconsistently amongst countries. Obvious discrepancies between the mICM indicator values and the national ratings are easily identified and adjusted. This option provides the opportunity to harmonize the national methods by implementing only minor, thus easily justified changes, rather than seeking wholesale changes in class boundaries which politically may be more difficult to achieve. However, in applying our approach, it is critical that national methods display an adequate level of relatedness from the outset. Without this the variation in EQR between different methods at any given site will be so high that the final relationship between the mICM and one or more national methods may be too poor for the index to have any practical value in terms of harmonization.

The mICM indicator scores were derived by Spearman rank correlation. In selecting this analysis we assumed that a linear model best describes the distribution of taxa abundance across the quality gradient represented by the mean index range. Though not further described in this article we also tested if a unimodal approach using weighted averaging of taxa abundances was a more suitable method to derive indicator scores. Due to lower correlations of the mICM with the national assessment methods this option was rejected. The linear model thus seems to better reflect the similarity amongst national sensitivity metrics that indicator taxa, when abundant, contribute proportionally more to the national EQR score.

We also dismissed alternative approaches based on relative macrophyte abundance data or downweighting rare taxa (e.g. the exclusion of taxa occurring in less than five surveys) since both options also provided weaker relationships. All national metrics use absolute instead of relative abundance data to calculate the national assessment results. The indicator scores assigned to rare taxa may be biased by their intrinsically low occurrence. It is possible that their scores would change when using a larger data set. This especially applies to the data sets of the lowland types that feature very few high quality sites. However, the better correlation of the mICM with the individual national EQRs justifies the inclusion of rare indicator taxa. And in general, rare taxa are not likely to play a significant role in intercalibration anyway, because they are either transient, relevant in one country, or they occur in low abundances and therefore make a low contribution to the site score in national assessments.

Stream type-specific macrophyte communities

The identification of common high status sites confirms our first hypothesis. For each stream type several surveys are classified in high quality status by most or all of the national methods. The smaller relative number of such surveys at lowland sites can be attributed to the generally more degraded condition of lowland rivers across Europe and emphasizes the difficulty in finding contemporary examples of minimally impacted sites on lowland rivers (Baattrup-Pedersen et al., 2006). The strong linear relation between the abundance of certain macrophyte taxa and the mean index supports the existence of common indicator species. These findings allow stream type-specific communities and their environment to be defined under near-natural and degraded conditions, as described above.

The method used here for deriving indicator scores may explain several apparent anomalies in values of selected species compared with previous reports concerning lowland rivers. For example, *Sparganium emersum* and *Elodea canadensis* tended to be associated with higher status sites in both lowland river types, and were also reported as having a high frequency in unimpacted lowland rivers by Baattrup-Pedersen et al. (2006, 2008); yet, both species are also regarded as indicators of impact on lowland rivers in Denmark (Baattrup-Pedersen & Riis, 2004). Thus, because our index is based on absolute cover, a species may decline with decreasing quality, reflecting a reduction in available habitat (whether due to physical change or expansion of a competitor), even if the relative contribution of that species to the aquatic vegetation as a whole actually increases due to the decrease in cover or loss of other species. Conversely, some species that increase in absolute cover in impacted sites may still be a normal and important yet subordinate component of the vegetation in high quality sites. For example, *Sagittaria sagittifolia* was reported to be a common species in historical data sets covering lowland streams in Denmark (Baattrup-Pedersen et al., 2008); yet, in the UK *Sagittaria* will commonly persist in highly enriched and canalized rivers (Preston & Croft, 1997). It may also be the case that the national view of some species relates to a specific type of impact. For example, on Danish streams *Sparganium emersum* consistently increases in abundance in response to weed cutting (Baattrup-Pedersen & Riis, 2004). However, weed cutting may represent a relatively 'specialist' impact on lowland streams on a European scale. A further consideration is benchmarking of the term 'impacted'; a moderate status site is impacted relative to a high status site yet is still less impacted than a poor or bad status site. Labels such as 'impact' species should therefore be used cautiously without qualification. It should also be noted that other descriptions of the macrophyte assemblages of unimpacted streams in Europe (e.g. Baattrup-Pedersen et al., 2006, 2008) have either been based on very small numbers of sites, have not followed the same typology as that used here, or have not drawn sites from the same population of countries. Comparisons may therefore be of limited value.

Many species identified here as being strongly associated with unimpacted conditions are consistent with previous studies, either of contemporary data sets or of historical change at individual sites (e.g. Braithwaite et al., 2006). For example, the decline of broad-leaved pondweed species in lowland rivers is a recurrent theme in numerous studies of long-term change in aquatic vegetation (Sand-Jensen et al., 2000; Braithwaite et al., 2006; Schütz et al., 2008). Thus, in the medium-sized lowland streams and/or the sandy lowland brooks *Potamogeton alpinus* and *P. praelongus* have large positive mICM values. Other species, such as *Potamogeton pectinatus*, *Ceratophyllum demersum* or *Lemna minor*, which tend to dominate in highly enriched or engineered sites and therefore attract a low score in all national methods, are assigned a large negative mICM value across all the types in which they occur. Meanwhile several other widely distributed species have a contrasting mICM value in different stream types. For example, *Fontinalis antipyretica*, which is often associated with moderate enrichment, is a negative indicator in mountain streams, as one would expect, yet is associated with higher quality lowland rivers when it occurs in association with other species.

Taking the population of common high status sites in each type a common value of ~0.8 at the 5th percentile of the mICM EQR provides some reassurance over the potential utility of our approach and demonstrates that there is sufficient commonality in interpretation of high status within each river type for this view to form a robust basis for testing national classifications. A 5th percentile of EQR of ~0.8 is consistent with reference site EQR variability in invertebrate-based classification tools, such as RIVPACS (Clarke et al., 1996). It also lends itself to a statistically based placement of class boundaries from High–Good, down to Poor–Bad, at unit intervals of 0.2.

Integrating additional aspects of national methods

Growth form metrics add a new dimension to macrophyte-based classification which departs strongly from structural assessments in which the indicator value of individual taxa takes priority. Thus, species which share the same growth form may have very different indicator values (and will thus tend to have different mICM scores), whilst other species representing different growth forms may have similar indicator values. Further study is required to determine how best to integrate elements of national classification methods not shared by other countries within the approach presented here. Thiebaut et al. (2002) reported that the performance of diversity-based measures was generally inferior to trophic indices for use in macrophyte classification of rivers. However, Willby et al. (2009) argued that some form of diversity index was desirable within classification in order to differentiate between results based on data of contrasting biological quality, and also to better reflect physical habitat heterogeneity. In this regard the effect of taxon richness on the mICM scores especially amongst common high status sites will need to be further investigated in the continuing intercalibration work.

Conclusions

This study represents an important contribution to the intercalibration of river macrophyte classifications in Europe. It defines common reference conditions for three widespread stream types and provides a means to compare the good ecological status of national methods. Furthermore, this study offers a general approach to harmonize the national assessment methods for biological elements of any water category. Based on the differing national assessments of similar transnational ecotypes the approach reveals the common ground of national quality classifications. Basic elements are a biological data set that is compatible with the application of a range of national methods, the emergent common high status sites and the mICM indicator list.

The description of the ecotype-specific communities and their environmental conditions goes beyond these outcomes. It amalgamates the national notions of biological communities at high and bad quality status on which national methods are contingent, and establishes an international guiding image that is not influenced by national specialities or biogeographical differences. This image will be of crucial importance in the follow-up process towards harmonization of ecological quality classification.

Acknowledgements The authors are grateful to the following colleagues that kindly provided macrophyte data: Karin Pall (Austria); Lenka Tajmrova (Czech Republic); Annette Baattrup-Pedersen (Denmark); Kris van Looy, Anik Schneiders, An Leyssen and Luc Denys (Flanders, Belgium); Christian Chauvin (France); Jochen Schaumburg and Christine Schranz (Germany); Jo-Ann Pitt (Great Britain); Normunds Kadikis (Latvia); Zoñja Sinceviciene (Lithuania); Alain Dohet (Luxembourg); Hugo Coops and Roelf Pot (Netherlands); Agnieszka Kolada (Poland) and Daniel Galoux (Wallonia, Belgium). Christian Chauvin is also thanked for his collaboration in the preparatory steps of this study. Our study significantly benefitted from the worthwhile discussions held amongst the experts of the working group for river macrophyte intercalibration of the Central-Baltic Geographical Intercalibration Group. We acknowledge useful comments on the manuscript by the reviewers. This study was financially supported by the German Bund-/Länderarbeitsgemeinschaft Wasser, the Scotland and Northern Ireland Forum for Environmental Research, and the Environment Agency of England and Wales.

References

Baattrup-Pedersen, A. & T. Riis, 1999. Macrophyte diversity and composition in relation to substratum characteristics in regulated and unregulated Danish streams. Freshwater Biology 42: 375–385.

Baattrup-Pedersen, A. & T. Riis, 2004. Impacts of different weed cutting practices on macrophyte species diversity and composition in a Danish stream. River Research and Applications 20: 103–114.

Baattrup-Pedersen, A., K. Szoszkiewicz, R. Nijboer, M. O'Hare & T. Ferreira, 2006. Macrophyte communities in unimpacted European streams: variability in assemblage patterns, abundance and diversity. Hydrobiologia 566: 179–196.

Baattrup-Pedersen, A., G. Springe, T. Riis, S. E. Larsen, K. Sand-Jensen & L. M. K. Larsen, 2008. The search for reference conditions for stream vegetation in northern Europe. Freshwater Biology 53: 1890–1901.

Bennett, C., R. Owen, S. Birk, A. Buffagni, S. Erba, N. Mengin, J. Murray-Bligh, G. Ofenböck, I. Pardo, W. van de Bund, F. Wagner & J.-G. Wasson, 2010. Bringing European river quality into line: an exercise to intercalibrate macro-invertebrate classification methods. Submitted to Hydrobiologia.

Birk, S., N. Willby, C. Chauvin, H. C. Coops, L. Denys, D. Galoux, A. Kolada, K. Pall, I. Pardo, R. Pot & D. Stelzer, 2007. Report on the Central Baltic River GIG Macrophyte Intercalibration Exercise, June 2007. University of Duisburg-Essen, Essen: 82 pp.

Birk, S., T. Korte & D. Hering, 2006. Intercalibration of assessment methods for macrophytes in lowland streams: direct comparison and analysis of common metrics. Hydrobiologia 566: 417–430.

BMLFUW, 2006. Leitfaden für die Erhebung der Biologischen Qualitätselemente. Arbeitsanweisung Fließgewässer. A4-01a Qualitätselement Makrophyten: Felderhebung, Probenahme, Probenaufbereitung und Ergebnismitteilung. Dezember 2006. Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasserwirtschaft (BMLFUW), Vienna: 66 pp.

Braithwaite, M. E., R. W. Ellis & C. D. Preston, 2006. Change in the British Flora 1987–2004. British Wildlife 17: 381–390.

Braun-Blanquet, J., 1928. Pflanzensoziologie. Grundzüge der Vegetationskunde. Springer, Berlin: 631 pp.

Buffagni, A., S. Erba, S. Birk, M. Cazzola, C. Feld, T. Ofenböck, J. Murray-Bligh, M. T. Furse, R. T. Clark, D. Hering, H. Soszka & W. van de Bund, 2005. Towards European Inter-calibration for the Water Framework Directive: procedures and examples for different river types from the E.C. project STAR. Quaderin dell'Istituto di Ricerca Sulle Acque 123: 1–467.

Buffagni, A., S. Erba & M. T. Furse, 2007. A simple procedure to harmonize class boundaries of assessment systems at the pan-European scale. Environmental Science and Policy 10: 709–724.

Butcher, R. W., 1933. Studies on the ecology of rivers I. On the vegetation distribution of macrophyte vegetation in the rivers of Britain. Journal of Ecology 21: 58–91.

Clarke, R. T., M. T. Furse, J. F. Wright & D. Moss, 1996. Derivation of a biological quality index for river sites: comparison of the observed with the expected fauna. Journal of Applied Statistics 23: 311–332.

Dawson, F. H., 1978. The seasonal effects of aquatic plant growth on the flow of water in a stream. EWRS 5th Symposium on Aquatic Weeds. EWRS, Amsterdam: 71–78.

Demars, B. O. L. & D. M. Harper, 2002. Phosphorus in a lowland calcareous river basin: a multiscale approach to understanding human impacts. International Journal of Ecohydrology and Hydrobiology 2: 149–164.

ECOSTAT (CIS WG 2.A Ecological Status), 2004. Overview of common intercalibration types. Final version for finalisation of the intercalibration network spring 2004. Version 5.1 – April 23rd, 2004. Joint Research Centre, Ispra: 37 pp.

EN 14184:2003. Water Quality – Guidance Standard for the Surveying of Aquatic Macrophytes in Running Waters. European Committee for Standardization, Brussels.

European Union, 2000. Directive 2000/60/EC of the European Parliament and of the council of 23 October 2000 establishing a framework for community action in the field of water policy. Official Journal of the European Communities L 327: 1–72.

Ferreira, M. T., P. M. Rodríguez-González, F. C. Aguiar & A. Albuquerque, 2005. Assessing biotic integrity in Iberian rivers: development of a multimetric plant index. Ecological Indicators 5: 137–149.

Furse, M., D. Hering, O. Moog, P. Verdonschot, R. K. Johnson, K. Brabec, K. Gritzalis, A. Buffagni, P. Pinto, N. Friberg, J. Murray-Bligh, J. Kokes, R. Alber, P. Usseglio-Polatera, P. Haase, R. Sweeting, B. Bis, K. Szoszkiewicz, H. Soszka, G. Springe, F. Sporka & I. Krno, 2006. The STAR project: context, objectives and approaches. Hydrobiologia 566: 3–29.

Haslam, S. M., 1978. River Plants of Western Europe: The Macrophytic Vegetation of Watercourses of the European Economic Community. Cambridge University Press, Cambridge: 512 pp.

Heiskanen, A.-S., W. van de Bund, A. C. Cardoso & P. Nõges, 2004. Towards good ecological status of surface waters in Europe – interpretation and harmonisation of the concept. Water Science and Technology 49: 169–177.

Hilton, J., M. O'Hare, M. J. Bowes & J. I. Jones, 2006. How green is my river? A new paradigm of eutrophication in rivers. Science of the Total Environment 365: 66–83.

Holmes, N. T., P. J. Boon & T. A. Rowell, 1999. Vegetation communities of British rivers: a revised classification. Joint Nature Conservation Committee, Peterborough: 120 pp.

Kelly, M. G., C. Bennett, M. Coste, C. Delgado, F. Delmas, L. Denys, L. Ector, C. Fauville, M. Ferreol, M. Golub, A. Jarlman, M. Kahlert, J. Lucey, B. Ni Chathain, I. Pardo, P. Pfister, J. Picinska-Faltynowicz, J. Rosebery, C. Schranz, J. Schaumburg, H. van Dam & S. Vilbaste, 2009. A comparison of national approaches to setting ecological status boundaries in phytobenthos assessment for the European Water Framework Directive: results of an intercalibration exercise. Hydrobiologia 621: 169–182.

Kleeberg, A. & M. Heidenreich, 2004. Release of nitrogen and phosphorus from macrophyte stands of summer dried out sediments of a eutrophic reservoir. Archiv für Hydrobiologie 159: 115–136.

Kohler, A., 1978. Methoden der Kartierung von Flora und Vegetation von Süßwasserbiotopen. Landschaft + Stadt 10: 73–85.

Lacoul, P. & B. Freedman, 2006. Environmental influences on aquatic plants in freshwater ecosystems. Environmental Reviews 14: 89–136.

Leyssen, A., P. Adriaens, L. Denys, J. Packet, A. Schneiders, K. Van Looy & L. Vanhecke, 2005. Toepassing van verschillende biologische beoordelingssystemen op Vlaamse potentiële interkalibratielocaties overeenkomstig de Europese kaderrichtlijn water: partim 'Macrofyten'. Instituut voor Natuurbehoud, Brussels: 178 pp.

Melzer, A., R. Harlacher, K. Held, R. Sirch & E. Vogt, 1986. Die Makrophytenvegetation des Chiemsees. Informationsbericht des Bayerischen Landesamts für Wasserwirtschaft 4/86. Bavarian Water Management Agency, Munich: 204 pp.

NF T90-395:2003, Water quality – determination of the macrophytes biological index for rivers (IBMR). Association Française de Normalisation (AFNOR), Saint Denis La Plaine.

O'Hare, M. T., A. Baattrup-Pedersen, R. Nijboer, K. Szoszkiewicz & T. Ferreira, 2006. Macrophyte communities of European streams with altered physical habitat. Hydrobiologia 566: 197–210.

Paal, J. & T. Trei, 2004. Vegetation of Estonian watercourses; the drainage basin of the southern coast of the Gulf of Finland. Annales Botanici Fennici 41: 157–177.

Preston, C. D. & J. M. Croft, 1997. Aquatic plants in Britain and Ireland. Harley, Colchester: 366.

Riis, T., K. Sand-Jensen & O. Vestergaard, 2000. Plant communities in lowland Danish streams: species composition and environmental factors. Aquatic Botany 66: 255–272.

Sand-Jensen, K., 1998. Influence of submerged macrophytes on sediment composition and near-bed flow in lowland streams. Freshwater Biology 39: 663–679.

Sand-Jensen, K., T. Riis, O. Vestergaard & S. E. Larsen, 2000. Macrophyte decline in Danish lakes and streams over the past 100 years. Journal of Ecology 88: 1030–1040.

Schaumburg, J., C. Schranz, J. Foerster, A. Gutowski, G. Hofmann, P. Meilinger, S. Schneider & U. Schmedtje, 2004. Ecological classification of macrophytes and phytobenthos for rivers in Germany according to the Water Framework Directive. Limnologica 34: 283–301.

Schaumburg, J., C. Schranz, D. Stelzer, G. Hofmann, A. Gutowski & J. Foerster, 2006. Instruction Protocol for the ecological Assessment of Running Waters for Implementation of the EC Water Framework Directive: Macrophytes and Phytobenthos. Bavarian Water Management Agency, Munich: 125 pp.

Schneider, S., 2007. Macrophyte trophic indicator values from a European perspective. Limnologica 37: 281–289.

Schneider, S. & A. Melzer, 2003. The trophic index of macrophytes (TIM) – a new tool for indicating the trophic state of running waters. International Review of Hydrobiology 88: 49–67.

Schütz, W., U. Veit & A. Kohler, 2008. The aquatic vegetation of the Upper Danube river – past and present. Archiv für Hydrobiologie, Supplement Large Rivers 18: 167–191.

Szoszkiewicz, K., T. Ferreira, T. Korte, A. Baattrup-Pedersen, J. Davy-Bowker & M. O'Hare, 2006a. European river plant communities: the importance of organic pollution and the usefulness of existing macrophyte metrics. Hydrobiologia 566: 211–234.

Szoszkiewicz, K., J. Zbierska, S. Jusik & T. Zgoła, 2006b. Opracowanie podstaw metodycznych dla monitoringu biologicznego wód w zakresie makrofitów i pilotowe ich zastosowanie dla części wód reprezentujących wybrane kategorie i typy. Etap II, tom II – rzeki. Institute of Environmental Protection, Agricultural Academy "A. Cieszkowicz", Warsaw.

Tansley, A. G., 1946: Introduction to Plant Ecology. Allen and Unwin, London: 228 pp.

Thiebaut, G., F. Guérold & S. Muller, 2002. Are trophic and diversity indices based on macrophyte communities pertinent tools to monitor water quality? Water Research 36: 3602–3610.

Van der Molen, D. T. & R. Pot, 2007. Referenties en maatlatten voor natuurlijke watertypen voor de Kaderrichtlijn Water. STOWA report number 2007-32. Dutch Foundation for Applied Water Research (STOWA), Utrecht: 375 pp.

Wiegleb, G., 1984. A study of habitat conditions of the macrophytic vegetation in selected river systems in Western Lower Saxony (Federal Republic of Germany). Aquatic Botany 18: 313–352.

Willby, N., J.-A. Pitt, & G. Phillips, 2009. The Ecological Classification of UK Rivers Using Aquatic Macrophytes. Environment Agency, Science Report.