JRC Scientific and Technical Reports

Revision of the consistency in Reference Criteria application in the Phase I of the European Intercalibration exercise

Isabel Pardo, Sandra Poikane, Wendy Bonne

EUR 24843 EN - 2011

JRC
EUROPEAN COMMISSION

Institute for Environment and Sustainability

The mission of the JRC-IES is to provide scientific-technical support to the European Union's policies for the protection and sustainable development of the European and global environment.

European Commission
Joint Research Centre
Institute for Environment and Sustainability

Contact information
Address: Via Enrico Fermi 2749, Ispra (VA), Italy I21020
E-mail: sandra.poikane@jrc.ec.europa.eu
Tel.: +39 0332 789720
Fax: +39 0332 789352

http://ies.jrc.ec.europa.eu/
http://www.jrc.ec.europa.eu/

JRC 65230

EUR 24843 EN
ISBN 978-92-79-20444-9 (PDF)
ISBN 978-92-79-20443-2 (print)

ISSN 1018-5593 (print)
ISSN 1831-9424 (online)

doi: 10.2788/27809

Luxembourg: Publications Office of the European Union

Printed in Italy

TABLE OF CONTENTS

This document is produced within the cross - GIG working group on Reference conditions and is a final report on the consistency check survey, with an assessment of implications and recommendations for the application of reference conditions during IC Phase II.

This document is drafted by Isabel Pardo for rivers, Sandra Poikane for lakes, Wendy Bonne for coastal waters, with contributions from Wojciech Uszko, Wouter van de Bund, Roger Owen, Martyn Kelly, Didier Pont, Sebastian Birk, Cathy Bennett, Carola Gómez-Rodríguez, Georg Wolfram, Frauke Ecke, Marcel van den Berg, David Ritterbusch, Geoff Phillips, Sandra Brucet and José Ortiz-Casas.

INTRODUCTION

The refinement of the methodology and criteria for setting reference conditions (RC) is a priority Intercalibration (IC) activity to be led by a cross GIG working group (RC WG) between 2009-2011 (phase II IC). The initial task (reported in this document) will be to investigate consistency in the application of RC criteria across Member States (MSs) during phase I of IC. Secondly, MS datasets for biological quality elements (BQE) and pressures (collected during phase II of IC) will be used to run further analyses on pressure-response relationships. The objective will be to confirm the underlying concepts of these relationships using MSs data and, if necessary, refine the limits of the reference thresholds that are used to screen candidate reference sites. This will facilitate MSs in more accurate and consistent reference site selection, improve the common understanding of RC across Europe and assure better consistency and comparability in RC setting between BQEs, water categories and MSs for future IC exercises.

This document reports the results of an analysis of RC setting by MSs during the first phase of IC. The analysis is based on a review of IC technical reports for each water category (Rivers (section 1), Lakes (section 2), Coastal (section 3) and Transitional Waters (section 4)) and broadly considers two aspects:

i) Firstly, we consider the approach taken by each water category for setting Reference Conditions and analyse how RC criteria were applied by MS within different water categories, GIGs, types and BQEs. For river macroinvertebrates, this analysis considered responses to the reference screening questionnaire within different GIGs and river types.

ii) Secondly, we examine the application of the quantitative thresholds for pressures by MSs for Rivers and Lakes (Coastal and Transitional GIGs have not progressed to this stage yet); pressures data associated with the reference sites used for the phase I IC was requested for this analysis.

The report ends with a chapter on conclusions and recommendations to help advance the work on the refinement of RC criteria for phase II of IC (2009-2011).

1. Rivers

1.1. Application of Reference Conditions criteria

1.1.1. Reference screening questionnaire

The methodology to compare how RC criteria were applied by MS was based on the analysis of the responses provided in the reference screening questionnaire for macroinvertebrates. See the Central/Baltic (CB) GIG reference screening questionnaire in Annex 1 (same as Annex 2.1.1.3 in van de Bund, 2009). The RC criteria used for the consistency analysis are highlighted in grey in Annex 1.

Reference Conditions criteria and screening thresholds

All GIGs (Central/Baltic, MEDiterranean, ALPine, EasternContinental), except the Northern (NO) GIG, used the reference screening questionnaire developed by the CB GIG. The NO GIG used the same maximum threshold for artificial land use (< 0.8%) as CB GIG, and maximum values for physicochemical parameters were similar to the ones defined for CB GIG river types (see annex 2.2.3 and 2.1.1.2 in van de Bund 2009). CB GIG chemical thresholds, used in conjunction with the reference screening questionnaire, were also implemented by MED and ALP GIGs. Major differences existed between GIGs in the reference threshold for agricultural land use. This threshold was more stringent in NO GIG (maximum of 25% of intensive agriculture and maximum of 30% of non intensive agriculture was allowed) compared to the 50% intensive agriculture threshold agreed in the CB GIG, even though the categories for intensive/non intensive agriculture differed somewhat to the ones assigned in the CB GIG screening questionnaire. The riparian zone was scarcely accounted for in NO GIG; only a general statement stated that the area should have natural vegetation - a criterion that was only adhered to by two MSs, UK and Sweden. Similarly CB GIG provided a more detailed characterisation of morphological alterations than NO GIG, but there was more general agreement on this criterion at the GIG level in NO GIG, even though the definition of the criterion was less specific.

Response to reference screening questionnaire

The CB GIG proposed a range of possible answers (missing info; not a relevant criterion, measured; estimated; field inspection; expert judgement; alternative criterion used; okay) to individual criteria in the reference screening questionnaire. A simplified version of these responses was adopted by MED GIG (missing info; measured; field inspection; expert judgement; okay). The ALP and NO GIGs used a tick box system for criteria that were applied by MSs. Meanwhile, EC GIG adopted the CB GIG screening questionnaire, but did not include the same categorical answers. EC GIG did not collate answers from MSs to the list of criteria. This analysis was based on the following scoring categories (missing info; measured; field inspection; expert judgement; alternative criterion; okay) for CB and MED GIG; by combining the categories “measured” and “estimated” into “measured” for CB GIG. For the ALP and NO GIGs, we only analysed whether or not the RC criteria were fulfilled. The ALP GIG used both the general criteria from the CB questionnaire and added more specific pressure criteria within the general pressures criteria (see ALP GIG Reference conditions criteria Annex in van de Bund 2009), being the specific criteria particular for the ALP GIG, not included in other GIGs. NO GIG modified some of the RC thresholds compared to those applied in CB GIG and thus could not be compared with the other GIGs. EC GIG screening questionnaires were not collected and could not be included in the consistency check.

The analysis is based on the answers provided by MSs and GIGs to the RC list of criteria reported in the IC technical reports in van de Bund (2009). In Annex 1 of this document we present the reference screening questionnaire containing the total list of criteria; the criteria selected for this analysis are highlighted in grey.

a/ Analysis of responses to reference screening questionnaire

The reference screening questionnaire included 42 RC criteria covering eight general types of pressures (Wallin et al. 2003). The percentage of responses provided for each possible answer by MSs in each GIG, excluding EC GIG was calculated (see Table 1 for the 42 criteria analysed in the CB GIG screening questionnaire). Figures 1 and 2 show the percentage of answers provided by CB GIG and MED GIG (Figure 1) and ALP and NO GIG (Figure 2). The ALP GIG applied the 42 criteria using a tick box system. NO GIG only applied 16 of the 42 criteria (which represented general statements) (Figure 2).1. Point source pollution
"Reference" threshold: < 0.4% of artificial land use in the catchment area.
"Rejection" threshold: 0.8% of artificial area in the catchment.
Between 0.4 and 0.8%, a validation with physico-chemical parameters at the site scale is necessary.
2. Diffuse source pollution
Intensive agriculture: < 20% of the catchment area as reference threshold. Rejection threshold: > 50% of intensive agriculture in the catchment.
Between 20% and 50% of intensive agriculture, a validation with physico-chemical parameters at the site scale is strongly recommended.
3. Riparian zone vegetation
In agricultural landscape (intensive agriculture between 20% and 50%), intensive agriculture land cover < 10% of the reach. Riparian corridor land use > 90% semi natural or low intensity agricultural areas.
In non agricultural landscape (intensive agriculture < 20%): valley floor and riparian corridor occupied by semi natural or low intensity agricultural areas.
Artificial areas: < 10% of the reach.

The riparian zone of the site is entirely bordered by the type specific natural vegetation or semi-natural land cover, with the possible exception of access to the river site.
Riparian vegetation zone continuity: uninterrupted or with few interruptions (access to the site).
The lateral connectivity between river and riparian corridor is maintained along the site.
No direct impact of cattle trampling.
4. Morphological alterations
Sediment transport: No dams which significantly modify the sediment regime (sediment retention) leading to morphological alterations, evidenced by signs of incision of the river bed (e.g. incision > 0.2 m * stream order, bare bed rock appearing…).
"Continuity" for fish should be related to the maintenance of river and stream continuity to facilitate movement of type specific species that should be present in reference state.
If this condition is not fulfilled and some migratory species have disappeared, these species should be added to the type -specific list of fish species.
Flow impedance: < 10% of the reach is affected by flow impedance, due to hydraulic effects of weirs, sluices, etc…
Channelisation: < 10% of the reach is affected by "hard works" (like modification of longitudinal and/or transverse profiles, narrow embankment, loss of lateral connectivity…).
Stabilisation: < 20% of the reach is affected by "soft works" (like bank protection on one side, distant dikes, bank maintenance, not affecting the longitudinal and/or transverse profile, and lateral connectivity globally maintained…).
If both types of works are combined (Anneex1, lines 134 and 135) < 10% of the reach must be affected.
Siltation: reaches with anomalous siltation suspected, due to agricultural soil erosion, should be avoided (expert judgment).
Connection to groundwater: Total lateral and vertical connection to groundwater.
Substrate conditions: Correspond to related typology.
River profile and variation in width and depth: Correspond to related typology.
River continuity: At the reach scale, the continuity of the river is not disturbed by anthropogenic barriers and allows undisturbed migration of aquatic organisms (including resident fish populations).
River continuity: At the reach scale, the continuity of the river is not disturbed by anthropogenic barriers and allows free sediment transport.
The site is not situated in a zone directly or indirectly impacted by a nearby artificial structure upstream or downstream.
Lacking any instream structural modifications (weirs or dams) that affect the longitudinal and lateral connectivity, and natural movement of river bed, sediment load, water and biota (except for natural waterfalls).
Only very small artificial constructions with very minor local effects can be accepted.
5. Water abstraction
No dams or water storage significantly altering the low flow regime; low flow alteration < 20% of the monthly minimum flow.
No significant water abstraction in the reach. The cumulative effect of water regulation and abstraction at the basin and reach scales is < 20% of low flow discharge.
6. River flow regulation
No dams which significantly modify the natural hydrological flow regime (flow regulation): e.g. suppression of frequent floods (< 5 years) with anomalous development of vegetation in the channel, or low flow alteration.
The total storage capacity of the reservoirs in the catchment is < 5% of the mean annual discharge at the site.
No change of the natural (type specific) annual flow characteristics (seasonality of high and low flow).
No by-passed section with residual flow (legal minimum discharge).
No significant hydropower peaking effect (ratio Q hydropeaking/Q baseflow < 2).
Absence of flow regulation (dam) on the reach itself.
7. Biological pressures
At the site scale, no invasive species, but alien species which are not at the invasive stage are tolerated.
No intensive (commercial) fishery.
No or very limited direct pollution by aquaculture plants.
No biomanipulation.
8. Other pressures
No intensive use of reference sites for recreation purposes (no intensive camping, swimming, boating, etc.).
No nearby intensive recreational use at the site scale: No regular bathing activities or motor boating. Occasional recreational uses (such as camping, swimming, boating, etc.) should lead to no or very minor impairment of the ecosystem.

Table 1. Forty two criteria analysed in the CB GIG reference screening questionnaire

The percentage of responses corresponding to the category "measured" or "field inspections" accounted for close to 40% for CB and MED GIGs (Figure 1). MSs answering "okay" accounted for up to 12% and 31% in CB and MED GIGs, respectively but this answer does not really give an indication of how well the MS fulfilled the criterion. The responses corresponding with "missing information" accounted for 38% in the CB, and for 27% in the MED GIG.

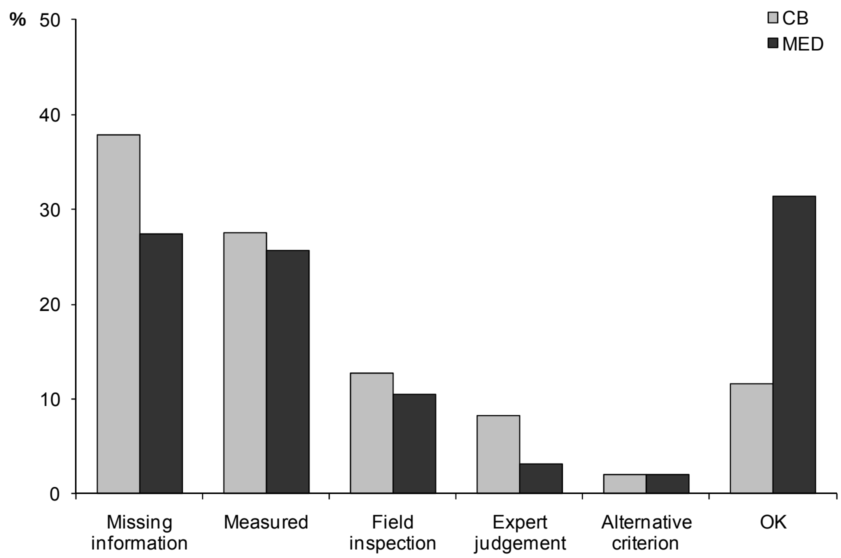


Figure 1. Responses provided by the CB and MED GIGs to the RC screening questionnaire

Both ALP and NO GIG used Xs to indicate the criteria applied for reference screening (Figure 2). Northern GIG showed the highest agreement in the application of the 16 general RC criteria (83%, indicating a mean of 5 countries out of 6 applied the same criteria), versus the ALP GIG, where there was 47% agreement (mean of 3 countries out of 6 applied the same criteria). As previously stated, the "okay" or X category does not provide any indication of how well the criterion has been fulfilled, but this type of response accounted for up to 17% and 53% in NO and ALP GIGs, respectively.

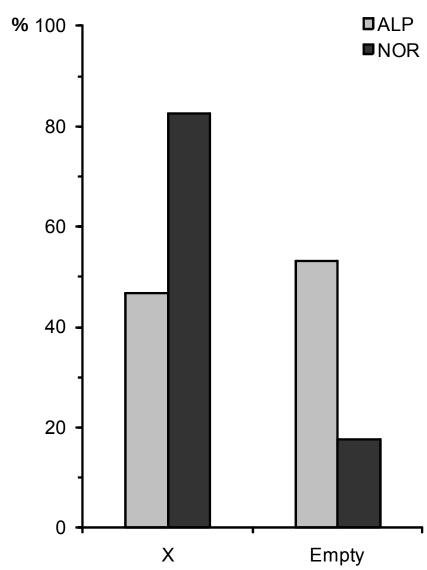


Figure 2. Percentage of response categories provided by the ALP and NO GIGs

Type of responses to the reference screening questionnaire differed among MSs within each GIG. Figure 3 represents the proportion of each response category: Cyprus, Greece and Italy responded similarly, indicating that reference screening was predominantly carried out using "measured" criteria; France and Spain replied mostly with either "okay" or "missing information" and most of the Portuguese responses were evenly distributed between "field inspections", "okay" and "missing information".

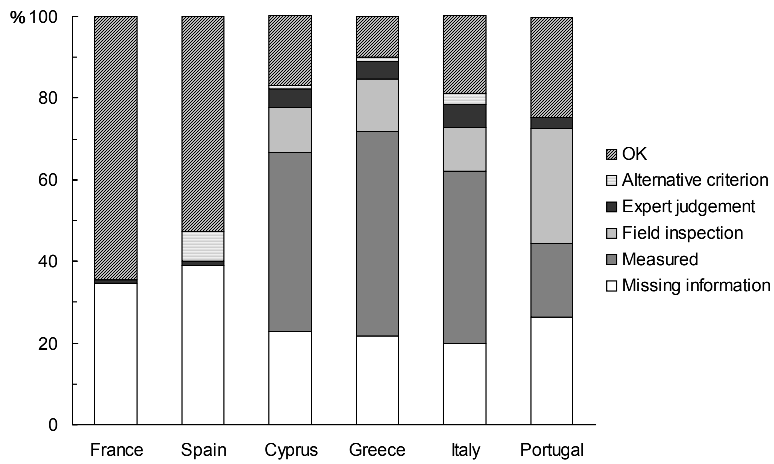


Figure 3. Percentage of responses replied by MSs in the MED GIG

In the CB GIG, the variety of responses suggests that reference screening was done using a variety of approaches (i.e. expert judgement, measured, etc.). Some countries did not answer at all while other countries responded using all of the response categories (Figure 4).

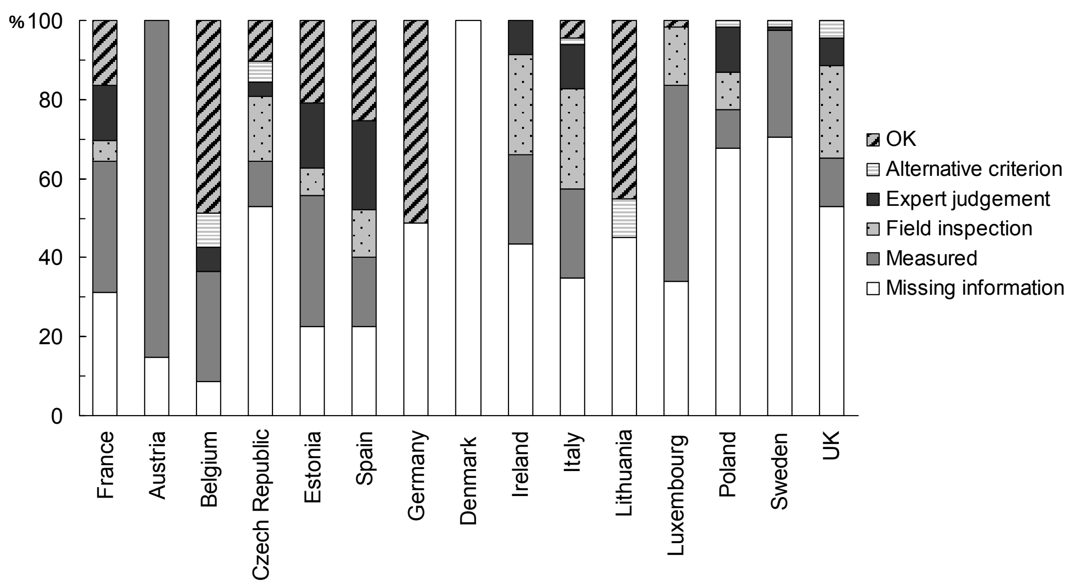


Figure 4. Percentage of responses provided by MSs in the CB GIG

b/ Comparability of responses to Reference Criteria for specific pressures

A more detailed analysis of the way the MSs responded to the eight general pressures was only possible using data from CB and MED GIG. It revealed different approaches when responding to the 42 reference criteria included in the screening questionnaire. Figures 5 to 8 show the percentage of responses provided by the CB and MED GIGs for 4 of the 8 general pressures.

Point source pollution and diffuse pollution criteria were generally “measured” by the MSs, but there was a high percentage of missing information in CB GIG and a high percentage of “okay” answers in the MED GIG (Figures 5 and 6). The evaluation of morphological alterations shows the highest value of missing information for both GIGs at the basin scale (Figure 7), meanwhile a high percentage of responses at the basin scale were evaluated with field inspections for river flow regulation (Figure 8).

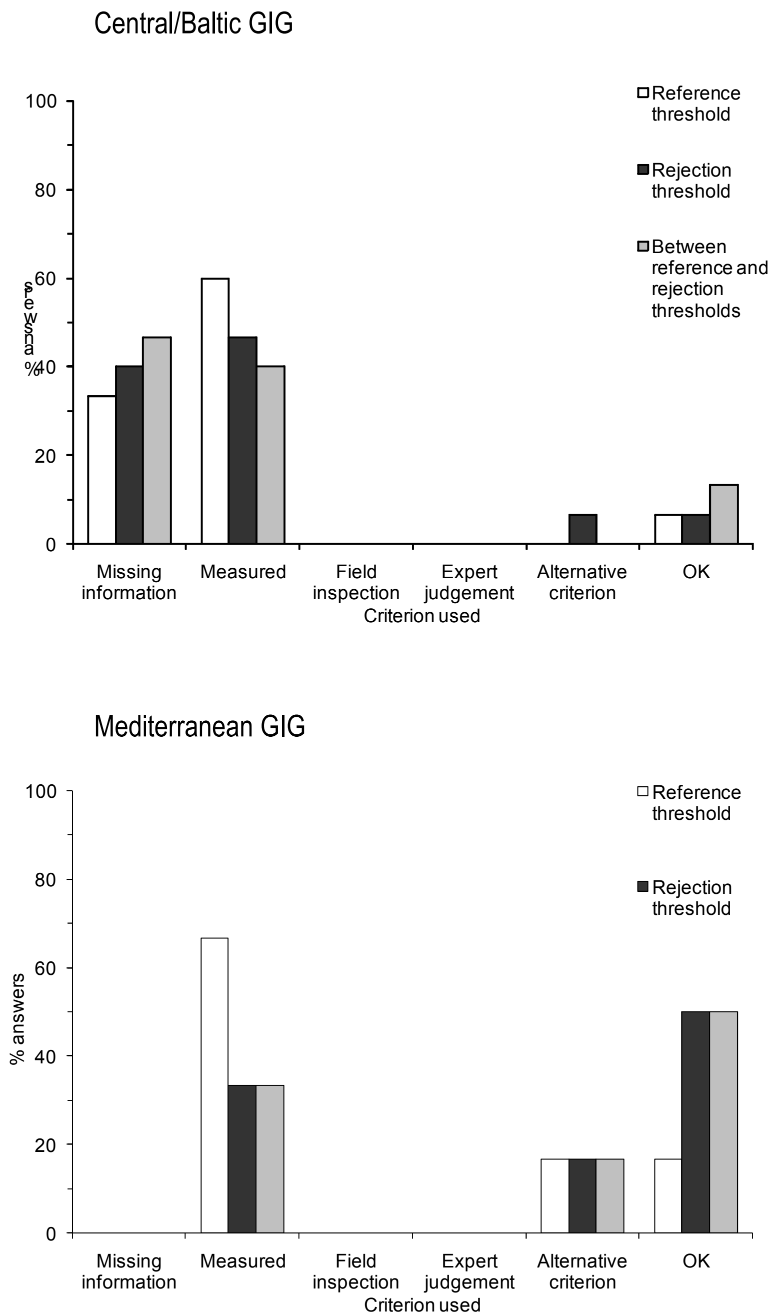
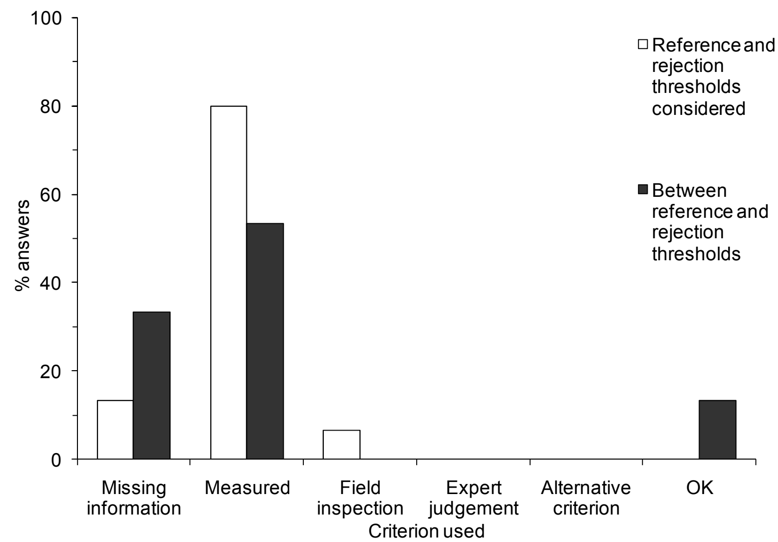


Figure 5. Point source pollution

Central/Baltic GIG



Mediterranean GIG

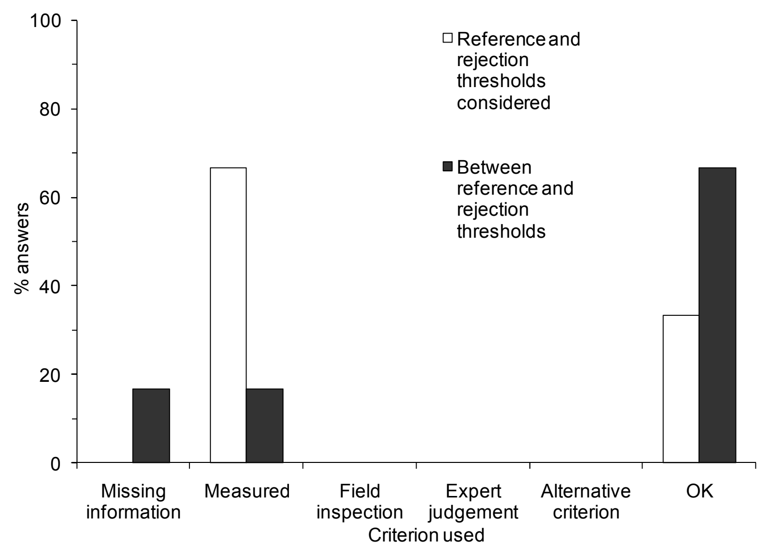
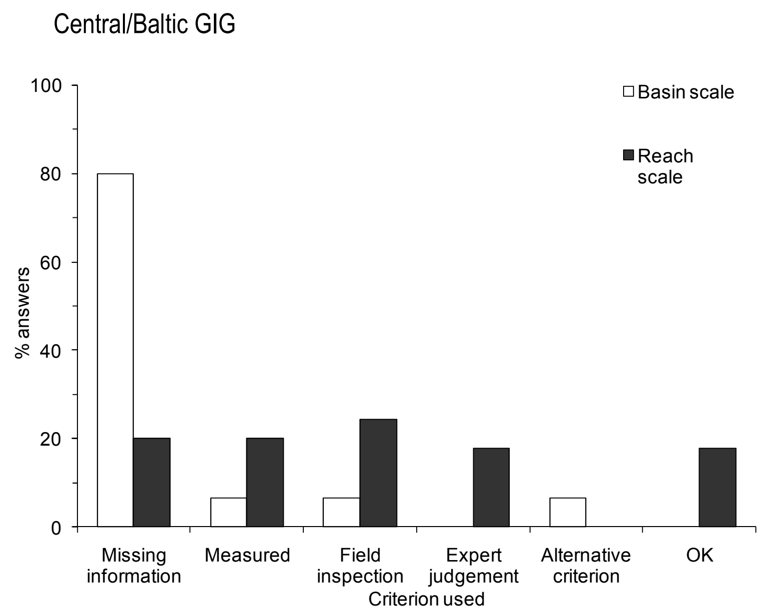


Figure 6. Diffuse source pollution



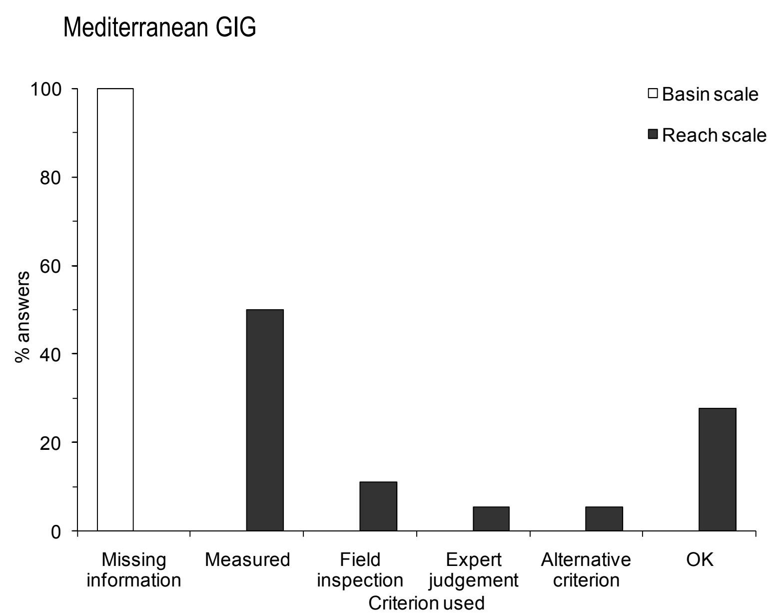
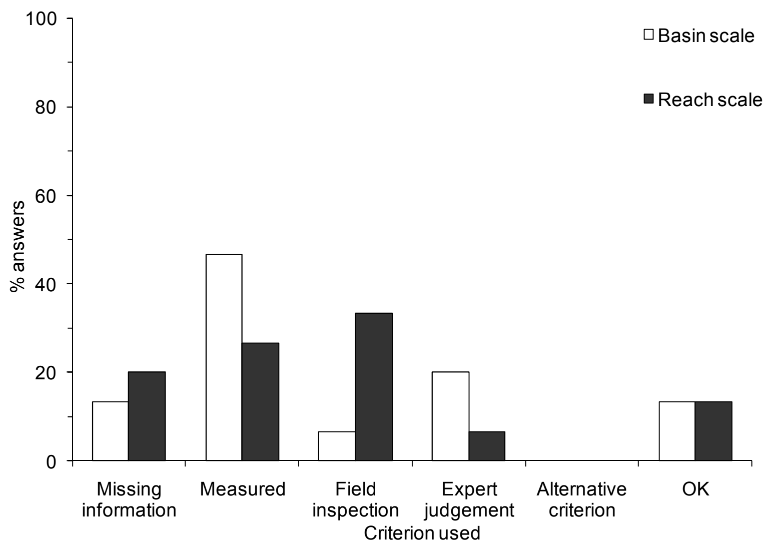


Figure 7. Morphological alterations

Central/Baltic GIG



Mediterranean GIG

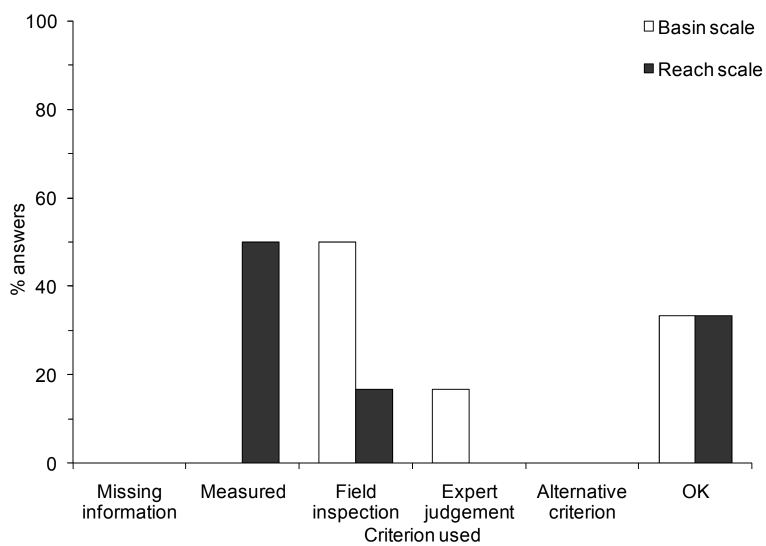


Figure 8. River flow regulation

1.1.2. Review of approaches for setting Reference Conditions across BQE

During phase I of IC, macroinvertebrates and phytobenthos were the only two BQE to be intercalibrated in Rivers. Intercalibration for fishes and macrophytes has not been finalised yet and therefore we can not account for the consistency in their approach for setting RC as yet. Table 2 shows the WFD approaches used by different BQEs in Rivers to set Reference Conditions. For macroinvertebrates and phytobenthos, during phase II of IC (2009-2011) we are reviewing how MS have applied the approach and the criteria used for reference setting but the approach will be kept. The approach for fishes and macrophytes is in general terms comparable with the approaches taken for invertebrates and phytobenthos, and during phase II of IC the consistency between the approaches taken for the different BQEs should be assured.

BQE	IC phase	WFD approach to Reference condition
Invertebrates	I	spatial network of minimally disturbed sites (Reference sites)
Phytobenthos	I	spatial network of minimally disturbed sites (Reference sites)
Macrophytes	II	spatial network of minimally disturbed sites & expert judgement
Fish	II	historical data & spatial network of sites minimaly disturbed* & expert judgement evaluation of pressures

*not necessarily reference sites according to macroinvertebrate and phytobenthos criteria

Table 2. WFD approach used to establish the Reference Conditions in Rivers for the Biological Quality Elements

Invertebrates

The RC criteria applied to macroinvertebrates has been analysed in the previous section (1.1.1). The approach followed was to apply pressure criteria to the screening of potential candidate/reference sites. Two MSs (Netherlands and Belgium-Flanders) did not have any reference sites.

Benthic macroinvertebrates and phytobenthos (diatoms) generally shared the same general approach to screening reference sites. It is assumed that complete screening of the pressures influencing these two BQE communities will result in a reduced spatial network of sites that can be considered to be reference sites (minimally disturbed sites), leading to increased precision of "expected" values of metrics. However, approaches to screening reference sites for macroinvertebrates and phytobenthos differed in several details meaning that not all reference sites were common to both elements.

Phytobenthos

Two GIGs (CB and NO) decided to abandon the use of the IC typology. This was for two reasons:

1. Preliminary studies showed that there was no consistent relationship between the IC typology and either diatom assemblage composition (evaluated by DCA) or metric values;

2. Several MSs had insufficient data for some river types to permit robust comparisons.

The final IC established a single relationship between the national metric and the IC common metric (ICM), from which values for national types could be derived. It was assumed that national types were better "tuned" to local conditions than the IC typology, whilst using a single relationship between national metric and the ICM was more robust than using separate relationships based on relatively small datasets.

In general during phase I of phytobenthos IC, all GIGs claim to have used a consistent approach for reference screening; however, the following points can be deduced from examination of the technical reports:

Northern GIG: Common thresholds for land use and chemical parameters were used with the aim of interpreting the WFD requirement of "very minor anthropogenic impact".

Central/Baltic GIG: Most countries followed the RC criteria established by CB GIG for macroinvertebrates and used common thresholds for land use criteria, but responses from MSs varied in the way they evaluated land use criteria. Chemical thresholds derived by CB GIG principally for macroinvertebrates were applied to phytobenthos sites but with more stringent values for N-NO3 and P-PO4 in some river types (R-C2, R-C3). Two MSs (Netherlands and Belgium-Flanders) did not have any reference sites.

Alpine GIG: The final IC technical report does not specify which approach, if any, was applied to the selection of a spatial network of minimally disturbed sites.

Mediterranean GIG: Common thresholds for some chemical parameters from CB GIG were applied but the MED GIG agreed new values for some nutrients (see van de Bund 2009); Spain provided a different set of threshold values for each river type. Land use data was also used for screening, but without reference to common thresholds and therefore the approach was not consistent.

Eastern Continental GIG: Expert judgement was used to select the best quality sites based on environmental variables.

The following conclusions correspond to the analyses that the CB GIG phytobenthos group produced in the final IC technical report (van de Bund 2009) on the issue of Reference Conditions:

Open issues and need for further work: Typology and Reference Conditions (notes from phytobenthos IC report)

It was not possible to derive a diatom-specific typology due to the lack of comparability of environmental data. However, the diatom IC expert groups believe that the present approach, with all types pooled is 'fit for purpose'. A variety of approaches were adopted by MSs for screening Reference Conditions. The IC typology did not discriminate between reference sites. In addition, there was a strong trophic gradient within the reference sites. It is not clear whether differences between MSs are due to screening procedures or to genuine ecological differences. Several samples with floras indicative of high nutrients came from those MSs which had apparently adopted comprehensive screening procedures. However, the protocol for reference site selection does not ascertain that actual pressures are determined on the same basis in all MSs. Land use categories can represent a wider range of effective nutrient loading, some types of point source pollution may be neglected and some MSs included a final screening involving (different) biological criteria, whereas others did not. This may reduce the overall effectiveness of the screening procedure.

The following recommendations were agreed upon to deal with these issues:

- Problems associated with reference site screening are shared by other IC exercises and a means of validating and publishing criteria used for reference site selection is needed in order to ensure that the IC process is open and transparent
- Testing the validity of the IC typology should be a priority in future phytobenthos IC exercises. Future work should improve the approach used for assessing the comparability of the results in order to confirm them
- Future phytobenthos IC exercises should consider developing a common format for collecting key environmental data in order to facilitate development of a diatom-specific typology

Fishes

National river fish assessment methods are effectively quite diverse and are based on different concepts. Some MSs consider historical data, expert evaluation and present data (e.g. Germany and Austria) when defining Reference Conditions but most MSs (e.g. France, UK, Wallonia and Flanders, Sweden and Finland) use the concept of minimally disturbed sites, in conjunction with expert judgement, to define Reference Conditions.

For IC, two points of view were considered when selecting reference sites:

First, the fish working group defined a list of variables to evaluate the "intensity" of the different types of pressures. The description of these pressures is available in the fishes IC exercise common database. Then a defined list of criteria was used to select "undisturbed sites".

In addition, each MS defined a list of "reference sites" using their national criteria.

Finally, the sites considered as IC reference sites must be both classified as national reference site using national criteria, and considered as undisturbed site using the list of pressure variables and common criteria outlined by the fish IC working group. In that way, "reference sites" are selected using similar criteria at the European scale, but also fulfil the criteria defined at the national level.

The next step involved the use of common European metrics to determine if reference condition was defined in a comparable manner between MSs (comparison of common metrics values between MSs when considering only the selected sites).

It is also possible to get an indication of the response of different national methods (and the common metrics) to different pressures as well as to an overall index of pressures. A description of an overall pressure index is available in the First Milestone Report (September 2009) and an analysis of the responses of the different national methods and of the common metrics to pressure will be integrated in the next milestone reports.

Macrophytes

The CB GIG river macrophyte benchmarking approach

The initial approach to define benchmarks for the IC of the national quality classifications using rivers macrophytes followed the procedure established by the diatom and macroinvertebrate exercise in the first round of IC screening for near-natural reference sites based on common criteria (i.e. “CB GIG reference criteria”).

However, the approach was not successful for two reasons: Firstly, the number of reference sites nominated by the MSs was very small, especially for sandy brooks (R-C1) and medium-sized streams in the lowlands (R-C4). Flanders and the Netherlands were not able to assign any reference sites for these types, while the number of sites identified by France, Wallonia, Germany and Poland did not allow for sound statistical treatment of the data. Here, the reference criteria seem too stringent to collate reasonable benchmark datasets from within the contemporary data yielded by the current national monitoring programs (see also Baattrup-Pedersen et al. 2009). This is probably due to the generally distorted character of the lowland river systems in the CB GIG, but also to the national focus on impacted instead of natural sites in their monitoring programmes. It has to be mentioned that in the design of national macrophyte assessment methods defining Reference Conditions based on existing reference sites was not the norm, and countries relied instead on geographical analogues and, historical data or modelling (e.g. Baattrup-Pedersen et al. 2008). National expertise and data availability may thus be generally scarce, probably forcing inconsistent and inexact screening procedures.

More reference sites passing the CB GIG criteria were nominated by the countries for the mountain brooks (R-C3). However, with only four to eight sites defined by Austria, Germany, the United Kingdom and Wallonia, the number per country was still small (Table 3). France defined 37 reference sites, allowing for preliminary analysis of the biological features: The French and Walloon sites showed high variability of the macrophyte communities, including sites classified in worse than moderate quality status according to the national assessment methods

(Figure 9). Since the macrophyte IC group shared a common notion of the common type environment and its biological setting, we precluded typological or biogeographical differences to explain these results. This highlights the second reason for the unsuccessful application of the CB GIG reference criteria: The selected parameters and thresholds seem inappropriate for the screening of macrophyte reference sites.

Country	AT	BE (WL)	DE	FR	UK
# surveys	6	8	4	37	6

Table 3: Number of macrophyte reference sites complying with the CB GIG criteria for the mountain brooks (R-C3)

Several pressure parameters relevant for the macrophyte communities were not included in the catalogue of reference criteria: Nutrient concentrations in the substrate, alteration of the land-water interface, extent of riparian tree cover and flow modification. Furthermore, macrophyte communities are dependent on natural factors showing high variability at small scales, e.g. light conditions and flow patterns. Macrophytes are long-term indicators and compositional changes are therefore more likely to reflect chronic change rather than short term fluctuations, or, possibly, ambient conditions. The response of the communities integrates various stresses among which the natural influences are either often hard to disentangle from the anthropogenic pressures or buffer the effects of anthropogenic pressure. Therefore, the pressure-impact relationship is not as clear as for other biological quality elements. Here, scientific knowledge is limited (e.g. Janauer 2001), and with regard to an improvement of the reference criteria catalogue the questions is: which screening parameters need to be measured and how? Against the background of an already problematic screening procedure revealed by the current revision of national reference delineations, we regard the efforts arising from a revised catalogue as being impracticable for this IC phase.

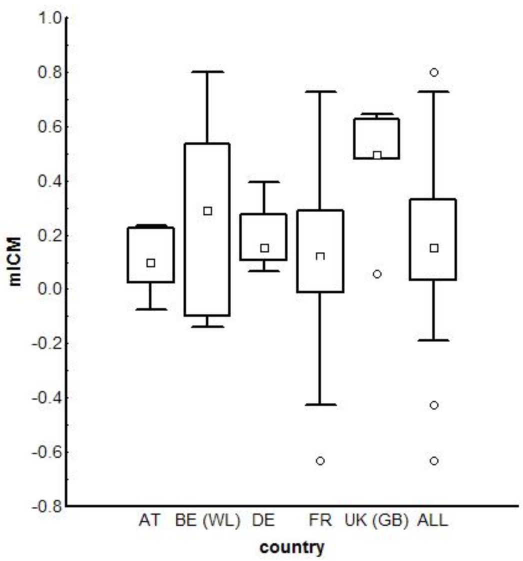


Figure 9. Range of macrophyte IC Common Metric (mICM) among national reference sites in R-C3

To find a common benchmark for intercalibrating the national quality classifications, the macrophyte group is thus resorting to the description of biological benchmark communities. The approach is to define a common level of biological deviation from natural reference communities. Since MSs are using comparable protocols for data acquisition and share a common notion of undisturbed macrophyte communities, these assemblages can generally be described on the international level for each common IC type. Using data on national surveys stored in the IC database, a pool of sites in minimally impacted conditions is selected. This is done by extracting "Common High Status" sites (CHS), i.e. surveys that are assessed at least in good status by all countries, but in high status by the majority of countries. These CHS represent the full spectrum of commonly rated high quality macrophyte communities representative of a river type, and, by definition, represent minimal departure from the biological conditions that are found or expected in the absence of detectable pressures. Following Baattrup-Pedersen et al. (2008, 2009) the CHS approach thus establishes a GIG-wide guiding image for macrophyte reference communities that goes beyond mere expert judgement, building also on empirical data. Differences in common metric values or macrophyte assemblages between geographical regions point to biogeographical or typological trends within the common IC type. Supporting environmental data for the CHS are currently being collected from the MSs. The requested information covers catchment land use data, the evaluation of general chemical, hydromorphological and hydrological pressure and water chemistry data. This benchmark dataset covers a broad geographical gradient from Northern Ireland to the Baltic countries (Figure 10) and comprises both reference and slight to moderately impaired sites. We intend to analyse the effects of biogeographical and typological differences on the macrophyte assemblages of these sites, as well as the macrophytic response to moderate levels of pressure. By quantifying the global stress exerted on the benchmark sites, we will link this alternative reference approach to the overall reference concept followed in the IC exercise.

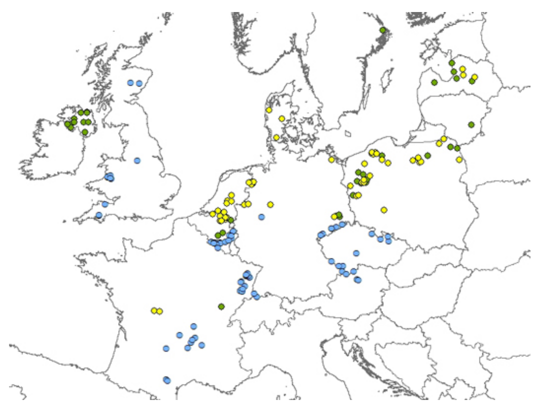


Figure 10. Location of benchmark sites used to establish an alternative reference for the intercalibration of river macrophyte assessments in the CB GIG (yellow dots = R-C1, blue dots = R-C3, green dots = R-C4)

1.2. Comparison of pressure data - Application of reference / rejection thresholds

It was not possible to derive a methodology to compare the application of reference/rejection thresholds (i.e. For some pressure criteria, two thresholds are defined: a reference threshold, below which a site is considered as "probably reference" and a rejection threshold, corresponding to a high probability of significant impact, above which a site is eliminated; for sites between reference and rejection thresholds, a posterior analysis on physico-chemical thresholds has to be done; see Figure 11). It was therefore not possible to assess the degree of comparability between MSs when applying either reference or between reference and rejection thresholds for screening candidate reference sites. Few countries indicated that they accepted sites that had 10% of pressures criteria between the reference and rejection threshold. At the same time it was not possible to aim at a comparison of the result of different pressures that may lead to a prioritisation of the importance of pressures corresponding to 10%.

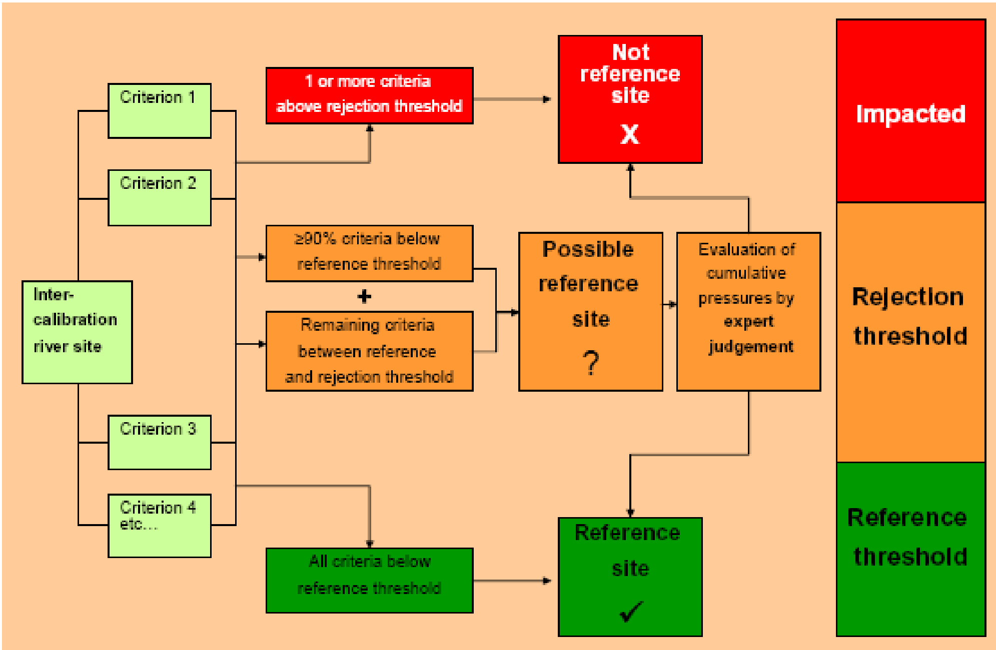


Figure 11. Flow diagram of the procedure for validating reference sites, extracted from van de Bund (2009)

In April 2009, ECOSTAT approved a request to MSs for pressures data from reference sites used for phase I of IC. As a result of the information collected, a comparison was performed on the values provided for the different variables and indicators of pressure for which reference thresholds had been established by the GIG.

Thirteen MSs responded to this request (see table 4), accounting for 615 samples including macroinvertebrates and phytobenthos samples. Not all MSs provided land use or nutrient values for their reference sites. The latter is important especially in cases where the application of land use thresholds has shown to be weak. Figures 12 to 15 show the distribution of land use values at reference sites for each MS, along with the thresholds agreed by CB GIG where appropriate.

Land use variables are generally assumed to be surrogates for point source pollution (artificial land use) and diffuse source pollution (intensive agriculture). Both reference and rejection thresholds for artificial and intensive agriculture are shown in Figures 12 and 13, respectively. MSs generally adhered to the rejection threshold for intensive agriculture, but there was a general lack of consistency with regard to the application of the artificial land use threshold among the MSs that responded.

	BQE	Country	Country	GIG	Number of river types	River type	Reference samples
1	invertebrates/diatoms	AT	Austria	ALP	2	RA1	6
						RA2	6
				CB	1	RC3	11
	phytobenthos			CB	1	RC3	7
				ALP	2	RA1	18
						RA2	24
2		BE-F	Belgium (Randers)				
3		BE-W	Belgium (Wallonia)				
4	invertebrates	CY	Cyprus	MED	1	RM4	8
5	invertebrates/diatoms	CZ	Czech Republic	CB	1	RC3	7
				EC	2	RE1	5
						RE3	2
6		DE	Germany				
7		DK	Denmark				
8	invertebrates	EE	Estonia	CB	1	RC4	12
	diatoms			CB	1	RC4	7
9	invertebrates	ES	Spain	CB	5	RC2	8
						RC3	35
						RC4	10
						RC5	10
						RC6	6
	diatoms			CB	1	RC3	18
10	invertebrates	FI	Finland	N	1	RN3	15
	phytobentos			N	N/A	N/A	57
11	invertebrates/diatoms	FR	France	ALP	2	RA1	4
						RA2	3
				CB	5	RC1	3
						RC2	15
						RC3	23
						RC4	2
						RC6	9
				MED	1	RM1	10
12		GR	Greece				
13		HU	Hungary				
14		IE	Ireland				
15		IT	Italy				
16	invertebrates	LT	Lithuania	CB	4	RC1	1
						RC4	6
						RC5	8
						RC6	5
17		LV	Latvia				
18		LU	Luxemburg				
19		NL	The Netherlands				
20	Invertebrates	NO	Norway	N	1	RN4	9
21		PL	Poland				
22	invertebrates	PT	Portugal	MED	3	RM1	33
						RM2	31
						RM5	18
23		RO	Romania				
24	phytobenthos	SE	Sweden	CB	4	RC1	5
						RC2	7
						RC3	2
						RC4	1
						No type	5
				N	4	RN1	1
						RN3	13
						RN5	1
						RN9	16
						No type	17
	invertebrates			CB	1	RC2	14
				N	2	RN3	15
						RN5	21
25		SK	Slovakia				
26	invertebrates/diatoms	SI	Slovenia	ALP	1	RA1	9
27	invertebrates	UK	United Kingdom	CB	2	RC1	23
						RC3	4
						RC4	9

Table 4. List of Member States that responded to pressures data request associated with the reference sites used for phase I IC; the 13 MSs that responded are shown in grey

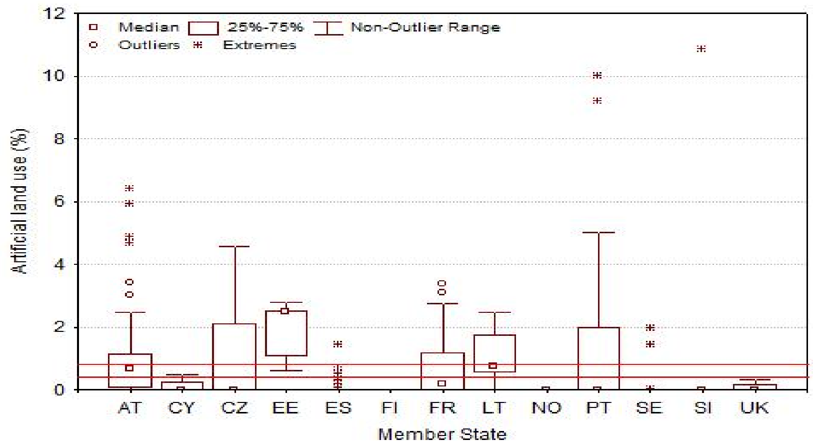


Figure 12. Artificial land use values provided for reference sites by MSs. Lines show the CB GIG agreed reference and rejection thresholds. Value of 54.5% from CZ (best available site) is not shown in figure. Data from the UK represents Landuse 2000 not CORINE land cover values

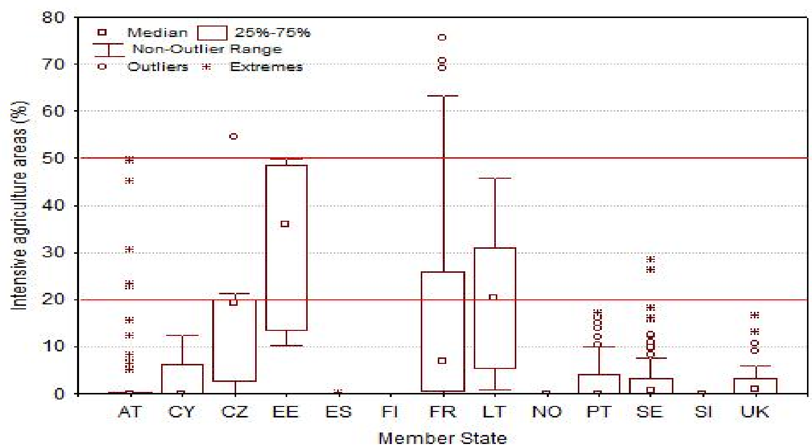


Figure 13. Intensive agriculture land use values provided for reference sites by MSs. Lines correspond to the CB GIG reference and rejection thresholds. Data from the UK represents Landuse 2000 not CORINE land cover values

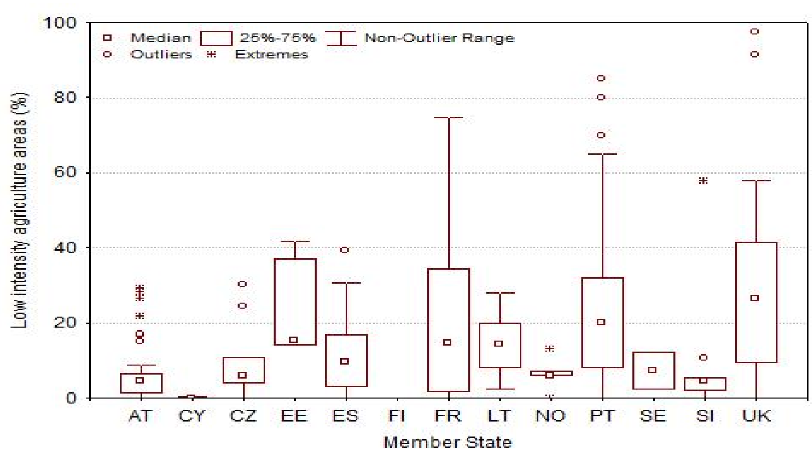


Figure 14. Low-intensity agriculture Land use values provided for reference sites by MSs. No reference and rejection thresholds were defined by CB GIG for this category. Data from the UK represents Landuse 2000 not CORINE land cover values

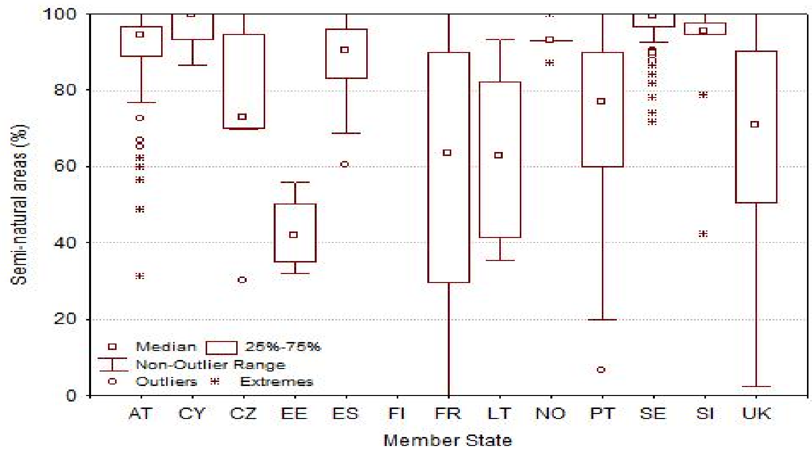


Figure 15. Semi-natural land use values provided for reference sites by MSs. No reference and rejection thresholds were defined by CB GIG for this category. Norway data was excluded for data incongruence. Data from the UK represents Landuse 2000 not CORINE land cover values

Physico-chemical variables characterising water composition are stressors influencing biological communities. The CB GIG agreed chemical values representative of Reference Conditions for Central/Baltic rivers for the following variables: Biological Oxygen Demand after five days (BOD5), dissolved oxygen (DO) and nutrients (see values in annex 2.1.1.2, in van de Bund 2009). Figures 16 to 21 present the distribution of mean and spot measurements of BOD5 and DO values from MSs. Approximately half of the 13 countries that answered this request for reference site data provided chemical values for this analysis. Mean values, or alternatively spot or punctual measurements, for the physico-chemical variables are presented.

Most MSs generally adhered to the threshold values for BOD5 and N-NO3 (Figures 16 and 19), even though there were some exceptions. But agreement between MSs was poor for levels of DO, N-NH4 and P-PO4 at reference sites (Figures 17, 18 and 20).

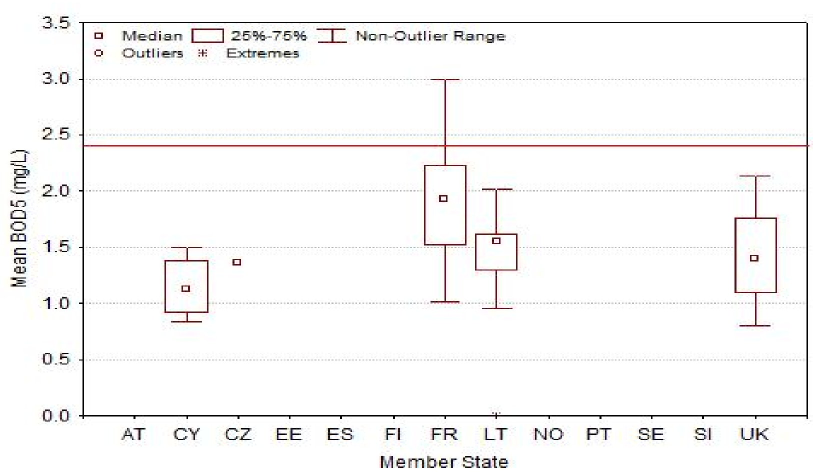


Figure 16a. Mean BOD5 values for reference sites. The line corresponds to the CB GIG mean BOD5 reference threshold.

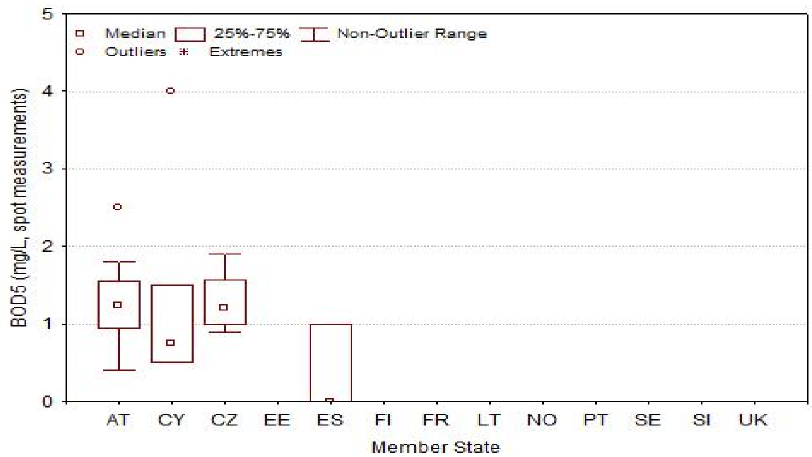


Figure 16b. BOD5 spot measurements for reference sites

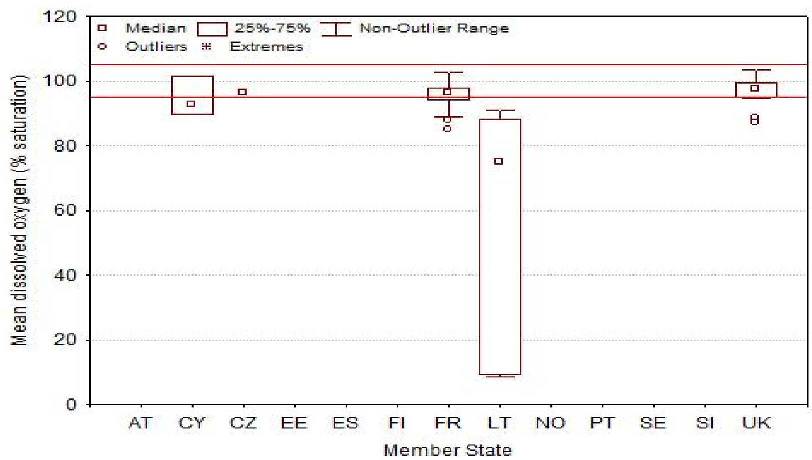


Figure 17a. Mean dissolved oxygen (% saturation) values. The lines correspond to the upper and lower CB GIG mean dissolved oxygen reference thresholds

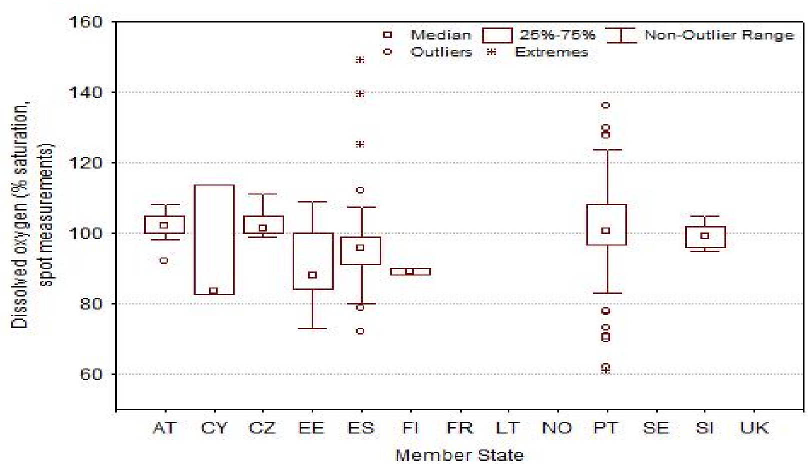


Figure 17b. Spot measurements of dissolved oxygen (% saturation)

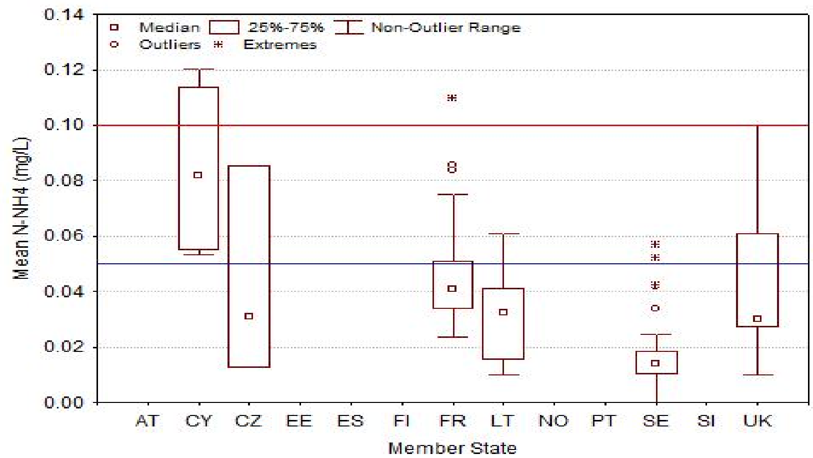


Figure 18a. Mean N-NH4 values. The lines correspond to the mean CB GIG N-NH4 reference thresholds for types RC2, RC3 (0.05 mg/L) and types RC1, RC4, RC5 and RC6 (0.1 mg/L)

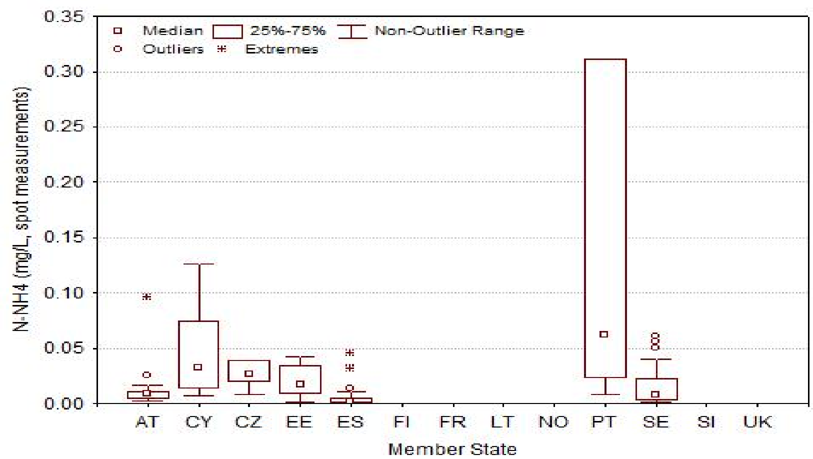


Figure 18b. Spot measurements of N-NH4

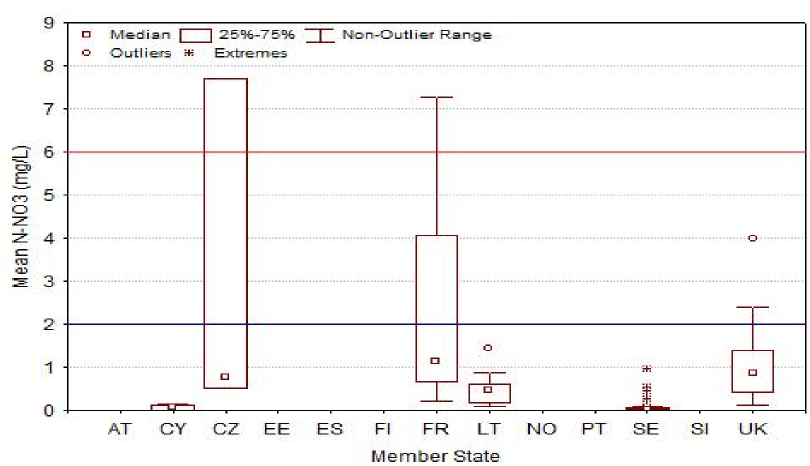


Figure 19a. Mean N-NO3 values. The lines correspond to the mean CB GIG reference thresholds for N-NO3 for type RC3 (2 mg/L) and types RC1, RC2, RC4, RC5 and RC6 (6 mg/L)

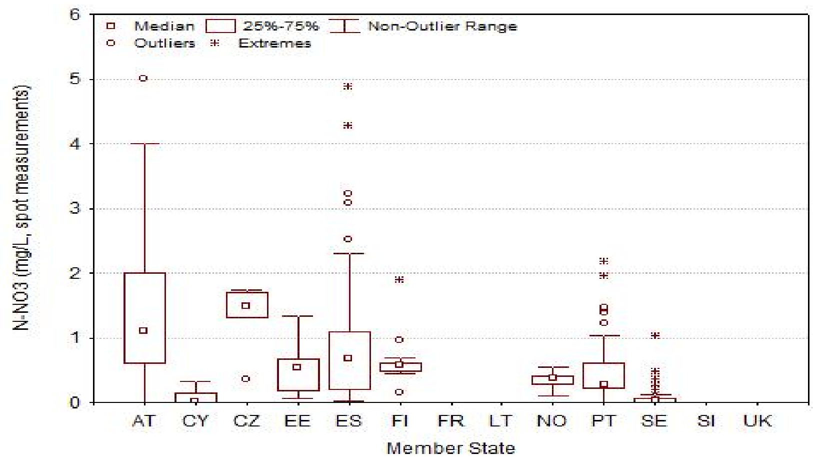


Figure 19b. Spot measurements of N-NO3

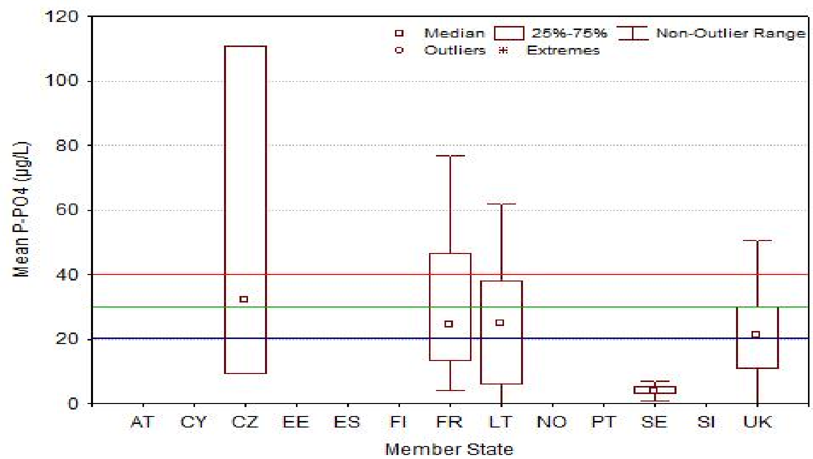


Figure 20a. Mean P-PO4 values. The lines correspond to the mean CB GIG reference thresholds for P-PO4 for type RC3 (20 µg/L), type RC2 (30 µg/L), and types RC1, RC4, RC5, RC6 (40 µg/L)

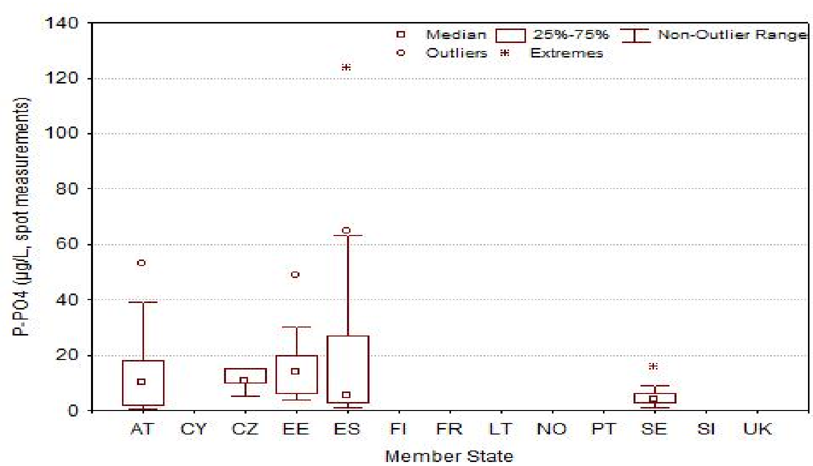


Figure 20b. Spot measurements of P-PO4

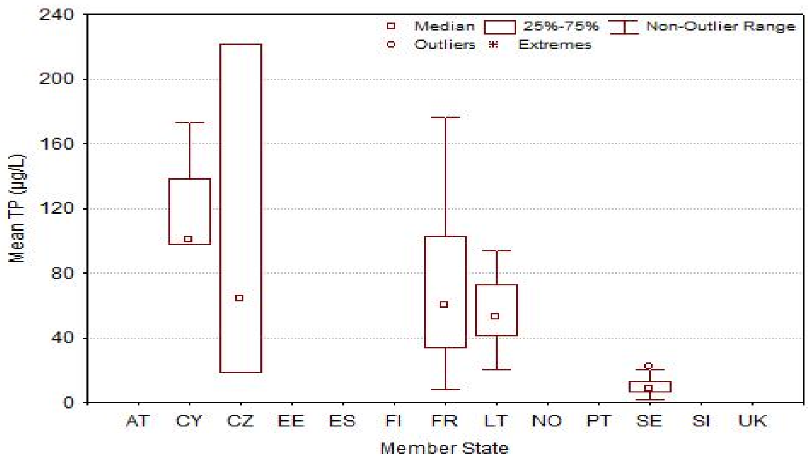


Figure 21a. Mean total phosphorus values (µg/L)

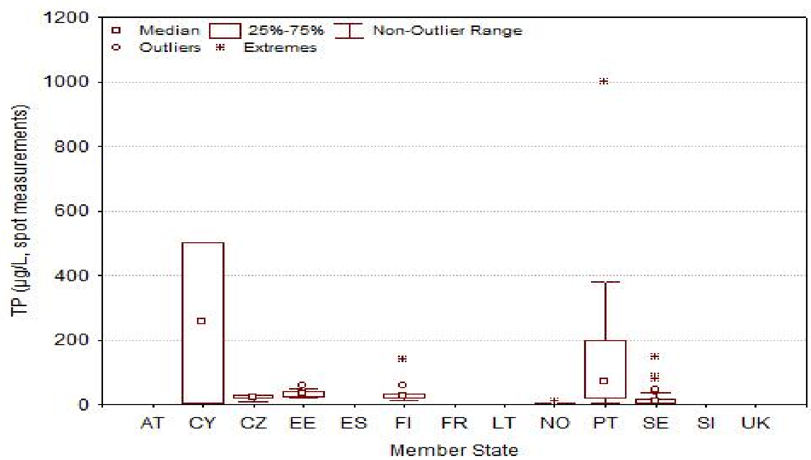


Figure 21b. Spot measurements of total phosphorus (µg/L)

1.2.1. Suitability of CB thresholds to select Reference Rivers across Europe (not limited to the CB GIG)

Following the *WFD CIS Guidance Document No. 14: Guidance on the intercalibration process 2008-2011*, "The level of very low pressure should be defined on the basis of statistical relationships demonstrating that the level of pressure accepted to select a reference site is unlikely to have a significant impact on the biological quality element". Thereby, the suitability of non-impact thresholds can be assessed by means of evaluating the relationship between EQR values and pressure levels within reference sites. The conceptual basis is that non-significant relationship will evidence the adequacy of the selected non-impact threshold. A significant and negative relationship will evidence that the accepted level of pressure may not be adequate and further refinement of the threshold is needed.

Data on MS EQRs and pressure levels for reference sites across Europe were collated (n = 327 for invertebrates; n = 143 for diatoms) in order to evaluate whether thresholds used by the CB rivers GIG could be useful for the rest of rivers in Europe. For each BQE and pressure, a univariate linear regression was conducted with EQR values as response variable and the pressure values as predictor. In the case of land use pressures, both reference and rejection thresholds were evaluated. In the case of water chemistry pressures, when threshold values depended on river type, one analysis was conducted for each threshold except for those thresholds exclusive of only one river type as its representativeness was more limited. Results are shown in Table 5.

Pressure (Threshold)	Regression results (Invertebrates)	Regression results (Diatoms)	Observations
Artificial land use (0.4%)	[-] $R^2 = 0.003$ $F_{1,196} = 0.618$, $p = 0.433$	[-] $R^2 = 0.001$ $F_{1,43} = 0.055$, $p = 0.816$[1]	Reference threshold
Artificial land use (0.8%)	[-] $R^2 = 0.003$ $F_{1,220} = 0.599$, $p = 0.440$	[-] $R^2 = 0.038$ $F_{1,45} = 1.762$, $p = 0.191$[1]	Rejection threshold
Intensive agriculture land use (20%)	[-] $R^2 = 0.030$ $F_{1,258} = 8.082$, $p = 0.005$	[-] $R^2 < 0.001$ $F_{1,78} = 0.001$, $p = 0.980$[2]	Reference threshold
Intensive agriculture land use (50%)	[-] $R^2 = 0.047$ $F_{1,289} = 14.433$, $p < 0.001$	[-] $R^2 = 0.015$ $F_{1,87} = 1.365$, $p = 0.245$[2]	Rejection threshold
Mean BOD_5 (2.4 mg/l)	[-] $R^2 = 0.003$ $F_{1,62} = 0.164$, $p = 0.686$	No data	
Mean O_2 (lower limit = 95%; upper limit =105%)	[+] $R^2 = 0.138$[3] $F_{1,32} = 5.109$, $p = 0.031$	No data	
Mean P-PO_4 (40 µg/l)	[-] $R^2 = 0.038$ $F_{1,48} = 1.875$, $p = 0.177$	Insufficient sample size (4 cases)	
Mean N-NH_4 (0.05 mg/l)	[+] $R^2 = 0.009$ $F_{1,47} = 0.930$, $p = 0.340$	Insufficient sample size (1 case)	River types: RC2, RC3
Mean N-NH_4 (0.10 mg/l)	[+] $R^2 = 0.033$ $F_{1,67} = 2.309$, $p = 0.133$	Insufficient sample size (4 cases)	River types: RC4, RC5, RC6
Mean N-NO_3 (6 mg/l)	[-] $R^2 = 0.005$ $F_{1,63} = 0.294$, $p = 0.589$	Insufficient sample size (4 cases)	

[1] Number of cases with non-null value is highly limited (< 7%). Results should be taken with caution.
[2] Number of cases with non-null value is limited (< 35%). Results should be taken with caution.
[3] A positive and significant relationship evidenced threshold value inadequacy because the lower the O_2 concentration, the higher its negative effect on BQEs.

Table 5. Univariate linear regression results (EQRs vs. pressure values). Sign of the relationship [- or +], coefficient of determination (R^2), F, degrees of freedom and p-values are shown for each Biological Quality Element (Invertebrates/Diatoms). The threshold value from CB used to select sites according to "very low pressure level" is also shown. Significant results are highlighted in red

Due to limited diatom sample size, the suitability of CB non-impact thresholds to select reference sites across Europe cannot be completely assessed. Invertebrate results evidenced that threshold values for artificial land use, mean BOD_5, mean P-PO_4, mean N-NH_4 and mean N-NO_3 might be also appropriate for the rest of rivers across GIGs. However, this suitability needs to be confirmed with diatom data. On the contrary, present results evidence the need of threshold refinement in the case of intensive agriculture land use and mean O_2 since, at least in the case of invertebrates, a negative effect of pressure levels is observed. However, it should be stressed that the effect is weak (very low R^2 values) and thus might be considered negligible. Figures 22a and 22b show the relationship between EQR values for invertebrate data and pressure levels.

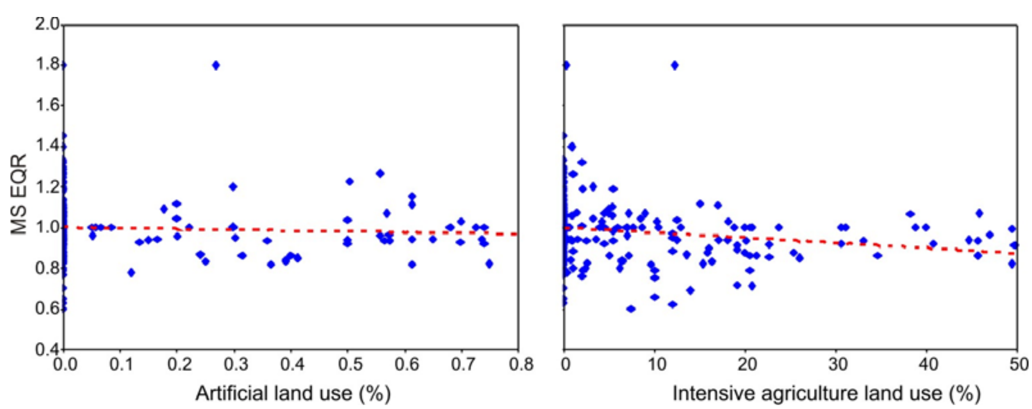


Figure 22a. Relationship between MSs EQRs and artificial and land use pressures

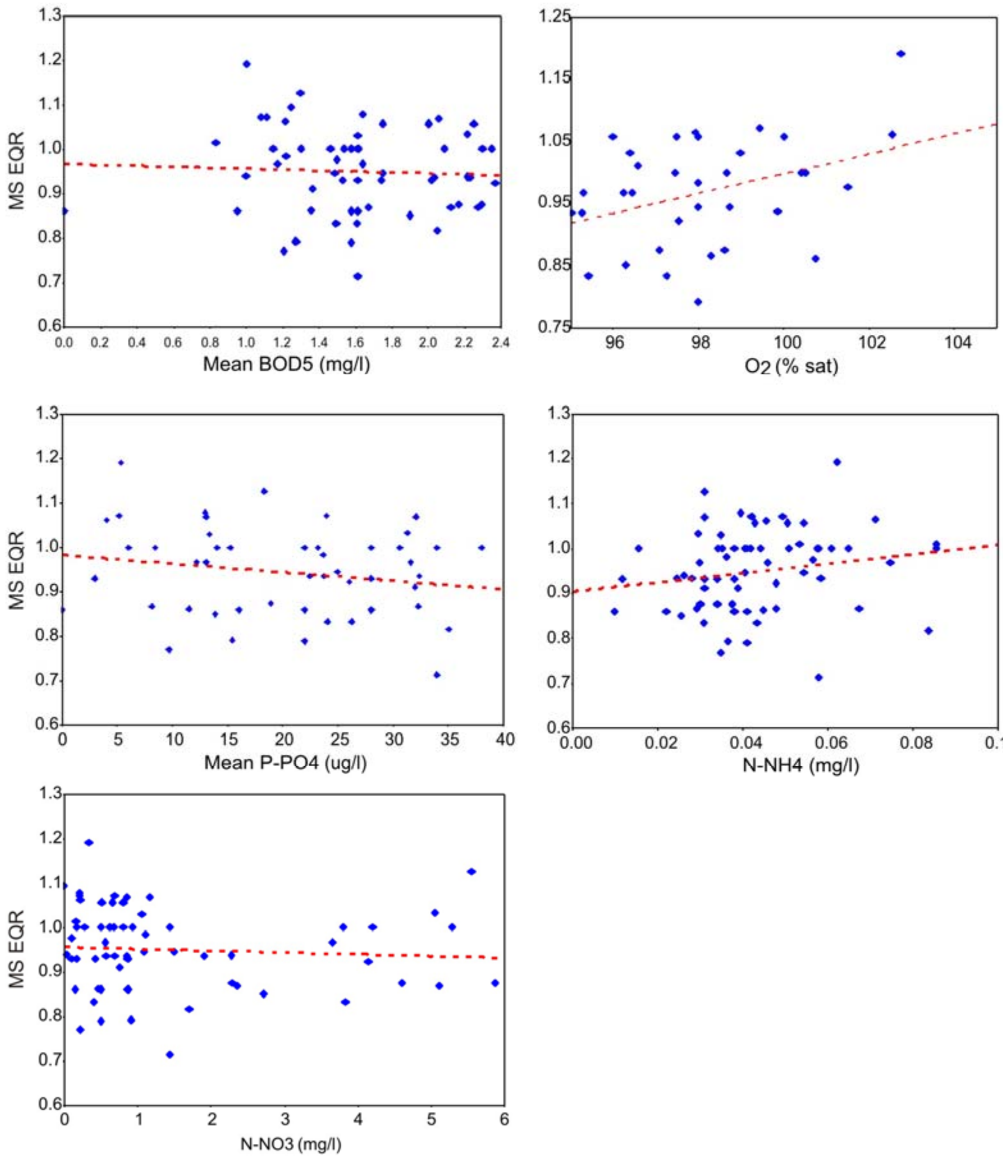


Figure 22b. Relationships between artificial MSs EQRs and water chemistry pressures

In sum, no robust conclusion can be drawn because present results were obtained from a limited data set. Firstly, only 12 MSs provided data and, thereby, it cannot be assured that all geographical peculiarities across Europe have been taken into account in the analyses. For that reason, if we aim to ensure that thresholds may be broadly applicable across Europe, all MSs should provide reference data. Secondly, sample size of diatom data was insufficient for statistical analyses. Thus, the utility of CB non-impact thresholds to identify reference sites for diatoms across Europe still needs to be evaluated.

1.2.2. Verification of Central/Baltic Geographical Intercalibration group (CB GIG) thresholds with STAR / AQEM data

Here, the adequacy pressure levels (hereinafter "CB thresholds") is evaluated using a large database kindly provided by the AQEM and STAR projects. Following the *WFD CIS Guidance Document No. 14: Guidance on the intercalibration process 2008-2011*, "The level of very low pressure should be defined on the basis of statistical relationships demonstrating that the level of pressure accepted to select a reference site is unlikely to have a significant impact on the biological quality element." Thereby, we can assess the suitability of CB thresholds by means of evaluating the relationship between EQR values and pressure levels using data compliant with such thresholds. The conceptual basis is that a non-significant linear relationship will show if the threshold is adequate. A significant and negative relationship will signify that the accepted level of pressure may not be adequate and further refinement of the threshold is needed.

In particular, three main questions are addressed here:

- ✖ Are land-use CB thresholds adequate?
- ✖ Is there evidence of anthropogenic impact within reference sites?
- ✖ Is it necessary to check water-chemistry even below CB land-use reference thresholds?

Database and reference sites screening

The data used from the AQEM/STAR database consisted of 1537 samples from 13 MSs (AT; CZ; DE; DK; FR; GR; IT; NE; PO; PT; SE; SK; UK). Forty-three (43) stream-types were represented (i.e. medium-sized calcareous streams 0-200 m in Southern Greece; medium-sized lowland streams; medium-sized lowland streams of Southern Portugal; medium-sized streams in lower mountainous areas of Southern Portugal; etc.), but the intercalibration type was not specified in the database. The assignment of CB GIG types was not possible at this stage. Each sample consisted of invertebrate taxa abundances (most of them identified to species level). Abiotic information (environmental description and pressure levels) was provided for each sample, along with spot measurements of water-chemistry. The Intercalibration Common Metric (Buffagni et al., 2006, 2007; Bennett et al., 2011) was computed for each sample, using only information at the family level, with the help of Asterics 3.1.1 software (for individual metrics computation).

A set of "operational reference sites" were selected following the strict application of all CB thresholds simultaneously (both water-chemistry and reference land-use thresholds) as well as screening for hydromorphological alterations. As IC types were not provided and water-chemistry thresholds are type-based, the least stringent value was applied in the case of water chemistry parameters (BOD_5 = 2.4 mg/l; O_2 = 95-105%; $N\text{-}NH_4$ = 0.10 mg/l; $N\text{-}NO_3$ = 6.00 mg/l; $P\text{-}PO_4$ = 0.040 mg/l). It should be noted that such CB thresholds are defined for mean values although they were applied to spot measurements here.The other criteria used were: eutrophication = false; acidification = FALSE; liming = false; mining = false; toxic substances = false; number of dams ≤ 10; the cumulative height of such dams ≤ 10m; urban land-use = 0%; crop-land (equivalent to intensive agriculture) ≤ 20% (note that land use values were classes representing 0%, 10%; 20%, etc; and thus the exact CB GIG threshold value for artificial use was not applicable). Additionally, the raw ICM values had to be 5.5 or higher to consider the site as reference. The rationale behind is that a reference site might be rejected if the biota indicates human-induced disturbance (Stoddard et al., 2006). It should be noted that the ICM consists on the weighted sum of EQR values of individual metrics. Here, a modified value was computed (raw ICM value), consisting on the weighted sum of individual metrics and thus, not requiring the previous definition of reference sites. A total of 270 "operational reference sites" were identified. To compute EQR values, the ICM raw value was divided by the median value of the reference sites corresponding to its type (STAR/AQEM types). Types with less than 3 reference sites were excluded from further EQR-based analyses.

✖ *Are land-use CB thresholds adequate?*

To evaluate differences in EQR values between land-use classes for urban and crop land (equivalent to intensive agriculture) (0%; 10%; 20%; 30%; …), an ANOVA analysis was conducted with all the samples. In the case of urban land-use, four categories were considered (0%; 10%, 20%; 30%). Though there was no significant difference in EQR values between urban land use classes (ANOVA $F_{3,\ 984}$ = 2.14; p = 0.094), there was an overlap of EQR values especially between classes 10% and 20% (Figure 23). In the case of crop-land, eight categories were considered (0% - 70%) and this factor was significant ($F_{7,1479}$ = 13.63; p < 0.001). Post-hoc Fisher test evidenced that classes 0% and 10% had significantly higher EQR values than the rest (Figure 23). Moreover, no significant differences were observed between 20%, 30%, 40% and 50% classes.

No clear conclusion can be drawn for urban land-use, since available data (0%, 10%;…) is not adequate to evaluate CB thresholds given their small value (0.4% and 0.8%). In the light of results, the rejection threshold for intensive agriculture (50%) may be too permissive since there is a clear reduction in EQR values in-between 20% and 50% crop land values.

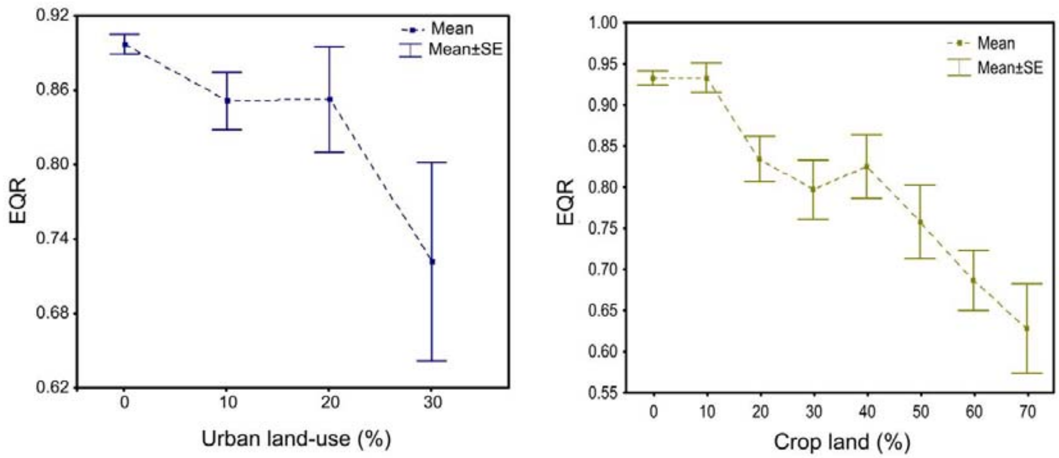


Figure 23. EQR mean value and standard error for each land use level. Land use categories are urban land use and crop land (equivalent to intensive agriculture)

✖ Is there evidence of anthropogenic impact within reference sites?

To evaluate the effectiveness of screening reference sites using CB thresholds, the response of EQR values to water-chemistry variables (linear regression) was assessed using 270 operational reference sites. Thus, if the screening is effective, no significant or marked relationship is expected. On the contrary, a significant negative relationship would suggest that all anthropogenic disturbances have not been accounted for by the screening. Using univariate regression, the results showed that the relationships were not significant except in the case of NO_3 (R^2 = 0.018; $F_{1,225}$ = 4.13; p = 0.043), for which a small significant negative tendency was observed that may be considered negligible given the amount of variance explained (Figure 24).

So, we can conclude that the screening of reference sites was effective. It should be highlighted that the followed approach is stricter than the CB GIG one, as all available thresholds are applied simultaneously. Besides, some quantitative thresholds, not considered in the CB GIG approach, were used to screen for hydromorphological pressures.

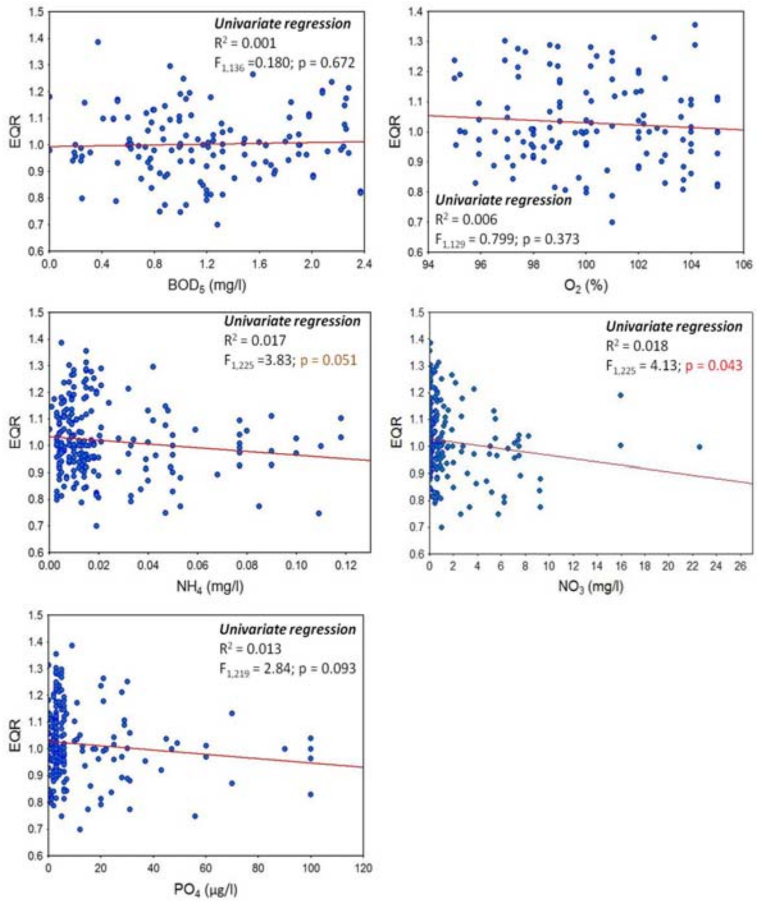


Figure 24. Relationship between EQR values and water-chemistry parameters within "operational reference sites"

✖ Is it necessary to check water-chemistry even below CB land-use reference thresholds?

To evaluate whether it is necessary to check water-chemistry even below CB land-use reference thresholds for urban and crop land (intensive agriculture), the response of EQR values to water-chemistry variables (linear regression) was assessed within all sites below CB reference and rejection thresholds. One regression was conducted for each threshold (i.e. urban and crop land thresholds were not applied simultaneously). In the case of urban land-use, categories of 0% and 10% were tested.

The results showed some anthropogenic impact in sites with 0% urban land-use since a significant relationship (although weak) was observed for all parameters (Figures 25-29). A significant relationship was also observed for nutrients (NH_4, NO_3, PO_4) below in sites with intensive agriculture (crop land) levels within the CB reference threshold (<20%, see figures 26-29).

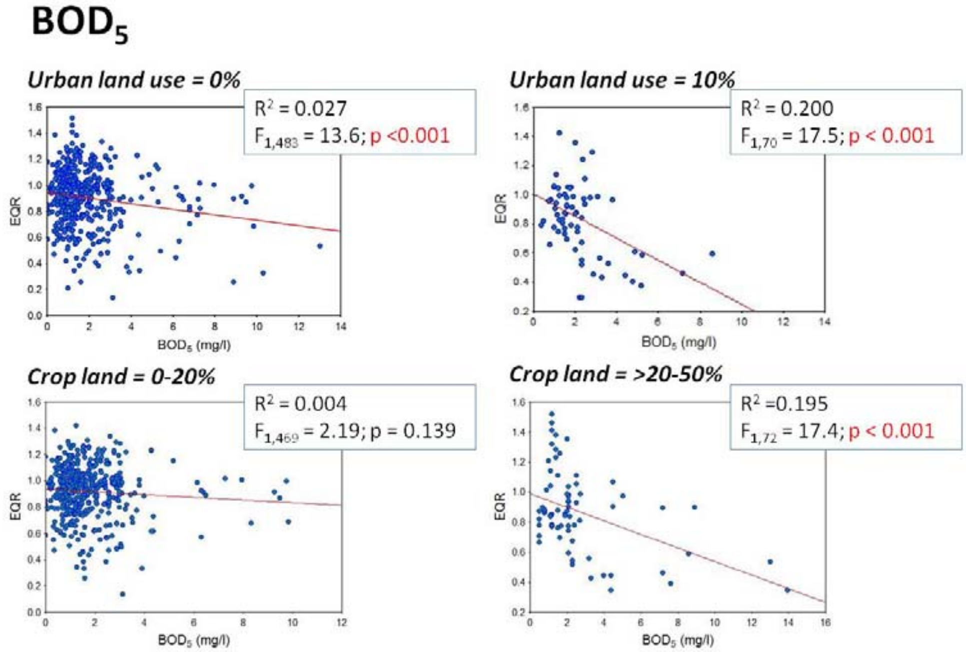


Figure 25. Relationship between EQR values and BOD_5 in sites with land-use values below the specified CB threshold

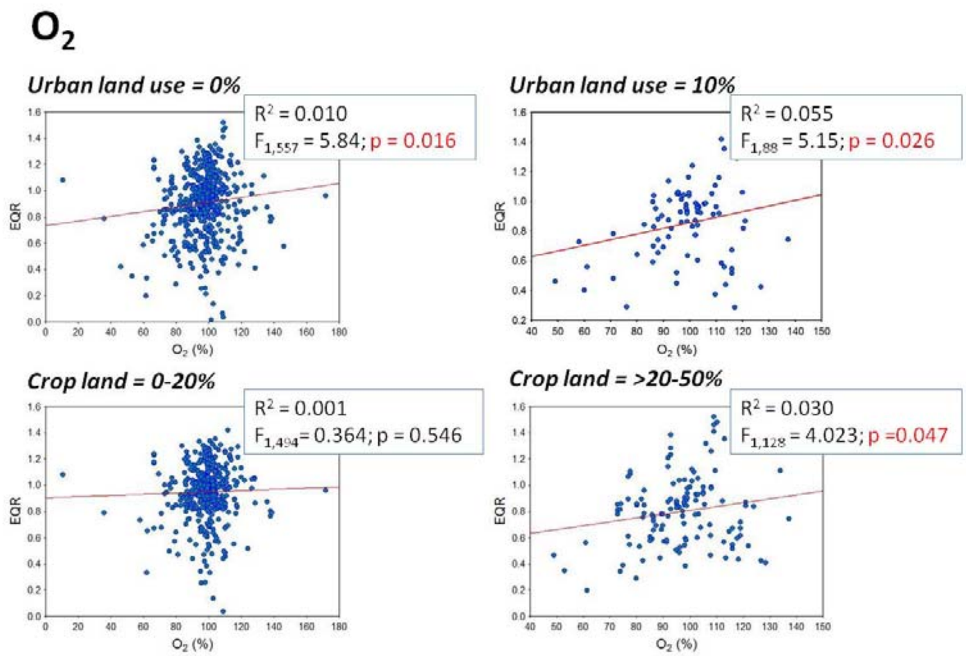


Figure 26. Relationship between EQR values and oxygen in sites with land-use values below the specified CB threshold

NH_4

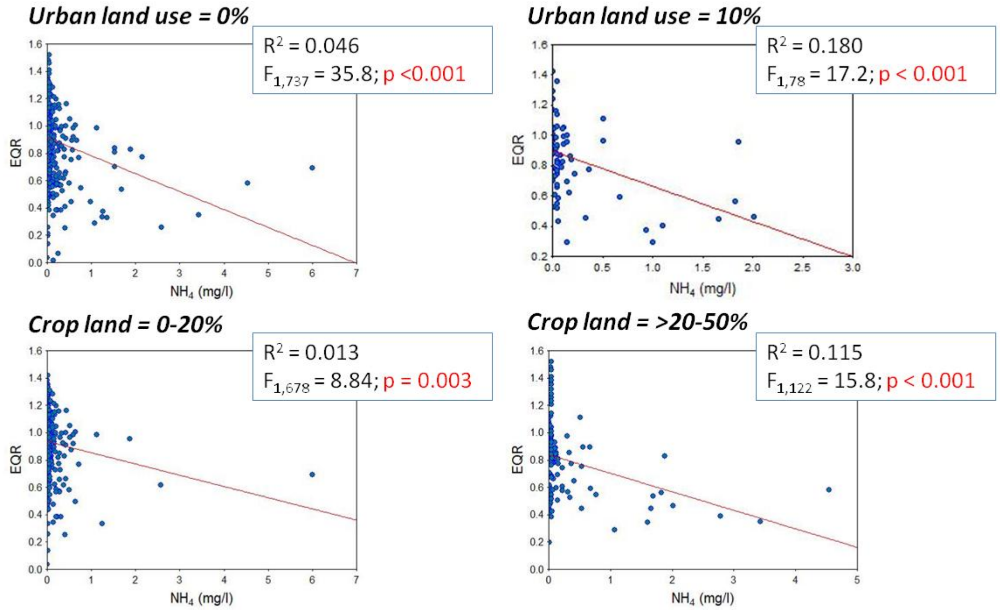


Figure 27. Relationship between EQR values and ammonium in sites with land-use values below the specified CB threshold

NO_3

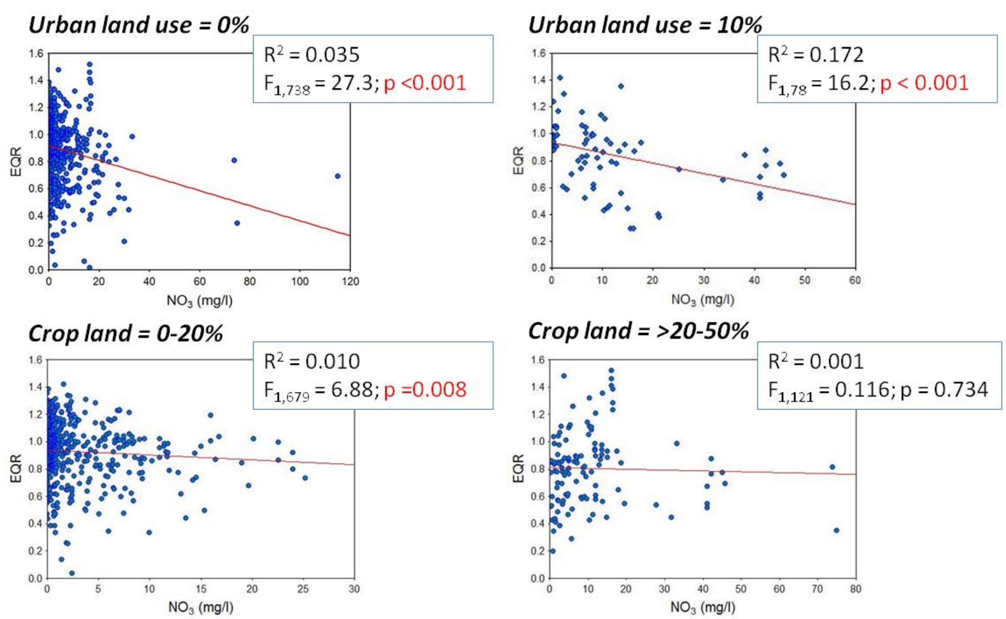


Figure 28. Relationship between EQR values and nitrate in sites with land-use values below the specified CB threshold

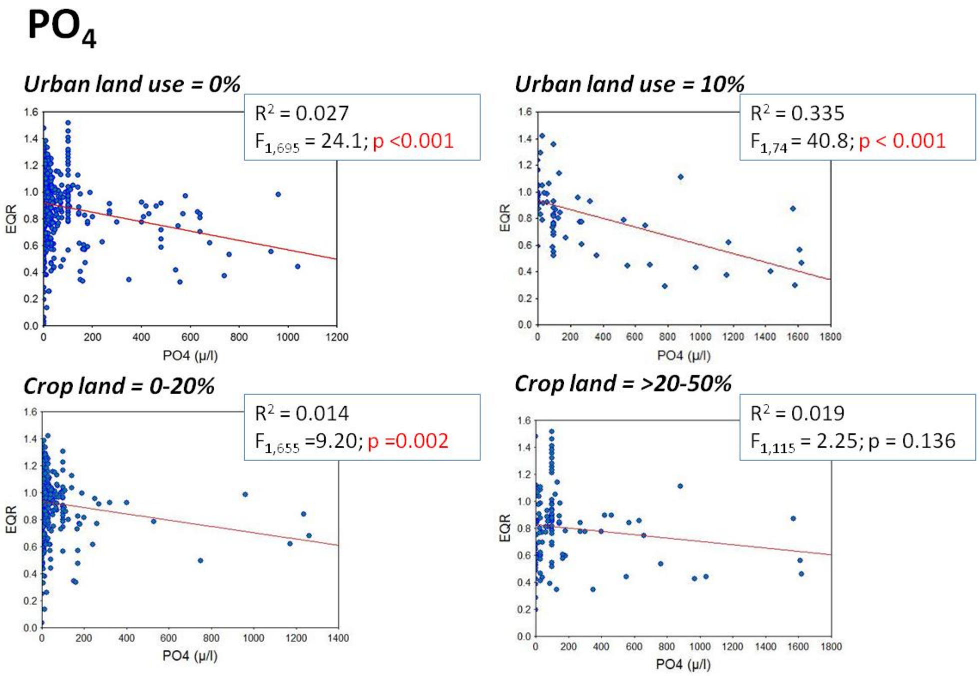


Figure 29. Relationship between EQR values and phosphate in sites with land-use values below the specified CB threshold

Conclusions

- The CB rejection threshold for intensive agriculture may be too permissive as a significant reduction in EQR is observed in sites with crop land values between 20% and 50%.
- An effective screening of reference sites can be accomplished when applying all CB thresholds (land use and water-chemistry) simultaneously as well as considering some quantitative criteria to evaluate hydromorphological pressures.
- Analyses conducted with an external database (AQEM/STAR) showed some anthropogenic impact (EQR_ICM) in sites with intensive agriculture (or crop land) below the CB reference thresholds.
- Checks with water-chemistry for reference sites that fulfil the CB proposed reference thresholds for land-use (artificial land-use < 0.4%; intensive agriculture < 20%) seems to be necessary because some negative tendencies between EQR and water-chemistry are evidenced even below such thresholds (i.e. urban: all parameters; crop land: NH_4, NO_3, PO_4).

In sum

Preliminary analyses (carried out with a limited data set from 12 MSs out of 27) evidenced that the thresholds set for “Intensive agriculture” and “Mean dissolved oxygen” yield a significant relationship within reference sites, so thresholds may not represent absence of biological impact and needs refinement. If CB GIG thresholds were refined, the discrimination of IC types may for land-use values might be considered.

Conclusions herein derived from an external database (AQEM/STAR) agree with CB GIG previous conclusions resulting from the analyses of databases from CB GIG MSs (national EQRs).

2. Lakes

2.1. Application of Reference Conditions criteria

The methodology to compare how MSs applied REFCOND criteria (EC, 2003a) is based on the responses provided by the lake GIGs in the final Lake IC technical report (Poikane et al. 2009). In Table 6 we present the RC criteria agreed by each Lake GIG for phytoplankton and macrophytes.

Different approaches were used by the GIGs and the MSs:

Alpine GIG: The Alpine GIG developed two sets of reference criteria to select reference lakes:
- General reference criteria – focusing on the level of anthropogenic pressure exerted on reference lakes (these criteria are not absolute exclusion criteria, but have to be proved by expert judgement depending on their relevance for the lake ecosystem);
- Specific reference criteria – focusing on ecological changes caused by the anthropogenic pressure; these criteria differ for phytoplankton and macrophytes (see Table 6).

Several additional approaches were also used by Alpine GIG:
- The use of historical data (data from 1930s);
- Modelling approach - reference lakes defined as "no deviation from the natural trophic state" and the type-specific natural trophic state were established by modelling.

Atlantic GIG: A two-step approach was used in the Atlantic GIG
- Reference sites were identified based on existing chemical and biological data, the absence of deleterious impact or land use and expert judgment;
- Furthermore, candidate reference lakes were confirmed by paleolimnological data – which was considered an overriding factor.

Central/Baltic GIG:
Central/Baltic GIG used three main pressure criteria: land use, population density and absence of point sources, but these rules can be overruled if: (1) there is paleolimnological evidence that the lake corresponds to reference state; (2) it is very likely that the use in the catchment is not reaching or affecting the lake.

Mediterranean GIG:
The situation was substantially different in the MED and NO GIGs where no common agreed criteria were set at the GIG level. In the MED GIG, every country selected reference lakes according to their interpretation of the REFCOND guidance and available information.

Northern GIG (NO/NORD): In the NO GIG both pressure and impact criteria were used:
- The main pressure criteria were < 10% agriculture (in total catchment area), and no major point sources. These were mainly judged from visual observation of GIS land use and population data. Due to the high number of lakes in the NO GIG area, it was not possible to quantify the pressure criteria for every single lake.
- Main impact criteria are total phosphorus and chlorophyll or biovolume excluding the worst classes of the present classification systems.

Alpine GIG	
GENERAL CRITERIA (not used as absolute exclusion criteria)	
Catchment area	> 80–90% natural forest, wasteland, moors, meadows, pasture No (or insignificant) intensive crops, vines No (or insignificant) urbanisation and peri-urban areas
	No deterioration of associated wetland areas No (or insignificant) changes in the hydrological and sediment regime of the tributaries
Direct nutrient	No direct inflow of (treated or untreated) waste water
input	No (or insignificant) diffuse discharges
Hydrology	No (or insignificant) change of the natural regime (regulation, artificial rise or fall, internal circulation, withdrawal)
Morphology	No (or insignificant) artificial modifications of the shore line
Connectivity	No loss of natural connectivity for fish (upstream and downstream)
Fisheries	No introduction of fish where they were absent naturally (last decades) No fish-farming activities
Other pressures	No mass recreation (camping, swimming, rowing)
Others	No exotic or proliferating species (any plant or animal group)
SPECIFIC CRITERIA for the selection of phytoplankton reference sites	
Historical data	Prior to major industrialisation, urbanisation and intensification of agriculture
Anthropogenic nutrient load	Insignificant contribution to total nutrient load
Trophic state	No deviation of the actual from the natural trophic state
	Natural trophic state of L-AL3: oligotrophic (threshold value for the pre-selection of reference sites: TP ≤ 8 µg L^{-1}) Natural trophic state of L-AL4: oligo-mesotrophic (threshold value for the pre-selection of reference sites: TP ≤12 µg L^{-1})
SPECIFIC CRITERIA for the selection of macrophyte reference sites	
Trophic state	No deviation of the actual from the natural trophic state
pH, salinity	No deviation from Reference Conditions
Hydrology	Artificial water level fluctuations not larger than the range between the natural mean low water level (MLW) and the natural mean high water level (MHW)
Transect	(At least 100 m shore length)
Surrounding	No intensive agriculture or settlements in the near surrounding
Nutrient input	No direct local nutrient input near the transect
Hydrology	No tributary near the transect
Morphology	No (or insignificant) artificial modifications of the shore line at the transect
Other pressures	No recreation area near the transect
AtlanticC GIG	
Hydromorphology and catchment use	Water level fluctuation: within natural range Absence of mineral abstraction Absence of shoreline alteration e.g. roads and harbours Absence of major modification to catchment e.g. intensive afforestation Groundwater connectivity within natural range No discharges present that would impair ecological quality Water abstraction at level that would not interfere with ecological quality

Water chemistry	Dissolved oxygen: within range 80 – 120% saturation. Oxygen depletion (66% of lake deoxygenated for a period > 2 months) absent pH within range 6- 9, salinity: < 100 mg Cl/l Nutrients: Total phosphorus value < 15 µg P l^{-1} (Irish lakes only, may not appropriate be for some GB lakes, GB lakes using MEI model and paleolimnological data) Temperature: within natural range Synthetic and non- synthetic pollutants: below limit of detection
Biological pressures	No impairment by invasive plant or animal species Stocking of non- indigenous fish not significantly affecting the structure and functioning of the ecosystem No impact from fish farming
Recreational pressures	No intensive use for recreation purposes
Paleo-limnology	Comparison of diatom assemblages with type specific reference data

Central/Baltic GIG

Pressure criteria:

- 90% of catchment land use natural (or semi-natural)
- Population density < 10 km^{-2}
- No point sources in the catchment

Criteria can be overruled if:

- clear and sound evidence from paleolimnological data, which is published or otherwise publicly available;
- the direct related catchment of the lake is surrounded is for more than 90% of the area by natural land use and there are no signs of any disturbance;
- the use of agricultural land is very extensive meaning, no artificial fertilizers are used;
- the whole population in the catchment is connected to waste water treatment plants while the discharge is not connected to the candidate reference lake;
- other reasons, to be specified in the database.

Mediterranean GIG

Cyprus: Based on CORINE Land Cover, 90% of land in the catchment area is covered by semi-natural coniferous forest; 8% is agricultural land. No industry, nor significant human settlements.

France: Reference sites have been defined using land cover types within different buffer zones (CORINE Land Cover analyses): an index based on coefficients allocated for cover types (including inputs of pesticides, phosphorus, hydrocarbons and heavy metals and soil impermeability) was calculated for each scale. For each site these indices were combined to form an overall impact index. Lakes with the lowest total index value were considered as reference sites (Lafage 2004).

Greece:
Land use: The coverage of natural areas is high (91%) and agriculture forms only 7% of the catchment area. There are no artificial surfaces upstream.
Pressures: There are no major pressures in the area. Nutrient loading is considered as very low.
Trophic status: Based on results of chl –*a* and biovolume, the reservoir is considered as oligotrophic.

Portugal:
Sites with less than 20% of the catchment for agricultural land use and the rest remaining as natural or semi-natural coverage (CORINE Land Cover). Additionally, historical records for chemical parameters and chlorophyll concentration were checked, as well as low/moderate level fluctuations (0-20 m) and historical absence of Cyanobacteria blooms. Low/moderate fishing and navigation pressures (expert opinion) were also taken into account. The Castelo de Bode Reservoir was considered as *Best available*, not Reference, due to navigation use, nutrient pressure and presence of upstream dams.

Romania:
More than 70% of the catchment classified as natural;
historical records of Cyanobacteria blooms taken into account;
historical records of total phosphorus and nitrogen forms taken into account;
low fishing and low navigation pressure.

Spain:
Demand of water for different uses, as indicator of the most important anthropic activities that can affect to water bodies. Upstream accumulated demand of water for agricultural irrigation being < 10% was used as indicator of agricultural use. Upstream accumulated demand of water for industrial use being < 1.5% was used as indicator of industrial use. Upstream accumulated demand of water for domestic being < 3% of annual loading was used as indicator of population upstream.
"Naturality" of the catchment according to CORINE Land Cover using 70% of the catchment area classified as "natural areas" (forest, autochthonous vegetation etc) as percentage for less alteration sites.

NORTHERN GIG					
Criteria	Finland	Sweden	Norway	UK	Ireland
Pressure criteria					
Agriculture*	In data sets at present mainly ≤ 10%	< 10% of catchment	< 5%	< 10% arable or intensive grazing	
Point sources	No major point sources	No major point sources	No major point sources		No major point sources
Urbanised area		< 0.1% of catchment			No urbanisation i.e. villages/ towns < 1%
Population density			< 5 p.e./km2	< 10 p.e./km2	
Other pressures	No significant water level regulation or morphological changes	Annual mean ≥ pH 6		No fish farms	No intensive use of lake i.e. abstractions
Impact criteria					
Total phosphorus		< 10 µg/L, or higher if high colour	< 11 µg/L, or higher if high colour		< 10 µg/L
Chlorophyll			< 4 µg/L (low alk. clear types) (< 6 for other types)		< 4 µg/L
Biovolume phytoplankton					
Paleodata				if available	some sites
Expert judgement	yes, partly	no	yes	yes	yes

* Agriculture: This is mainly judged from visual observation of GIS land use data

Table 6. Reference criteria used by the Lake GIGs for selection of reference lakes

Reference Conditions criteria classified according to REFCOND guidelines

Firstly, we classified each of the RC criteria provided by the GIGs into one of eight general pressures identified by the REFCOND guidance (Wallin et al. 2003). For Lakes, we only considered six of the eight general types of pressures, having excluded pressures that specifically address rivers, i.e. "river flow regulation" and "riparian zone vegetation". For Lakes, we combined this criterion with "morphological alterations", creating a new pressure named "hydromorphological alterations" for Lakes. Table 7 shows the lake criteria classification according to six general pressures provided by REFCOND Guidance.

1	Point source pollution
RC	No or very local discharges with only very minor ecological effect
ALP	No direct inflow of (treated or untreated) waste water
ATL	No discharges present that would impair ecological quality
CB	No point sources in the catchment
MED	No industry, nor significant human settlements (CY)
NO	No major point sources (FI, IE, NO, SE)
2	Diffuse source pollution
RC	Pre-intensive agriculture or impacts compatible with pressures pre-dating any recent land use intensification
ALP	> 80–90% natural forest, wasteland, moors, meadows, pasture No (or insignificant) intensive crops, vines in catchment No (or insignificant) urbanization and peri-urban areas in catchment No (or insignificant) diffuse discharges
ATL	Absence of major modification to catchment e.g. intensive afforestation
CB	90% of catchment land use natural (or semi-natural) Population density < 10 km^{-2}

MED	90% of land in the catchment area is covered by semi-natural coniferous forest, agricultural land 8% (CY) Lakes with the lowest total overall impact index value (FR) The coverage of natural areas is high (91%) and agriculture forms only 7% of the catchment area (GR) Sites with less than 20% of the catchment for agricultural land use and the rest remaining as natural or semi-natural coverage (PT) More than 70% of the catchment classified as natural (RO) 70% of the catchment area classified as “natural areas” (forest, autochthonous vegetation etc) (ES)
NO	Agriculture: < 10% of catchment, NO < 5%, IE not described Urbanized area: < 0.1% of catchment (SE), 1% of catchment (IE) Population density < 10 km^{-2} (UK), < 5 km^{-2} (NO)
3	Morphological alterations
RC	Level of direct morphological alteration compatible with ecosystem adaptation and recovery to a level of biodiversity and ecological functioning
ALP	No (or insignificant) changes in the hydrological and sediment regime of the tributaries No (or insignificant) change of the natural regime (regulation, artificial rise or fall, internal circulation, withdrawal) No (or insignificant) artificial modifications of the shore line No deterioration of associated wetland areas No loss of natural connectivity for fish (upstream and downstream) For macrophyte BQE: Artificial water level fluctuations not larger than the range between the natural mean low water level (MLW) and the natural mean high water level (MHW)
ATL	Water level fluctuation: within natural range Absence of mineral abstraction Absence of shoreline alteration e.g. roads and harbours Groundwater connectivity within natural range
MED	Low/moderate level fluctuations (0-20m) (PT)
NO	No significant water level regulation or morphological changes (FI)
4	Water abstraction
RC	Levels of abstraction resulting in only very minor reductions in flow levels or lake level changes having no more than very minor effects on the quality elements
ALP	No (or insignificant) change of the natural regime (regulation, artificial rise or fall, internal circulation, withdrawal)
ATL	Water abstraction at level that would not interfere with ecological quality
MED	Upstream accumulated demand of water for agricultural irrigation being < 10% was used as indicator of agricultural use. Upstream accumulated demand of water for industrial use being < 1.5% was used as indicator of industrial use. Upstream accumulated demand of water for domestic being < 3% of annual loading was used as indicator of population upstream (ES)
NO	No intensive use of lake i.e. abstractions (IE)
5	Biological pressures
5.1	Introductions of alien species
RC	Introductions compatible with very minor impairment of the indigenous biota by introduction of fish, crustacea, mussels or any other kind of plants and animals
ALP	No exotic or proliferating species (any plant or animal group)
ATL	No impairment by invasive plant or animal species
5.2	Fisheries and aquaculture
RC	Fishing operations should allow for the maintenance of the structure, productivity, function and diversity of the ecosystem (including habitat and associated dependent and ecologically related species) on which the fishery depends Stocking of non indigenous fish should not significantly affect the structure and functioning of the ecosystem No impact from fish farming
ALP	No introduction of fish where they were absent naturally (last decades) No fish-farming activities
ATL	Stocking of non- indigenous fish not significantly affecting the structure and functioning of the ecosystem No impact from fish farming
MED	Low/moderate fishing pressures (PT, RO)
NO	No fish farms (UK)
5.3	Biomanipulation
RC	No biomanipulation
6	Other pressures e.g. Recreation uses

RC	No intensive use of reference sites for recreation purposes (no intensive camping, swimming, boating, etc.)	
ALP	No mass recreation (camping, swimming, rowing)	
ATL	No intensive use of reference sites for recreation purposes	
MED	Low moderate fishing and navigation pressure (PT, RO)	
ADDITIONAL APPROACHES USED		
ALPINE GIG		
Historical data	Prior to major industrialisation, urbanisation and intensification of agriculture	
Modelling nutrient load	Contribution of anthropogenic nutrient load insignificant comparing to total nutrient load	
Modelling natural trophic state	Trophic state - no deviation of the actual from the natural trophic state	
ATLANTIC GIG		
Chemical data	Oxygen, pH, salinity, chlorides, salinity, temperature, total phosphorus, synthetic and non-synthetic pollutants	
Paleo data	Comparing diatom with the sediment reference state	
CB GIG		
Chemical and biological data	Conforming with total phosphorus, chlorophyll a, Secchi depth	
MEDITERRANEAN GIG		
Historical records used for checking	TP, Total Nitrogen, cyanobacteria (RO) Cyanobacteria blooms (PT)	
NORTHERN GIG		
Chemical and biological data	Total phosphorus (SE, NO, IE), chlorophyll a (NO, IE)	
Paleo data	Comparing diatom with the sediment reference state (UK, IE, if available)	

Table 7. Lake reference criteria classified according to six general pressures identified in the REFCOND Guidance

Assessment of Lake GIGs along REFCOND Guidance

The second step was to evaluate REFCOND criteria fulfilment by Lake GIGs (Table 8) using the following terms:

- “Complete”: the GIG has interpreted and implemented criteria completely;
- “Incomplete”: the GIG has covered only part of the criterion or not interpreted it according to REFCOND Guidance;
- “Partly” - only some MSs have used this criterion (in case when there was no common agreement among MSs about the reference criteria);
- “Diverse” – all MSs have used this criterion but their interpretation (i.e. the thresholds used) was different (in cases where there was no common agreement among MSs about the reference criteria).

GIG	REFCOND pressure criteria					
	1	2	3	4	5	6
ALP	Complete	Complete	Complete	Complete	Complete	Complete
ATL	Complete	Incomplete	Complete	Complete	Complete	Complete
CB	Complete	Complete	Not Considered	Not Considered	Not Considered	Not Considered
MED	Partly	Diverse	Partly, Incomplete	Partly, Incomplete	Partly, Incomplete	Partly
NO	Partly	Diverse	Partly, Incomplete	Partly, Incomplete	Partly, Incomplete	Not Considered

Table 8. Evaluation system used to compare Reference Condition criteria fulfilment between Lake GIGs

The situation clearly differs between the GIGs:

- Alpine and Atlantic GIGs fulfil all criteria though Atlantic GIG has incomplete interpretation of diffuse source pollution, describing it vaguely by “absence of major modification to catchment”;

- Central/Baltic GIG has covered only point source and diffuse source pollutions;

- Mediterranean and Northern GIGs have not derived common reference criteria, therefore different MSs have used different criteria, interpreted it differently, and have used diverse threshold values (e.g. 70 or 90% of agriculture from catchment area);

- It is clear that all GIGs have focused on those anthropogenic pressures which are the most important for eutrophication, while other pressures were mostly neglected;

- Diffuse/point source pollution was covered by all GIGs, as it is recognised as the most important driver for nutrient loading;

- Whereas hydromorphological, water abstraction and biological pressures were not considered mainly because of two reasons: (1) it was not considered important for eutrophication, (2) there was a lack of data/information in the central databases;

- It must be stressed that Table 8 should be interpreted with caution (*the best GIGs are not perfect and the worst GIGs are not as bad as it can seem from the first glance!)*;

- In the Alpine GIG, “the general criteria” were not used as strict exclusion/inclusion criteria, especially those of minor relevance for trophic state and phytoplankton such as connectivity to tributaries or presence of non-indigenous species, instead the main focus was paid to the current trophic state of lake and nutrient loading as opposed to natural trophic state and nutrient loading;

- Even though Atlantic GIG included all REFCOND criteria in their reference criteria list, the final selection was based exclusively on paleo data (and only three lakes were found to confirm the requirements);

- Although CB GIG has not included several pressures, as morphological alterations and water abstraction, in their “official” reference criteria list, expert judgement was widely used based on all information and data available, therefore the GIG is very sure that its criteria, and all other information available, guarantees that human pressure do not have a significant effect on the reference state for the indicators considered in this exercise.

Reference Conditions criteria for different BQE

- Basically reference criteria were developed and used for the selection of reference lakes for phytoplankton and eutrophication pressure because the IC approach used required common selection of reference lakes and setting of common Reference Conditions;

- In some GIGs, the same reference criteria were used both for phytoplankton and macrophytes (e.g. NO GIG);

- Only Alpine GIG developed different sets of specific reference criteria for phytoplankton and macrophytes;

- Phytoplankton reference criteria focused on the most important aspect for phytoplankton: in-lake nutrient concentrations;

- Macrophyte reference criteria included hydrology, morphology and recreational pressure as essential for macrophyte vegetation;

- Now the development of reference criteria for other BQEs is in progress in all GIGs.

Representativeness of reference lakes

It is important to examine how representative the selected reference sites are for all lake populations. To address these issues, selected descriptors (altitude, depth, area, alkalinity, conductivity, total phosphorus, chlorophyll-a, Secchi depth) of type-specific reference lake populations were compared with impacted lake populations. If reference lakes are type representative, there should be no significant differences between impacted and non-impacted lakes by basic characteristics, e.g., depth, area, altitude. On the other hand, significant differences are expected for impact indicators, e.g., chlorophyll-a and total phosphorus concentrations.

In general, the reference lake population represented all lake populations; there were no significant differences between reference and non-reference lakes by hydromorphological and physico-chemical (alkalinity, colour) parameters (Table 9). Nevertheless, there were some exceptions where reference lake selection or type characteristics must be reconsidered:

- CB1 type reference lakes were significantly deeper than CB1 non-reference lakes (median mean depth values, 7.7 and 5.9 m respectively), less alkaline (median alkalinity values, 2.0 and 2.5 meq/l respectively) and with lower humic content (median for reference lakes, 18 mg Pt/l, and for non-reference lakes, 38 mg Pt/l);

- Also, N2a reference lakes were significantly deeper and less alkaline than non-reference lakes of this type, and N2b reference lakes possessed lower alkalinity than non-reference lakes;

- There were also differences in CB2 and CB3 lake types, but the small number of available reference lakes hinders drawing of meaningful conclusions.

As expected, most reference lake types differed significantly from non-reference lakes in chlorophyll-a, total phosphorus (TP) and Secchi depth (Table 9). Nevertheless, in several types (AL4, CB3, N2b, N6a, N8a), reference lake chlorophyll-a distribution did not differ significantly from impacted lake chlorophyll-a distribution (N2b median value for reference lakes 2.0 µg/l, for non-reference lakes 2.5 µg/l; N6a median value for reference lakes 3.8 µg/l, for non-reference lakes 3.3 µg/l). In fact, in some lake types, reference lake populations outnumber impacted lakes; e.g., there are 71 reference and 25 non-reference lakes within the N2b type population. Even if some sound reasons for such homogeneity could be supposed (e.g., the whole type is relatively unimpacted), it is necessary to review the reference lake selection criteria and the sensitivity of the selected indicators (TP, chlorophyll-a, Secchi) to pressure factors occurring in these lake types.

Type code	Lake type characterisation	Altitude (m a.s.l.)	Mean depth (m)	Alkalinity (meq/l)	Additional characteristics
		Lake Alpine Geographical Intercalibration Group			
AL3	Lowland or mid-altitude, deep, high alkalinity, large	50 - 800	> 15	> 1	Lake size > 50 ha
AL4	Mid-altitude, shallow, high alkalinity, large	200 - 800	3 - 15	> 1	Lake size > 50 ha
		Lake Atlantic Geographical Intercalibration Group			
A1/2	Lowland, shallow, calcareous	< 200	3-15	> 1 meq/l	Non-humic
		Lake Central Geographical Intercalibration Group			
CB1	Lowland, shallow, calcareous	< 200	3 - 15	> 1	Residence time 1-10 years
CB2	Lowland, very shallow, calcareous,	< 200	< 3	> 1	Residence time 0.1-1 years
CB3	Lowland, shallow, small, moderate alkalinity	< 200	3 - 15	0.2 - 1	Residence time 1-10 years
		Lake Mediterranean Geographical Intercalibration Group			
Msw	Reservoirs, deep, large siliceous, lowland, "wet areas"	0 - 800	> 15	< 1	Area > 50 ha; annual mean precipitation > 800 mm or annual mean T < 15°C
Mc	Reservoirs, deep, large, calcareous	0 - 800	> 15	> 1	Lake size > 50 ha
		Lake Northern Geographical Intercalibration Group			
N1	Lowland, shallow, moderate alkalinity, clear	< 200 m	3 - 15	0.2 - 1	Colour < 30 mg Pt/l
N2a	Lowland, shallow, low alkalinity, clear	< 200 m	3 - 15	< 0.2	Colour < 30 mg Pt/l

N2b	Lowland, deep, low alkalinity, clear	< 200 m	> 15	< 0.2	Colour < 30 mg Pt/l
N3a	Lowland, shallow, low alkalinity, humic	< 200 m	3 - 15	< 0.2	Colour 30-90 mg Pt/l
N5a	Mid-altitude, shallow, low alkalinity, clear	200-800 m	3 - 15	< 0.2	Colour < 30 mg Pt/l
N6a	Mid-altitude, shallow, low alkalinity, humic	200-800 m	3 - 15	< 0.2	Colour 30-90 mg Pt/l
N8a	Lowland, shallow, moderate alkalinity, humic	< 200 m	3 - 15	0.2 - 1	Colour 30-90 mg Pt/l

Table 9. Description of Lake Geographical Intercalibration Group (GIG) types included in analysis. Lake type codes: AL -Alpine, A -Atlantic, CB - Central Baltic, M- Mediterranean, N – Northern GIG

2.2. Comparison of pressure data- Application of reference/rejection thresholds

An initial data request for reference lakes was sent to the ECOSTAT and IC representatives. Eighteen MSs (AT, CY, DE, DK, ES, FR, IE, IT, LT, LV, NL, NO, PL, PT, RO, SE, SI, UK) provided requested data for 437 lakes (346 of them are the reference lakes from the IC Phase I), additional information was retrieved for 8 reference lakes from 5 countries using IC datasets (AT, EE, GR, IE, UK).

Not all MSs have provided all requested pressure data. For example, we don't have population data or detailed land use data from Finnish reference lakes which represent 29% of the total reference lake population. So, the consistency analysis presented below is still preliminary, based on information submitted by 19 MSs for 427 reference lakes (Table 10).

	Chl data	TP data	Land use data	Population data	Other pressure data
AT	14	20	22	1	24
CY	1	1	1	1	1
DE	9	9	8	0	3
DK	2	2	2	0	0
EE	3	2	0	0	0
ES (Spain)	4	4	4	1	0
FI	81	81	81	0	0
FR	1	1	1	1	1
GR	1	0	0	0	0
IE	15	15	15	15	13
LT	0	3	3	6	14
LV	14	14	6	6	14
NL	5	5	5	5	5
NO	134	134	134	134	0
PL	7	7	7	7	7
PT	2	2	2	2	2
RO	1	1	1	1	1
SE (Sweden)	17	17	17	17	0
SI (Slovenia)	1	1	1	0	1
UK	41	41	75	75	75
Total	353	360	385	272	161

Table 10. Overview of reference sites provided by MSs up to November 2009, for the revision of the consistency on RC criteria application during IC phase 1

Population density

Population density as reference criteria was used only in the Central/Baltic region (threshold < 10 p.e. km^{-2}) and Northern GIG (thresholds: Norway < 5 p.e. km^{-2}, UK < 10 p.e. km^{-2}), other regions used this criteria indirectly as a "very low occurrence of anthropogenic pressure in the catchment area". Data analyses reveal considerable differences between regions and MSs (see Figure 30).

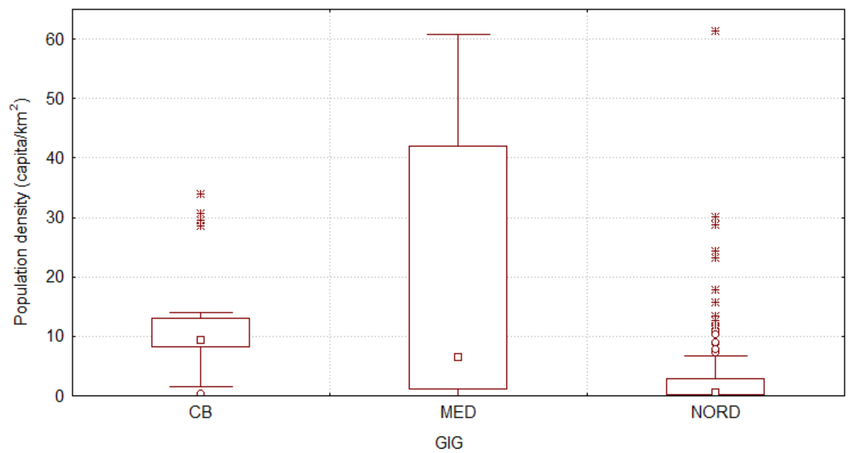


Figure 30. Population density of reference lakes in the Mediterranean (MED), Northern (NORD) and Central/Baltic (CB) GIGs (the highest value of the CB GIG (Lake Busnieku 171 p.e. km^{-2} not included in the graph)

Northern reference lakes have sparsely populated catchments:

- all Swedish reference lakes have population density < 10 km^{-2};
- 80% reference lakes of Norway comply with predefined threshold level < 5 p.e. km^{-2} (Figure 31);
- also catchments of UK and Irish reference lakes have population density mostly < 5 p.e. km^{-2};
- the median value of population density for Northern region reference lakes is low (1 p.e. km^{-2}).

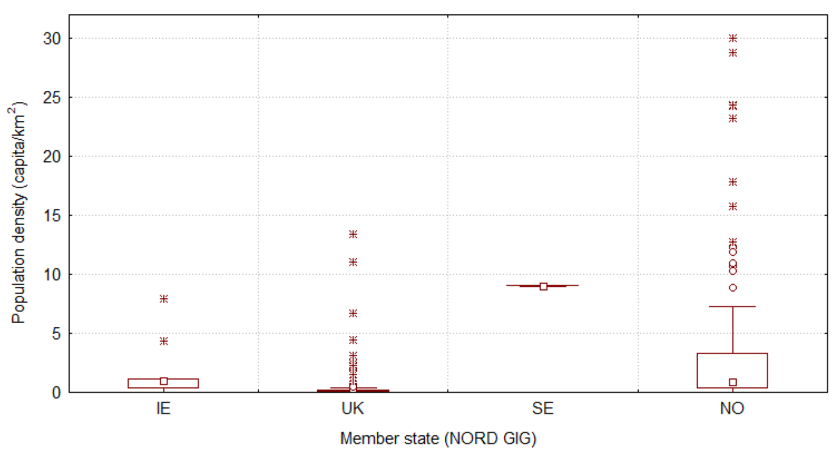


Figure 31. Population density in catchment of reference lakes in the Northern region (IE – Ireland, SE – Sweden, NO – Norway, no data provided from Finland)

Conversely, 32% of Central/Baltic reference lakes exceed the threshold level 10 p.e. km^{-2} (see comparison of MSs in Figure 32), the most outstanding difference occurs at the lake of Busnieku in Latvia (171 p.e. km^{-2}).
In the Mediterranean region, the median value of population density is low (6.5 p.e. km^{-2}: Figure 30) and comparable with median values for Northern region, but there are several reference lakes with very high populated catchment areas (e.g. Saint Cassien from France: 61 p.e. km^{-2}).

Conclusions:

- There are considerable information gaps i.e. no population data from several countries (Finland, Austria), and a low number of reference lakes with population density data (6 in the Mediterranean region);
- There are considerable differences in population data both between MSs and between GIGs: lowest population density values are found in the Northern region (median value 1 p.e. km^{-2}), highest – in the Central region (median value 9.3 p.e. km^{-2}) where 32% of reference lakes exceed the agreed threshold value;
- Population density is very variable in the catchments of Mediterranean reference lakes; the average value (median value 6.5 p.e. km^{-2}) is low but there are several highly populated lake catchments.

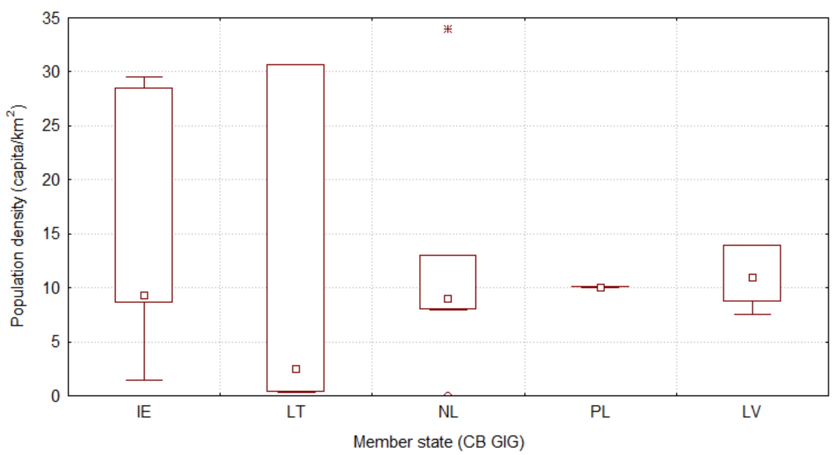


Figure 32. Population density of reference lakes in the Central/Baltic region (IE – Ireland, LT – Lithuania, NL – the Netherlands, PL - Poland, LV - Latvia, no data from AT, DK, DE, EE, ES, FR, GR, IE, IT, LV, SI, UK)

Land use data

Artificial land use was not used as a quantitative reference criteria except in some countries in the Northern region (< 0.1% in UK and < 1% in Ireland), however most of the regions included it as a descriptive criteria, e.g., “no (or insignificant) urbanization or peri-urban areas” (Alpine region).

Comparison of the proportion of artificial land use in reference lake catchments shows considerable differences between regions and MSs (Figures 33 and 34):

- As with population density, most of the Northern reference lakes have very low values of artificial land use – in general < 1% (95th percentile 1.4%; only 5 lakes out of 253 have values > 2%);
- On average, artificial land use is also low in Celtral/Baltic and Mediterranean reference lakes, however some reference lakes exhibit high values (e.g. 15% for Almind Sø in Denmark, 13.5% for Zevenhuizerplas in the Netherlands, 6.6% for Saint Cassien in France);
- The highest share of artificial land use is found in the Alpine region where the median value equals 3.2%, with some lakes reaching more than 20% (Pressegger See in Austria).

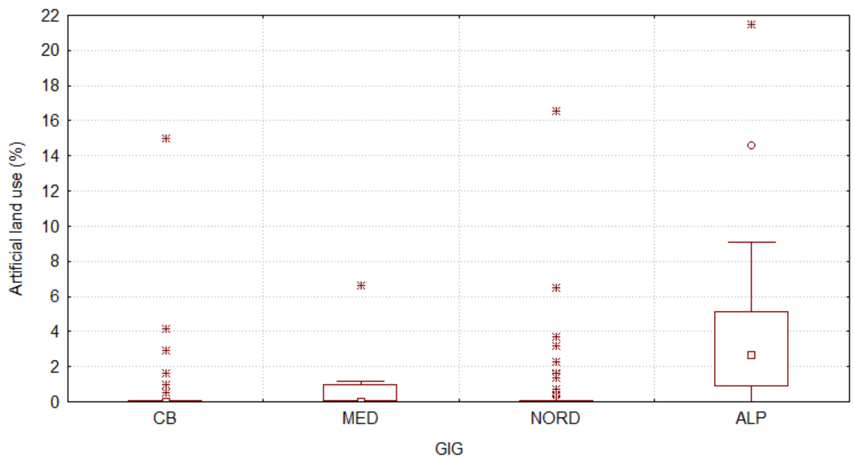


Figure 33. Artificial land use (% from catchment area) in catchments of reference lakes in Central/Baltic (CB), Mediterranean (MED), Northern (NORD) and Alpine (ALP) region

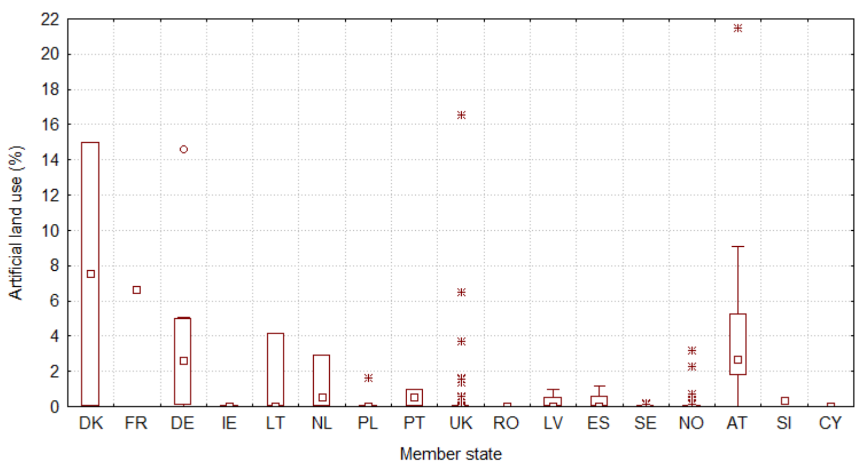


Figure 34. Artificial land use (% from catchment area) in catchments of reference lakes in Member States

Agriculture

Agricultural land use was included in the Northern region reference criteria (< 10% in catchment, < 5% Norway). In other regions, the criteria was included in a descriptive (e.g., the use of agricultural land is very extensive, no artificial fertilizers are used – Celtral/Baltic region) or an indirect way (e.g., absence of major modification to catchment – Atlantic region).

An analysis of the data revealed that the percentages of intensive agriculture areas are low in the catchment area of Northern reference lakes, but considerably higher in other regions (see Figures 35 and 36) – the highest being in Poland (median value 9.4%), Portugal (7.4%) and Latvia (median value 2%, range from 0 to 38%).

Intensive agriculture areas are in some MSs considerably higher, suggesting that the pressure is higher. It is, however, important to note that the agricultural practice differ in the intensively used agricultural areas amongst the MSs. Thus, 0.2% in the UK or NL may cause a comparable pressure as 2% in LV. In other words, 'pressure' is quite difficult to "catch" even with Corine land cover, especially comparing different regions.

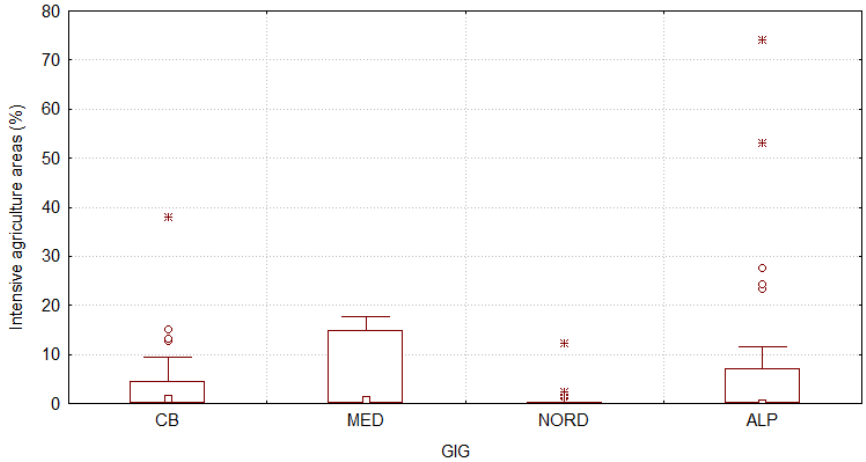


Figure 35. Intensive agriculture areas (% from catchment area) in catchments of reference lakes in Central/Baltic (CB), Mediterranean (MED), Northern (NORD) and Alpine (ALP) region

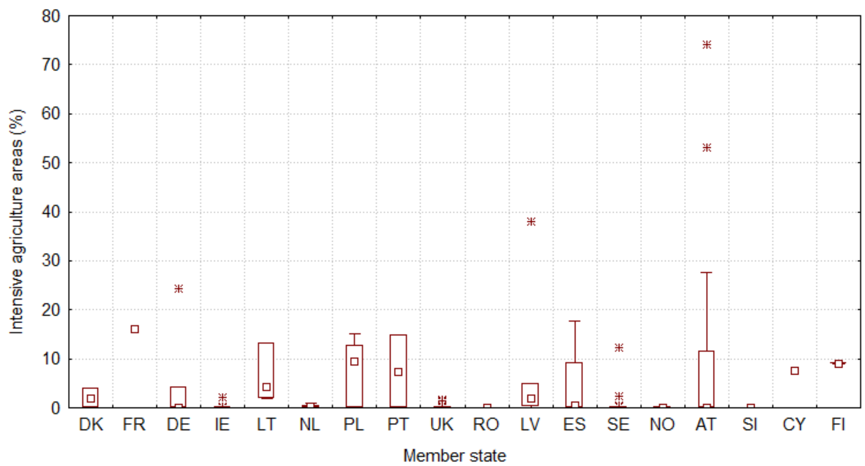


Figure 36. Intensive agriculture areas (% from catchment area) in catchments of reference lakes in Member States

Natural land cover

An important reference criteria is the share of natural land cover which was used as one of the core criteria in the CB GIG (> 90%), as well as in the Alpine GIG (80-90%) and Mediterranean GIG (70-90%). However, the initial analysis (Figures 37 and 38) revealed that the agreed reference thresholds were frequently exceeded in most MSs from the CB and ALP GIGs (the most striking examples –the Netherlands, Latvia, Ireland). Conversely, natural land cover was high in general in the Northern reference lakes (median value 94.1%) and Mediterranean region (median value 89.5%).

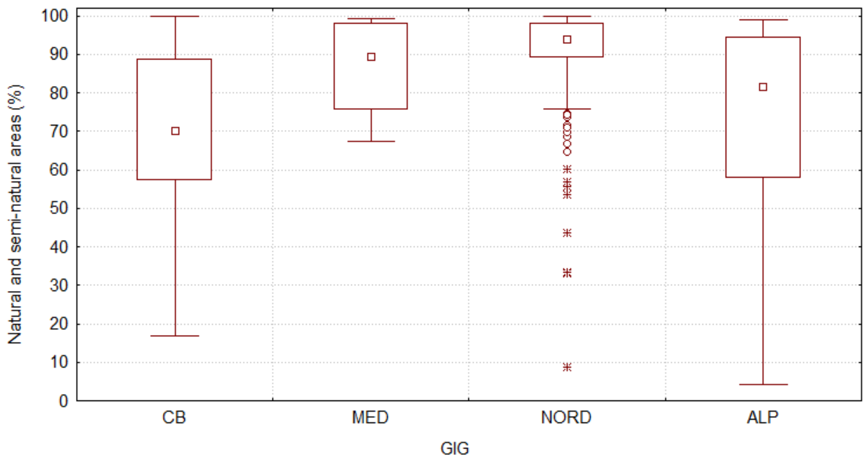


Figure 37. Natural and semi-natural areas (% from catchment area) in catchments of reference lakes in Central/Baltic (CB), Mediterranean (MED), Northern (NORD) and Alpine (ALP) region

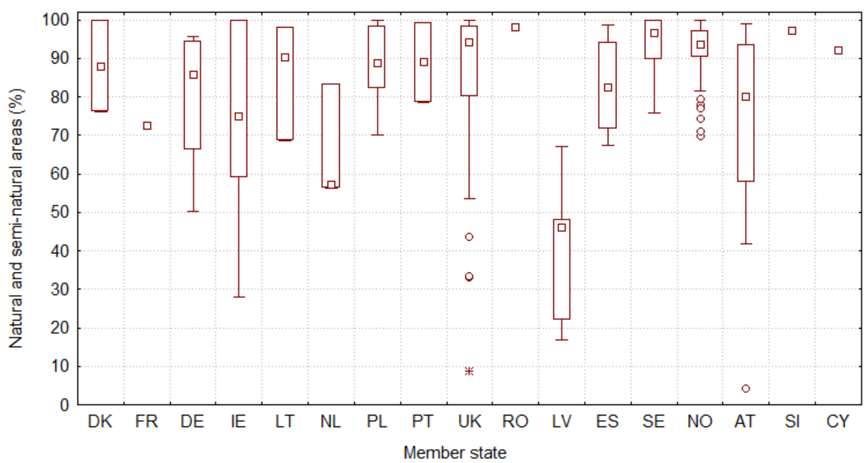


Figure 38. Natural and semi-natural areas (% from catchment area) in catchments of reference lakes in Member States

Conclusions:

- There are considerable differences in reference lake catchment land use patterns both between MSs and between GIGs;
- Northern reference lakes show a low proportion of artificial land cover (95th percentile 0.46%) and intensive agriculture (95th percentile 0.13%), as well as high proportion of natural and semi natural land cover (median value 94.1%), mostly corresponding to reference criteria;
- Also Mediterranean reference sites exhibit low proportion of artificial land cover and intensive agriculture, as well as high proportion of natural and semi natural land cover (median value 89.5%);
- Central/Baltic reference lakes have variable and on average low proportions of natural and semi natural landcover, which do not correspond to the predefined reference criteria (e.g. median value of natural land cover is only 70%, while the reference threshold is 90%);
- Also Alpine GIG reference lakes have variable land use patterns with a high average artificial land use and variable natural land use (e.g. median value of natural land cover is only 80.1%, range from 4.3 % to 98.8% while reference threshold requires 80-90%). Other reference criteria were used in Alpine GIG to ensure that the reference lakes are not impacted by anthropogenic pressure.

Total phosphorus

Total phosphorus (TP) was used as a reference criterion in the Alpine region (defined thresholds corresponding to natural trophic state), in some Nordic countries (Sweden and Norway: TP < 10 µg/l or higher in humic lakes) and in the Atlantic region (TP < 15 µg/l).

Type-specific TP values found in reference lakes were similar among countries in the Northern region (e.g. see Figure 39 for LN2a type) and Alpine region (e.g. see Figure 40a for LAL3 type, Fig 40b for both Alpine intercalibrated lake types together), but considerably different in the CB GIG (Figure 41) where the difference between median values of MSs reference lake populations varied considerably (e.g. median for Latvia lakes 0.012 mg/L, for the Netherlands lakes 0.049 mg/L).

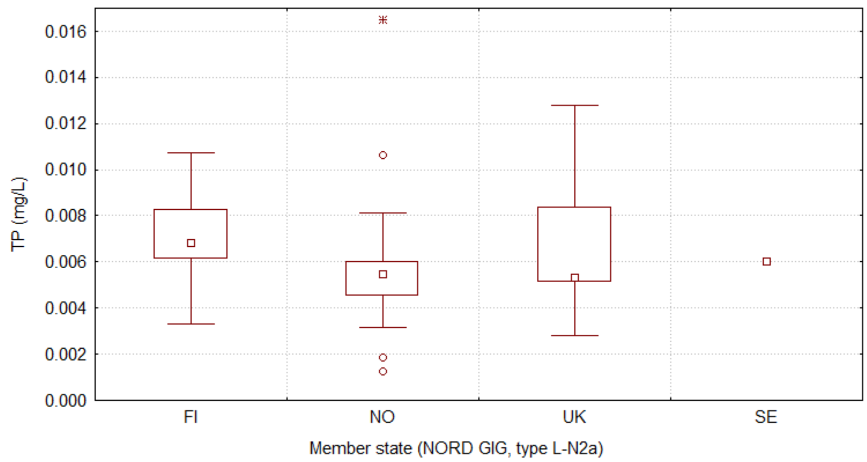


Figure 39. Total phosphorus concentrations in reference lakes of Northern region (type LN2a)

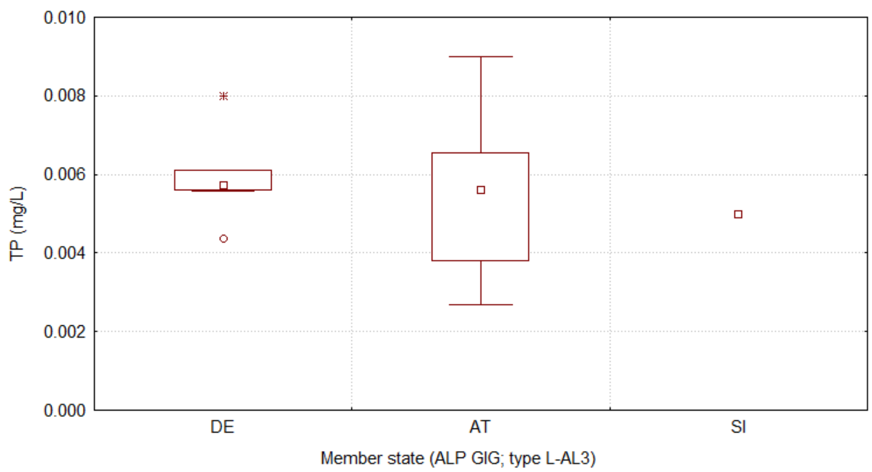


Figure 40a. Total phosphorus concentrations in reference lakes of Alpine region (type LAL3)

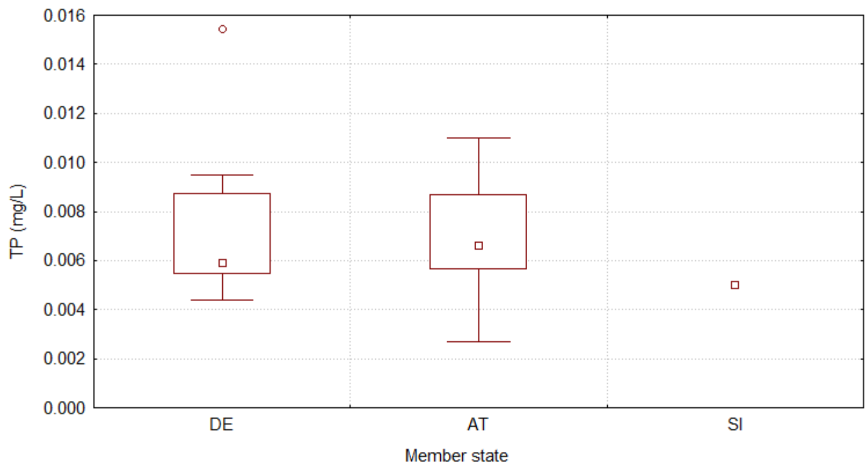


Figure 41b. Total phosphorus concentrations in reference lakes of Alpine region (type LAL3 and LAL4 together)

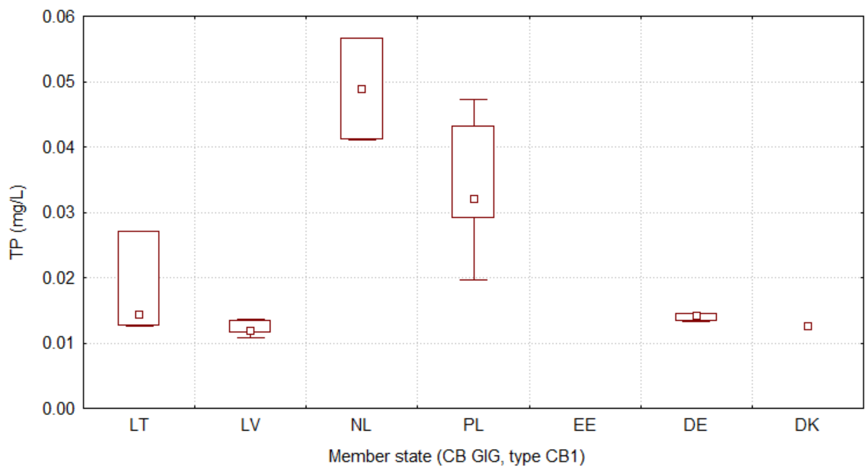


Figure 42. Total phosphorus concentrations in reference lakes of Central/Baltic region (type CB1)

Comparison of phytoplankton (PH) and macrophyte (M) reference lakes

It was possible to compare characteristics of reference lakes selected for setting Reference Conditions for phytoplankton and macrophytes in two regions (ALP and NO).

The results show that:

- Total phosphorus concentrations are broadly similar for Alpine GIG phytoplankton and macrophyte reference lakes (Figure 43);
- Also there was no significant difference between pressure characteristics of Northern phytoplankton and macrophyte reference lakes (Figure 44a; phytoplankton reference lakes: TP median value 0.007 mg/l, 95th percentile: 0.023 mg/l, macrophyte reference lakes: TP median value 0.009 mg/l, 95th percentile: 0.031 mg/l). It has to be considered that Northern Macrophyte IC types do not correspond directly to the Northern Phytoplankton IC types, for example, shallower lakes with mean depth less than 3 meters were included. It can explain for some lakes with higher TP values (see Figures 44a and 44b).

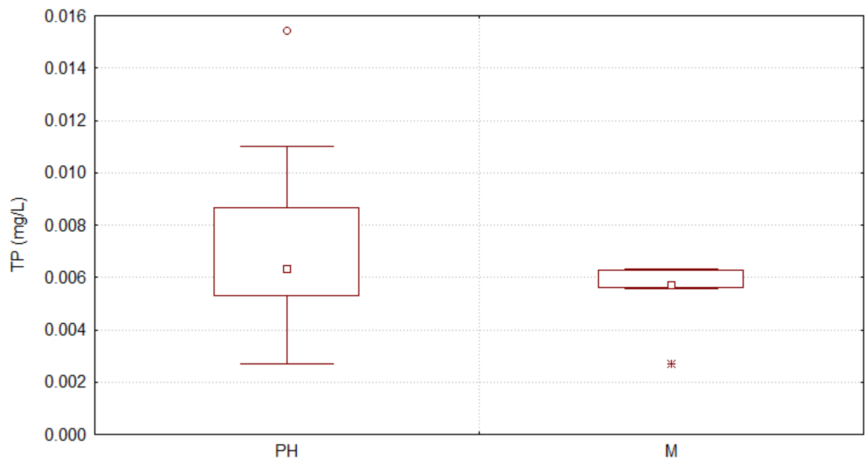
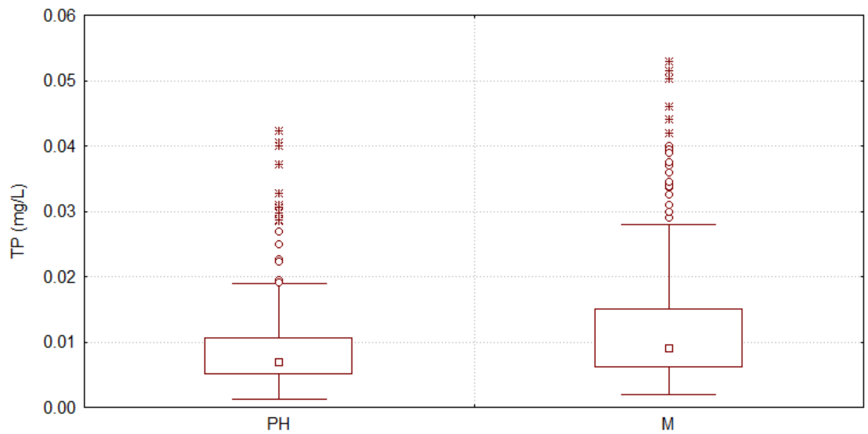
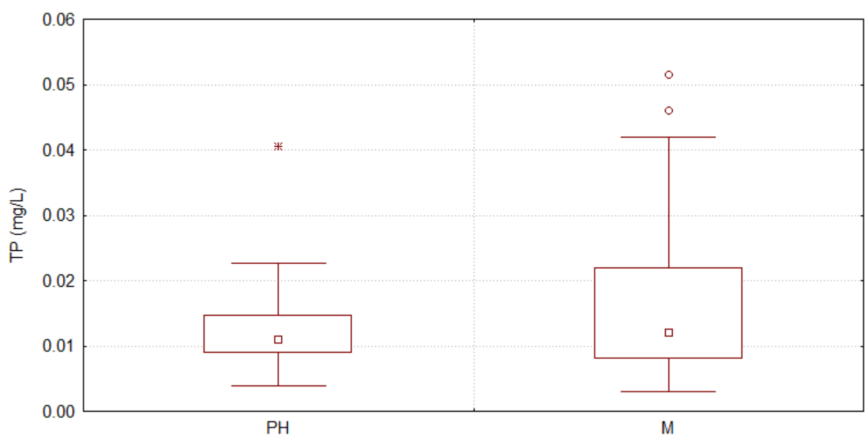


Figure 43. Total phosphorus concentrations in phytoplankton (PH) and macrophyte (M) reference lakes of Alpine region



(M: without types 301 and 302)

Figure 44a. Total phosphorus concentrations in phytoplankton (PH) and macrophyte (M) reference lakes of Northern region (all lake types, except Macrophyte IC lake types 301 and 302 are excluded because they were not included in the Phytoplankton IC)



(PH: types LN3a and LN6a; M: type 102)

Figure 44b. Total phosphorus concentrations in phytoplankton (PH) and macrophyte (M) reference lakes of Northern region (Phytoplankton IC types LN3a and LN6a vs corresponding Macrophyte IC type 102)

2.3. Conclusions

Pressure data

- There are considerable differences **in population density** data both between MSs and between GIGs: the lowest values were generally found in the Northern region (median value 1 p.e. km^{-2}) and mostly complied with the reference thresholds. The highest values were found in the Central/Baltic region (median value 9.3 p.e. km^{-2}) where 32% of reference lakes exceeded the agreed threshold value;
- **Artificial land use** is very low in the catchments of Northern reference lakes and variable in the catchments of CB and MED reference lakes. However, the highest share of artificial land use is found in the Alpine region where median value equals 3.2%, with some lakes reaching more than 20% (e.g. Pressegger See in Austria);
- **Intensive agriculture** is low in the catchment area of Northern reference lakes (complying to reference threshold < 5-10%), while considerably higher in other regions, being highest in Poland (median value 9.4%), Portugal (7.4%) and Latvia (median value 2%, range from 0 to 38%);
- **Natural land cover** was high in general in the Northern reference lakes (median value 93.8%) and Mediterranean region (median value 89.5%), while agreed reference thresholds were frequently exceeded in the CB (median value of natural land cover 70%, reference threshold 90%) and ALP GIGs (median value of natural land cover 80.1%, reference threshold 80-90%);
- **Total phosphorus values** found in reference lakes were homogenous within lake type for the Northern region and Alpine region, but considerably different in the CB GIG where the difference between median values of MSs reference lake populations varied by a factor of four;
- Total phosphorus concentrations are broadly similar in reference lakes for **phytoplankton and macrophytes** in the Alpine GIG as well as in the Northern phytoplankton and macrophyte reference lakes.

The summary:

Three regions present different responses and also different outcomes (see Table 11)

– Northern GIG reference lakes: values for pressure data are low, similar total phosphorus values within type;

– Central/Baltic GIG reference lakes: high variability of pressure data and high variability of total phosphorus values among countries, thresholds exceeded in many cases;

– Alpine GIG reference lakes: high values for pressure data but stable and low total phosphorus concentrations. This is because the lakes that exceeded the threshold values were included as the reference lakes for other reasons: i.e. no deviation from historical or paleolimnological data in terms of taxa composition of phytoplankton, no significant anthropogenic contribution to total nutrient load as exerted from nutrient budget calculations, no deviation from natural trophic state as defined by nutrient concentrations in the lake.

Region	Availability of reference lakes	Criteria used for selection of reference lakes	Outcome
Northern	High	Both pressure data and total phosphorus data used	Reference lakes with low pressure characteristics and low total phosphorus values (similar between MS)
Central/Baltic	Low	Mostly pressure data used, many exemptions allowed based on expert judgement	Reference lakes with variable pressure characteristics and diverse total phosphorus values (different between MS)
Alpine	Medium	Focus on total phosphorus data, based on assumption that land use and population density are not the right tools to define pressure for Alpine lakes	Reference lakes with variable pressure data, but low and homogenous total phosphorus concentrations (similar between MS)

Table 11 . Comparison of approaches to select reference lakes in three intercalibration regions.

Limitations of the study:

- There is the risk that we might have missed some important pressure criteria within the different GIGs. For example forestry (clear-cutting and drainage ditching) is an important pressure that affects both the physico-chemical and biological conditions in lakes. Especially in the Nordic countries, forestry is probably the most important land use management type. The intensity of forest management should thus be included in the selection procedure for reference lakes. Forestry-related variables are difficult to extract from land use databases, thus such information often needs to be digitized manually which is a time-consuming approach;
- Further distinguish between natural and semi-natural area is important, e.g., reference conditions (in lakes of which the catchment is dominated by mires are probably different from those in lakes dominated by grassland in the catchment);
- Also agricultural practices differ in MS, so loads from “intensively used agricultural areas” or “agricultural areas” can differ substantially between different MS – in other words, similar proportions of land use values are not representing equal level of pressure across Europe;
- Atmospheric deposition was not included as pressure but can be important for some lake types;
- Some important factors were not taken into consideration, as area (or volume) of lake and flushing rate, (e.g., it is well known that the water bodies with faster cycling of water could assimilate a larger phosphorus load with no adverse eutrophication responses than slower-flushing water bodies);
- In our study we expressed land use pressures in percent of catchment area, while the most correct approach would be to relate land use area to the area of lakes or preferably to the volume of the lakes;
- A further potential problem with land use data is that they do not generally take into account the proximity of particular land uses to the water body. It is found that attenuation of nutrients due to subsequent transport can be significant and that this significance increases with catchment size;
- It has to be stressed that consequences of some of the conclusions about the reference consistency are very different for the different IC options used:
  - Option 2 (comparison of assessment systems via Intercalibration common metrics) results are much more sensitive for the reference value used by the MS than most of the option 3 or option 1 results where the reference value is not affecting anymore the intercalibration results in the comparability between MSs;
  - In the Lake IC, mainly IC Option 1 (Phytoplankton biomass – all GIGs) or Option 3 (Macrophytes in the CB Gig and ALP GIG) were used, so this notion makes the consequences of the conclusions less pessimistic than as they are now.

Relationship between land-use and chlorophyll-a concentrations in reference lakes

To further explore the relationship between pressure criteria and biological response in reference lakes, we analysed:

- Chlorophyll-a values in reference lakes vs natural land cover (Figure 45);
- Chlorophyll-a values in reference lakes vs agricultural land cover (Figure 46).

Results show that there is generally weak relationship between chlorophyll-a values and land-use categories:

- For natural land cover the only significant differences (Mann-Whitney test) were between the following groups: 50-60% vs 60-70%, 80-90% and 90-100%;
- For agricultural land use the only significant differences (Mann-Whitney test) were between the following groups: 0-10% and 40-50%.

In other words, there is no significant impact to chlorophyll-a values in reference lakes in ranges of 60-100% natural land cover and 0-40% agricultural land use.

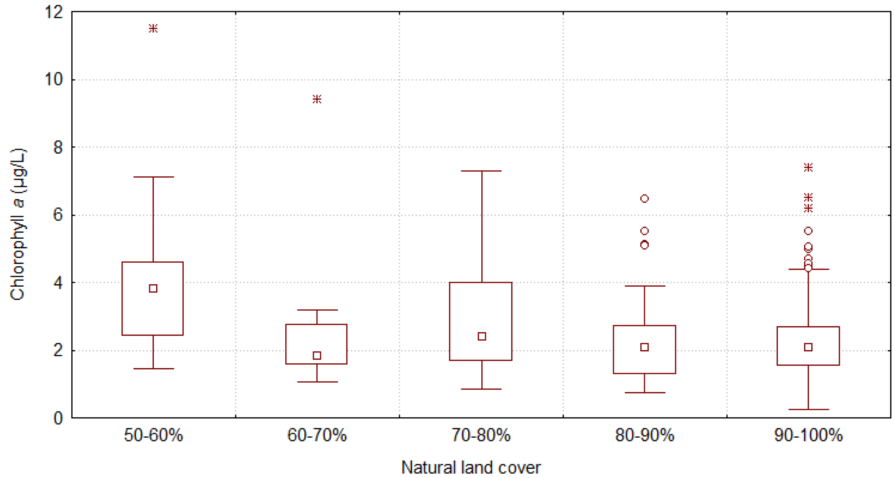


Figure 45. Comparison of Chlorophyll-a values in reference lakes with different natural land cover in catchments (all regions and all lake types). The only significant differences (Mann-Whitney test) are between the following groups: 50-60% and 60-70%: P-value = 0.00148; 50-60% and 80-90%: P-value = 0.019; 50-60% and 90-100%: P-value = 0.0001

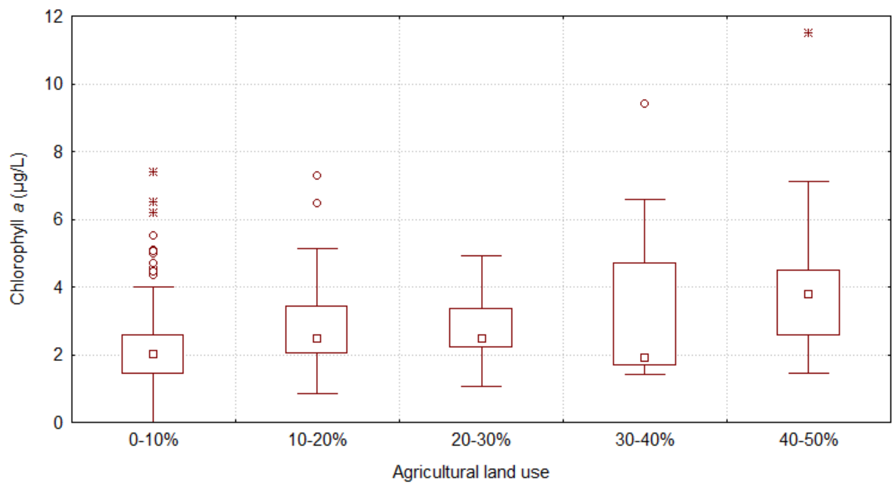


Figure 46. Comparison of Chlorophyll-a values in reference lakes with different agricultural land cover in catchments (all regions and all lake types). The only significant differences (Mann-Whitney test) are between the following groups: 0-10% and 40-50%: P-value = 0.0002

The study highlights the question: what is the right tool to select reference lakes - pressure data (land use, population density) or total phosphorus data?

It is well known that land use in the catchment of lakes is an important determinant of water quality, trophic status and biological communities (Knoll *et al.* 2003, Arbuckle and Downing 2001). However, at the scale of entire watersheds, relationships between land use and water quality are surprisingly variable (Hunsaker et al. 1992, Osborne and Kovacic 1993). Although the proportional areas of different land use types in watershed explain some variance in water quality parameters, much remains unexplained (Rechkow et a.l 1980, Osborne and Wiley 1988). Many studies have demonstrated that effects of land cover on water quality and biological communities are affected by numerous other factors, e.g., drainage ditching (Ecke 2009), presence/absence of WWTPs (Hill 1981), spatial arrangement of land cover (King et al. 2005), watershed size (Strayer 2003), etc.

Lake IC analyses (e.g. Wolfram et al. 2009) show that there is a correlation between land use and trophic state, but usually over a wide range and not in the very narrow lower end of the trophic gradient where the ecological status changes. If we want to define pressure from the catchment area, we have to include a lot of other information, e.g., quantitative and detailed information on spatial distribution of land use, intensity of agriculture and wastewater treatment which is not easily accessible information (Figure 47).

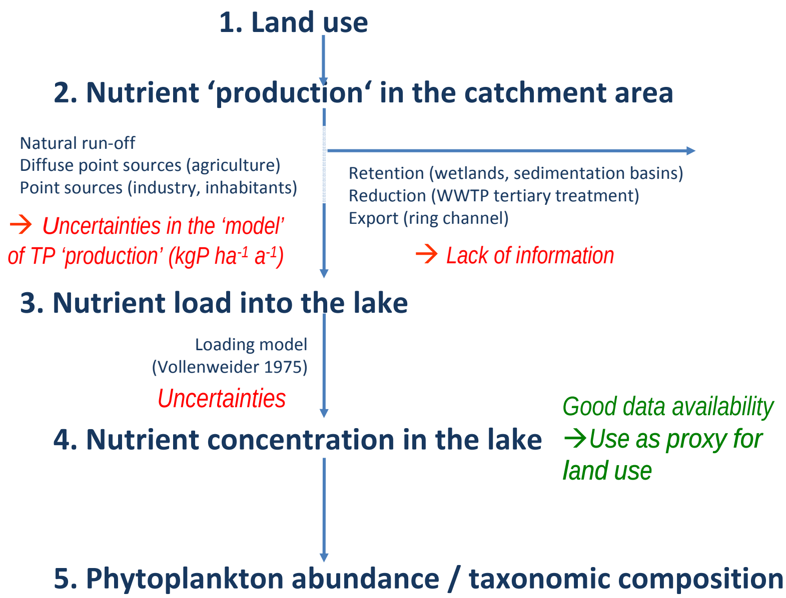


Figure 47. Factors affecting nutrient load and concentrations in the lake

Since the response of phytoplankton to pressure criteria like land use is not a simple and direct correlation, another way was agreed in the Lake IC groups for the 2nd phase of the IC - to use directly nutrient levels within the lake which are correlated to the anthropogenic impact.

The unanimous common opinion of all Lake BQE-GIG leaders at the Lake IC meeting (JRC, Ispra, 2009) was:

- Land use in the catchment area and population density may not be the right tools to define pressure in lakes;
- Similar values are not representing equal quantity of pressure throughout Europe;
- BQE and thus the ecological status of a lake will react on the pressures (nutrient levels) and not on the drivers (land use);
- The WFD stipulates the assessment of the ecological status of a lake, not of the catchment area;
- And summarized: land use is not a suitable proxy for anthropogenic pressures on a lake.

Nevertheless, it must be acknowledged that there is a considerable risk to rely only on TP levels:

- At first, we need accurate and well-grounded TP reference criteria (there is a risk that we decide that e.g. 0.05 TP mg/l is a reference level - which could be wrong);
- Secondly, a comprehensive information base why the particular lake is not impacted by pressures in the catchment.

Alpine GIG used a number of different approaches to set TP reference values: historical data, modeling of natural nutrient loads, paleoreconstruction (form literature) + data from unimpacted lakes. If we have such a profound background, we can be sure that the defined TP levels reflect "true reference" and can be used as a proxy for eutrophication pressure. We have different situations in other GIGs where TP reference levels are taken arbitrarily, without a sufficient base – so there is a risk that these TP reference levels do not reflect reference but some (unknown) level of human impact.

Discussion on TP as eutrophication proxy can be summarized as follows:

- TP can be used as eutrophication proxy to select reference lakes and define biological reference conditions;

– Still this approach has to be used with a caution, including a thorough analysis of reference TP levels (best approached through combination of available data, hindcasting using export coefficient models (Johnes 1996), historical data, paleoreconstruction and expert judgment). Additional information should be included why lake is not impacted by pressures in the catchment (e.g., because there is no deviation from historical or paleolimnologicla data, no significant anthropogenic contribution to nutrient load as exerted from nutrient budget calculations, etc).

About definitions and using alternative benchmarks

Tiered approach to reference screening would be recommendable:

- Tier 1 - "true" reference sites – sites with no or minimal anthropogenic pressure that fulfill all criteria proposed in REFCOND Guidance for all pressures;
- Tier 2 - "reference conditions" sites or "partial" reference sites – impacted by some level of anthropogenic pressure but (some) biological communities corresponding to the reference conditions (e.g. "phytoplankton reference sites" with no or minimal eutrophication pressure but significant hydromorphological pressure which still is not affecting phytoplankton community in a significant manner);
- Tier 3 - "alternative benchmark" sites – sites with some pressure and some level of impairment to biology (can be used for setting benchmark, see EC, 2010).

These findings highlight the importance and necessity to harmonize reference criteria between the MSs and geographical regions, and between BQEs, and at the same time take regional peculiarities and factors influencing lake ecological status into account.

3. Coastal water bodies

Introduction

The section on the description of reference conditions for coastal waters differs considerably from the approach used for rivers and lakes due to a number of problems in the use of reference sites for coastal and transitional waters. The following points had already been made in 2003 in the Guidance nr. 5 on Typology, Reference Conditions and Classification Systems for Transitional and Coastal Waters (EC 2003b):

- In the marine environment there is a lack of biological and chemical data for high status sites as the focus for monitoring programmes has been centred on polluted areas.
- There are very few sites across the whole of Europe at high status because of wide spread human pressures and impacts.
- Deriving and making complete descriptions of reference conditions for coastal and transitional waters was not possible in the 1st phase, as there are few or no data for some of the biological quality elements. Derivation of reference conditions that encompass the full natural variability found within a water body type is likely to take many years. It will be an iterative process and will be assisted by the collection of monitoring data for the purposes of the Directive over the forthcoming years (since not existing for some BQEs). Defining correct reference conditions will probably take a duration at least as long as the first and second river basin management plans. At least in the short term, expert judgement is essential because of the lack of good data sets. Over the forthcoming years as understanding increases it may be possible to develop sound predictive models, thus reducing the degree of expert judgement.

Reference sites have been used only in a few cases of derivation of reference conditions. For these cases, a general agreement on the accepted level of human pressures within the coastal GIGs or reference criteria to select reference sites (or minimally disturbed sites) have not been documented. Most of the reference conditions used during the IC exercise are biologically derived, by either models or by expert judgement. No common checking of the absence of pressures or a common evaluation of ecological responses in relation to gradients of pressures has been documented, so the revision of the consistency between MS when quantifying pressures can not be made. This review on reference conditions aims at giving an overview of the different approaches used to define minimally disturbed conditions (since real reference conditions were often not available).

In the technical report on intercalibration of the first phase, the terminology of a reference (condition) is often not used in a consistent way. The word reference is often not pointing to the high status, but is used in the meaning of the starting point on the ecological gradient from high to moderate status where enough data are available to define a starting point that can be the good status or H/G or G/M boundary, to compare a recent assessment with. In the overview below the word reference conditions is used consistently for the high status only. In all the cases where reference sites were not available and modeling and expert judgment were not able to estimate the high status in a direct way but the good status or the H/G or the G/M boundary, the word alternative benchmark should be used for these boundaries and the good status.

3.1. Phytoplankton

NE Atlantic, Baltic Sea and Mediterranean Sea GIG have intercalibrated **single metrics based on chlorophyll *a* concentration**, while Romania and Bulgaria in the Black Sea GIG agreed on the phytoplankton boundaries based on **seasonal average biomass** (mg/m^3). NEA and MED GIG used 90% percentile chlorophyll *a* concentration of the growing season whereas the Baltic Sea GIG used summer mean chlorophyll *a* concentration as common metric for phytoplankton biomass assessment. Only NEA GIG intercalibrated also other metrics (besides chlorophyll *a*) indicative for phytoplankton composition. Those were Indicator Taxa (Frequency of *Phaeocystis* Cell counts) and Taxa Cell Counts (Frequency of phytoplankton taxa cells counts).

In all the GIGs the reference concentration and boundary setting for chlorophyll *a* was defined based on separate national datasets. There was no common approach applied among GIGs or MSs on how to fix reference conditions and no common database was built for benchmarking, to verify and validate the high status, the H/G or the G/M boundary at the scale of the common intercalibration type, a clear requirement of the updated IC guidance for the second IC phase. The summary below clarifies in which cases reference sites have been used and where expert judgment or a mixture of modeling, statistics on historical data and empirical relationships with other metrics (e.g. Secchi depth, see Baltic GIG) have been applied. Reference sites have been selected in the Mediterranean Sea GIG, and in Ireland, UK, France and Spain in the NEA GIG for type 1/26a and 1/26e only. A checking procedure on the pressure criteria used, in order to verify a consistent selection of reference sites, has not been performed.

Remarkably, the reference or alternative benchmark for chlorophyll *a* seems in many cases (especially in the MED GIG), not to be established based on illustrated relationships with physico-chemical parameters like nutrients. These relationships have only been illustrated for the Baltic Sea GIG and for The Netherlands in the NEA GIG. For several countries in the NEA GIG expert judgments were performed on available measurements, without explanation on the level of absence/presence of a pressure.

A. Baltic Sea GIG

In the Baltic Sea no reference sites exist and the relationship between Secchi depth and chlorophyll *a* or nutrient concentration was considered of major importance to define the high status or an alternative benchmark. Denmark and Germany also used hind-casted modeled estimates of chlorophyll *a* concentrations.

For **Denmark** and **Finland** the starting point to define the alternative benchmark was the Secchi depth of the early 1900s (1925-1934 for Finland, 1903–1959 for Denmark). **Estonia** used the summer Secchi depth (high status = 8 m) and nutrient and chlorophyll *a* concentrations of the period 1993-2005. **Sweden** considered a summer Secchi depth in the Baltic of 10 m as corresponding to the high status in the Baltic proper and **Poland** 6 m as the H/G boundary. These Secchi depths were converted into chlorophyll *a* concentration of high status or H/G boundary respectively, based on relationships between Secchi depth and chlorophyll *a* obtained from recent monitoring data from coastal waters. Boundaries between good and moderate status for chlorophyll *a* were then defined as high status plus 50 % in Denmark and Estonia and H/G boundary plus 50 % in Poland, in accordance with the HELCOM Eutrophication approach. The H/G boundary was defined as reference conditions plus 20% in Estonia and as the 95% confidence limits of variation of calculated historical values at its maximum in Finland. Denmark discovered that the approach using Secchi depth resulted in less reliable and higher reference concentrations than the calculated results based on hind-casted estimates of reference loading and nutrient inputs, relating historical nitrogen inputs with total nitrogen levels and then with chlorophyll *a* concentrations. So the latter approach was selected by Denmark. For Poland, the H/G boundary estimate from the summer period relationship between chlorophyll *a* concentration and TN (from data from 1999–2005) was lower than the result from the relationship between chlorophyll *a* concentration and Secchi depth, but Secchi depth remained the preferred approach.

Germany did not consider the Secchi depth as the most important determining factor to derive reference conditions, but they modeled the relationship between TN and chlorophyll *a* with data from 1978 to 2004, the marine high status value of 10 µM TN taken from literature and adjusting for all salinities by combining this marine reference value with freshwater references of relevant rivers in mixing diagrams. Chlorophyll *a* high status values were calculated using regressions between TN and chlorophyll *a* during the growing season.

From the relationship between TN and Secchi depth Sweden defined the reference TN concentration (from 1900-1920) as 15,3 µM and the chlorophyll *a* concentration as 1,2 µg/l for the Baltic open coastal waters, correcting individual water body reference conditions for background concentrations in freshwater by means of mixing models. Only the southern part of the type of the Gulf of Bothnia (= B0 with salinity 0,5 - 3) was considered to be influenced by the Baltic proper and therefore got the same reference conditions as the Baltic proper. For the northern part of the Gulf of Bothnia the reference is higher due to the influence of humic substances from freshwater increases.

Table 12 summarises the methodology used to derive the reference high status or the H/G boundary (in the case of Poland)

Member State	Historical abiotic data		Hind-casting of abiotic data	Recent abiotic / biotic relationship
	Secchi depth	TN		
Finland	1925 - 1934			Chlorophyll *a* and Secchi depth, Depth limit of Fucus vesiculosus, occurence of cyanobacteria
Sweden	1900 – 1950: 10 m	15,3 µM		Secchi depth related to TN and TN related to chlorophyll *a*
Denmark	1903 - 1959		Nutrient loading and inputs related to TN	TN related to chlorophyll *a* with recent data (May-Sept)
Germany		10 µM	TN loading	TN related to chlorophyll *a* with recent data (March-Oct 1978 – 2004)
Poland	Summer 6 m			Secchi depth and TN related to chlorophyll *a* with recent data (May-Sept 1999 – 2005)
Lithuania				
Latvia				
Estonia	8 m	10,6 µM	1.1	TN related to chlorophyll *a* with recent data (June-Sept 1993 – 2005)

Germany and Sweden used a reference value for TN for marine waters that differed considerable (DE 10 µM North Sea or SE 15,3 µM Baltic proper). The TN reference value derived by Estonia was close to the German one, so lower than the Swedish TN reference value. However, Estonian secci depth should then be deeper than the Swedish one, which is not the case.

B. North-East Atlantic GIG

1. Biomass: chlorophyll a 90^{th} percentile

Table 13: overview of the methodology used to derive reference conditions in each type or subtype of NEA 1/26

Type	Member State	Reference conditions	Boundary		Geographical delineation
			H/G	G/M	
NEA 1/26a	Norway	No long term historical series, expert judgment and the Norwegian Pollution Control Authority 1997 guide used, < 2.5 µg/l	2.5 µg/l	5 µg/l	Eastern Cantabrian Coast, Canary Islands, Atlantic Coast, Western Irish Sea, Scandinavian Coast: North Sea Norway + Norwegian Sea
	Ireland	Reference sites in undisturbed waterbodies by expert judgment (Art.5 report and environmental data)			
	UK	Reference concentration 3.33 µg/l derived from the high/good boundary	5 µg/l as 50% elevated value by expert judgment for Atlantic Coast offshore waters	10 µg/l	
	France Channel + Atlantic coast	Reference from 1.3 up to 4.4 µg/l based on reference sites (1992 – 2006) lowest values Northern and Western parts of Brittany up to highest in Normandy			
	Spain Eastern Cantabrian Coast	Reference 1 - 3 µg/l, derived from concentrations at reference sites ranging between 0.8 µg/l Pais Vasco and 2.8 µg/l Asturias 90^{th} percentile (narrow shelf, no large plumes are formed and minor upwelling impact)			
	Spain Atlantic Coast	Reference < 5 µg/l (reference stations values from 0.8 µg/l until 2.8 µg/l)	Expert judgment of H/G = 5 µg/l	Expert judment 10 µg/l	
	Spain	Canary Islands: maxima typically < 1 µg/l			
NEA 1/26b	UK	Reference concentration 6.7 µg/l derived from the high/good boundary	10 µg/l as 50% elevated value by expert judgement for chlorophyll a in offshore North sea waters	15 µg/l	Eastern Irish Sea, North Sea (UK East Coast, Channel, NL Wadden Coast, Zeeland Coast, Belgium coast
	Netherlands	Same reference concentrations as UK			
	Belgium		G/M boundary estimated from natural diatom bloom and pre-bloom *Phaeocystis* biomass		
	France	Expert judgment for reference of 6.7 µg/l based on offshore measurements from 1992 – 2006 at Dunkerque (9.9 µg/l) and Boulogne (6.2 µg/l)			
NEA 1/26c	Denmark				German Bight & Jutland
	Germany				
NEA 1/26d	Denmark				Scandinavian Skagerrak
NEA 1/26e	Spain Western Cantabrian Coast	Reference = 2 - 5 µg/l, derived from the concentrations at a reference site in Asturias 2.2 µg/l 90^{th} percentile			Western Iberian Coast upwelling, Western Cantabrian Coast
	Spain Western Iberian upwelling coast	Reference = 4 - 11 µg/l, derived from measurements at Coruna station (1992 – 2006): 4.73 µg/l up to 10.55 at Vigo station, + values of 4 µg/l in southern part and 5 µg/l in northern part (1992 – 2004) of this Spanish area			
	Portugal Iberian	Atlantic Coast: reference < 4 µg/l	H/G = 8 µg/l derived from measurements 5 µg/l to 10 µg/l	G/M = 12 µg/l	

Type	Sali-nity psu	Member State Reference period	Reference conditions	H/G boundary	G/M boundary	Geographical delineation of the type
NEA 3/4	18-30	Germany				Wadden Sea type polyhaline
		Netherlands	9.3 µg/l derived from H/G boundary	14 µg/l deduced from freshwater discharges model simulations and expert judgment as high/good boundary, which is set as 1.5 times the reference, from this G/M 21 µg/l derived for NEA3/4	Cadee and Hegeman (2002) estimated annual average in the seventies (already elevated but below 1.5 times the back-ground) chlorophyll a between 2 and 5 µg/l, 90th percentile as between 5-13 µg/l, elevated values between 11 to 17 µg/l (1.7 higher than UK and 2.1 higher than Germany because of higher fresh water runoff)	
NEA 7	>30	Norway		Expert judgment and the Norwegian Pollution Control Authority 1997 guide used		Deep fjordic and sea loch systems
		UK				
NEA 8	18-30	Norway		Idem NEA 7		Kattegat + small area of Skagerrak (Inner Arc) sheltered, shallow
		Sweden		Reference based on offshore data. Temporal changes of Secchi depth, TN and chlorophyll a used to set reference for each type		
		Denmark				
NEA 9	18-30	Sweden		Idem as NEA 8		Skagerrak fjords with a shallow sill
		Norway		Idem as NEA 7		
NEA 10	18-30	Sweden		Idem as NEA 8		Skagerrak Outer Arc, exposed, deep
		Norway		Idem as NEA 7		

Levels of difference between high/good and good/moderate thresholds in the NEA GIG

In the NEA GIG, the three levels of difference between high/good and good/moderate thresholds are set at 100% for type 1/26a, 8, 9, 10; 50% for type 1/26b, 1/26e, 3 and 4; and 33% for type 1/26c and 1/26d. These typologies reflect the different way the phytoplankton reacts in these waterbodies. In the clear open Atlantic waters, the phytoplankton respond quickly to small changes in nutrients while those areas of more enclosed seas and upwelling's tend to have higher turbidity and nutrients, so may respond more slowly initially but bigger natural blooms can be maintained. This perspective explained in the technical report needs to be discussed among all GIGs.

2. Frequency of blooms: *Phaeocystis* counts

UK, Germany, the Netherlands, Belgium: bloom frequency (percentage of monthly samples with exceedence of bloom threshold) of 8.3% (1 out of 12 samples) as high status, 9% as H/G boundary and 17% as G/M boundary (2 out of 12 samples). The bloom threshold is defined as 1 000 000 *Phaeocystis* cells/l. No explanation was provided in the technical report why this threshold and these bloom frequency percentages are applied.

3. Frequency of blooms: Elevated taxa cell counts (agreed by UK, Ireland, France, Spain and Portugal)

Table 14 gives the specifications for each Member State for the metric on taxa cell counts

Type	Member State	Reference conditions		Boundary		Geographical delineation of the subtype
		Bloom treshold	Bloom frequency			
			H	H/G	G/M	
NEA 1/26a	Ireland	250 000 cells/l large phytoplankton	16.7% = 2/12 samples	20%	39% = 5/12 samples	Eastern Cantabrian Coast, Canary Islands, Atlantic Coast, Western Irish Sea, Scandinavian Coast: North Sea Norway + Norwegian Sea
	UK					
	France Channel + Atlantic coast	100 000 cells/l large phytoplankton, 250 000 cells/l small phytoplankton				
	Spain Eastern Cantabrian Coast	750 000 cells/l (based on phyto-plankton counts since 2002 in the Basque Country)				
	Spain Atlantic Coast + Canary Islands	500 000 cells/l				
NEA 1/26b	UK	Same as in 1/26a	16.7%	20%	39%	North Sea
	France	Same as in 1/26a				
NEA 1/26e	Spain Western Cantabrian Coast	750 000 cells/l	25% = 3/12 samples	30% = 4/12 samples	49% = 6/12 samples	Western Iberian Coast upwelling, Western Cantabrian Coast
	Spain Western Iberian upwelling coast	1 000 000 cells/l				
	Portugal Iberian	100 000 cells/l large phytoplankton, 1 000 000 cells/l small phytoplankton				

No explanation was provided in the technical report why these bloom thresholds and bloom frequency percentages are applied.

C. Mediterranean Sea GIG

Reference conditions have been calculated by selecting High status stations from monitoring programs (1997 – 2005) based on expert judgement.

D. Black Sea GIG

Bulgaria analysed a phytoplankton biomass dataset with earliest data from 1954 (up to 1970) corresponding to high status and a long-term dataset (1983-2005) to derive lowest spring and summer values between 10^{th} and 25^{th} percentile to compare with the thresholds for high status of historical phytoplankton biomass. Based on the relationship between these two datasets, the reference for chlorophyll *a* was defined as the range between 10^{th} and 25^{th} percentile of the data available for chlorophyll *a* per season (1990-2006). Values from the period 1983-1998 correspond to bad status.

Romania analysed a phytoplankton biomass dataset with earliest data from 1960 (up to 1970) corresponding to high status and a long-term dataset (1986-1997) corresponding to bad status. Only values for phytoplankton biomass were derived per season, chlorophyll *a* data not being separately analysed.

Overview of reference concentration values and the H/G and the G/M boundary for chlorophyll *a*

The table 15 shows the reference concentrations and the H/G and the G/M boundary for chlorophyll *a* for all types in the GIGs. The procentual relation between the reference values and these boundaries is also given in order to illustrate the level of difference. The difference between the background (=reference) concentration and the elevated assessment concentration (=good/moderate boundary) was previously set at 50% in the marine conventions like OSPAR and HELCOM. Deviations from these principles are thus also illustrated in the table. Also the time period (when reported) is given from which data originated that have been used to derive reference values for each type. This information should be made available for all reference conditions definitions.

Table 15: Chlorophyll *a* reference values and H/G and G/M boundary for each type in each GIG

Type	Salinity psu	MemberState Reference period	Reference conditions	H/G boundary	G/M boundary	H/G boundary	G/M boundary	Geographical delineation of the type
Baltic Sea CW B0 sheltered	0.5-3	Finland 1925-1932	1.3	1.8	2.7	Ref + 38%	H/G + 50% Ref + 108%	Sites in Botnian Bay (Northern Quark) More influence of humic substances
		Sweden 1900-1920	1.1 – 1.4	1.5 – 1.8	2.0 – 2.3	Ref + 29-36%	H/G + 28-33% Ref + 64-82%	
		ECOM DEC boundary		1.7 (1.5 - 1.8)	2.3 (2.0 - 2.3)		H/G + 35 %	
Baltic Sea CW B2 sheltered	3-6	Finland 1925-1934	1.4	1.8	2.6	Ref + 29%	H/G + 44% Ref + 86%	Sites in Bothnian Sea More influence of humic substances
		Sweden 1900-1920	1.4	1.8	2.3	Ref + 29%	H/G + 28% Ref + 64%	
		ECOM DEC boundary		1.8	2.5 (2.3 - 2.6)		H/G + 39 %	
Baltic Sea CW B3 a sheltered	3-6	Finland 1925-1934	1.8	2.2	2.9	Ref + 22%	H/G + 32% Ref + 61%	Sites in the area extending from the southern Bothnian Sea to the Archipelago Sea and the western Gulf of Finland
		Sweden 1930-1950	1.3	1.6	1.9	Ref + 23%	H/G + 19% Ref + 46%	
		ECOM DEC boundary		2.4 (2.2-2.6)	3.5 (2.9-4.0)		H/G + 46 %	
Baltic Sea CW B3 b exposed	3-6	Finland 1925-1934	1.5	2.6	4.0	Ref + 73%	H/G + 54% Ref + 167%	
		Sweden 1930-1950	1.2	1.5	1.8	Ref + 25%	H/G + 20% Ref + 50%	
		ECOM DEC boundary		1.6 (1.5 - 1.6)	1.9 (1.8 - 1.9)		H/G + 19%	
Eastern Baltic Sea CW B12 a sheltered	5-8	Estonia	1.8	2.2	2.7	Ref + 22%	H/G + 23% Ref + 50%	Sites in the Gulf of Riga
Western Baltic Sea CW B12 b sheltered	8 - 22	Germany?	< 1.1	1.1	1.9		H/G + 73%	Sites at the Southern Swedish coast and the South western Baltic Sea open coast along Denmark and Germany
		Denmark 1903-1912	1.4	1.5	1.9	Ref + 7%	H/G + 27% Ref + 36%	
		Sweden?	1.2	1.5	1.9	Ref + 25%	H/G + 27% Ref + 58%	
		ECOM DEC boundary		1.3 (1.1 - 1.5)	1.9		H/G + 46%	
Baltic Sea CW B13 Exposed	6-22	Estonia ?	1.1	1.3	1.65	Ref + 18%	H/G + 27% Ref + 50%	Sites along the coast of the Estonia, Latvia and Lithuania, the Polish coast and the Danish island "Bornholm"
		Latvia ?	1.1	1.3	1.65	Ref + 18%	H/G + 27% Ref + 50%	
		(Lithuania) ? not in EC DEC	3.2	4	5	Ref + 25%	H/G + 20% Ref + 56%	
		(Poland) ? not in EC DEC	<1.5	1.5	3		H/G + 50%	
		Denmark 1958-1959	1.2	1.3	1.6	Ref + 8%	H/G + 23% Ref + 33%	
		EC DEC		1.3	1.6		H/G + 23%	
Baltic Sea CW B 14 sheltered	6-22	Denmark ?	0.9	1.1	1.6	Ref + 22%	H/G + 45% Ref + 78%	Lagoons
		(Poland) ? not in EC DEC	4 - 5	4 - 5	7.5 - 10		H/G + 88% - + 100%	
Baltic Sea TW B 13 Exposed	6-22	Lithuania ?	3.85	4,8	6,0	Ref + 25%	H/G + 25% Ref + 56%	Transitional water. Sites along the coast of Lithuania and Poland
		Poland ?	< 3.7	3.7	5.6	Ref + 0,1%	H/G + 51%	
		EC DEC		4.2	5.8		H/G + 38%	

Type	Salinity psu	Member State Reference period	Reference conditions	H/G boundary	G/M boundary	H/G boundary	G/M boundary	Geographical delineation of the type
NEA 1/26a	>30	Norway	1.67 (1.7)	2.5	5	Ref + 47%	H/G + 100% Ref + 194%	Open oceanic Eastern Cantabrian Coast, Canary Islands, Atlantic Coast, Western Irish Sea, Scandinavian Coast: North Sea Norway + Norwegian Sea
		Ireland	3.33 (Atlantic Sea 3.4 – Irish Sea 3.7)	5	10	Ref + 35%	H/G + 100% Ref + 200%	
		UK	3.33	5	10	Ref + 50%	H/G + 100% Ref + 200%	
		France	3.33 (1.4 - 4.4)	5	10	Ref + 50%	H/G + 100% Ref + 200%	
		Spain Eastern Cantabrian	2.33 (1 – 3)	3.5	7	Ref + 50%	H/G + 100% Ref + 200%	
		Spain Canary Islands	<1 (0.67)	1	2		H/G + 100%	
		Spain Atlantic coast	3.33 (<5)					
NEA 1/26b	>30	UK	6.7	10	15	Ref + 50%	H/G + 50% Ref +125%	Enclosed seas Eastern Irish Sea, North Sea (UK East Coast, Channel, NL Wadden Coast, Zeeland Coast, Belgium coast)
		Netherlands	6.7	10	15			
		Belgium	6.7	10	15			
		France	6.7	10	15			
NEA 1/26c	>30	Denmark	3.3	5	7.5	Ref + 51%	H/G + 50% Ref + 127%	Enclosed seas partly stratified, German Bight & Jutland
		Germany	3.3	5	7.5		Agreed H/G + 33%	
NEA 1/26d	>30	Denmark	2	3	4	Ref + 50%	H/G + 33% Ref + 100%	Scandinavian Coast: Skagerrak
NEA 1/26e	>30	Spain Iberian	5.33 (4 – 11)	8	12			Areas of upwelling: Western Iberian Coast upwelling, Western Cantabrian Coast
		Portugal Iberian	5.33 (< 4)	8	12	Ref + 100%	H/G + 50% Ref + 200%	
		Spain Cantabrian	4 (2 – 5)	6	9	Ref + 50%	H/G + 50% Ref + 125%	
NEA 3/4	18-30	Germany	3.3				H/G + 50%	Wadden Sea type polyhaline
		Netherlands	9.3	14	21	Ref + 50%	H/G + 50% Ref + 126%	
NEA7	>30	Norway						Deep fjordic and sea loch systems
		UK						
NEA8	18-30	Norway	1	1.5	3	Ref + 50%	H/G + 100% Ref + 200%	Kattegat + small area of Skagerrak (Inner Arc) sheltered, shallow
		Sweden	1	1.5	3			
		Denmark	1	1.5	3			
NEA9	18-30	Sweden	1.7	2.5	5			Skagerrak fjords with a shallow sill at the mouth
		Norway	1.7	2.5	5	Ref + 47%	H/G + 100% Ref + 194%	
NEA10	18-30	Sweden	2	3	6			Skagerrak Outer Arc, exposed, deep
		Norway	2	3	6	Ref + 50%	H/G + 100% Ref + 200%	
NEA11	0-35	No results in 1st IC phase						Transitional waters

Type	Salinity psu	Member State Reference period	Reference conditions	H/G boundary	G/M boundary	H/G boundary	G/M boundary	Geographical delineation of the type
Mediterranean Sea Type I		France 1997-2006	?	?	?			Highly influenced by freshwater inputs
		Italy 2001-2004	1.8	2.4	3.5	Ref + 33%	H/G +46% Ref +94%	
Mediterranean Sea Type II		Spain 1991-2006	1.9	2.3	3.5	Ref + 21%	H/G +52% Ref +84%	Moderately, not directly affected by freshwater inputs (continent influence)
		France 1997-2006	<2	2	4	Ref + 0.1%	H/G +100% Ref +?%	
		Italy 2001-2004	0.77	1	1.24	Ref + 30%	H/G +24% Ref +61%	
		Slovenia 1997-2004	0.99 annual geomean	1.28	1.62	Ref + 29%	H/G +27% Ref +64%	
		Harmonised in EC DEC	1.9	2.4	3.6	Ref + 26%	H/G +50% Ref +89%	
Mediterranean Sea Type III Western		Spain 1991-2006	1.1	1.3	1.8	Ref + 18%	H/G +38% Ref +64%	Not affected by freshwater inputs, continental coast
		France 1997-2006	<1	1	2	Ref + 0.1%	H/G +100% Ref +?%	
		Italy 2001-2004	0.4	0.51	0.64	Ref + 28%	H/G +25% Ref +60%	
		Croatia 2000-2004						
		Harmonised in EC DEC	0.9	1.1	1.8	Ref + 22%	H/G +64% Ref +100%	
Mediterranean Sea Type III Eastern		Greece 2000-2004	0.08 annual average	0.1	0.4	Ref + 25%	H/G +300% Ref +400%	Not influenced by freshwater input
		Cyprus 2005						

Type	Salinity psu	Member State Reference period	Reference conditions	H/G boundary	G/M boundary	H/G boundary	G/M boundary	Geographical delineation of the type
Black Sea CW-BL 1 Bulgaria 1954-1970		winter	1.6	1.9	3.9	Ref + 19%	H/G + 105% Ref + 144%	
		spring	2.3	3.0	6.2	Ref + 30%	H/G + 107% Ref + 170%	
Black Sea Type 2 Bulgaria 1954-1970								
		spring	1.7	2.3	4.6	Ref + 35%	H/G + 100% Ref + 171%	

Table 16: Phytoplankton biomass reference values and H/G and G/M boundary for the spring season in the Baltic Sea GIG

Type	Sali-nity psu	Member State Reference period	Reference condi-tions	H/G boun-dary	G/M boun-dary	H/G boundary	G/M boundary	Geographical delineation of the type
Black Sea CW-BL 1		Romania ?		3000	5700	Ref + ?%	H/G + 90% Ref +?%	
		Bulgaria 1954-1970	2700	3700	5200	Ref + 37%	H/G + 40% Ref + 93%	
		ECOM DEC boundary	< 3000	3000	5000		H/G + 67%	
Black Sea Type 2		Romania?						
		Bulgaria 1954-1970	2500	2900	4100	Ref + 16%	H/G + 41% Ref + 64%	

Conclusions phytoplankton

- In all the GIGs the reference concentration and boundary setting for chlorophyll *a*, taxa cell counts or *Phaeocystis* blooms was defined based on separate national datasets. There was no common approach /benchmarking applied among GIGs or MSs on how to fix reference conditions and no common database was built to verify and validate the high status, the H/G or the G/M boundary in relation to an abiotic characterization at the scale of the common intercalibration type, a clear requirement of the updated IC guidance for the second IC phase.

- The criteria to select the reference sites have not been explained. A checking procedure on the pressure criteria used, in order to verify a consistent selection of reference sites, has not been performed. A map of the location of reference sites has to be included in the following report. Description on the selection criteria of reference sites, where applied to define reference conditions, has to be given, including the illustration of the absence or which level of presence of a pressure has been allowed, not leading to a significant ecological impact.

- Mentioning time periods of measurements/description of dataset used for derivation of the reference conditions or the alternative benchmark is often still lacking, but needed to understand which situation in time and space the assessment of the present situation is compared with. Addition of this information is still necessary.

- An explanation on the levels of difference between Reference, H/G and G/M boundary in relation to the provisions of OSPAR – HELCOM has to be provided. In the NEA GIG the difference between the reference and the G/M treshold is often 2,5 to 4 times more the difference as was accepted in the OSPAR marine convention.

- The Netherlands explained that Cadee and Hegeman (2002) estimated the annual average chlorophyll *a* concentration in the seventies between 2 and 5 µg/l and considered this as an elevated value but below 1.5 times the background concentration. It is not clear how it can be concluded that these values are below 1.5 times the reference concentration and included in a reference condition derivation, since this leads to a circular reasoning.

- Illustration of relationships with physico-chemical parameters like nutrients (1-1 relationship is not possible) should be included, since sensitivity to eutrophication is envisaged. If there are abiotic variables complicating the impact-response relationships, water bodies should be selected for intercalibration that have a similar abiotic characterisation.

- Relationship between hind-casted estimates of nitrogen inputs – TN – chlorophyll *a* should be a wider applied approach in the future, also trying to make the link with reference nutrient concentrations in rivers and the Marine Strategy Framework Directive.
- It is not clear if a normalisation is performed of the chlorophyll *a* values for a specific salinity or if the thresholds are applied at any measurements along the salinity gradient along the coast of a specific type.

3.2. Macroalgae

The NE Atlantic GIG and the Mediterranean Sea GIG have followed two different approaches for assessment of macroalgae and to fix reference conditions. Only the Mediterranean Sea GIG has been able to intercalibrate the full BQE, where every Member State included macroalgae (relative) cover and species composition parameters in their assessment method. In the NEA GIG not every Member State has defined how to evaluate macroalgae cover and/or composition. Results for the Baltic Sea GIG were not finalised in the first phase of intercalibration and no results were available for the Black Sea GIG.

Macroalgae are assessed on rocky substrates in the MED GIG and the NEA GIG, with the exception of Germany that assesses the total intertidal area in the NEA GIG. In the NEA GIG macroalgae are assessed in the intertidal area at the coasts of its open seas (Atlantic coasts, Irish Sea, North Sea), where only Spain assesses also the subtidal area. In the Kattegat and Skagerrak the subtidal area is considered, where for the moment only a depth limit metric is used and intercalibrated, but work is in progress on coverage and composition metrics.

A. North-East Atlantic GIG

The NE Atlantic GIG have intercalibrated several assessment methods for macroalgae that include metrics on diversity (species richness) and composition like percentage of sensitive/opportunistic species (Norway, Ireland, UK and Portugal) and in addition characteristic species (Spain) in the macroalgae composition of type NEA 1/26. The parameter on abundance (=coverage of macroalgae) was reflected partly in the metric „Depth limit of macroalgal species“ agreed by Norway and Sweden for the types NEA 8, 9 and 10 only; in the Macroalgae Blooming tool for opportunistic species in Ireland, UK and Germany; and in the perennial algae tool of Spain and Portugal (only cover of opportunists in PT). The results were obtained for specific different habitats within the types. For NEA 1/26 all countries developped an assessment for the intertidal rocky shore, except Spain that applied assessments to both intertidal (subdivided in exposed and semi-exposed) and subtidal communities with their method. Germany did not have information for specific rocky substrates separately, but developes reference values for the total intertidal, including soft sediments. Definition of reference conditions was based on expert judgment of the maximum possible values for the selected indices/metrics.

B. The Mediterranean Sea GIG

The Mediterranean Sea GIG intercalibrated two different methods, the CARLIT, based on cartography of littoral and upper-sublittoral rocky-shore communites, and the EEI, based on the division of the species in two categories of sensitive/opportunistic species and a spatial assessment of their coverage. The macroalgae assessment with both methods is restricted to hard substrate of 0.2 m to 3.5 m depth. For the intercalibration of these two methods, appropriate descriptions have been provided of the reference macroalgae communities' compositions, including the possible variation related to some environmental variables, based on selected reference sites and a common metric. By means of the common metric BENTHOS, common views for reference conditions within the Mediterranean Sea have been described. However, a checking procedure on the pressure criteria used, in order to verify a consistent selection of reference sites, has not been documented.

The common views for reference conditions can be summarized as follows:

1. Macroalgal communities of high diversity should be dominated quantitatively by brown algae mainly of the order Fucales in high irradiance sites and red algal Corallinales in vertical cliffs.

2. Dense well-developed macroalgal communities thriving in the upper infralittoral zone with most characteristic species belonging to the genera Cystoseira, Sargassum, Lithophyllum, Peyssonnelia, Corallina and Padina. Other common species belong to the genera Halopteris, Stypocaulon, Dictyota, Dictyopteris, Laurencia, Cladophora and Jania.

3. In shadow zones (exposed steep vertical cliffs) Lithophyllum byssoides develops, forming important organogenic structures (trottoir). In marine caves with scarce light conditions a sciaphilic vegetation of red and green algae is dominant.

4. Spatio-temporal variability of the community's composition and abundance affected by hard substrata availability, intense and frequency of natural disturbances, e.g. hydrodynamism, grazing, by seasonal cycle of light period and intense, and by limiting factors like nutrients.

Table 17 gives an overview of the description of derivation of the reference conditions for macroalgae and the H/G and G/M boundaries. A central key issue is the percentages and/or coverage of opportunistic species that can be accepted in the macroalgae communities.

Conclusions macroalgae

- Mentioning time periods of measurements/description of dataset used for derivation of the reference conditions or the alternative benchmark is often still lacking (like for the NEA GIG type 1/26 and the MED GIG), but needed to understand which situation in time and space the assessment of the present situation is compared with. Addition of this information is still necessary.
- Description on the selection criteria of reference sites, where applied to define reference conditions, has to be extended, including the illustration of the absence or which level of presence of a pressure has been allowed, not leading to a significant ecological impact. A checking procedure on the pressure criteria used, in order to verify a consistent selection of reference sites, has not been performed.
- The community descriptions for macroalgae in the MED-GIG are a nice example of a common understanding of the ecological meaning of a reference condition.
- The method description of the CARLIT still needs amendment.
- There was no distinction of different types in the NEA 1/26, although there seem to be significant differences in reference conditions. Portugal applies a higher proportion of red algae species (>70%) as a reference condition than the other countries (40-55%), but requires a lower species richness (>25) in relation to the others (>33-35) as a reference for diversity. Would it not be more appropriate to distinguish a different type in the southern part of the NEA, consistent with the phytoplankton types in this area?
- The information in the technical report needs to be extended with an illustration of the level of pressure corresponding to the good and the moderate status expert judgment.

Table 17: Macroalgae reference conditions descriptions and H/G and G/M boundary derivation for each type in each GIG

<table>
<tr><th>Memb.State Reference period</th><th colspan="7">Reference conditions NEA 1/26
Salinity psu >30</th><th>H/G boundary</th><th colspan="5">G/M boundary</th></tr>
<tr><td>Norway Intertidal rocky shore</td><td></td><td rowspan="3"></td><td colspan="5" rowspan="2">1.Reduced Species list (RSL) = Perennial intertidal algae
Diverse community of red, green and brown seaweeds with high levels of species richness. Cover variable depending on local physical conditions but species richness relatively constant temporally. Red species present as richest group along with a high proportion of long-lived spp. Opportunist and green species should constitute a lower proportion of the algal present. The reduced species list has a maximum of 70 species that should be present for UK& IE shores, and 68 for Norway.
ESG: ratio of perennial to annual or ephemeral forms
Reference values are expressed as values higher or lower then the H/G boundary values (see below)</td><td rowspan="2">H/G boundaries were established using historic data, historic reports and publications and expert judgement (see column to the left).</td><td colspan="5">1.Reduced Species list (RSL) = Perennial intertidal algae
No direct relationship with a pressure gradient could be established to set the G/M boundary and step 8 of the boundary setting procedure was invoked.
Norway
S <20 | Green species > 30% | Red species <30% | ESG Ratio < 0.6 | %Opportunists >25
Northern Ireland (UK)
S <20 | Green species > 30% | Red species <35% | ESG Ratio < 0.6 | %Opportunists >25</td></tr>
<tr><td>Ireland RSL Intertidal rocky shore</td><td rowspan="2">2.Macroalgae Blooming (MAB) = Opportunistic Macroalgae blooms of anthropogenic origin should be absent or if present should cover less than 5% of the available intertidal habitat and total biomass/m2 < 100 (both for AIH as affected area only.
Total area coverage of opportunist macroalgae < 100 hectares with no effect on quality status.
Entrained algae should only be present in less than 5% of quadrats.</td><td colspan="5"></td></tr>
<tr><td>UK RSL Intertidal rocky shore</td><td colspan="5">Norway
S >33 | Green species < 20% | Red species >40% | ESG Ratio > 0.8 | %Opportunists <15
Northern Ireland (UK)
S >34 | Green species < 20% | Red species >45% | ESG Ratio > 0.8 | %Opportunists <15
UK (England, Wales, Scotland) and Republic of Ireland
S >35 | Green species < 15%, Scotland: <12% | Red species >55% | ESG Ratio > 1 | %Opportunists <10</td><td></td><td colspan="5">UK (England, Wales, Scotland) and Republic of Ireland
S <25 | Green species > 20% | Red species <45% | ESG Ratio < 0.8 | %Opportunists >15
2.Macroalgae Blooming (MAB) = Opportunistic Macroalgae (IE + UK)
% cover of available intertidal habitat (AIH) >15 = moderate | Total Area Coverage (ha) >500 = moderate | Biomass of AIH >500 = moderate | Biomass of affected area > 500/m2 = moderate | Presence entrained algae (% quadrats) in > 20% = moderate</td></tr>
<tr><td>Denmark</td><td colspan="7"></td><td></td><td colspan="5"></td></tr>
<tr><td>Germany Soft sediments Total intertidal</td><td colspan="7">Macroalgae Blooming (MAB) = Opportunistic Macroalgae blooms of anthropogenic origin should be absent or if present should cover less than 1% of the total intertidal of the waterbody. Coverages with densities ≥1 % rarely ever exceed 15 % of the intertidal; normalised to 100 % density they stay below 5 % in most of the cases. Generally directed at intertidal sedimentary shores in both transitional and coastal waters.</td><td></td><td colspan="5">Opportunistic macroalgal blooms of anthropogenic origin should cover less than 1% of the total intertidal of the waterbody</td></tr>
</table>

MembState Reference period	Reference conditions NEA 1/26 Salinity psu >30							H/G boundary	G/M boundary						
Belgium	No macroalgae naturally present in this area														
France															
Spain Intertidal rocky shore	Quality of Rocky Bottoms (CFR) = Perennial intertidal algae								Quality of Rocky Bottoms (CFR) = Perennial intertidal algae						
	Characteristic Macroalgae Cover		Populations Richness	Relative Coverage of Opportunists to total vegetated surface					Characteristic Macroalgae Cover	Populations Richness	Relative Coverage of Opportunists to total vegetated surface				
	Semi-exposed shore	> 70%	= 6	= 9% (for semi-exposed and exposed)					< 40% = moderate	<4 = moderate	> 9% = moderate (for semi-exposed+exposed)				
	Exposed shore	> 50%	= 4						< 30% = moderate	<3 = moderate					
Spain Subtidal rocky shore	Characteristic Macroalgae Cover		Populations Richness	Relative Coverage of Opportunists to total vegetated surface					Characteristic Macroalgae Cover	Populations Richness	Relative Coverage of Opportunists				
	5 - 15 m	> 70%	= 6 (for both depth ranges)	= 4% (for both depth ranges)					< 40% = moderate	<4 = moderate (for both depth ranges)	> 9% = moderate (for both depth ranges)				
	15 - 25 m	> 50%							< 30% = moderate						
Portugal Intertidal rocky shore	Portugese Marine Macroalgae Assessment Tool (P-MarMAT) = Perennial intertidal algae								P-MarMAT values						
	S >25	Green species < 10%	Red species >70%	ESG Ratio > 2.5	Proportion of opportunists < 10%	Shore description < 7	Coverage opportunists < 10%		S <17 = mode-rate	Green species >20%	Red species <55%	ESG Ratio <2	Proportion Opportu-nists >20%	shore >11	Coverage opportunists >20%

Type	Sali-nity psu	MembState Referenc e period	Reference conditions	H/G boundary	G/M boundary	Geographical delineation of the type
NEA 3/4	18-30	Germany				Wadden Sea type polyhaline
		Netherlands				
NEA7	>30	Norway				Deep fjordic and sea loch systems
		UK				
NEA8 Subtidal rocky shore	18-30	Norway 1947-1952, 1953, 1988-1989, 1990-2005	**Subtidal algae** 'depth limits of selected macroalgal species' Reference depth limits of nine selected macroalgal species have been defined with the help of Norwegian historical data. All depth limits (ranging from 10 to 25 m) are set by expert judgement (including recent Norwegian and Swedish surveys) for each of the nine selected macroalgal species since the few available historical data from dredging do not allow a good comparison with the recent results from diving.	The High/Good boundaries (ranging from 8 to 18 m) vary between 18 and 37% of the estimated reference value.	The dose-response relationships developed so far seem to show no discontinuities/thresholds. Good/Mode-rate boundaries (ranging from 5 to 12 m) vary between 44 and 58% of the estimated reference values.	Kattegat + small area of Skagerrak (Inner Arc) sheltered, shallow
		Sweden 1994-1998				
		Denmark				

Type	Sali-nity psu	Member State Reference period	Reference conditions	H/G boundary	G/M boundary	Geographical delineation of the type
NEA 9 Subtidal rocky shore	18-30	Norway 1990, 1998-1999	Subtidal algae 'depth limits of selected macroalgal species' Reference depth limits of nine selected macroalgal species has been defined. All depth limits (ranging from 12 to 17 m) are set by expert judgement (including 2 recent Norwegian surveys) for each of the nine selected macroalgal species since the few available Swedish historical data do not allow any good statistical treatment.	The H/G boundary for depth limits of nine selected macroalgal species (ranging from 8 to 13 m) has been defined as a 17-33% deviation from the estimated reference levels.	The dose-response relationships developed so far seem to show no discontinuities/thresholds. Boundaries (ranging from 6 to 9 m) set by expert judgement and represent a 42 to 50% deviation from reference levels.	Skagerrak fjords with a shallow sill
		Sweden 1941 Gullmar Fjorden				
NEA10 Subtidal rocky shore		Norway 1947-1952, 1989, 1990-2005	Subtidal algae 'depth limits of selected macroalgal species' Reference depth limits of nine selected macroalgal species (ranging from 14 to 30 m) has been defined based on Norwegian historical data and expert judgement of Norwegian and Swedish recent data.	The H/G boundaries (ranging from 10 to 22 m) vary between 24 and 29% of estimated reference level.	G/M boundaries (ranging from 8 to 18 m) between 40 and 50% of reference levels	Skagerrak Outer Arc, exposed, deep
		Sweden 1994-1998				
NEA11	0-35		Macroalgae Blooming (MAB) = Opportunistic Macroalgae for Germany and UK and IE Same reference as in coastal water types			Transitional waters

Type	Salinity psu	Member State Reference period	Reference conditions	H/G boundary	G/M boundary
Mediterranean Coastal and transitional waters No types have been distinguished for macroalgae		Greece EEI	10>EEI>8 Reference sites have been identified according to the low pressures and impacts they receive in accordance with Annex V of WFD. In all methods (EEI, BENTHOS, CARLIT) the reference sites are real sites (existing) and this allows the application of the tested methodologies in these places. The rocky upper infralittoral zone reference conditions are based on Greek coastal waters 62 samples from 26 putatively pristine Aegean sites dominated by Cystoseira cf. crinita community as part of the Hellenic "NATURA 2000" data-base. Community description available.	EEI = 8 Boundaries are set according to biotic index and/or combined with the results of or multivariate analysis. No statistical analysis exclusively to set boundaries. No discontinuities. Continuum of possibilities with gradual disappearance/ appearance of different indicator species. The late successionals taxa, especially species of Cystoseira genus, represented by the ecological group ESG I account for more than 60% of the mean macroalgae abundance-coverage and the early successionals taxa represented by the groups ESG II account for 0-30% of the macroalgae coverage.	EEI = 6 At the good status as is indicated by the EEI, the ESG I group may range from 30 to 60% while the ESG II from 0 to 30% of the macroalgae coverage, or the combination may thus be that ESG I accounts for over 60% and ESG II between 30 and 60% of the total macroalgae coverage. At the moderate status as is indicated by the EEI, the two groups may equally share the macroalgae coverage accounting for equally low, moderate or high percentages.
		Cyprus & Slovenia EEI	Reference community description confirmed in pristine sites of Cyprus and Slovenian coasts		
		Spain (2001 May to June) & Italy & France CARLIT	Three reference zones outside Catalonia: Façade maritime du Parc Naturel Régional de Corse (France), Parc Natural de Ses Salines (Balearic Islands, Spain) and Reserva Marina del Nord de Menorca (Balearic Islands, Spain) + historical data in the Catalan coast before 1980's and in the adjacent Albères coast. Community description available		

3.3. Angiosperms

For the BQE angiosperms only results for seagrasses were worked out in the 1st IC phase and not yet for salt marshes. The Baltic Sea GIG and the NE Atlantic GIG have followed two different approaches for assessment of angiosperms and to fix reference conditions. The Baltic Sea GIG has only intercalibrated angiosperm abundance partly by 'depth limit of eelgrass', whereas the NEA GIG has considered seagrass bed extent, bed density and taxonomic composition (1 or 2 species). Results for the Mediterranean Sea GIG were not finalised in the first phase of intercalibration and no results were available for the Black Sea GIG. The Mediterranean Sea GIG has submitted results for seagrasses, which have been reviewed and are being improved to include in the next European Commission Decision.

A. Baltic Sea GIG

The selected metric 'depth limit of eelgrass' is only intercalibrated between Denmark and Germany in one type. It describes the depth extension of eelgrass (*Zostera marina*), which is the dominant seagrass in Scandinavian coastal waters. The metric represents the eelgrass main distribution rather than the maximum depth limit and is affected by nutrient concentration and water transparency (Nielsen *et al.* 2002, Krause-Jensen *et al.* 2005). Angiosperms are not intercalibrated in the other types, because the vegetation is scarce and distribution scattered.

Two approaches have been used for deriving reference conditions and setting the boundaries of the classification for Denmark: 1) percent deviation from reference conditions based on historical data (4 scenarios) and 2) modeling the relationship between TN and the depth limit. Denmark used the maximum depth of 5 % eelgrass (*Zostera marina*) cover to define the depth limit. The reference depth limits are defined as the values >90% of the historical values. In order to define reference levels of benthic vegetation metrics by modeling, dose-response relationships between physico-chemical variables (nutrient concentration, transparency) and vegetation metrics have been estimated. The reference levels of nutrient concentration and/or transparency are entered in the model and corresponding levels of the vegetation metrics are calculated. In this approach the reference levels of nutrient concentration and transparency have been identified within Danish waters, where vegetation data are available i.e. Øresund. The model by Nielsen *et al.* (2002) was used to hind-cast reference depth limits based on reference TN levels. If reference TN levels are defined as e.g. 14 μM along open Danish coasts, then the corresponding reference eelgrass (*Zostera marina*) depth limits are 7.7 m. If the high-good boundary is defined as 25% of reference levels, then the high/good boundary for eelgrass is 5.8 m. The modeled values defined a lower depth than the data based on historical data (7.5 – 10.4).

Germany used historical records of *Zostera marina* depth limit to define the reference and light modeling to define depth limits for boundary setting. In Germany from historical records eelgrass "stands" were defined as >50 shoots/m², which is the minimum end of a range of 50->2500 shoots/m² for the Baltic Sea (Schories *et al.* 2006). The historical depth limit of *Zostera marina* was assessed as 10 m for stands.

B. North-East Atlantic GIG

UK, Ireland and the Netherlands have intercalibrated the seagrass abundance parameter (areal bed extent and density) and the taxonomic composition parameter. Germany could only consider bed extent so far. To define reference conditions the assumption is made that these occur in unimpacted areas with unpolluted water quality and no hydromorphological alterations to the shore or seabed. However, Dutch waterbodies are embanked and may be classed as heavily modified. Although the waterbodies are managed and protected by engineering works, habitats such as seagrass beds have established naturally within them. Potential Reference Conditions (P-REF) and Potential Good Ecological Status (P-GES) are the highest two classes heavily modified waterbodies can attain, and scientists in the Netherlands have set values for these by focusing on the current situation in the waterbodies concerned (de Jong, 2004).

Table 18: Macroalgae reference conditions descriptions and H/G and G/M boundary derivation for each type in each GIG

<table>
<tr><th>Type</th><th>Sali-
nity
psu</th><th>Member State
Reference
period</th><th>Reference conditions</th><th>H/G boundary</th><th>G/M boundary</th></tr>
<tr><td rowspan="6">Western
Baltic Sea
CW
B12 b
Sheltered
Sites at the
Southern
Swedish
coast and
the South
western
Baltic Sea
open coast
along
Denmark
and
Germany</td><td rowspan="6">8 -
22</td><td>Denmark
(1908) +
Germany

‘depth limit
of eelgrass
Zostera
marina’</td><td>Based on historical data: values above 90% of the historical maximum, (calculated from the most extensive historical data set of eel-grass depth distribution reported by Ostenfeld (1908) around year 1900) represent a high ecological status/reference situation.
These estimations were compared with modelling exercise: Reference levels of total-nitrogen concentrations (TN) and water trans-parency were defined, based on historical data and modeling: TN concentration DK: 16.6 µM (Fakse & Hjelm Bay), 14 µM (open Danish coasts); DE coast = 9 µM and the corresponding Zostera marina reference depth limit (m) modeled based on TN-concentration DK: 8.3 (Fakse & Hjelm Bay), 7.7 (open Danish coasts); DE coast = 10</td><td>Based on historical data: H/G boundary is defined to represent 90% of the historical maximum: Zostera marina depth limit (m) = 7.5 – 10.4
These estimations were compared with modelling exercise: H/G TN concentration DK = 17.6-17.7 (Fakse & Hjelm Bay) and DE coast = 12 corresponding with Zostera marina H/G boundary depth limit (m) modeled based on TN-concentration DK = 7.9-8.0 (Fakse & Hjelm Bay) and DE coast = 8.0</td><td>Based on an expert judgment the 15% and the 20% deviations from the H/G boundary (estimated from historical data) are the best scenarios to comply with the normative definition of good ecological state (Krause-Jensen 2006). Populations of eelgrass growing at depths of 4.4 – 4.0 m in Limfjorden and a population of eelgrass in Kattegat growing at 8.1-7.6 m depth are examples of good growing conditions. Accordingly, it was concluded that 25 % deviation from the reference condition is not congruent with “slight changes of disturbance”. The model scenario (to compare with the estimation from historical data) enters the total nitrogen concentration, believed to represent the boundary between good and moderate into the empirical relationship between nutrient concentration and eelgrass depth limit (Nielsen et al. 2002). The TN concentrations were estimated describing the reference condition and the G/M boundary for a number of Danish and German coastal waters.
For the modelling exercise, TN concentration DK = 19.6 - 19.8 (Fakse & Hjelm Bay) and DE coast = 16 corresponding with Zostera marina depth limit (m) modeled based on TN-concentration DK = 7.3 (Fakse & Hjelm Bay) and DE coast = 7.0
Finally, the good-moderate boundary was defined as a 25-30% deviation from H/G boundary: DK: Depth limit (m): 4,7 - 5 (Fakse Bay) and 6,6 - 7,1(Hjelm Bay)</td></tr>
<tr><td rowspan="2">Germany
‘depth limit
of eelgrass
Zostera
marina’</td><td rowspan="2">Reference = 8 – 10 m</td><td colspan="2">Through light-modelling and under the assumption that 10% of the incident light is necessary to maintain Zostera stands, Schories et al. (2006) calculated depth limit for the WFD classes. They arbitrarily set certain percentages of light reduction (compared to pristine conditions) by enhanced attenuation and calculated border depth limits for Zostera stands</td></tr>
<tr><td>H/G boundary = 8</td><td>The eelgrass depth limit for good status is defined as 7.0 - 8.0 m and moderate status as 4.5 – 7.0 m,
so G/M boundary = 7</td></tr>
<tr><td>Harmonised
values DK-DE</td><td>9.4 (8 – 10.4)</td><td>8,5 (8 – 9.4)</td><td>7 (6.6 – 7.1)</td></tr>
<tr><td>Sweden ?</td><td></td><td></td><td></td></tr>
<tr><td>Poland ?</td><td></td><td></td><td></td></tr>
</table>

Type	Member State Reference period	Habitat if distin-guished	Reference conditions NEA 1/26 Salinity psu >30	H/G boundary	G/M boundary
NEA 1/26	Norway				
NEA 1/26	Denmark				
NEA 1/26	UK and Ireland	Intertidal seagrass *Zostera noltii* and *Z. angustifolia* & *Ruppia* sp.	The assumption is made that these occur in unimpacted areas with unpolluted water quality and no hydromorphological alterations to the shore or seabed. **Seagrass abundance:** Change in bed extent: No loss in seagrass bed extent – at maxi-mum potential and in equilibrium (within natural variability) Change in density: Bed density at or above ~highest previously recorded **Change in taxonomic composition:** No loss of species (Most seagrass beds comprise only 1 or 2 species) The trend for an individual bed and the loss or gain, as compared with a maximum recorded density, can be used to identify whether the seagrass bed is in a state of degradation or recovery. the Member States agree that the ideal period over which to consider the trend in abundance is ~6 year, designed to coincide with the WFD reporting cycles.	No distinction can be made between High/Ref and Good for the UK metric because, for example, there are sublittoral beds of *Z. marina* that are naturally mono-specific and are at High status.	**Seagrass abundance:** Change in bed extent: < 30% loss/decrease in bed extent from highest recorded, between High/Ref and Good Change in density: Density <30 % loss/diffe-rence between High/Ref and Good **Change in taxonomic composition:** ¼ to ⅓ loss of species
NEA 1/26	Germany	Intertidal seagrass beds; *Zostera noltii* & intertidal *Z. marina*	Change in bed extent Reference conditions explanations still to be added from Kolbe, 2007		
NEA 1/26	Netherlands		No seagrasses in this type in the reference situation ?		
NEA 1/26	Belgium		No seagrasses in this type in the reference situation		
NEA 1/26	France				

Type	Sali-nity psu	Member State Reference period	Habitat if distin-guished	Reference conditions			H/G boundary	G/M boundary		
NEA 3/4 Wad-den Sea type poly-haline	18-30	Germany	*Zostera noltii* & intertidal *Z. marina*	Change in bed extent Reference conditions explanations still to be added from Kolbe, 2007						
		Netherlands	Intertidal seagrass beds; *Zostera noltii* & intertidal *Z. marina*	Seagrass abundance		Change in taxonomic composition: Number of species present: 2		Seagrass abundance		Change in taxonomic composition: Number of species present: 1 species
				Change in bed extent: Wadden Sea total: 250 hectares Oosterschelde: 1000 hectares	Change in density: >=30% of *Zostera marina*, >=60% *Zostera noltii*) based on modelling and expert judgment			Change in bed extent: The average difference between REF and G/M boundary is ~30%. The mean difference between the NL's Moderate/ Poor and REF for all water-bodies is ~50%. Wadden Sea total good: 150 hectares Oosterschelde good: 750 hectares	Change in density: difference between NL's REF and G/M boundary for both species is ~30%, *Zostera marina* >=20%, *Zostera noltii* >=40% = good	
				Sub-metric to support Metrics 2 & 3: Trends in seagrass abundance over a period of 5/6 years (seagrass acreage (bed extent) and coverage (% density)): positive, neutral (reference) or negative (moderate)						
NEA 7	>30	Norway								
		UK		The same as in NEA1/26 ?						
NEA 11	0-35	Ireland & UK		The same as in NEA1/26						
		Germany	*Zostera noltii* & intertidal *Z. marina*	Change in bed extent Reference conditions explanations still to be added from Kolbe, 2007						
		Netherlands	Intertidal seagrass beds; *Zostera noltii* & intertidal *Z. marina*	Seagrass abundance		Change in taxonomic composition: Number of species present: 2		Seagrass abundance		Change in taxonomic composition: Number of species present: 1 species
				Change in bed extent: Ems-Dollard: 100 hectares Westerschelde: 3 hectares	Change in density: >=30% of *Zostera marina*, >=60% *Zostera noltii*			Change in bed extent: The average difference between REF and GES is ~30%, The mean difference between the NL's Moderate and REF for all four waterbodies is ~50%, Ems-Dollard: 50 hectares Westerschelde: 2 hectares	Change in density: difference between NL's REF and GES for both species is ~30%, *Zostera marina* >=20% = good, *Zostera noltii* >=40% = good	
				Sub-metric to support Metrics 2 & 3 Trends in seagrass abundance (seagrass acreage (bed extent) and coverage (% density)): positive, neutral or negative						
		Spain								
		Portugal								

Conclusions angiosperms

- Mentioning time periods of measurements/description of dataset used for derivation of the reference conditions or the alternative benchmark is often still lacking (like for the NEA GIG), but needed to understand which situation in time and space the assessment of the present situation is compared with. Addition of this information is still necessary.
- Baltic Sea GIG:
    - Reference depth for *Zostera marina* is 8 – 10.4m, while reference Secchi depth was set at 6 m (coastal) – 10 m (offshore) to derive the phytoplankton reference. Consistency comparisons can be done between the reference abiotic characterisation for phytoplankton in relation to the defined reference conditions for other aquatic flora.
    - 25% deviation from the H/G boundary was considered not congruent with the normative definitions by Krause-Jensen 2006, but finally the G/M boundary was defined as a 25-30% deviation from the H/G boundary, which therefore seems to be an inconsistent result.
- The relationship with physico-chemical parameters like nutrients (like TN) is tested for the Baltic Sea GIG, not for the NEA GIG where general degradation of seagrass fields is described, but no description is provided what level of pressure corresponds with the good or moderate status.

3.4. Benthic invertebrates

All four coastal water GIGs have been able to produce intercalibration results for the benthic invertebrate quality element. Baltic Sea GIG and NEA GIG included parameters on (relative) abundance, diversity, species richness and composition (community or sensitive taxa percentages) in their methods. In the Mediterranean Sea GIG only Slovenia included parameters on diversity and species richness, whereas Greece, Cyprus and Spain didn't.

A. Baltic Sea GIG

The finalized results produced for the Baltic Sea GIG (types types B0, B2, B3) involved two methods of which the Finnish method was an extension of the Swedish method and so quite similar.

B. North-East Atlantic GIG

In the NEA GIG various methods including different concepts of quality assessment have been used. Most Member States followed the concept to classify the species in different classes of sensitivity and assessing their relative abundance, in addition to diversity and species richness. The Netherlands and Belgium assess changes in the whole community composition, species richness and absolute abundance. Also the way reference conditions are derived differs in a large extent between all Member States as illustrated in table 19.

C. Mediterranean Sea GIG

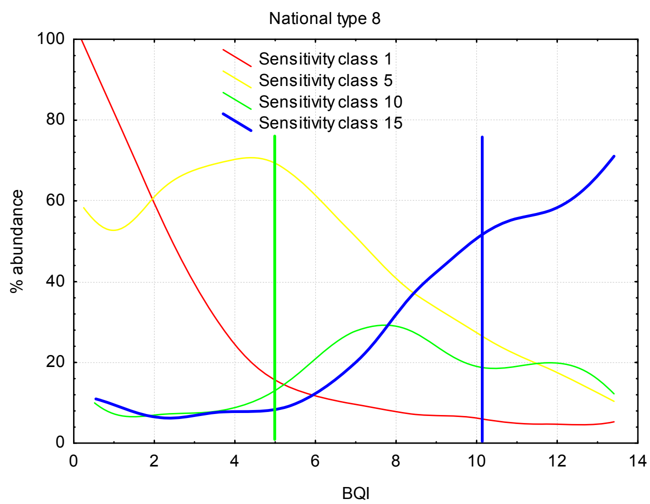
In the Mediterranean Sea GIG all the methods involved a similar approach, classifying the species in different categories of sensitive/tolerant species. However, the number of sensitivity categories differs between the methods (BENTIX has 2, MEDOCC has 4 and AMBI has 5) and between MEDOCC and AMBI also the designation of the species to a specific category differs. Therefore a comparison of the sensitivity of species and their classification in categories should be carefully considered when comparing reference conditions set for the contribution of the different classes to the community. The MED GIG defined a common understanding of diversity and species richness for comparison puposes in the intercalibration but, except from Slovenia, did not include these parameters in their assessment method.

D. Black Sea GIG

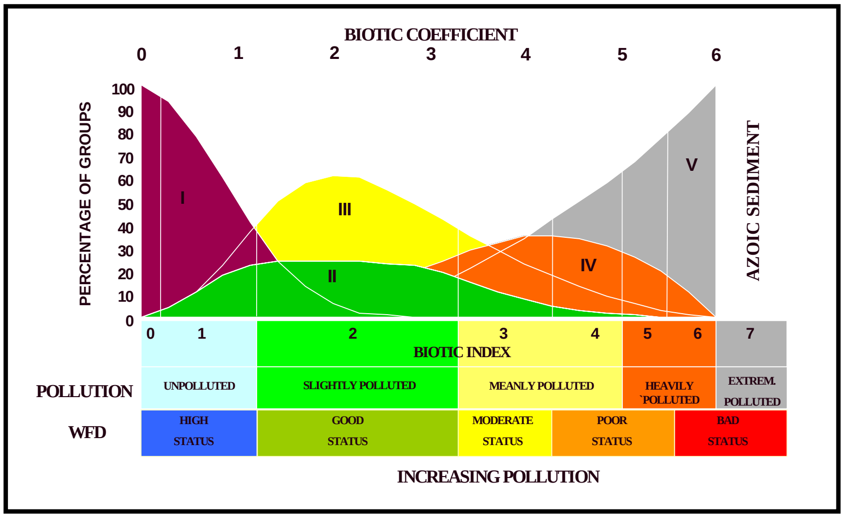
In the Black Sea GIG both countries have used the same methods for sampling, laboratorial analyses, reference conditions and EQRs. The reference conditions were derived using literature, historical data and expert judgment.

It is important to note that the methods intercalibrated are habitat specific. All methods are for use in soft sediment habitats, and more specifically for subtidal shallow muddy sand, except the BEQI that includes an evaluation of different habitats. Methods for other habitats are largely still under development and cannot be compared at this stage, as was mentioned in the technical report.

Concerning the reference conditions derived for the procentual contributions of different categories of sensitive and tolerant species, it is important to compare the conceptual framework used in the different methods to evaluate sensitive species and how the percentages relate to each-other at the class boundaries. The figures below show the boundaries used in the different national methods in the Baltic Sea (BQI method used by Sweden and included in the method of Finland), the NEA GIG (AMBI method) and the Mediterranean GIG (Bentix by Greece and Cyprus and MEDOCC by Spain).

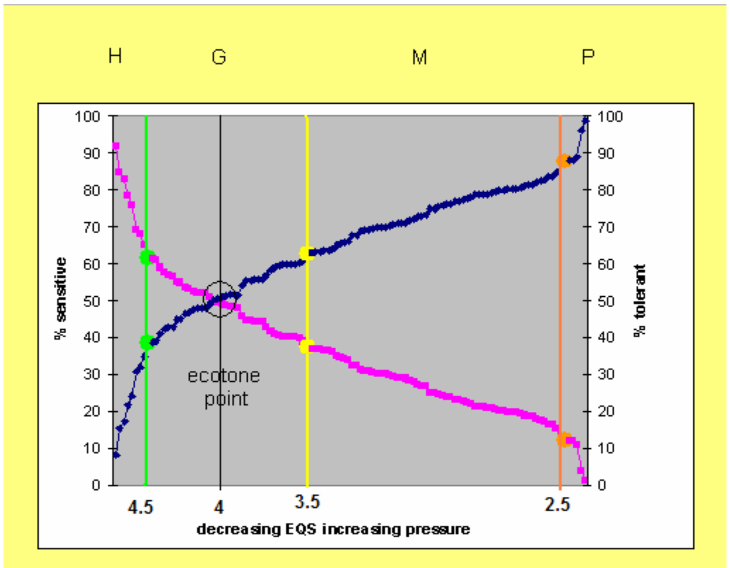


BQI (Baltic Sea)
Figure 48: Distribution of the relative abundance of the four macroinvertebrate sensitivity groups as a function of the BQI values. Group 15 is the most sensitive and group 1 the most tolerant group of species. Good-moderate boundary is indicated with a green and the High-good boundary with a blue vertical line.



AMBI (NEA GIG)

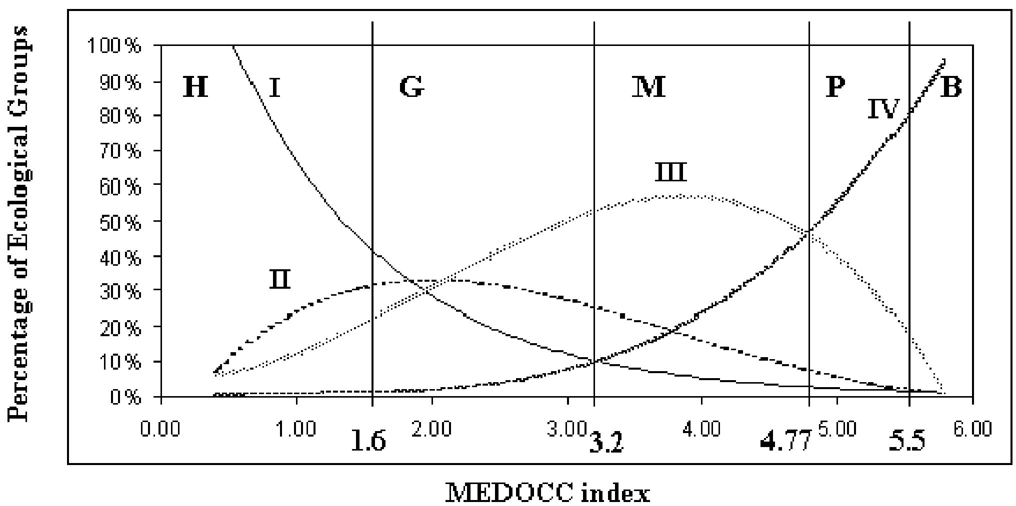
Figure 49: The AMBI biotic coefficient, relating the ecological groups present in a sample to an assessment of the benthic invertebrate community (Borja *et al*, 2003).
EG I – sensitive taxa,
EG II – indifferent taxa,
EG III – tolerant taxa,
EG IV – opportunistic taxa,
EG V – indicator taxa, taxa indicative on pollution



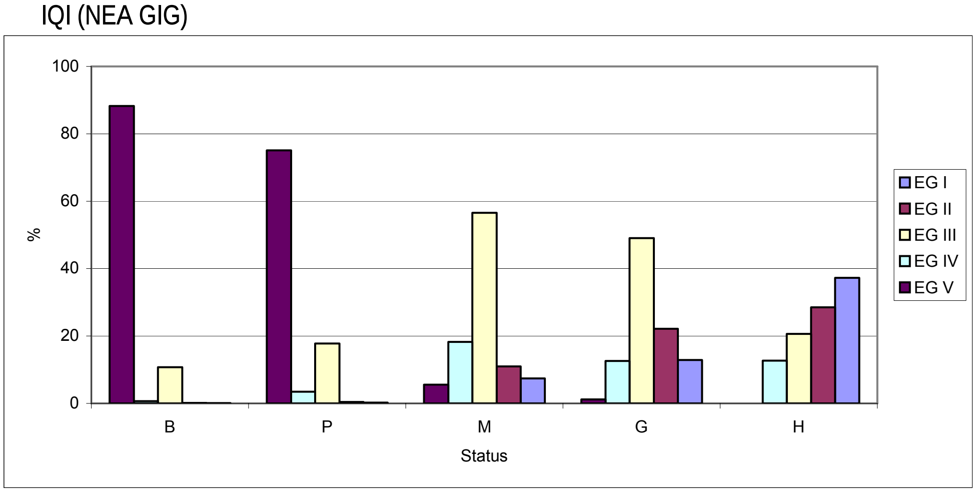
BENTIX (Mediterannean Sea)
Figure 50: Bentix values (from Greece and Cyprus IC data) plotted against the percentages of the two ecological groups. P: poor class, M: moderate class, G: good class, H: high class. GS includes the sensitive and indifferent taxa, GT includes all tolerant taxa: the tolerant and second order opportunistic and the first order opportunistic taxa.

MEDOCC (Mediterranean Sea)

Figure 51: MEDOCC values relating to the percentages of the ecological groups for the Catalonia and Balearic Islands dataset. Vertical lines show boundaries of the different ecological status. I: Sensitive species; II: indifferent species; III: tolerant species; IV: opportunistic species. GT includes all tolerant taxa: the tolerant and second order opportunistic and the first order opportunistic taxa.



IQI (NEA GIG)



IQI (NEA GIG)
Figure 52: AMBI group proportions within each status based on the Garroch Head data using the following boundaries: High/Good =0.8; Good/Moderate = 0.65; Moderate/Poor = 0.43; Poor/Bad = 0.2.

Table 19: Macrobentic invertebrate fauna reference conditions descriptions and H/G and G/M boundary derivation for each type in each GIG

<table>
<tr><th>Type(s)</th><th>Sali-nity psu</th><th>Member State / method</th><th>Habitat if distin-guished</th><th>Reference conditions period / surface</th><th>Reference conditions description for method parameters / derivation</th><th>H/G boundary derivation</th><th>G/M boundary derivation</th><th>Geographical delineation of the type</th></tr>
<tr><td rowspan="4">Baltic Sea CW B0 sheltered, CW B2 sheltered, CW B3 a sheltered and CW B3 b exposed</td><td rowspan="4">B0: 0.5-3, B2-B3: 3-6</td><td rowspan="2">Finland BBI</td><td rowspan="2">Separate assess-ment 0 –10 m and +10 m</td><td rowspan="2">1990-present, one area since 1964, few data from 1920 0.1 m²</td><td>The median of the 10% highest BBI-values recorded. when S < 9, individuals < 20, method result cannot be high or good.
No description available what high status means concerning faunal composition.</td><td>Set at the 10%-percentile of the reference EQR-values</td><td>Values below the H/G boundary are divided into 5 equal classes: Good = 2/5, Moderate = 1/5, Pollution sensi-tive and very sensitive species should be dominant above G/M boundary</td><td rowspan="4">Sites in Botnian Bay (Northern Quark), + Bothnian Sea, + area extending from the southern Bothnian Sea to the Archipelago Sea and the western Gulf of Finland</td></tr>
<tr><td><table><tr><td colspan="3">BQI parameters</td><td>H´= ?</td><td>+ S</td><td>species</td></tr><tr><td>parameter Number of individuals</td><td>S = ?</td><td>% sen-sitive specie s = ?</td><td>(log2-base)</td><td></td><td>abun-dance</td></tr></table></td><td><table><tr><td>Type</td><td>0-10m</td><td>+10m</td></tr><tr><td>B0 outer</td><td>0.71</td><td>0.64</td></tr><tr><td>B2 outer</td><td>0.67</td><td>0.60</td></tr><tr><td>B2 inner</td><td>0.52</td><td>0.71</td></tr><tr><td>B3 outer</td><td>0.74</td><td>0.62</td></tr><tr><td>B3 middle</td><td>0.70</td><td>0.53</td></tr></table></td><td><table><tr><td>Type</td><td>0-10m</td><td>+10m</td></tr><tr><td>B0 outer</td><td>0.43</td><td>0.38</td></tr><tr><td>B2 outer</td><td>0.40</td><td>0.36</td></tr><tr><td>B2 inner</td><td>0.31</td><td>0.42</td></tr><tr><td>B3 outer</td><td>0.44</td><td>0.37</td></tr><tr><td>B3 middle</td><td>0.42</td><td>0.32</td></tr></table></td></tr>
<tr><td rowspan="2">Sweden BQI</td><td rowspan="2">Deeper than 5 m</td><td rowspan="2">1981-2006 0.1 m²</td><td>Best available data from areas without local discharges for each type were assumed represent-ting at least good ecological status: the upper third of the span exceeding the good–moderate boundary was reserved for the status of high</td><td rowspan="2">Two-thirds (2/3) of the span exceeding the good–moderate boundary was assigned a status of good</td><td rowspan="2">Expert judgment on G/M boundary from available data: significant differences from comparison material</td></tr>
<tr><td><table><tr><td>Number of individuals parameter</td><td>S = ?</td><td>Percentage of sensitive species = ?</td></tr></table></td></tr>
</table>

NEA 1/26							
Member State / method	Habitat if distinguished	Reference conditions period / surface	Reference conditions description for method parameters / derivation			H/G boundary derivation	G/M boundary derivation
Ireland IQI		0.1 m2	Simpsons = 0.97	S = 68	1-AMBI/7 = 0.96	Initially equidistant boundaries were adjusted, plotting the ecological group proportions within the first and third quartiles of each status and adjusting the boundaries until agreement with the Normative Definitions was maximised. The validity of the adjusted boundaries was then assessed by analysis of the composition of the taxa in each status class. SIMPER analysis (PRIMER©) was carried out, assessing the top 90% of contributing taxa (family level taxonomic discrimination). Changes in composition between the status classes were evaluated to determine whether the adjusted ranges mirrored the Normative Definitions in terms of contributing taxa.	
UK IQI		0.1 m2	Simpsons = 0.97	S = 68	1-AMBI/7 = 0.96		
Norway NQI		0.4 m2	Combined reference value 0.78				
Denmark DKI			H'(logbase2) = 5		AMBI = 0 = 100% sensitive species		
Germany M-AMBI	Subtidal 18m, fine sand, sand		H' = 2.66	S = 31	AMBI = 0.107		
	low littoral sand		H' = 2.22	S = 17	AMBI = 0.393		
Netherlands BEQI		1983-1990 1 m2 six nautical miles from the coast	Based on permutation calculations, reference values are determined for each component metric. The reference values are calculated per habitat over increasing sampling surfaces. This allows for the estimation of the reference value for any given sampling surface. The reference for a 1m2 sampling surface is based on a set of 2000 artificial random samples out of the reference dataset. The H/G boundary is the median for species number and species composition similarity and the 25th and 75th percentile for density and biomass.				Out of the randomisation procedure, for each component metric (indicators: density, biomass, species richness, species composition changes) a 5th percentile value is selected as the value that has to be reached to achieve good status (the value of the G/M boundary). For the parameters density and biomass, a two side deviation from the reference values is scored (2.5th and 97.5th percentile).
			Minimum sampling surface 1m2				Minimum sampling surface 1m2
			Similarity =	S =	Density =		Similarity = 0.74; Density = 2584 - 7975 S = 60; Biomass = 14.2 - 52.4
Belgium BEQI	Shallow fine muddy sand, *Abra alba* community	1994 – 2004 1.5 m2	Sampling surface OK (2m2)				Sampling surface OK (2m2)
			Similarity > 0.82	S > 85	Density = 4908-7384		Similarity < 0.78; S < 76; Density < 3443 or > 10698 = moderate
			Minimum sampling surface (1.8m2)				Minimum sampling surface (1.8m2)
			Similarity > 0.8	S > 79	Density = 4678-7545		Similarity < 0.75; S < 69; Density < 3130 or > 11255

Member State / method	Habitat if distinguished	Reference conditions period / surface	Reference conditions description for method parameters / derivation			H/G boun-dary	G/M boundary derivation		
Belgium BEQI	well-sorted mobile sands, *Nephtys cirrosa* community		Sampling surface OK				Sampling surface OK		
			Similarity > 0.73	S > 47	Density = 272 - 396		Similarity < 0.65	S < 38	Density < 204 or > 544
			Minimum sampling surface				Minimum sampling surface		
			Similarity > 0.67	S > 40	Density = 264 - 403		Similarity < 0.57	S < 31	Density < 179 or > 620
	shallow sandy mud, *Macoma balthica* community		Sampling surface OK				Sampling surface OK		
			Similarity = 0.72	S = 46	Density = 496 - 998		Similarity = 0.63	S = 38	Density = 280 - 1657
			Minimum sampling surface				Minimum sampling surface		
			Similarity = 0.66	S = 39	Density = 421 - 1041		Density = 0.55	S = 29	Density = 227 – 1807
France M-AMBI									
Spain M-AMBI	Iberian, Canta-brian and Atlantic		H' = 4	S = 42	AMBI = 1				
Portugal Iberian P-BAT		0.1 m2	H' = 4.1 Margelef = 5		AMBI = 0				

Type	Sali-nity psu	Member State / method	Habitat if distinguished	Reference conditions period / surface	Reference conditions description for method parameters / derivation				H/G boundary derivation	G/M boundary derivation	Geographical delineation of the type
NEA 3/4	18-30	Germany M-AMBI 3m3	High littoral mud		Similarity	S	AMBI		High/Good boundary (0.85)	The boundary G/M was set as 0.7	Wadden Sea type polyhaline
					2.16	18	2.7				
			Middle littoral muddy sand		2.34	23	0.947				
			Low littoral sand		2.22	17	0.393				
			Brackish sublittoral		2.178	16	1.541				
		Netherlands BEQI	High littoral mud	1969 to 1983 for littoral Wadden Sea 1988 to 1990 for Wadden and Eemscoast	Similarity	S	Density	Biomass			
					0.68	13	448 - 7643	4.1 – 20.6			
			Middle littoral muddy sand		0.7	17	269 - 12063	18.4 – 58.9			
			Low littoral sand		0.6	13	106 - 7384	4.3 – 24.3			
			Brackish sublittoral		0.82	26	1810 - 103353	18.7 – 88.8			

Type	Salinity psu	Member State / method	Habitat if distinguished	Reference conditions period / surface	Reference conditions description for method parameters / derivation	H/G boundary derivation	G/M boundary derivation	Geographical delineation of the type
NEA 7	>30	Norway NQI		0.4 m2	Combined reference value 0.78			Deep fjordic and sea loch systems
		UK IQI			The same as in type 1/26			
NEA 8 / 9 / 10	18-30	Norway			The 90-percentile (the border value between the 90% lower and 10% higher values among all the grab samples from the reference stations) was used to quantify the reference value for the method.			NEA8:Kattegat + small area of Skagerrak (Inner Arc) sheltered, shallow NEA9: Skagerrak fjords with a shallow sill NEA10: Skagerrak Outer Arc, exposed, deep
		Sweden BQI			Comparative data was chosen for each national type from regions lacking local discharges; in practice areas with the highest mean BQI values existing for that type. The upper third of the span exceeding the G/M boundary was reserved for the status of high.		Sequential tests identifying the level of BQI (20th percentile) where a water body significantly differs from the comparison material were made to assist in the setting of G/M boundary. Once the G/M boundary was determined, it was deemed acceptable to consider the area from the G/M boundary up to the highest observed index value in the existing type as mainly constituting a status of good. Two-thirds (2/3) of the span exceeding the G/M boundary was assigned a status of good.	
		Denmark			The reference state is the highest diversity (Hmax) and the lowest AMBI found in the material having at least Good status from the type in question.	Set midway between the G/M value and the theoretical highest index value of 1.	Determined by the 5th percentile of samples/sites classified as being in at least Good status	

Mediterranean Sea							
Member State / method	Habitat if distin-guished	Reference conditions period / surface	Reference conditions description for method parameters / derivation			H/G boundary derivation	G/M boundary derivation
Greece and Cyprus Bentix	Mixed soft sediments	1985-1997 20 stations in the Aegean & Ionian Seas 0.1 m2	The benthic fauna is usually very diverse and evenly distributed with no one species naturally dominating over 10%. The probability of one species randomly picked up from the fauna to belong to a “tolerant” over a “sensitive” group is 3:1			Bentix = 4.5 (sensitive species = 62,5% with this number, in the Good to High boundary sensitive species become over 60% and the tolerant species less than 40%), (S > 80, H > 5 for intercalibration, not for assessment)	Bentix = 3.5 (sensitive species 37,5%) class centre good bentix = 4 (sensitive and tolerant species = 50%) = ecotone point of the transitional zone from sensitive to tolerant species. At the G/M boundary, the percentage of tolerant species becomes over 60% (roughly 2/3 of the fauna) and the sensitive taxa less than 40% (1/3 of the fauna)
			(Hmax = 6 for intercali-bration, not for assess-ment)	(Smax = 110-120 for intercali-bration, not for assess-ment)	Bentix > 5, composition of the fauna corresponds to sensitive species over 75%		
	Muddy bottoms	1992-1997 5 stations in the Aegean sea, 0.1 m2	(Hmax = 5 for IC, not for assessment)	(Smax = 40 for IC, not for assessment)	Bentix > 4, Sensitive species % over 50%	Bentix = 4 (H > 4 - 4.5, S > 30, for intercalibration, not for assessment)	Bentix = 3
Croatia		?					
Slovenia M-AMBI		1997-2004	Average values from two sampling sites from the area with minimal known human impact and adding 15% to these values.			Natural variability, presumed to be around 20%, defines width of High class, so upper and lower limit of High class differ for 20%. H/G boundary (lower limit) was calculated by taking median from EQR values of the two stations used in calculating reference conditions and subtracting additional 5% from this value. Subtracting 5% was needed because median of the actual data lays 15% from upper limit, so to get the lower limit, which differs from upper for 20% this subtraction must be done.	Other boundaries were set equidistantly from the H/G boundary (0,83): between G/M on 0,62, between M/P on 0,41 and between P/B on 0,20
			H' = 5.8	S = 110	AMBI = 1.3/6		
Italy AMBI, M-AMBI, Bentix		2001-2004 (50 stations)	Option 1: Samples were selected with individuals in ecological group I >70% (considered more important than R and S), samples with species richness R> 20, H> 2. The median value of AMBI, R and H in these samples, is used as reference.			Option 1: The 10th percentile of EQR values obtained for the reference samples. Option 2: The 75th percentile of EQR values (obtained by applying M-AMBI in all the samples)	Option 1, 2 and 3: The width of the four remaining classes was evenly spaced over the remaining interval after setting the H/G boundary.

Member State / method	Habitat if distin-guished	Reference conditions period / surface	Reference conditions description for method parameters / derivation	H/G boundary derivation	G/M boundary derivation
Italy AMBI, M-AMBI, Bentix		2001-2004 (50 stations)	Option 2: the 90th percentile for R and H and the 10th percentile for AMBI was taken from the entire dataset Option 3: the 75th percentile for R and H, and the 25th percentile for AMBI results were taken from selected control sites Option 1 H' = 4 R = 30 AMBI = 0.5 Option 2 H' = 4 R = 42 AMBI = 0.4 Option 3 H' = 3.3 R = 33 AMBI = 0.8	Option 3: The 30th percentile of AMBI, R and H calculated in the reference sites	
France Multi-metric approach AMBI, H´, BQI trophic index	stratified sampling per major community type (4 particle types) and per reference station (undisturbed zone)	? - 2005	170 samples (79 % of 214) in Languedoc – Roussillon region were considered undeteriorated by expert judgment having high and good status. H' enabled an efficient discrimination of deteriorated stations, AMBI and TI were not stringent enough and BQI was too stringent for these set of stations, but for the 7 reference stations AMBI and TI were too stringent and BQI was as good but less efficient than H'. For 15 stations in Corsica the high and good status conditions of Languedoc-Roussillon could not be applied. 7 reference sites in Languedoc Roussillon (2 stations) and Provence-Alpes-Côte d'Azur (5 stations) were defined	H' = 4 AMBI = 1.2 BQI = 18.8 at depth < 20 m BQI = 26.4 at depth > 20 m IT = 80	H' = 3 AMBI = 3.3 BQI = 14.1 at depth < 20 m BQI = 19.8 at depth > 20 m IT = 60
Spain MEDOCC		1991-2006	Selecting the best situation (sample) where most species belong to EGI (sensitive species) and EGII (indiferent species). From these samples, all tolerant (EGIII) and opportunistic species (EGIV) have been excluded and a new theoretical situation was created where the fauna is composed of only sensitive (EGI: 90%) and indifferent species (EGII: 10%). (H'=5.54) (R=100) (AMBI = 0.14) (for use in M-AMBI comparison in intercalibration, not for assessment)	MEDOCC = 1.6: sensitive ecological group (EGI) accounting for more than 40% of total abundances	MEDOCC = 3.2: tolerant ecological group (EGIII) accounts for 20-50%, but sensitive taxa (EGI) are also present (10-40%)

Type	Member State / method	Habitat if distin-guished	Reference conditions period / surface	Reference conditions description for method parameters / derivation			H/G boundary derivation	G/M boundary derivation
Black Sea CW BL 1	Romania One out-all out of H´, AMBI and M-AMBI		2002-2006 (21 samples)	Expert judgment			Expert judgment	Expert judgment
				Shannon diversity index > 4	S = 50	AMBI and M-AMBI same as NEA	Shannon diversity index: 4 AMBI = 1.2 M-AMBI = 0.85	Shannon diversity index: 3 AMBI = 3.3 M-AMBI = 0.55
	Bulgaria (95 samples for 2 habitats) One out-all out of H´, AMBI and M-AMBI	Sandy and mixed sedi-ments	2002-2006	Derived from areas as unaffected by human activities as possible using the data of slightly disturbed benthic communities (having at least good status)			Deviation of 10% from the High status	Deviation of 30% from the High status
				H´= the average community diversity index of stations reaching good ecological status (3.6) constitutes 75 % from the high status (so add ¼ to the average = 4.5)	S = 50	AMBI > 0.2 and M-AMBI same as NEA	The ecotone of the good status = the ecotone of the high status (H´ 4.5) * 80% = 3.6 for H´, the H/G boundary is then defined as the middle between the ecotone of high (4.5) and good status (3.6) H/G for H´= 4 (middle between 4.5 and 3.6) AMBI = 1.2 M-AMBI = 0.85	G/M for H´= 3.1 (middle between 3.6 and 2.7) The ecotone of the moderate status = the ecotone of the high status (H´ 4.5) * 60% = 2.7 AMBI = 3.3 M-AMBI = 0.55
		Muddy sediments (considered more sensitive?)	2002-2006				Deviation of 10% from the High status	Deviation of 30% from the High status
				H´= average community diversity index of stations having more than 25 species per sample (2.9) is 70 % from reference = 3.6		AMBI > 0.2 and M-AMBI same as NEA	For H´= 3.3 The ecotone of the good status = the ecotone of the high status (H´ 3.6) * 80% = 2.9 for H´, the H/G boundary is then defined as the middle between the ecotone of high (3.6) and good status (2.9) = 3.3 AMBI = 1.2 M-AMBI = 0.85	For H´= 2.5 (middle between 2.9 and 2.2) The ecotone of the moderate status = the ecotone of the high status (3.6) * 60% = 2.2 AMBI = 3.3 M-AMBI = 0.55

Conclusions benthic invertebrate fauna

- Mentioning time periods of measurements/description of dataset used for derivation of the reference conditions or the alternative benchmark is often still lacking, but needed to understand which situation in time and space the assessment of the present situation is compared with. Addition of this information is still necessary, since it is related to the illustration of the absence or which level of presence of a pressure has been allowed, not leading to a significant ecological impact to define the alternative benchmark of good status.
- Concerning the procentual contributions of different categories of sensitive and tolerant species, it is important to compare the conceptual framework used in the different methods and how the percentages relate to each-other at the class boundaries. A nice comparison of this kind between different methods has been done for the MED GIG in a separate scientific publication (Simboura & Argyrou, 2010) and should also make part of the description on the common understanding of reference conditions in the technical report. In the technical report such comparison is only included for the MEDOCC and the AMBI, which should be extended with the BENTIX as illustrated in the paper. This will clarify how the ecotones and class boundaries from different methods relate to each-other concerning percentages of specific sensitivity categories of species and if there is not an inconsistency in setting these boundaries.
- The use of multimetrics for which a single value is given as the result of a formula renders the ecological understanding of the method very difficult. The understanding of the characteristics of a community classified as high, good or moderate cannot be understood from the values of the method. Further exchange between the Member States and JRC is needed to find the least common denominator how to describe the ecological meaning of the class boundaries for the macrobenthic invertebrate fauna. Getting an understanding of the results for the different required parameters in the different status classes is crucial.
- Only Germany and the Netherlands considered littoral areas for assessment in the Wadden Sea type, only the sublittoral area is considered in all the other types and GIGs.
- More care should be taken in the use of sensitivity classifications of species in other areas than where they have been investigated. Significant differences have been found in sensitivity analyses of species like between the Baltic Sea and the NEA. Species have been assigned to different sensitivity categories in the MED GIG using the MEDOCC or the AMBI. It has not been verified if the sensitivity of the species as defined in the AMBI can be applied in the Black Sea without any adjustments.
- In the Mediterranean Sea GIG, Greece, Cyprus, Slovenia and Spain agreed on a maximal species richness of 100-110 for intercalibration purposes, but Italy defined a maximal species richness of 30 to 42. This inconsistency still needs to be clarified and adjusted.
- Established relationships between the methods and physico-chemical parameters have been illustrated for the Baltic, the NEA and the MED GIG, but not for the Black Sea.
- The Black Sea GIG does not have sufficient data yet for proper statistical treatment (RO 21 and BU 95 samples, mostly not replicated) and will have to increase the confidence on the derivation of reference conditions.

4. Transitional Water (TW)

For Transitional water bodies there has not been an official exercise of IC during the first phase, and the Reference Conditions issue has not been considered.

A document has been distributed within the MED TW GIG. It aims to be a supporting proposal to increase the consistency in the evaluation of criteria on pressures by MS; the criteria arising from a priori and a posteriori analyses of human activities influencing transitional coastal lagoons. The purpose of the paper is to start a process that will help in the further development and agreement of common pressures and thresholds for reference sites for Transitional Waters, in a similar way as has been done for Rivers and Lakes. The idea is to provide a spatial network of reference sites (minimally disturbed sites) for the types that will be possible.

CONCLUSIONS ON CONSISTENCY IN THE APPLICATION OF REFERENCE CONDITIONS CRITERIA FOR IC PHASE I

This review revealed poor consistency in the way RC criteria were applied by MSs. This fact can be mainly attributed to the way the IC work was organised and structured into GIGs without common guidance on the establishment of horizontal Reference Conditions criteria. MSs used a variety of criteria to identify reference sites that differed both within and between GIGs. The information derived from the technical reports and summarised below suggests that there is an insufficient level of comparability in how RC were derived:

1. There was a broad variety of responses when evaluating pressures within river GIGs and river types. Inconsistencies were due to MSs' misinterpretation of the answer system, the absence of pressures data and differences in RC criteria between BQE, among others.

2. In the river GIGs, there were different interpretations about how the RC criteria from CB GIG should be applied; this included the use of different thresholds and different interpretations of the answering/scoring system.

3. In general, there was weak fulfilment of the Rivers RC criteria by MSs and GIGs. This has been concluded by checking the application of reference and rejection thresholds in relation to the pressures data supplied by MS for their reference sites.

4. River and Lake GIGs used similar approaches for quantifying pressures, but clear differences were evident when these were analysed at the water category level.

5. Despite some differences in the choice of reference criteria and their values caused by different data availability and geographic conditions between the Lake GIGs, a common understanding was reached on Reference Conditions in the Lake IC: no industrialization, urbanization or intensive agriculture in the catchment and no or only minor ecological effects caused by humans. However, the GIGs used slightly different criteria/ approaches for selecting reference lakes so there is no guarantee for comparable Reference Conditions across Europe.

6. In the Northern and Mediterranean Lake GIGs, MSs used slightly different criteria that were not harmonized within the GIG.

7. For many lakes pressure data were not available (e.g., land use in Finland), due to several causes, e.g., a lack of data/information in the central databases. Also not all REFCOND pressure criteria were used as not considered important for specific pressure as eutrophication.

8. Similarly, difficulties to find "true" reference lakes have to be acknowledged.

- There are few reference lakes left in Europe, mainly in the Northern region. Therefore some flexibility has to be allowed, e.g. if we are looking for reference lake for eutrophication pressure, there might be some deviation from the strict hydromorphological criteria (as used in several GIGs).

- There is not always a clear link between pressure criteria (e.g. land use) and in-lake nutrient concentrations that are the sum of different factors, mainly the processes of the lake ecosystem (inner nutrient load, top-down control, etc). So in many cases the lake nutrient levels can be used as a proxy of eutrophication pressure to select reference lakes and define biological reference conditions (as demonstrated by ALP GIG). Still this approach has to be used with a caution, including a well-grounded setting of reference TP levels (best approached through combination of available data, hindcasting using export coefficient models, historical data, paleoreconstruction and expert judgment).

- An analysis of representativeness should be carried out to ensure that reference lakes are representative of the type (i.e. to ensure that there are no significant differences in basic characteristics between impacted and non-impacted lakes).

9. Coastal Reference Conditions were mostly biological and not consistent with the general approach from Rivers and Lakes, which was based on a spatial network of minimally disturbed sites for macroinvertebrates and diatoms. We refer to the section 3 on Coastal waters for conclusions on each BQE.

10. The IC work on Transitional Water bodies is starting, but due to the intermediate position of these water bodies in the catchment they should be a key type of ecosystem linking approaches for the identification of inland freshwater and Coastal RC.

11. Tiered approach to reference screening would be recommendable:

- Tier 1 - "true" reference sites – sites with no or minimal anthropogenic pressure that fulfill all pressure criteria proposed in REFCOND Guidance;
- Tier 2 - "reference conditions" sites or "partial" reference sites – impacted by some level of anthropogenic pressure but (some) biological communities corresponding to the reference conditions (e.g. "phytoplankton reference sites" with no or minimal eutrophication pressure but significant hydromorphological pressure which still is not affecting phytoplankton community in a significant manner;
- Tier 3 - "alternative benchmark" sites – sites with some pressure and some level of impairment to biology (can be used for setting benchmark, see Intercalibration Guidance, 2010).

RECOMMENDATIONS TO IMPROVE CONSISTENCY IN DEFINING REFERENCE CONDITIONS

1. To establish a common framework, as a guidance document for MSs and GIGs, for the characterisation and quantification of pressures for different water categories (Rivers, Lakes, Transitional and Coastal), types and BQE. Within the common framework, specific criteria addressing different ecosystem types should also be considered.
2. Improve the characterisation of pressures and adopt consistent methodologies for the assessment of pressures. Identify relevant indicators of pressures within water categories that are of relevance to the biota.

3. Components of the same BQE should have consistent RC criteria. Develop reference concepts that embrace all BQE to avoid deviations.

4. Reinforce MSs' consistent application of agreed RC criteria.

5. Further work to improve our understanding of biological responses to pressures and the establishment of relevant quantified and standardised thresholds for reference and other status classes.

GLOSSARY

Term	Explanation
BQE	Biological quality element (Annex V of the WFD)
CIS	Common Implementation Strategy of the Water Framework Directive
CW	Coastal waters
Class boundary	The EQR value representing the threshold between two quality classes: H/G Boundary – Boundary between "high" and "good" status, G/M Boundary - Boundary between "good" " and "moderate" status
Ecological status	One of two components of surface water status, the other being chemical status. There are five classes of ecological status of surface waters (high, good, moderate, poor and bad)
ECOSTAT CIS	Common Implementation Strategy (CIS) Working Group A Ecological Status
EQR	Ecological Quality Ratio
GIG	Geographic Intercalibration Group (a geographical area assumed to have comparable ecological conditions): ALP - Alpine; ATL – Atlantic; CB – Central Baltic; EC – Eastern Continental; MED - Mediterranean; NEA- North East Atlantic; NO – Northern
IC	Intercalibration - Benchmarking exercise to ensure that good ecological status represents the same level of ecological quality everywhere in Europe
IC Phases	IC phase I - the first part of the Intercalibration process (2003 – 2008) resulting in the EC Decision (2008) on the values of the MSs classification systems IC phase II - the second part of the Intercalibration process (2008 – 2011) with the aim to close the gaps of the IC and improve the comparability of the results
JRC	Joint Research Centre which provides research support for EU policy-making
MS	Member State (of the European Union)
Pressures	Physical expression of human activities that changes the status of the environment (nutrient loading, hydromorphological modifications, water abstraction, etc...)
RC	Reference conditions: The benchmark against which the effects on surface water ecosystems of human activities can be measured and reported
REFCOND	Guidance on setting Reference conditions, see EC 2003a
TW	Transitional waters
WFD	Water Framework Directive 2000/60/EC

REFERENCES

Arbuckle, K. E. and J. A. Downing, 2001. The influence of watershed land use on lake N:P in a predominantly agricultural landscape. *Limnology and Oceanograpy* 46: 970-975.

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

Baattrup-Pedersen, A., E. A. Kristensen, J. Jørgensen, J. Skriver, B. Kronvang, H. E. Andersen, C. C. Hoffmann and L. M. Kjellerup-Larsen, 2009. Can *a priori* defined reference criteria be used to select reference sites in Danish streams? Implications for implementing the Water Framework Directive. *Journal of Environmental Monitoring* 11: 344-352.

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 and J. G. Wasson, 2011. Bringing European river quality into line: an exercise to intercalibrate macro-invertebrate classification methods. *Hydrobiologia* 667: 31-48.

Borja, A., I. Muxika and J. Franco, 2003. The application of a Marine Biotic Index to different impact sources affecting soft-bottom benthic communities along European coasts. *Marine Pollution Bulletin* 46: 835-845.

Buffagni, A., S. Erba, M. Cazzola, J. Murray-Bligh, H. Soszka and P. Genoni, 2006. The STAR common metrics approach to the WFD intercalibration process: full application for small lowland rivers in three European countries. *Hydrobiologia* 566: 379-399.

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

van de Bund, W. J. (ed.), 2009. Water Framework Directive intercalibration technical report. Part 1: Rivers. JRC Scientific and Technical Reports.

EC, 2003a. Common implementation strategy for the Water Framework Directive (2000/60/EC). Rivers and Lakes - Typology, Reference Conditions, and Classification Systems. Luxembourg, Office for Official publications of the European Communities. http://circa.europa.eu/Public/irc/env/wfd/library

EC (2003b) Common implementation strategy for the water framework directive (2000/60/EC). Towards a guidance on establishment of the intercalibration network and on the process of the intercalibration exercise. Luxembourg, Office for Official publications of the European Communities. http://circa.europa.eu/Public/irc/env/wfd/library

EC, 2010. WFD CIS Guidance Document No14. Guidance on the Intercalibration Process 2008-2011.

Ecke, F., 2009. Drainage ditching at the catchment scale affects water quality and macrophyte occurrence in Swedish lakes. *Freshwater Biology* 1: 119-126.

Hill, A. R., 1981. Stream phosphorus exports from watersheds with contrasting land uses in southern Ontario. *Water Resources Bulletin* 17: 627-634.

Hunsaker, C. T., D. A. Levine, S. P. Timmins, B. L. Jackson and R. V. O'Neill, 1992. Landscape characterization for assessing regional water quality. pp. 997-1006. In McKenzie, D., E. Hyat, and J. McDonald (eds.). Ecological Indicators. Elsevier, New York.

Janauer, G. A., 2001. Is what has been measured of any direct relevance to the success of the macrophyte in its particular environment? Scientific and legal aspects of biological monitoring in freshwater. *Journal of Limnology* 60: 33-38.

Johnes, P. J., 1996. Evaluation and management of the impact of land use change on the nitrogen and phosphorus load delivered to surface waters: the export coefficient modelling approach. *Journal of Hydrology* 183: 323-349.

King, R. S., M. E. Baker, D. F. Whigham, D. E. Weller, T. E. Jordan, P. F. Kazyak and M. K. Hurd, 2005. Spatial considerations for linking watershed land cover to ecological indicators in streams. *Ecological Applications* 15(1): 137-153.

Knoll, L. B., M. J. Vanni and Renwick, 2003. Phytoplankton primary productivity and phytosynthetic parameters in reservoirs along a gradient of watershed land use. *Limnology and Oceanograpy*, 48: 608-617.

Nielsen, S. L., K. Sand-Jensen, J. Borum and O. Geertz-Hansen, 2002. Depth colonization of eelgrass (*Zostera marina*) and macroalgae as determined by water transparency in Danish coastal waters. *Estuaries* 25: 1025-1032.

Nijboer, R. C., R. K. Johnson, P. F. M. Verdonschot, M. Sommerhauser and A. Buffagni, 2004. Establishing reference conditions for European streams. *Hydrobiologia* 516: 91-105.

Nõges, P., W. van de Bund, A. C. Cardoso, A. G. Solimini and A. S. Heiskanen, 2009. Assessment of the ecological status of European surface waters: a work in progress. *Hydrobiologia* 633: 197-211.

Osborne, L. L. and D. A. Kovacic, 1993. Riparian vegetated buffer strips in water-quality restoration and stream management. *Freshwater Biology* 29: 243-258.

Osborne, L. L. and M. J. Wiley, 1988. Empirical relationships between land use/cover and stream water quality in agricultural watershed. *Journal of Environment Management* 26: 9-27.

Pardo, I., C. Gómez-Rodríguez, J.-G. Wasson, R. Owen, W. van de Bund, M. Kelly, C. Bennett, S. Birk, A. Buffagni, S. Erba, N. Mengin, J. Murray-Bligh and G. Ofenböeck (in prep). The European reference condition concept: A scientific and technical approach to identify minimally impacted aquatic ecosystems. Submitted to *STOTEN*.

Poikane, S. (ed.), 2009. Water Framework Directive intercalibration technical report. Part 2: lakes. JRC Scientific and Technical Reports. EUR 23838 EN/2.

Poikane, S., M. van den Berg, J. Ortiz-Casas, G. Phillips, A. L. Solheim, D. Tierney, G. Wolfram and P. Noges, 2009. Lake assessment strategy in European Union (EU): Case study of European large lakes. *Verhandlungen der Internationalen Vereinigung für Theoretische und Angewandte Limnologie*, vol. 30, Part 7

Reckhow, K. H., M. N. Beaulac and J. R. Simpson, 1980. Modeling Phosphorous. Loading and Lake Response Under Uncertainty: A Manual and Compilation of export coefficients. US EPA/440/5-80/011, U.S. Environmental Protection Agency, Washington, DC.

Schories, D., U. Selig and H. Schubert, 2006. Testung des Klassifizierungsansatzes Mecklenburg-Vorpommern (innere Küstengewässer) unter den Bedingungen Schleswig-Holsteins und Ausdehnung des Ansatzes auf die Außenküste. Küstengewässer-Klassifizierung deutsche Ostsee nach EU-WRRL. Teil A: Äußere Küstengewässer.

Stoddard, J.L., D.P. Larsen, C. P. Hawkins, R. K. Johnson and R. H. Norris, 2006. Setting expectations for the ecological condition of streams: The concept of reference condition. *Ecological Applications* 16: 1267-1276.

Strayer, D. L., L. C. Beighley, L. C. Thompson, S. Brooks, C. Nilsson, G. PInay and R. J. Naiman, 2003. Effects of land cover on stream ecosystems: roles of empirical models and scaling issues. *Ecosystems* 6: 407-423.

van de Bund, W. J., 2009. Water Framework Directive intercalibration technical report. Part 1: Rivers, JRC Scientific and Technical Reports.

Wallin, M., T. Wiederholm and R. K. Johnson, 2003. Guidance on establishing reference conditions and ecological status class boundaries for inland surface waters. Produced by CIS Working Group 2.3 - REFCOND. Available via the internet at http://www-nrciws.slu.se/REFCOND/7th_REFCOND_final.pdf

Wolfram, G., C. Argillier, J. De Bortoli, G. Buzzi, M. T. Dokulil, E. Hoehn, A. Marchetto, P. J. Martínez, G. Morabito and M. Reichmann, 2009. Reference conditions and WFD compliant class boundaries for phytoplankton biomass and chlorophyll-a in Alpine lakes. *Hydrobiologia* 633: 45-58.

European Commission

EUR 24843 EN – Joint Research Centre – Institute for Environment and Sustainability
Title: **Revision of the consistency in Reference Criteria application in the Phase I of the European Intercalibration exercise**

Author(s): Isabel Pardo, Sandra Poikane, Wendy Bonne
Luxembourg: Publications Office of the European Union
2011 – 94 pp. – 21 x 29,7 cm
EUR – Scientific and Technical Research series – ISSN 1018-5593 (print), 1831-9424 (online)

ISBN 978-92-79-20444-9 (PDF)
ISBN 978-92-79-20443-2 (print)

doi: 10.2788/27809

Abstract
This document is produced within the cross GIG working group on Reference conditions and is a final report on the consistency check survey, with an assessment of implications and recommendations for the application of reference conditions during IC phase two

How to obtain EU publications

Our priced publications are available from EU Bookshop (http://bookshop.europa.eu), where you can place an order with the sales agent of your choice.

The Publications Office has a worldwide network of sales agents. You can obtain their contact details by sending a fax to (352) 29 29-42758.

The mission of the JRC is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

LB-NA-24843-EN-N

EUROPEAN COMMISSION

ISBN 978-92-79-20444-9

Publications Office