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

WFD Intercalibration Phase 2: Milestone 6 report

Water category/GIG/BQE/horizontal activity:	**EC GIG**
Information provided by:	Sebastian Birk With contributions from Lukács Balázs András, Peter Balazi, Gana Gecheva, Mateja Germ, Karin Pall, Lívia Tóthova, Gorazd Urbanic, Nigel Willby

1. Organisation

1.1. Responsibilities

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

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

Lead, DE – Sebastan Birk

SK – Peter Baláži, Lívia Tóthova
SI – Gorazd Urbanic, Mateja Germ
AT – Karin Pall, Franz Wagner
HR – Antun Allegro
HU – Lukács Balázs András
BG – Gana Gecheva
RO – Serban Iliescu

1.2. Participation

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

Austria, Bulgaria, Croatia, Hungary, Romania, Slovakia, Slovenia

1.3. Meetings

List the meetings of the group:

Vienna (AT), January 20, 2011
Plovdiv (BG), October 20-21, 2011

2. Overview of Methods to be intercalibrated

Identify for **each** MS the national classification method that will be intercalibrated and the status of the method

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

Member State	Method	Abbr.	ID[1]	Status
Austria	Austrian Index Macrophytes for Rivers	AIM	69	1
Bulgaria	Reference Index	RI-BG	355	2
Hungary	Reference Index[a]	RI-HU	328	3
Slovakia	Biological Macrophyte Index for Rivers	IBMR-SK	167	2
Slovenia	River Macrophyte Index	RMI	81	1

[a] Hungary will adopt the class boundaries resulting from the intercalibration exercise using the reference index developed for large to very large lowland rivers (R-E3).

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

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

Do all national assessment methods meet the requirements of the Water Framework Directive? (Question 1 in the IC guidance)

Do the good ecological status boundaries of the national methods comply with the WFD normative definitions? (Question 7 in the IC guidance)

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

Compliance criteria	Compliance checking conclusions
1. Ecological status is classified by one of **five classes** (high, good, moderate, poor and bad).	Yes – all methods classify ecological status by one of five classes.
2. High, good and moderate ecological status are set in line with the WFD's **normative definitions** (**Boundary setting procedure**)	All methods except Hungarian classification set the high, good and moderate ecological status in line with the WFD's normative definitions. National boundary setting protocols are attached to this report. Hungary will adopt the harmonized class boundaries derived from the intercalibration exercise.

[1] [http://www.wiser.eu/programme-and-results/data-and-guidelines/method-database/detail.php?id=] plus ID

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

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

All national methods using river macrophytes are WFD compliant. The combination of sub-components to full BQE assessment is not addressed in this intercalibration exercise.

4. Methods' intercalibration feasibility check

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

4.1. Typology

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

Common IC type	Type characteristics	MS sharing IC common type
Lowland rivers of the Plains	Abbreviation: R-E2 and R-E3 Catchment area: 100 - >10,000 km^2 Altitude: <200 m Geology: mixed Channel substrate: Sand, silt and gravel	Bulgaria, Hungary, Romania, Slovakia, Slovenia
Upland streams of medium-size, mid-altitude streams in the Plains	Abbreviation: R-E4 Catchment area: 100-1,000 km^2 Altitude: 200 - 500 m Geology: mixed Channel substrate: Sand and gravel	Austria, Bulgaria, Hungary, Romania, Slovakia, Slovenia

Remark

Statistical analysis revealed no difference between macrophyte communities of medium-sized lowland (R-E2) and large to very large lowland rivers (R-E3). Thus, we merged both types to a single intercalibration exercise for Eastern Continental **lowland rivers**.

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

Method	Appropriate for IC types / subtypes	Remarks
Austrian Index Macrophytes for Rivers (AT)	R-E4	none
Biological Macrophyte Index for Rivers (IBMR-SK)	R-E2 and R-E3, R-E4	
River Macrophyte Index (SI)	R-E2 and R-E3, R-E4	
Reference Index (BG)	R-E2 and R-E3, R-E4	
Reference Index (HU)	R-E2 and R-E3	

Conclusion

Intercalibration is feasible in terms of typology.

4.2. Pressures

Describe the pressures addressed by the MS assessment methods

Method	Pressures	Remarks
Austrian Index Macrophytes for Rivers (AT)	Eutrophication, General degradation	none
Biological Macrophyte Index for Rivers (IBMR-SK)	Eutrophication	
River Macrophyte Index (SI)	Eutrophication, General degradation	
Reference Index (BG)	Eutrophication, General degradation	
Reference Index (HU)	Eutrophication, General degradation	

Demonstration of empirically tested pressure-impact relationships

Country	Pressure	Indicator tested	Sample size	R	p
Austria	Hydromorphology and land use	River channelization, % intensive agriculture (*multiple regression*)	51	0.49	<0.01
Bulgaria	General	Ammonium, river channelization, % natural land use in catchment (*multiple regression*)	45	0.47	<0.05
	Water pollution/ eutrophication	Orthophosphate, Biological Oxygen Demand (*multiple regression*)	47	0.40	<0.05
Hungary	Water pollution/ eutrophication	Biological Oxygen Demand, Chemical Oxygen Demand (*multiple regression*)	236	0.51	<0.01
		Orthophosphate, BOD (*multiple regression*)	236	0.24	<0.01
		Total phosphorus, Nitrate, Ammonium (*multiple regression*)	236	0.48	<0.01
	Hydromorphology	Damming, river channelization (*multiple regression*)	236	0.83	<0.01
	Land use	Catchment land use (*multiple regression*)	236	0.25	<0.01
Slovakia	Eutrophication	Total phosphorus	92	0.56	<0.01
		Ammonium	92	0.42	<0.01
		Nitrite	92	0.67	<0.01
		Nitrate	92	0.42	<0.01
		Total nitrogen	92	0.45	<0.01
Slovenia[2]	General	Natural areas in the sub-catchment	208	0.76	<0.001
		Agricultural areas in the sub-catchment	208	0.67	<0.001

Conclusion

The multi-pressure environments of the lowlands preclude strong pressure-impact relationships. With regard to the pressure addressed by the individual methods intercalibration is generally feasible (but see remark about the Slovene assessment method below).

Remark

Due to differences in the pressure focus the Slovene assessment method was only weakly correlated with the average assessment score (i.e. PCM) of all other methods. However, we intercalibrated the Slovene method by transferring its class boundaries into units of the Slovak assessment method prior to the analyses of boundary comparison and harmonisation ("***satellite intercalibration***", see below for details).

4.3. Assessment concept

Do all national methods follow a similar assessment concept?

Examples of assessment concept:

- **Different community characteristics -** structural, functional or physiological - can be used in assessment methods which can render their comparison problematic. For example, sensitive taxa proportion indices vs species composition indices.

[2] published in:
Kuhar U, Germ M, Gaberščik A, Urbanič G (2010) Development of a River Macrophyte Index (RMI) for assessing river ecological status. Limnologica 41:235-243.

- Assessment systems may focus on **different lake zones -** profundal, littoral or sublittoral **-** and subsequently may not be comparable.
- Additional important issues may be the **assessed habitat type** (soft-bottom sediments versus rocky sediments for benthic fauna assessment methods) or **life forms** (emergent macrophytes versus submersed macrophytes for lake aquatic flora assessment methods)

Method	Assessment concept	Remarks
Austrian Index Macrophytes for Rivers (AT)	Focus on taxonomic structure; assessment based on indicator taxa	none
Reference Index (BG)		
Reference Index (HU)		
Biological Macrophyte Index for Rivers (IBMR-SK)		
River Macrophyte Index (SI)		
Conclusion Intercalibration is generally feasible in terms of assessment concept.		

5. Collection of IC dataset

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

Make the following table for each IC common type

IC Type	Country	Number of surveys	Number of surveys	Number of pressure-data
Lowland rivers (R-E2 and R-E3)	BG	27	27	20
	CZ	6	5	5
	HU	23	23	8
	RO	8	8	4
	SI	14	14	14
	SK[a]	33	29	12
	Total	***111***	***106***	***63***
Upland streams (R-E4)	AT	19	19	19
	BG	7	7	7
	SK[a]	4	4	4
	Total	***30***	***30***	***30***

[a] The national class boundaries of the Slovene method were translated into units of the Slovak method using the global relationship of data covering R-E2, R-E3 and R-E4 (number of surveys = 83, number of sites = 79, number of pressure-data = 83, national data used: AT, BG, SI, SK; see also Figure 7.2).

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

Data acceptance criteria	Data acceptance checking
Data requirements (obligatory and optional)	Only those data were used that meet the following national criteria for assessability of the survey site: • ***Slovene criteria:*** at least three RMI indicator taxa or abundance sum of RMI-indicators > 5 • ***Slovak criteria:*** at least three IBMR indicator taxa, average value of E_i (stenoecie coefficient) > 1 and abundance sum of IBMR-indicators > 5[a] • ***Bulgarian criteria:*** Lowland rivers (R-E2 and R-E3) – abundance of scoring taxa > 50% and abundance sum > 16 Upland rivers (R-E4) – expert evaluation of data suitability • ***Austrian criteria (only upland streams):*** at least three AIM indicator taxa with an abundance sum ≥ 16[b] To account for the diversity in pressure-focus we included only surveys that revealed standard deviations of national EQRs < 0.275 (lowland rivers) and < 0.3 (upland streams), respectively.
The sampling and analytical methodology	
Level of taxonomic precision required and taxalists with codes	
The minimum number of sites / samples per intercalibration type	
Sufficient covering of all relevant quality classes per type	
Other aspects where applicable	

[a] One Austrian survey at upland streams showed only two IBMR indicator taxa, $E_i = 2$ and an abundance sum of 8.
[b] One Austrian survey and one Slovak survey at upland streams showed an abundance sum of 10.

6. Benchmarking: Reference conditions or alternative benchmarking

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

Clarify if you have defined
- common reference conditions (*No*)
- or a common alternative benchmark for intercalibration (*Yes*)

6.1. Reference conditions

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

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

not applicable

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

not applicable

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

not applicable

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

not applicable

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

not applicable

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

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

Continuous benchmarking: Theoretical background

Benchmarking of national assessment methods is an important precondition for the comparison and harmonisation of ecological status class boundaries in intercalibration. National boundaries are expressed as relative deviations from reference conditions at which the aquatic biota show no (or insignificant) impact from anthropogenic disturbance. These reference conditions are defined differently for individual assessment methods. Intercalibration thus requires the standardisation of assessment methods using common benchmarks. Data from reference sites meeting harmonised criteria, for instance, provide such benchmarks. Since these data are generally scarce for most European water types, alternative benchmarking based on sites impacted by similar (low) levels of anthropogenic pressure (i.e. benchmark sites) can be applied.

However, both options rely on the homogeneous distribution of undisturbed or similarly disturbed sites among countries within a common type. If certain countries feature more pronounced anthropogenic disturbance than others, the common benchmark is imbalanced or unachievable. Balancing the selection of benchmark sites will be a common problem if, for instance, countries featuring contrasting population densities or land use practices, like Austria and Bulgaria, are involved in the same exercise.

In such cases the approach of "**continuous benchmarking**" represents a solution (Figure 6.1). Continuous benchmarking refers to data from sites featuring various levels of disturbance and thus accounts for the actual data availability of individual countries. However, all countries need to provide the same (set of) pressure-variables for selected sites along with the biological data. Depending on the available data countries may cover different parts of the pressure gradient ranging from undisturbed to highly perturbed conditions. This makes continuous benchmarking different from the approaches described above: the standardisation of national assessment methods is carried out not using a preselected small range of the gradient but the full continuum.

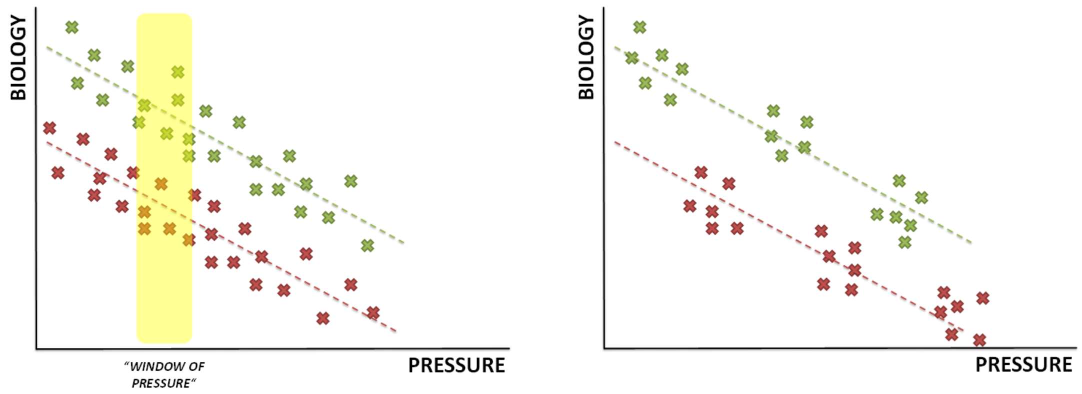

Figure 6.1: The two principal approaches of benchmarking in intercalibration

Left: "Benchmark sites" approach: Sites at similar (low) level of pressure ("window of pressure") are available from all countries (red and green).

Right: "Continuous benchmarking" approach: Countries' data (red and green) cover different levels of pressure.

The principle aim of benchmarking in intercalibration is to identify and remove differences among national assessment methods that are not caused by anthropogenic pressure (but by systematic discrepancies due to different methodology, biogeography, typology etc.). If such differences are ignored they may have an overriding effect on the comparability exercise. Therefore, the pressure effects on the assessment scores (i.e. national EQRs) have to be controlled to disclose any remaining discrepancies.

At reference or benchmark sites differences between countries due to pressure effects are minimised, because these sites do not cover a gradient of pressure. In continuous benchmarking the effects of pressure that differ between countries are removed by statistical modelling. Generalised linear models or comparable approaches (e.g. linear mixed effects models) can be applied for this purpose. The model yields the **offsets** (i.e. intercepts) of individual assessment methods which are then used in the benchmark standardisation of national EQRs (Figure 6.2).

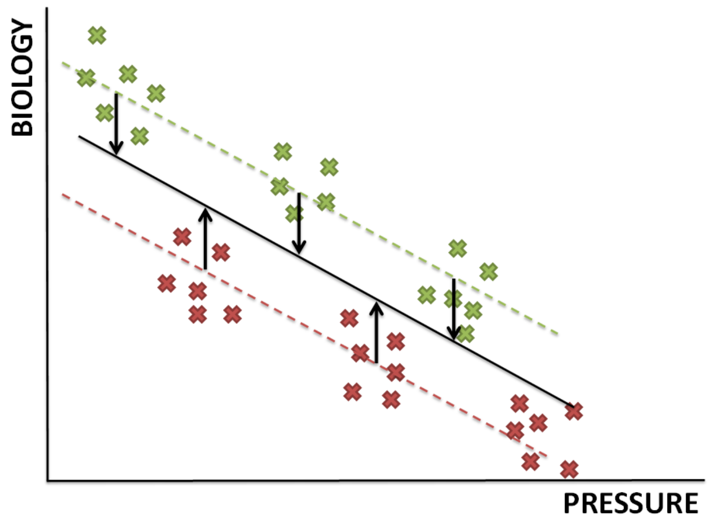

Figure 6.2: Adjustment of national EQRs (red and green) based on offset results from the statistical model

Results of benchmark standardisation

Upland streams (R-E4)

Country-effect of national EQRs was tested against single pressure-parameters applied together in a General Linear Model (GLM), i.e. percent natural and agricultural land use in the catchment, influence of flood protection dykes at survey site, total phosphorus and ammonium concentration, water conductivity – it was not possible to design a common pressure factor that all national EQRs responded to. The analysis revealed no significant country effects in the dataset. Therefore, benchmark standardisation was not applied to the upland dataset.

Lowland rivers (R-E2 and R-E3)

Country-effect of national EQRs was tested against a combined pressure factor (explaining 59 % of data variability) including degree of hydropeaking at survey site (factor loading: 0.47), influence of flood protection dykes at survey site (factor loading: 0.83), percent natural land use in the catchment (factor loading: -0.81) and total phosphorus concentration (factor loading: 0.88) (see also Figure 6.3). The GLM yielded significant country effects for the Bulgarian EQRs assessing data from Bulgaria and the Slovak EQRs assessing data from Hungary and Romania. Therefore, benchmark standardisation was applied to the complete lowland dataset by subtracting the values given in Table 6.1 (below) from the national EQRs assessing the different national data.

Relationship of Slovene and Slovak assessment methods ("Satellite intercalibration")
Country-effect of national EQRs was tested against single pressure-parameters applied together in a GLM, i.e. percent natural and agricultural land use in the catchment, influence of flood protection dykes at survey site, total phosphorus and ammonium concentration, water conductivity – it was not possible to design a common pressure factor that both national EQRs responded to. The analysis revealed no significant country effects in the dataset. Therefore, benchmark standardisation was not applied to the dataset.

Table 6.1: Offsets used to benchmark standardise individual datasets in the intercalibration exercise of lowland rivers

	EQR BG	EQR HU	EQR SK	EQR SI
BG	-0.058	-0.031	-0.012	-0.012
CZ	-0.016	0.005	-0.022	-0.022
HU	-0.009	-0.013	-0.083	-0.083
RO	0.087	0.073	0.097	0.097
SI	-0.007	-0.042	0.039	0.039
SK	0.004	0.008	-0.019	-0.019

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

Continuous benchmarking refers to data from sites featuring various levels of disturbance. The effects of pressure that differ between countries are removed by statistical modelling. See also section above.

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

The number of benchmark sites per country are specified in the above given table on the IC dataset (column: "number of pressure data").

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

Continuous benchmarking yields country-specific benchmarks that differ in degree of pressure and (thus) biological communities. The validation of the benchmark is carried out *post-hoc* investigating sites at harmonised high status (i.e. their pressure-levels and biological communities – see also section 8.2).

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

Figure 6.3 shows the relationship of three national assessment methods (left) and their PCMs (right) against the combined pressure gradient at lowland rivers. While the regression slopes were significantly different before benchmark standardisation, no significant differences were observed between regression lines of PCMs, indicating successful benchmarking.

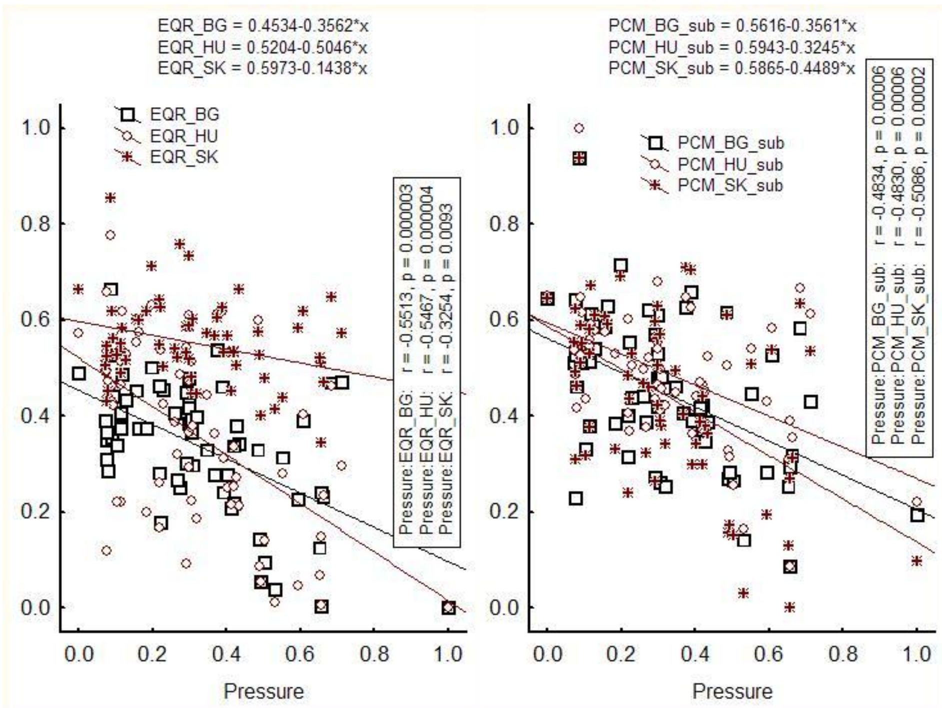

Figure 6.3: Pressure-impact relationship of the combined pressure index against Bulgarian, Hungarian and Slovak EQRs (left) and PCMs (right) for lowland rivers

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

See section on *biological communities representing the "borderline" conditions*

7. Design and application of the IC procedure

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

Which IC option did you use?

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

Explanation for the choice of the IC option:

We performed an IC Option 3 (direct comparison) using pseudo-common metrics (PCM). After adjusting and standardising the national EQRs (only necessary for lowland rivers) the values were normalised to a range of 0 to 1. Based on the survey data commonly assessed by each method we related the national EQRs to a PCM constructed by averaging the EQR values of the remaining methods using ordinary least squares **regression**. Thus, in an exercise involving three national assessment methods, for method A we related the national EQR of method A (the 'test' method) to the average EQRs of methods B and C (the 'assessors'); this was repeated for method B versus the average of methods A and C etc.
The analyses were undertaken using the Intercalibration Spreadsheets described in Birk et al. (2011)[3]

Satellite intercalibration

Due to low correlations of the Slovene assessment method with the PCM ($R_{Pearson}$=0.206 for upland streams and $R_{Pearson}$=0.072 for lowland rivers) we related this method only to Slovak EQRs following the scheme depicted in Figure 7.1. The correlation coefficient was $R_{Pearson}$=0.552. Based on linear regression analysis considering the SI-EQR interval >0.5 the Slovene quality class boundaries were translated into units of the Slovak assessment method (Figure 7.2). This allowed for including the Slovene boundary information in the subsequent comparison and harmonisation process of IC Option 3.

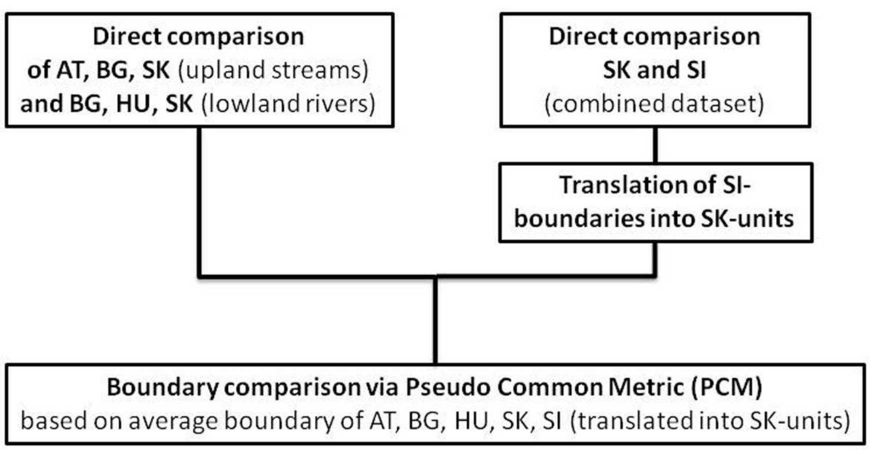

Figure 7.1: Scheme of the "satellite intercalibration" to translate the Slovene class boundaries into units of the Slovak assessment method

[3] Birk S, Willby NJ, Nemitz D (2011) User's Manual of the Intercalibration Spreadsheets. January 2011. Essen/Stirling, 16 pp.

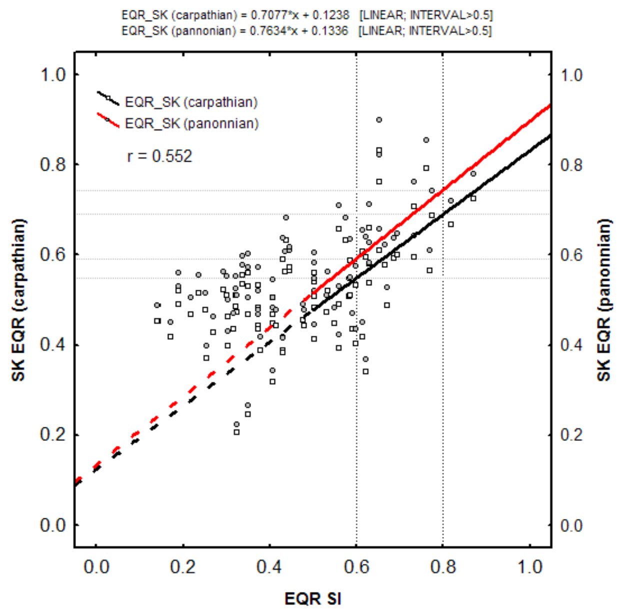

Figure 7.2: Linear regression of the Slovene assessment method against the Carpathian and Pannonian assessment scores of the Slovak method. Regression equations are based on the EQR SI interval >0.5 that best covers the common response pattern of the two methods to degradation.

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

not applicable

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

Describe the IC Common metric:

see above

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

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

Member State	$R_{Pearson}$	p
Lowland rivers (R-E2 and R-E3)		
Bulgaria	0.706	<0.01
Hungary	0.670	<0.01
Slovakia	0.586	<0.01
Slovenia	0.552[a]	<0.01
Upland streams (R-E4)		
Austria	0.845	<0.01
Bulgaria	0.791	<0.01
Slovakia	0.870	<0.01
Slovenia	0.552[a]	<0.01

[a] obtained from “satellite intercalibration” of the Slovenian and Slovak methods

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

No method had to be excluded due to application of "satellite intercalibration" – see explanations above

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

Clarify if

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

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

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

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

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

not applicable

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

not applicable

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

not applicable

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

Lowland rivers (R-E2 and R-E3)

At high status the macrophyte communities of the Eastern Continental lowland rivers feature a high diversity of growth forms, dominated by myriophyllids (e.g. *Myriophyllum spicatum*), small amphibious taxa (e.g. *Myosotis scorpioides* or *Glyceria fluitans*) and, to a lesser extent, bryophytes (e.g. *Fontinalis antipyretica*) (Figure 8.1). The share of these groups constantly declines towards moderate status, with the good-moderate boundary position characterised by the disappearance of myriophyllids and bryophytes. In contrast, large emergents such as *Sparganium erectum* or *Phragmites australis* increase in abundance from high throughout the good status, but also decline when conditions are worsening. Parvopotamids (e.g. *Potamogeton crispus*), elodeids (e.g. *Ceratophyllum demersum*) and filamentous algae (e.g. *Cladophora* sp.) show increasing abundance over the full condition gradient. Similar patterns are revealed by the dominance structure of macrophyte species occurring at sites that are classified at least in good status (CGS) by all national assessment methods (Table 8.1).

Average concentrations of selected nutrients at CGS-sites show least-disturbed conditions with regard to water quality. A simple model of the harmonised biological condition gradient against the gradient of combined pressure allows for estimating the level of anthropogenic pressure corresponding to the harmonised high-good and good-moderate boundary positions (Figure 8.2).

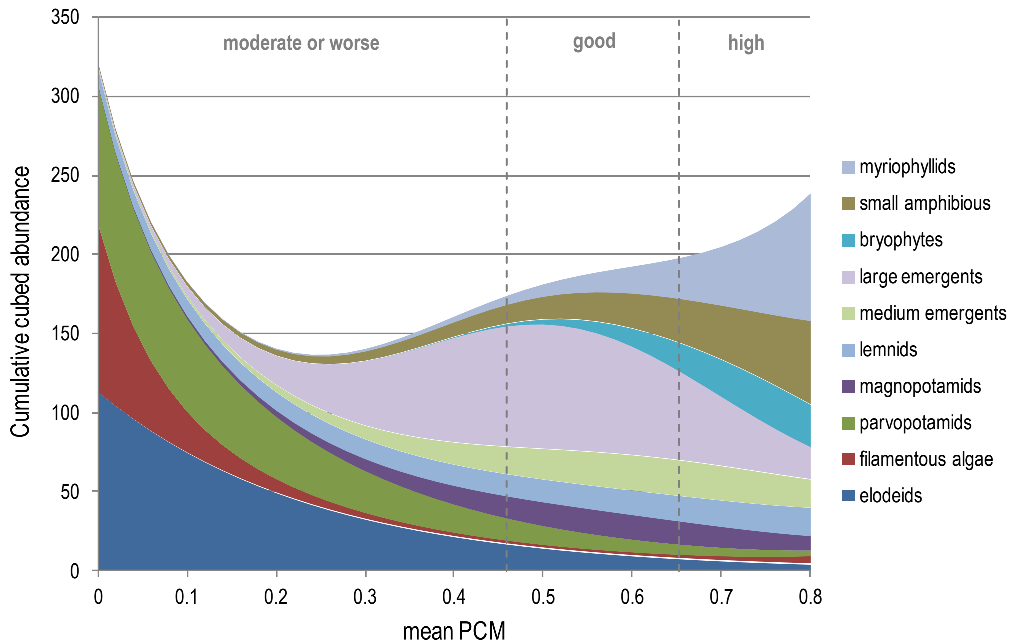

Figure 8.1: Graphical representation of modelled changes in major aquatic plant growth forms across the harmonised biological condition gradient (mean PCM) in lowland rivers of Eastern Europe
Vertical lines indicate harmonised position of high-good and good-moderate boundaries. Growth forms are stacked from the base in order of decreasing contribution to the total vegetation recorded in that river type.

Table 8.1: Dominant macrophyte species of the ECrivGIG lowland rivers
The table lists number of occurrences (#) and relative cubed abundance (%) of taxa recorded by ten surveys that were classified as at least in good status by all national assessment methods (CGS – "common good status sites"). The table only includes taxa with #>2 and %>2.
Average (median) concentrations of selected chemical parameters at CGS were: NH_4 = 0.060 mg N/l, NO_3 = 1.72 mg N/l, TP = 0.097 mg/l, PO_4 = 0.073 mg P/l, BOD_5 = 1.8.

Taxon	#	%
Sparganium erectum	5	13.3
Myosotis scorpioides	7	7.1
Scirpus lacustris	4	6.3
Lemna minor	3	6.2
Phragmites australis	4	5.7
Berula erecta	4	5.4
Butomus umbellatus	5	5.0
Veronica anagallis-aquatica	3	4.8
Lythrum salicaria	8	2.8
Veronica beccabunga	3	2.3

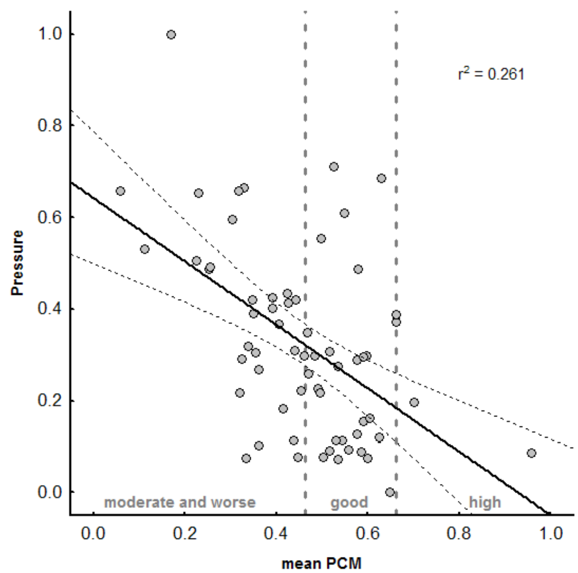

Figure 8.2: Values of the average PCM related to the combined pressure factor for 63 surveys at lowland rivers.

The harmonised high-good boundary (*dashed grey line to the right*) corresponds to 41% natural land use in the catchment, no influence of flood protection dykes at survey site and TP-concentrations of 70 µg/l.

The harmonised good-moderate boundary (*dashed grey line to the left*) corresponds to 24% natural land use in the catchment, slight influence of flood protection dykes at survey site and TP-concentrations of 147 µg/l.

Upland streams (R-E4)

Undisturbed conditions at Eastern Continental upland streams are characterised by the occurrence of bryophytes such as *Rhynchostegium riparioides* and *Fontinalis antipyretica*, indicating coarse substrates, higher current velocities and shaded river banks (Figure 8.3). Along with the constant presence of mosses small amphibious taxa (e.g. *Equisetum palustre*, *Rorippa amphibia*, *Veronica* sp.) increase in abundance towards good status. These two groups are most dominant at good status, while large emergent taxa, mainly *Phragmites australis*, start to contribute to the overall abundance as quality declines. At moderate and worse status myriophyllids (e.g. *Myriophyllum verticillatum*), parvopotamids (e.g. *Potamogeton crispus*) and magnopotamids (e.g. *Potamogeton nodosus*) represent significant shares of the macrophyte assemblage.

Like for the lowland rivers the data from CGS-sites generally confirm the growth-form model described above (Table 8.2). The taxalist highlights the specific ecotonal character of this upland stream type that links the Alpine/Carpathian mountains and the Pannonian lowlands. The water quality at CGS reflects good conditions (see Table 8.2).

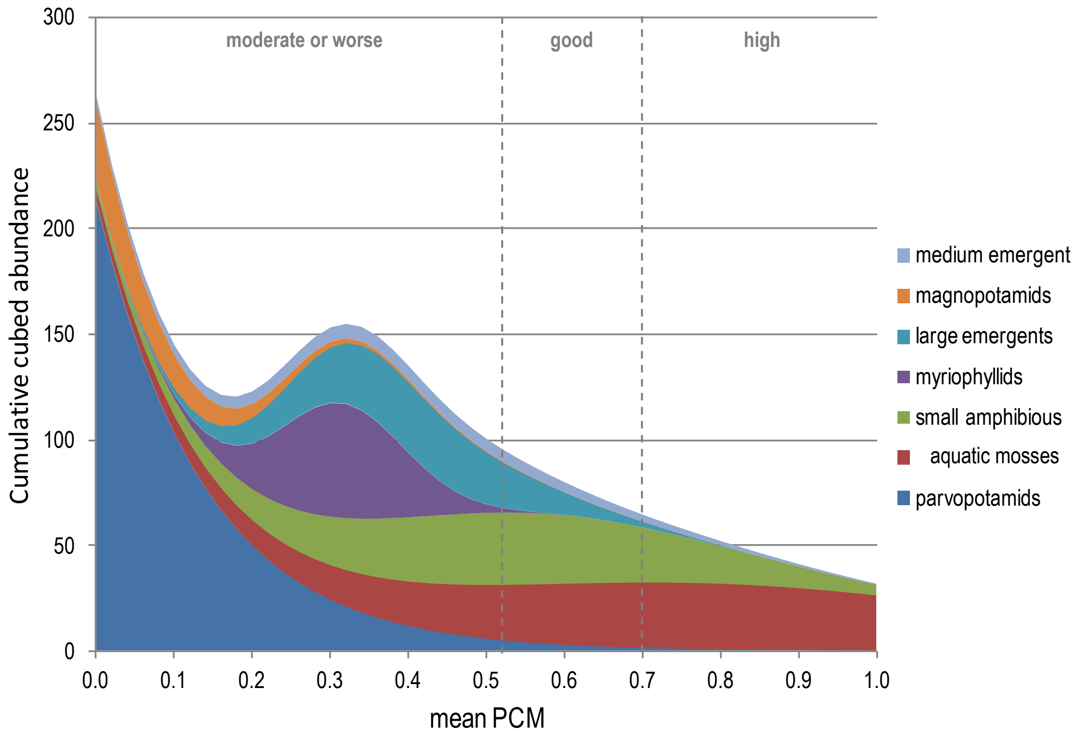

Figure 8.3: Graphical representation of modelled changes in major aquatic plant growth forms across the harmonised biological condition gradient (mean PCM) in upland streams of Eastern Europe

Vertical lines indicate harmonised position of high-good and good-moderate boundaries. Growth forms are stacked from the base in order of decreasing contribution to the total vegetation recorded in that river type.

Table 8.2: Dominant macrophyte species of the ECrivGIG upland streams

The table lists number of occurrences (#) and relative cubed abundance (%) of taxa recorded by ten surveys that were classified as at least in good status by all national assessment methods. The table only includes taxa with #>1 and %>2.

Average (median) concentrations of selected chemical parameters at CGS were: NH_4 = 0.023 mg N/l, NO_3 = 1.86 mg N/l, TP = 0.074 mg/l, PO_4 = 0.020 mg P/l, BOD_5 = 1.5.

Taxon	#	%
Fontinalis antipyretica	5	20.9
Phragmites australis	2	12.4
Zannichellia palustris	2	4.8
Rhynchostegium riparioides	2	3.8
Ranunculus trichophyllus	2	3.8
Veronica beccabunga	3	2.3
Polygonum hydropiper	2	2.2
Equisetum palustre	2	2.2

8.3. Boundary comparison and harmonization

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

Translating national boundaries

We took the model formula for each regression and determined the PCM value that equated to the upper class boundaries for each national method. Therefore, for method A, if $y = mx + c$ where y = the PCM value, m = the regression slope, x = the EQR value of method A and c = the regression intercept, we derived the value on the PCM scale for values of x corresponding to the High-Good and Good-Moderate class boundaries. This was repeated for each national method's regression. All regression characteristics were checked, i.e. the relationship had to be significant ($p \leq 0.05$), the correlation between each method and the PCM should be $r \geq 0.5$ (Pearson's correlation coefficient), the slope of the regression was tested to be significantly different from 0.

Defining a harmonisation guideline

By fitting national class boundaries to each of the national EQR versus PCM relationships we established the predicted values on a PCM scale for each upper class boundary. This yielded a mid-point represented by the global average of the predicted values. This mid-point was considered to represent the 'harmonisation guideline'. The more closely national class boundaries approached this guideline the greater the resulting level of harmonisation of their classifications and the lower the level of bias between methods. This principal applied irrespective of the error associated with the projection of each national class boundary onto a common scale.

Evaluating the level of boundary bias

We defined the difference between the harmonisation guideline and the national boundary values on the PCM scale and converted these differences into class equivalents. First, the widths of national classes high, good and moderate were determined by subtracting the lower from the upper boundary value of each corresponding class translated into PCM units. The boundary bias was then calculated by subtracting the boundary position on the PCM scale from the harmonisation guideline. Finally, we related this difference to the width of the respective national class intersected by the harmonisation guideline (Figure 8.4, Table 8.3).

Harmonising national class boundaries

National boundaries exceeding a bias of 0.25 class equivalents were adjusted to fall inside this permitted level of deviation. We raised or lowered the original boundary EQRs and observed the effects on the boundary bias. Thus, each adjusted boundary underwent the full analytical procedure as described above (benchmark standardisation, translation into PCM, evaluation of boundary bias) but was not allowed to change the position of the harmonisation guideline.

Analysis of class agreement

To check the class agreement after the boundary harmonization we calculated the unsigned (i.e. absolute) differences in classification between each method and every other method for all commonly assessed sites and defined the average absolute class difference (Table 8.4).

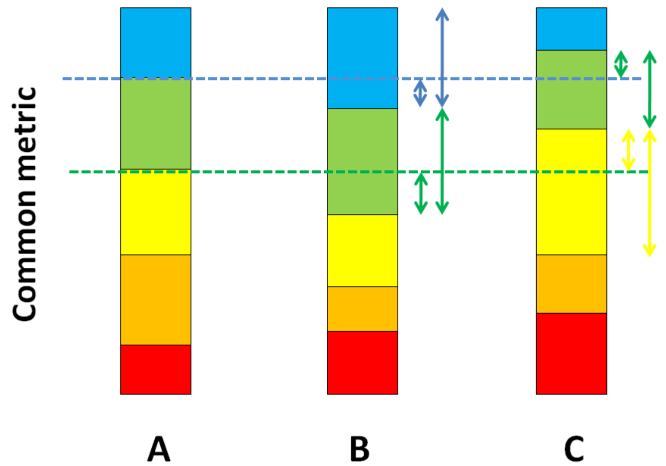

Figure 8.4: National classification schemes (A, B, C) intersected by the average boundary positions derived in intercalibration ("harmonisation guideline" – broken blue and green horizontal lines). Small arrows depict the distance of relevant national boundary to the harmonisation guideline. Large arrows define relevant national class width. The relation of small to large arrow-lengths specifies the boundary bias in class equivalents.

Table: 8.3: National class boundaries and boundary bias (adjusted boundaries, if bias >|0.25|)
Proposed adjustments: ↑ boundary to be raised, ↓ boundary can be lowered

IC type	Country		Original		Adjusted			
			HG	GM	HG		GM	
Lowland rivers (R-E2 and R-E3)	Bulgaria	Boundary	0.570	0.370				
		Bias	-0.14	-0.20				
	Hungary[a]	Boundary	n.a.	n.a.	0.700		0.370	
		Bias	-	-	0.03		0.24	
	Slovakia	Boundary	0.800	0.600				
		Bias	0.25	0.10				
	Slovenia	Boundary	0.800	0.600				
		Bias	-0.02	0.07				
Upland streams (R-E4)	Austria	Boundary	0.875[b]	0.625	0.845	↓		
		Bias	0.33	-0.09	0.23			
	Bulgaria	Boundary	0.430	0.270	0.510	↑		
		Bias	-0.35	0.02	-0.25			
	Slovakia	Boundary	0.800	0.600	0.749	↓		
		Bias	0.27	0.20	0.25			
	Slovenia	Boundary	0.800	0.600				
		Bias	-0.16	-0.08				

[a] Hungary adopts the class boundaries resulting from the intercalibration exercise.
[b] Austria does not intend to adjust the boundary to the recommended common boundary.

Table 8.4: Average absolute class difference (AACD) of national methods after boundary harmonisation, including the number of pairwise comparisons (count)

IC type	Country	Count	AACD
Lowland rivers (R-E2 and R-E3)	Bulgaria	333	0.33
	Hungary		0.37
	Slovakia		0.27
	Slovenia		0.27
Upland streams (R-E4)	Austria	90	0.34
	Bulgaria		0.39
	Slovakia		0.32
	Slovenia		0.34

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

The intercalibrated boundary values for the Hungarian method assessing lowland streams are specified in Table 8.3.

For the upland streams some national methods do not comply with the comparability criteria. Boundary bias is exceeded by the methods of

- Austria – High-good boundary too stringent Π[4]

- Bulgaria – High-good boundary too relaxed #[5]

- Slovakia – High-good boundary too stringent Π

The required boundary adjustments are specified in Table 8.1. The average absolute class difference after boundary harmonisation meets the comparability criteria for all national methods.

 # Bulgaria is requested to implement the adjusted boundaries to allow for completing the intercalibration exercise.

Π Austria and Slovakia are not obliged to lower the boundaries that have been identified as being too stringent.

[4] Austria does not intend to adjust the boundary to the recommended common boundary because national experts are confident that the national high-good boundary is not too stringent within the national system. Lowering the boundary would result in a high status class that includes sites that deviate substantially from the reference condition.

[5] Bulgaria will implement the adjusted boundary value into its classification of upland streams.

Annex: National boundary setting protocols

Austria

At high status the reference species and high quality indicating species (species adapted to chemical/physical and hydromorphological conditions at reference status) clearly dominate the vegetation. Only few indifferent species are present, disturbance indicating species are more or less absent. The high-good boundary demarks an equal ratio of reference species (together with high quality indicating species) and indifferent species and/or species adapted to conditions differing slightly from chemical/physical and hydromorphological situation at reference status. The abundance of disturbance indicating species is very low.

At good status reference species and high quality indicating species are still present, but they are not dominant. Indifferent species and/or species adapted to conditions differing slightly from chemical/physical and hydromorphological situation at reference status are present. However, disturbance indicating species are still rare. The good-moderate boundary is defined as the an equal ratio of reference species and disturbance indicating species with lots of indifferent species present.

At moderate status still lots of indifferent species are present, but species adapted to conditions differing moderately from chemical/physical and hydromorphological situation at reference status are dominant. The moderate-poor boundary demarks an equal ratio of species adapted to conditions differing moderately from chemical/physical and hydromorphological situation at reference status and disturbance indicating species.

At poor status disturbance indicating species dominate. Reference species and/or high quality indicating species are more or less absent. The poor-bad boundary is defined as the disappearance of macrophyte vegetation.

Hence, at bad status the macrophyte vegetation is lacking due to anthropogenic pressure.

Bulgaria

Upland streams (R-E4)

The high-good boundary is identified as the point at which sensitive species (Group A) are in quantity approximately more than 50% of the quantity of all groups. Because of the lack of such sites in the dataset, the average value of sites, where Group B was dominant and/or Group A and C were with equal quantities, was applied.

The good-moderate boundary is the minimum value at which Group B is dominant and/or the sensitive species (Group A) and taxa representing deviation of reference conditions are in equal quantities.

The moderate-poor boundary is set as the average where the community is dominated by tolerant species (Group C).

The poor-bad boundary is defined as the point at which macrophyte species are extinct due to anthropogenic pressures.

Lowland rivers (R-E2 and R-E3)

The high-good boundary is identified as the point at which sensitive species (Group A) are in quantity approximately more than 50% of the quantity of all groups. Because of the lack of sufficient number of such sites in the dataset the maximum value of sites, where Group B was dominant and/or Group A and C were with equal quantities was applied.

The good-moderate boundary is the average point at which Group B was dominant and/or the sensitive species (Group A) and taxa representing deviation of reference conditions are in equal quantities.

The moderate-poor boundary is set as the average where the community is dominated by tolerant species (Group C).

The poor-bad boundary is defined as the point at which macrophyte species are extinct due to anthropogenic pressures.

Slovakia

Due to the absence of existing near-natural monitoring sites reference values for the macrophyte assessment were derived based on expert judgement and comparisons with other BQE assessments. In addition, macrophyte species lists were generated resembling resembling undisturbed biological communities. These different approaches allowed for the definition of reference values.

The good-moderate boundary was defined using linear regression models of the IBMR and the assessment scores of others BQEs. An EQR of 0.6 was allocated to this boundary value. The remaining class boundaries were identified by equidistant division of the biological condition gradient using the EQR values 1.0, 0.8, 0.4 and 0.2. Both techniques to derive reference values (expert judgement and extrapolation from good-moderate boundary setting) showed similar outcomes, confirming the selected benchmark for the macrophyte assessment.

Slovenia

Boundary values between five ecological status classes were defined based on the changes in ratio between the number of »good« (ecological groups A and AB) and »bad« (ecological groups C and BC) taxa (Figure A1) using following criteria:

Boundary	**RMI**	**Boundary setting**
High/Good	0.86	»Good« ≈ »Bad«
Good /Moderate	0.68	»Good« < »Bad«
Moderate/Poor	0.48	»Good« << »Bad«
Poor /Bad	0.31	»Good« = 0

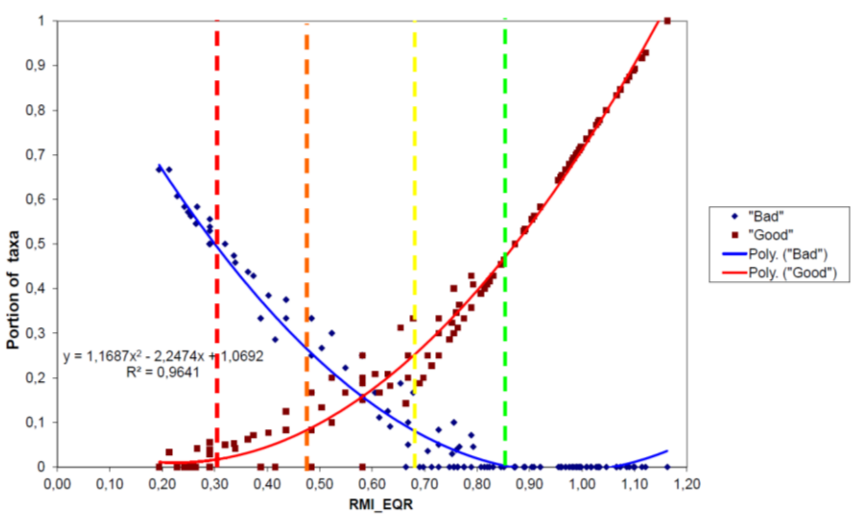

Figure A1: Boundary setting between ecological status classes using changes in portion of »good« and »bad« macrophyte taxa along the River Macrophyte Index (RMI_EQR).