Annex 2.1.2 Class boundary setting procedure for national methods

Information provided from the following MS:
Austria
Belgium Flanders
Belgium Wallonia
Czech Republic
Denmark
Estonia
France
Germany
Ireland
Italy
Lithuania
Luxembourg
Netherlands
Poland
Spain
Sweden
UK

Country	AT
Classification System:	Austrian Quality Assessment System
General Description	Selection of reference sites according to REFCOND Guidance, National Strategy paper ("Criteria for the identification of potential reference sites") and criteria used by AQEM/STAR. The Austrian classification scheme consists of three modules (figure 1): 1. Module "Organic Pollution" (Saprobic Index in relation to stream type specific reference value) 2. Module "General Degradation" consisting of two sub-modules (2 multimetric indices) 3. Module "acidification" index (Braukmann & Biss, 2004; applied only in bioregions at risk of acidification) Metrics used for the multimetric indices are standardised in relation to the 95th percentile of metric values under stream type specific reference conditions. These standardized values are termed as "scores". Indices are calculated by averaging these scores. The benchmark value between reference (High) and good status conditions is defined as the 25th percentile of index values under reference conditions and set to a value of 0.8. That means, observed index values are divided by the benchmark value and multiplied by 0.8. Values > 1 are set to 1. Class boundaries for the ecological quality classes are defined as follows: Quality Class 1: ≥ 0.8 Quality Class 2: ≥ 0.6 < 0.8 Quality Class 3: ≥ 0.4 < 0.6 Quality Class 2: ≥ 0.2 < 0.4 Quality Class 2: < 0.2 The Final Ecological Quality Class is determined by the worst case applying all relevant modules.

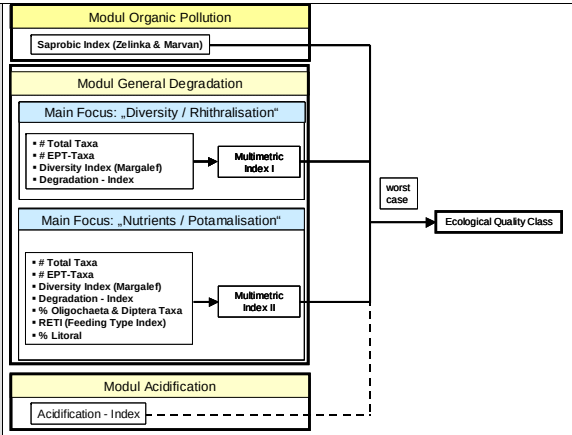


Figure 1: Scheme for the evaluation of ecological quality classes

Criteria for Boundary Setting	**High/Good boundary**	**Good/Moderate boundary**
Taxonomic composition and abundance	25th percentile of reference sites taxonomic composition represented in indices by: # taxa, % Oligochaeta and Diptera taxa, # EPT	25% deviation (of indices) from reference conditions comment: Major taxonomic groups (defined at the level of order - cannot be used for setting good/moderate

	Abundance included in Saprobic Index (# Individuals/m²) and RETI	boundary – see Appendix!): no groups missing
Ratio of disturbance sensitive to insensitive taxa	25th percentile of reference sites sensitive to insensitive represented in MMI by: # EPT, % Oligochaeta and Diptera taxa, RETI, % litoral, degradation index, acidification index	25% deviation (of indices) from reference sites comment: crossover points sensitive/insensitive taxa were not used for setting good/moderate boundary (depending too much on which taxa are selected as sensitive/insensitive)
Level of diversity	25th percentile of reference sites diversity is represented in indices by: Margalef diversity index, # taxa	25% deviation (of indices) from reference sites

Appendix (AT): Figures 2-4: Missing major taxonomic groups: number of orders vs. national MMI's

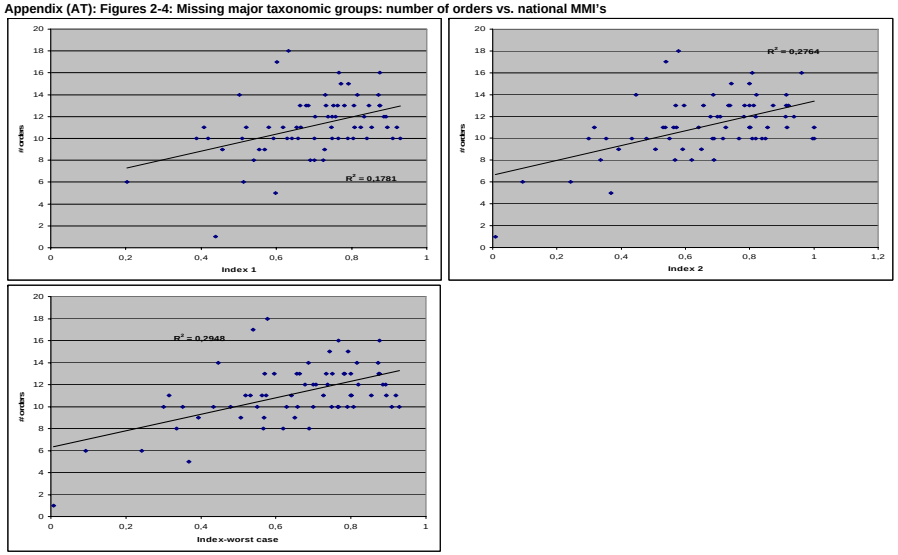
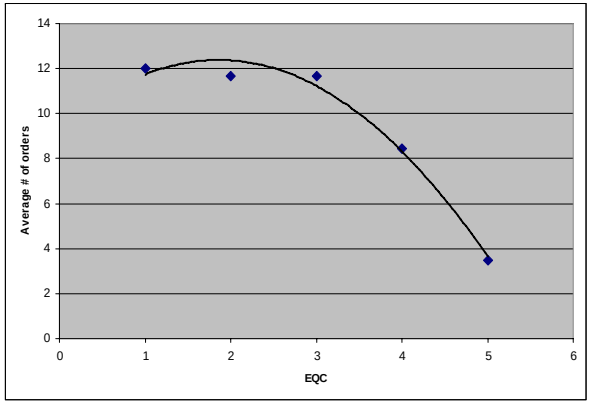


Figure 5: Average number of orders vs. Ecological Quality classes



A significant decline in the number of orders can only be observed for poor and bad status class. The number of orders can therefore not be used for setting good/moderate boundaries.

EPT-Taxa can be seen as the most sensitive taxonomic groups in aquatic ecosystems. There is a strong correlation between the number of EPT-taxa and anthropogenic alteration. The decrease in the number of EPT is used in the national Multi-metric-Indices
A total disappearance of E, P and T-Taxa can only be found below good/moderate class boundary. The disappearance of EPT can therefore not be used for defining good/moderate class boundary.

Figures 6-8: Number of E, P and T –Taxa vs. national MMI (EQC)

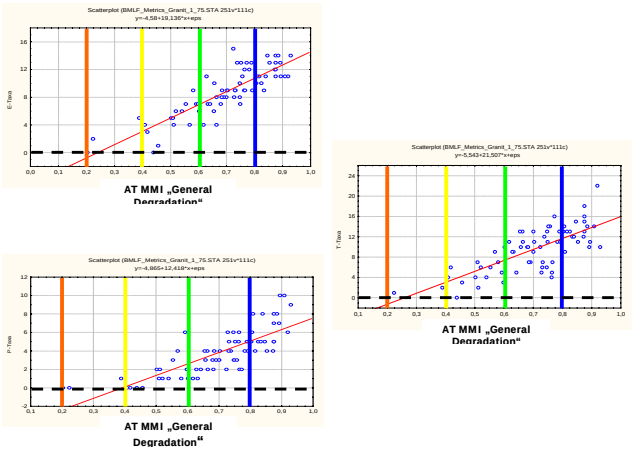
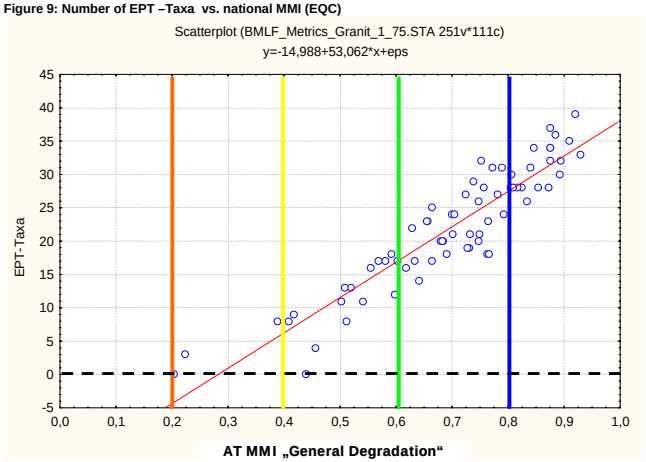


Figure 9: Number of EPT –Taxa vs. national MMI (EQC)



Country	Flanders (Belgium)
Classification System:	Multimetric Macroinvertebrate Index Flanders (MMIF)
General Description	The proposed national macroinvertebrate index for Flanders (Belgium), the MMIF, is a multimetric index for macroinvertebrates in rivers and lakes in Flanders. It was developed based on experience with the Belgian Biotic Index (BBI) method. Both methods provide a similar indication of general degradation but the MMIF tackles a number of problems regarding WFD-compliancy of the BBI. The identification levels used are the same as those used for the BBI, which means genus or family level for all taxa except Chironomidae (which are divided in two subgroups) and watermites (which are treated as a single taxon). The metrics used are: 1. Total number of taxa; 2. Number of EPT taxa; 3. Number of other (non-EPT) sensitive taxa; 4. Shannon-Wiener index; 5. Mean tolerance score (the mean of the tolerance scores of all taxa encountered) (similar to ASPT). For calculating metrics 3 and 5, a list of tolerance scores for all MMIF-taxa was developed, ranging from 1 to 10, 10 corresponding to the most sensitive taxa. The taxa included in metric 3 are all taxa with a tolerance score of at least 6, except the EPT taxa. The obtained values of these 5 metrics are attributed to a score ranging from 0-4. For this purpose, a scoring system was developed by setting a target value for all 5 metrics by expert judgement (with contributions from a panel of Belgian and Dutch macroinvertebrate experts). The scoring system consists of treshold values for assigning the scores 0, 1, 2, 3 and 4, 4 being assigned to the metric values that are nearest to the target value. The metric target values (and hence the scoring systems) are type-specific, which means a separate set of target values was developed for each Flemish river and lake type (8 and 4 types, respectively). The MMIF is then calculated as the sum of the 5 scores divided by 20, hence a score ranging from 0 to 1 (with intervals of 0.05). This means that the maximum MMIF value of 1.00 can only be attained when all metrics are close to their respective target value. For this reason, the maximum MMIF value is considered as a "surrogate" for the reference value and the MMIF score is considered as an EQR scale.

	The quality class boundary values were developed by equally dividing the total range of index values. 0.80 – 1.00: high quality; 0.60 – 0.75: good quality; 0.40 – 0.55: moderate quality; 0.20 – 0.35: poor quality; 0.00 – 0.15: bad quality. Because the EQR is a combination of five metrics, each related to one or more of the boundary setting criteria discussed in the next section, the high/good and good/moderate boundaries are not explicitly related to values of corresponding metrics. Rather, they are based on treshold values (0,8 and 0,6, respectively) of the EQR (MMIF) value based on general degradation which is assumed to correlate with all metrics (and hence the boundary setting criteria).	
Criteria for Boundary Setting	**High/Good boundary**	**Good/Moderate boundary**
Taxonomic composition and abundance	This criterium is related to the following metrics: -Total number of taxa; -Number of EPT taxa; -Number of other (non-EPT) sensitive taxa; -Shannon-Wiener index. The high/good boundary is set at an EQR value of 0,8. For further explanation, see previous section.	This criterium is related to the following metrics: -Total number of taxa; -Number of EPT taxa; -Number of other (non-EPT) sensitive taxa; -Shannon-Wiener index. The good/moderate boundary is set at an EQR value of 0,6. For further explanation, see previous section.
Ratio of disturbance sensitive to insensitive taxa	This criterium is related to the following metrics: -Number of EPT taxa; -Number of other (non-EPT) sensitive taxa; -Mean tolerance score. The high/good boundary is set at an EQR value of 0,8. For further explanation, see previous section.	This criterium is related to the following metrics: -Number of EPT taxa; -Number of other (non-EPT) sensitive taxa; -Mean tolerance score. The good/moderate boundary is set at an EQR value of 0,6. For further explanation, see previous section.
Level of diversity	This criterium is related to the following metrics: -Total number of taxa; -Number of EPT taxa; -Number of other (non-EPT) sensitive taxa; -Shannon-Wiener index. The high/good boundary is set at an EQR value of 0,8. For further explanation, see previous section.	This criterium is related to the following metrics: -Total number of taxa; -Number of EPT taxa; -Number of other (non-EPT) sensitive taxa; -Shannon-Wiener index. The good/moderate boundary is set at an EQR value of 0,6. For further explanation, see previous section.

Country	*Belgium Wallonia*	
Classification System:	Indice Biologique Global Normalisé (IBGN)	
General Description	The IBGN is a semi-quantitative method. The IBGN score is obtained by crossing two metrics: (1) the Faunistic Indicator Group (GFI) whose values range from 1 to 9 according to an increasing gradient of pollution sensitivity from group 1 to group 9; and (2) the number of families collected (Taxonomic Variety, VT). The IBGN is not completely WFD compliant since, among the four metrics recommended by the DCE (taxonomic composition, abundance, proportion of sensitive taxa and diversity), absolute abundances are not included in the calculation of these metrics. An agreed boundary setting protocol to select a pool of reference sites being not available at the moment, in a first step, the median of the "high" and "upper good" status sites defined according to the existing national boundaries was used in order to set up potential reference conditions (and other quality classes). The methodology used is described in Vanden Bossche (2004). Briefly, (1) the "high status" lower limit for a given metric was defined as the median value of the metric in all sites (related to the river-type group) reported in the "upper good" and "high status" selection; (2) the "reference condition value" was defined as the median of the metric in all "high status" sites; (3) the "good status" lower limit was defined as the high status lower limit multiplied by 0.75. After publication, the sites of high status have been checked against the criteria proposed by the GIG coordinators. This checking was successfull.	
Criteria for Boundary Setting	**High/Good boundary**	**Good/Moderate boundary**
Taxonomic composition and abundance	The IBGN score was used as indicator of the element "taxonomic composition and abundance". Calculated threshold values for high/good boundaries are 17 for R-C3 (Tab. 1).	Calculated threshold values for good/moderate boundaries are 13 for R-C3 (Tab. 1).
Ratio of disturbance sensitive to insensitive taxa	The GFI (Faunistic Indicator Group) was used as indicator of the element "ratio of sensitive to insensitive taxa".	Calculated threshold values for good/moderate boundaries are 5 for R-C3 (Tab. 1).

	Calculated threshold values for high/good boundaries are 8 for R-C3 (Tab. 1).	
Level of diversity	The VT (Taxonomic Variety) was used as indicator of the element "level of diversity". Calculated threshold values for high/good boundaries are 10 for R-C3 (Tab. 1).	Calculated threshold values for good/moderate boundaries are 7 for R-C3 (Tab. 1).

IC Types	Description	Reference values			High status			Good status		
R-C3		**IBGN**	**GFI**	**VT**	**IBGN**	**GFI**	**VT**	**IBGN**	**GFI**	**VT**
	Small mid-altitude siliceous	>= 18	>= 8	>=10	>= 17	>= 8	>= 10	>= 13	>= 5	>= 7
	Small mid-altitude siliceous (some very oligotrophic brooks)	>= 18	>= 8	>=10	>= 17	>= 8	>= 10	>= 13	>= 8	>= 6

Tab. 1: threshold IBGN, GFI and VT reference and ecological status values for R-C3 rivers in Belgium Wallonia GFI : Faunistic Indicator Group; VT : Taxonomic Variety.

Reference

Vanden Bossche, J. P. 2004. High status definition and intercalibration pilot exercise in Wallonia (Belgium) for R-C3 type rivers (Invertebrate benthic fauna). Central and Baltic Rivers Geographical Intercalibration Groupe. Report. November, 19th 2004. Centre de Recherche de la Nature, des Forêts et du Bois. DGRNE. Ministère de la region wallonne. B-5030 Gembloux (Belgium). 8pp.

Country	*Czech Republic*
Classification System:	PERLA (general quality assessment system)
General Description	The model generally follows the published mathematical principles of RIVPACS and represents the site specific and stressor non-specific approaches. Calculations are based on full taxa lists, animals are determined mostly to species level. Software HOBENT is used for calculations. The software allows the prediction of the target assemblage of benthic macroinvertebrates for any site based on a set of environmental variables (latitude, longitude, distance from source, altitude, slope, catchment area, and stream order) which characterise the site. Then the predicted assemblage can be compared with the fauna observed at the same site. The comparison makes it possible to evaluate the extent of disturbance, expressed by index B. For the categorization of an observed site into groups of the basic database, a discriminant analysis (Klecka, 1980; Deichsel & Trampisch, 1985) is applied. The SPSS package is used for the computation of discriminant equations and covariance matrices. Discriminant scores, Mahalanobis distance, and prior probability are used for the computation of probabilities that the checked site belongs to of the groups of the basic database. For the final computation, the same formulas as in SPSS are used (Anonymus, 1997). For every species of the reference database, the probability of capture at the observed site is computed according to the formula (Clarke et al., 1996): $Cs = \sum_{g=1}^{G} Fsg * Pg$ s = species, Cs = species probability captured at the observed site, g = group, G = number of groups, Fsg = frequency of occurrence of species s in group g, Pg = probability with which the observed site belongs to group g. All species are ordered according their Cs and the number of species expected at the observed site is computed as: $NE = \sum_{s=(Cs \geq CsL)}^{S} Cs$ NE = number of species expected at the observed site,

	S = number of species in the reference database, CsL = optional low limit of Cs (obviously 0.5). Finally, index B is computed as: B = NO / NE NO = number of species with Cs ≥ CsL found at the observed site. The low limit of Cs is an essential number. The B index computed using the limit is a type of similarity index. When the CsL is set to zero, only the simple number of taxa is compared and the result is not very useful. Besides index B and the basic ecological indices, the ASPT, BMWP, saprobic index, EPT, and other indices were incorporated into the software. Their expected values can be predicted. It makes it possible to express these metrics in the form of EQR.	
Criteria for Boundary Setting	**High/Good boundary**	**Good/Moderate boundary**
General	The national class boundary setting follows Option B of the REFCOND Guidance. The scale of B-index values was established by expert judgement as equal intervals from high to bad ecological status. Thus the boundary B – index values are: - 0,8 high/good - 0,6 good/moderate - 0,4 moderate/poor - 0,2 poor/bad The boundary values were checked on tested data set:	
Taxonomic composition and abundance	The median value of number of families in tested data set was 25, which corresponds to the B – index value of high – good boundary (0,8). (Fig. 1)	The lower quartile value of number of families in tested data set was 22, which corresponds to the B – index value of good - moderate boundary (0,6). (Fig. 1)
Ratio of disturbance sensitive to insensitive taxa	not assessed	not assessed
Level of diversity	The median value of Margalef index in tested data set was 2,8, which corresponds to the B – index value of high – good boundary (0,8). (Fig. 2)	-

Fig 1: Number of families vs. B – index value.

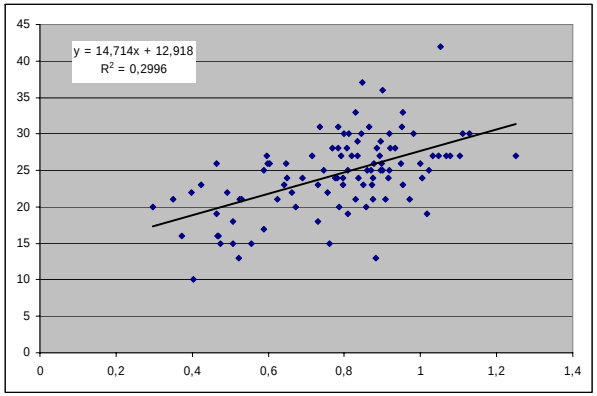
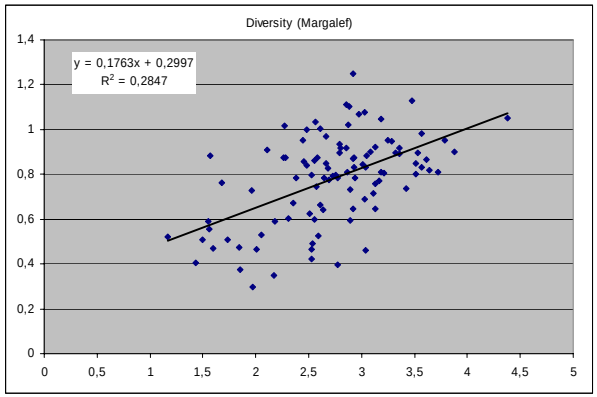


Fig 2: B - index value vs. diversity expressed as a value of Margalef index



Country	Denmark	
Classification System:	Danish Stream fauna Index (DSFI)	
General Description	DSFI is an index based both on sensitivity of macroinvertebrate taxa and diversity. DSFI is also classification system and seven quality classes (fauna-classes) can be calculated. An index value of 1 characterises a strongly impacted macroinvertebrate community. An index value of 7 characterises an unimpacted or only slightly impacted macroinvertebrate community. The value calculated should be compared with the optimum fauna-class for the type of stream investigated as well as fauna-class set as the objective for that given stream by the water authorities. The optimum fauna-class corresponds to the reference situation and will be 7 in the majority (>95%) of Danish streams (but not all streams with fauna-class 7 are necessarily reference streams). DSFI was introduced in 1998 and was not originally developed to fulfil the requirements of the WFD. However, it is almost fully in compliance with WFD, except that abundance is not used for all taxa when calculating the index value. It has been decided to continue with the use of DSFI for stream classification and that calculation will be based on full taxa lists (mostly species level) with abundance for each taxa. DSFI is used in the following way: High quality = DSFI index value 7, Good quality = DSFI index values 6 and 5 Moderate quality = DSFI index value 4 Poor quality = DSFI index value 3 Bad quality = DSFI index values 2 and 1	
Criteria for Boundary Setting	**High/Good boundary**	**Good/Moderate boundary**
Taxonomic composition and abundance	No missing groups and only minor changes in abundance	Typically, most major taxonomic groups (orders) are found. But several families especially in important groups like Ephemeroptera, Plecoptera, Trichoptera and Coleoptera (EPTC) are missing. Abundance of some insensitive taxa could increase.

Ratio of disturbance sensitive to insensitive taxa	No or only a very minor loss in sensitive taxa	The number of species and individuals of sensitive taxa of Plecoptera (genus-level), Ephemeroptera (family-level) and other sensitive groups are highly reduced at the good/moderate boundary. As a consequence the proportion of insensitive taxa becomes higher compared to the reference state.
Level of diversity	No or only a very minor loss in diversity	Loss in species diversity has been estimated for the EPTC families (see above). As a mean only about 45% of the EPTC **species** present at high quality will be found at the good/moderate boundary. At the **family** level (all families) as a mean about 70% of the families can be found at the Good/moderate boundary.

Country	Estonia
Classification System:	
General Description	Selection of reference sites according to REFCOND Guidance and criteria used by AQEM/STAR. The Estonian preliminary classification scheme is based on five indices (according to Johnson 1999, with modifications), indicating: 4. Biodiversity (taxa richness, based on standardized taxa list) 5. General quality, responding to hydromorphological stress alone, or combined with unspecified stress(es) (British Average Score Per Taxon, EPT richness, Shannon diversity) 6. Organic pollution (Danish Stream Fauna Index) 7. In certain cases, some additional metrics (such as acidity index, or natural value) may be taken into consideration The high/good boundary is preliminarily defined as 85%, the good/moderate boundary as 70%, and the moderate/bad boundary as 55% of the reference values of each index. However, it seems that 1) the current values are too low (i.e. there occurs risk to estimate quality higher than it actually is), and 2) different indices must have different boundaries. The adjustment of the boundaries is in process. We have not created universal multimetric sets, in order to better understanding of different aspects of possible stress types. The conclusive quality class is the median value of the single indice quality estimations. In doubtful cases (for example, when differences between single estimations exceed one class), expert opinion is applied. "One out, all out" principle is not practicized.

Country	**France**
Classification System:	**IBGN - Classification française DCE Indice Biologique Global Normalisé**
General Description	Description of the IBGN Index. (Norme AFNOR NF T 90-350, 1992) The index is based on macroinvertebrate fauna identified at the family level. Eight individual samples are taken from different habitats. The index is semi-quantitative, but quantitative data are available in many cases, allowing the comparison with the ICM index. The IBGN is a combination of 2 metrics: - the total number of taxa (at the family level for Insecta, Crustacea, Mollusca, Acheta; class for the other groups), is expressed in 14 classes of taxonomic richness. The Metric CV : Classes de Variété, varies from 14 (> 50 taxa) to 1 (1-3 taxa) . - the Indicator Faunistic Group representing the presence/absence of 39 indicator taxa, grouped in 9 classes of sensitivity to pollution. The Metric GFI : Groupe Faunistique Indicateur, varies from 9 (very sensitive taxa present) to 1 (only very tolerant taxa remaining). The IBGN index is given by the formula : IBGN = (CV + GFI) -1, varying in integer values between 20 (set as maximum value) and 1. The index is sensitive to pollution (including toxic), and to general degradation (including habitat alteration). Principles of the classification. A ***provisional*** classification derived in 2004 was endorsed in 2005 by the French Ministry of Environment (Circulaire MEDD/ DE / MAGE / BEMA 05 / n° 14 / 28 juillet 2005). The same principles are applied to all the river types in France. The classification is based on the level of biological alteration evaluated by the EQR-IBGN values, and compliant with the WFD normative definitions for what concerns a "slight deviation" of taxa richness (CV) and sensitive taxa (GFI), and for the disappearance of major taxonomic groups. **1**: definition of Reference values. Reference sites were selected in the monitoring network on the basis of the national reference criteria, following a procedure compliant with the REFCOND Guidance. A set of reference criteria, officially endorsed by the French Ministry of Environment (Circulaire MEDD/ DE/ DCE 08 du 23 décembre 2004) has been applied at the national level. These criteria are compliant with the CB GIG reference criteria, including chemical values when available. (See separate annex on reference criteria). The statistical distributions of observed biological values (i.e. IBGN index and its two metrics, GFI and CV) were analysed for all the types with a sufficient number of reference sites. For each type, a **reference value** was derived as ***the median of observed values in reference sites*** for the IBGN and its separate metrics. For some types, reference values were tested against an independent dataset provided by the Cemagref. All the IBGN values were then transformed in EQR-IBGN; for this calculation, the minimum IBGN value is set at IBGN = 1.

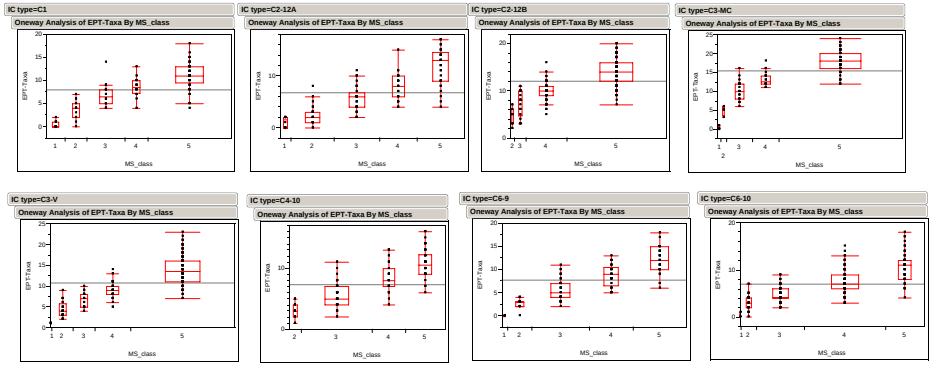
	2: definition of High /Good boundary The H/G boundary is based on the IBGN value corresponding to the combination of the *25th percentile* of the metrics values (CV and GFI) observed in reference sites. **3**: definition of the Good / Moderate boundary For each type, the remaining range below the H/G boundary and the minimum IBGN value was split into 4 equal classes following a procedure suggested in the REFCOND guidance, to derive a preliminary limit; then, following a pressure / impact analysis at the national scale, this limit was adjusted at a higher level, for almost all types, to set the G/M boundary. An official table (circulaire MEDD/ DE / MAGE / BEMA 05 / n° 14) gives the provisional IBGN values, corresponding to the Reference, H/G and G/M boundaries for all types with sufficient reference values. Some of the types corresponding to the large and very large rivers have no reference sites, and thus reference values could not be derived so far. Taking into account that any change in reference values will be reflected in the future classification, the concept of the G/M boundary for France represent a deviation of the EQR-IBGN from the reference values. At the national level, the average EQR-IBGN value for the G/M boundary is 0.80 ; **the range for the different types corresponding to the CB GIG is 0.78 – 0.82, according to the variability of the reference conditions**. **4:** Harmonisation of the High /Good boundary In the Central Baltic GIG intercalibration process, following the comparison of the national classifications and the definition of acceptable range, the average value initially proposed by France for the H/G boundary (corresponding to EQR-IBGN=0.92) appeared at the lower limit of the harmonised range, thus indicating that some individual types could lay outside of the band. As consequence, the H/G boundary of the IBGN classification has been adjusted. In the **harmonised classification**, at the national level, the average EQR-IBGN value for the H/G boundary is 0.94 ; **the range for the different types corresponding to the CB GIG is 0.93 – 0.94, ensuring that all the types have a H/G boundary fitting in the agreed IC range.** Future revision of the classification: A new reference network implemented in 2005 will provide a more consistent reference dataset at the end of 2007. At this date, all the reference values will be recalculated for all the types, and a definitive classification will be established. The future classification will take into account both - the revision of the preliminary reference values according to the data coming from the new reference network; - the deviation from reference conditions (as EQR-ICM) corresponding to the H/G and G/M boundaries agreed during the IC process.	
Criteria for Boundary Setting	**High/Good boundary**	**Good/Moderate boundary**

Taxonomic composition and abundance	**For the number of taxa,** 25th percentile of the values observed in the reference samples, transformed into the corresponding class of the metric CV.	**For the number of taxa,** the range below the H/G boundary and the minimum value (number of taxa =1) is split into 4 equal classes.
Ratio of disturbance sensitive to insensitive taxa	**For the sensitive taxa**, 25th percentile of the values observed in the reference samples, expressed as the corresponding class of the metric GFI.	**For the sensitive taxa**, loss of one class of the metric GFI from H/G boundary (i.e. GFI H/G minus 1).
Level of diversity	The diversity is reflected by the **number of taxa (**no quantitative diversity index).	The diversity is reflected by the **number of taxa (**no quantitative diversity index).
Global Index IBGN	The combination of the values corresponding to H/G boundary for the two metrics CV and GFI. Generally equivalent to the 25th percentile of the IBGN values observed in reference samples. EQR-IBGN at H/G boundary **used for the comparison** : R-C1: 0.87 R-C2: 0.87 & 0.94 R-C3: 0.89 &0.93 R-C4: 0.94 R-C6: 0.94 & 0.93 National average : 0.92 EQR-IBGN at H/G boundary **after harmonisation** : Range 0.93-0.94 for all the IC types National average : **0.94**	The combination of the values corresponding to G/M boundary for the two metrics CV and GFI as described above gave a preliminary limit, compliant with the REFCOND Guidance. Following a pressure / impact analysis, this limit was then adjusted at a higher level for almost all the types to derive the G/M boundary. EQR-IBGN at G/M boundary: R-C1: 0.80 R-C2: 0.80 & 0.81 R-C3: 0.78 &0.80 R-C4: 0.81 R-C6: 0.81 & 0.80 National average : **0.80**

Appendix (FR)- Compliance with the normative definition : missing major taxonomic groups
note : graphs and calculations correspond to the classification before harmonisation of the H/G boundary.

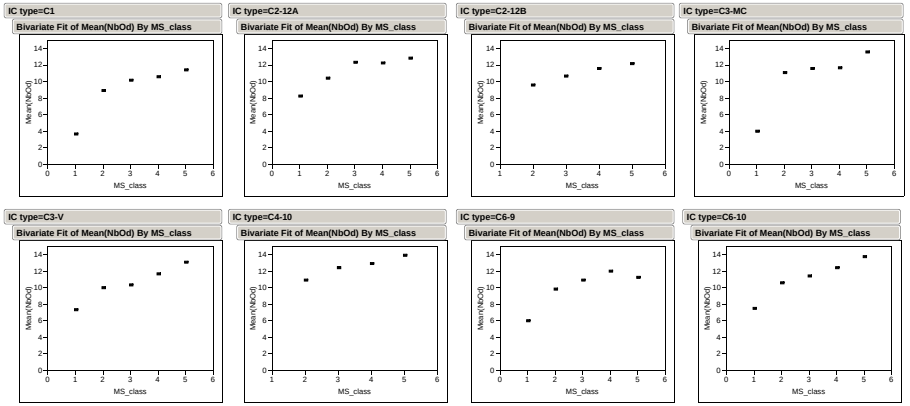
EPT-Taxa can be seen as the most sensitive taxonomic groups in CB rivers types. A positive value of this metric indicates that at least one of the three most sensitive major taxonomic groups is still present in the community.

In the Central datasets, a minimum of 3 EPT taxa is encountered at "good satus" (Figures 1). The disappearance of EPT-Taxa can only be suspected in the classes poor and bad. Therefore the most sensitive major taxonomic groups are always present at good status.



Figures 1: Distribution of the number of EPT taxa per ecological status classes based on the French IBGN. (1 = bad status to 5 = high status)

The number of orders represents the total number of "major taxonomic groups". The mean number of orders per sample observed at different ecological status is represented in figures 2 and table 1 for the different datasets. The difference in the total number of orders between high and good status is generally one order, and no more than two (Figure 2 and table 1). Thus it appears clearly that at good status there is no or only very limited loss of "major taxonomic groups".



Figures 2: Mean numbers of orders per ecological quality classes in the different datasets

Table 1: Mean numbers of orders per ecological quality classes in the different datasets

		EQC				
		Bad	Poor	Moderate	Good	High
FR-Datasets	C1	4	9	10	11	11
	C2-12A	8	10	12	12	13
	C2-12B		10	11	12	12
	C3-MC	4	11	12	12	14
	C3-V	7	10	10	12	13
	C4-10		11	12	13	14
	C6-9	6	10	11	12	11
	C6-10	8	11	11	12	14

Country	DE
Classification System:	German Quality Assessment System for Benthic Macroinvertebrates
General Description	The German classification scheme consists of three modules: 8. Module “Organic Pollution” (Saprobic Index in relation to stream type specific reference value); 9. Module “General Degradation” consisting of type-specific combinations of single metrics into a multimetric index. In general only metrics belonging to the metric types: composition/abundance, richness/diversity, ratio sensitive/insensitive, functional have been selected that show best correlation to land use and hydromorphological gradient (Table 1); 10. Module “Acidification Index” (Braukmann & Biss, 2004; applied only for stream types at risk of acidification). Table 1: List of single metrics included in the Module “General Degradation” per national stream type relevant in the CB GIG intercalibration exercise. The multimetric index is composed of 50% Fauna-Index of the specific stream type and 50% arithmetic mean of all other metrics marked by “x” (metric types: ca – composition and abundance, rd – richness and diversity, s – ratio sensitive to insensitive taxa, f – functional).

single metric	metric type	national type number					
		5	5.1	14	15k	15g	17
EPT [%]	ca	x	x	x	x	x	x
Fauna-Index Typ 14/16	s/rd			x			
Fauna-Index Typ 15g	s/rd					x	
Fauna-Index Typ 15k/17	s/rd				x		x
Fauna-Index Typ 5	s/rd	x	x				
Litoral [%]	ca/f			x	x	x	x
Pelal [%]	ca/f				x		
Rheoindex	ca/f	x	x				
# Trichoptera	ca/rd			x	x	x	x

Definition of type-specific reference values for relevant assessment metrics

The assessment of ecological quality of running waters in Germany using macrozoobenthos is based on type-specific reference conditions expressed in reference values of relevant assessment metrics (see Table 1). These values have individually been derived by the following procedures:

- type-specific “true” reference sites available: calculation of reference values of relevant assessment metrics on the basis of a dataset including “true” type-specific reference sites following REFCOND and AQEM/STAR criteria. Reference value is the 95th percentile of all metric values.
- only sites slightly deviating from type-specific reference state available: correlation of metric values against abiotic structure index or land use index (see Annex 1) and extrapolation of best-fit-straight-line to reference values including expert judgement.

Class boundaries for the individual modules are defined as follows:

A. Module “Organic Pollution”

Table 2: Type specific absolute German Saprobic Index values for reference and class boundaries. Reference values were derived from saprobic reference sites.

national type number	basic saprobic condition	high-good	good-moderate
5	1,35	1,45	2,00
5.1	1,45	1,60	2,10
14	1,65	1,80	2,25
15 and 17	1,75	1,85	2,30
all types	-	deviation of ~5%	deviation of ~25%

B. Module “General Degradation”: EQR class boundaries.

Quality Class 1 (high): ≥ 0.8
Quality Class 2 (good): ≥ 0.6 - < 0.8
Quality Class 3 (moderate): ≥ 0.4 - < 0.6
Quality Class 4 (poor): ≥ 0.2 - < 0.4
Quality Class 5 (bad): < 0.2

The Final Ecological Quality Class is determined by the worst case applying all relevant modules (still under discussion).

Criteria for Boundary Setting	**High/Good boundary**	**Good/Moderate boundary**
General	The national class boundary setting follows Option B of the REFCOND Guidance (Anonymous 2003): Scale of EQR values was established by expert judgement proposing appropriate intervals from high to bad ecological status. Application of scale to real data sets confirmed proposed boundary setting.	
	High status of the Module “General Degradation” corresponds to very low level of anthropogenic impact concerning land use und hydromorphological pressure (see Annex 1).	-
Taxonomic composition and abundance	EPT [%], Litoral [%], Pelal [%], Rheoindex	
Ratio of disturbance sensitive to insensitive taxa	Saprobic Index, Fauna-Index, Rheoindex	
Level of diversity	Fauna-Index, # Trichoptera	

Annex 1

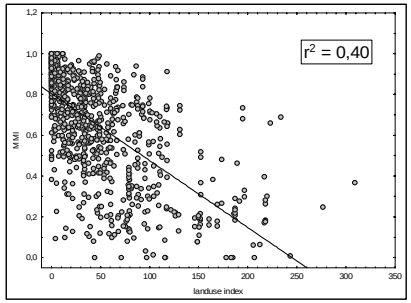


Figure 1: Correlation of land use index with multimetric index for samples belonging to national stream type 5 (=R-C3). The "landuse index" is calculated according to the formula:
landuse index = 4*%urban area + 2*%cropland area + %pasture area.
Sites in high ecological quality generally show low share of anthropogenic land use.

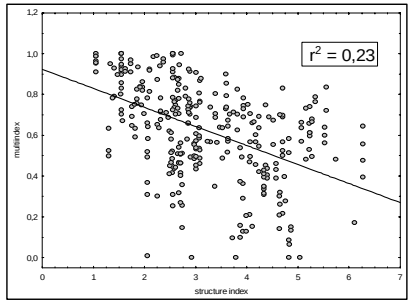


Figure 2: Correlation of structure index with multimetric index for samples belonging to national stream type 5 (=R-C3). The "structure index" is an abiotic metric including relevant hydromorphological stream features for the macroinvertebrate occurrence in the stream.

Country	IRELAND
Classification System:	Quality Rating System (Q-value)
General Description	The scheme of classification is based on macroinvertebrates but also takes macrophytes and phytobenthos, particularly the filamentous alga *Cladophora* sp., into account. This expert judgement system takes account also of time of year (seasonality) and was designed primarily for application in summer/autumn period when biological sampling for monitoring purposes is carried out. The system has 5-macroinvertebrate faunal Groups (A-E) based on established taxa sensitivity. Rules based on presence and abundance of indicator-taxa at family or species level. Q-values (Q1-Q5 with intermediate ratings) are ascribed which can be converted to quality classes and EQR banding as follows:

Q-VALUE	EQR	Quality - Status
Q5	**1**	**High – Reference conditions**
Q4-5	**0.9**	**High**
Q4	**0.8**	**Good – Slight deviation from reference condiitions**
Q3-4	**0.7**	**Moderate – Moderate deviation from reference conditions**
Q3	**0.6**	**Poor – Significant deviation from reference conditions**
Q2-3	**0.5**	**Poor – Significant deviation from reference conditions**
Q2	**0.4**	**Bad – Major deviation from reference conditions**
Q1-2	**0.3**	**Bad – Major deviation from reference conditions**
Q1	**0.2**	**Bad – Severe deviation from reference conditions**

To provide confidence limits for the divisions proposed above, e.g. t-tests comparisons of means could be carried out to show that each category can be separated on the basis of, e.g. BOD, NH_4, PO_4, NO_3 or on the basis of other biological elements, e.g. fish composition. The Q-value also shows a statistical relationship with catchment pressures defined as Corine land cover categories, e.g. urban fabric, arable and pasture.

Criteria for Boundary Setting	High/Good boundary	Good/Moderate boundary
Taxonomic composition and abundance	Loss of some Group A taxa	Group A taxa poorly represented with Groups B and C becoming well established
Ratio of disturbance sensitive to insensitive taxa	Lying between Q4-5 and Q4 - equivalent to EQR of 0.85	Lying between Q4 and Q3-4 - equivalent to EQR of 0.75
Level of diversity	May increase or decrease across the boundary	May increase or decrease across the boundary

Country	**Italy**
Classification System:	**STAR Intercalibration Common Metric Index (STAR_ICMi), type adapted**
General Description	Selection of Reference sites for the IC exercise The selection of reference sites was done according to REFCOND Guidance criteria (see REFCOND criteria table provided) and is based on pressure analysis, which included information on: water chemistry, (hydro)-morphology, general degradation, land use. In CB_GIG, Italy is involved in R-C1 type only, that corresponds to small streams in the Po valley. These streams mainly belong to spring-fed systems and are located in the very lowland areas of Northern Italy. This implies that some chemical compounds (e.g. Nitrates, Phosphates) are usually at higher level than in other geographic contexts. After a detailed and unsuccessful search for sites with lower concentrations, because it is not expected to impact on invertebrate communities, a slightly higher level than fixed for CB_GIG screening of reference sites was accepted for $P\text{-}PO_4$ (average value for reference sites: 0.07 mg/l). All other CB_GIG chemical and pressure criteria are met. In Italy, bio-monitoring based on aquatic invertebrates has a long hystory. The method in use before the Water Framework Directive is the Indice Biotico Esteso (IBE: Ghetti, 1997; APAT-IRSA/CNR, 2004) that only partly satisfies the WFD requirements. New, type-adapted assessment modules are under development to fully comply with the WFD and fulfill the needs of the different WFD monitoring types. Most of them will be ready by the beginning of 2007, to meet the terms of the WFD monitoring plan. Because they are in a refinement stage, i.e. they are being directly related one to each other to guarantee a comparable assessment of ecological quality, the relation to Intercalibration metric(s) (ICMi) is highly beneficial to finally set class boundaries. Thus, for the CB type(s), Italy has decided to adopt the STAR_ICMi, formally in use in the CB_GIG for the IC process, as the official Italian method for setting quality class boundaries. The boundaries for each of these original assessment modules being provided for WFD monitoring – and possibly adapted to monitoring aims, stressor acting, local tradition and expertise – will be derived starting from those set for the STAR_ICMi. Actually, the STAR_ICMi is used by Italy as: a) 'an interim common WFD assessment method for the purposes of the intercalibration exercise' (see the IC Guidance: EC, 2005) and b) the legitimate way to determine class boundaries for any other method more explicitly devoted to standard monitoring for invertebrates. Description of the STAR Intercalibration Common Metric Index (STAR_ICMi) (Buffagni et al., 2005; 2006) The STAR_ICM index was explicitly designed for European IC purposes and it represents one of the indices used in various GIGs for the comparison and harmonization of class boundaries of different MSs. For the CB_GIG, it is presently the only one in use. Its WFD compliancy has been discussed and demonstrated elsewhere (Wasson & Buffagni, 2005). The index was built to assess the overall (i.e. general) degradation of a river site, not being aimed at detecting the impact of single stressors on invertebrates (i.e. it is not a stressor-specific system). The STAR_ICMi is directly calculated in the form of Ecological Quality Ratio (EQR), in accordance with WFD requirements for classification systems. Three aspects of the used methodology to derive class boundaries have to be considered for intercalibration purposes and to

check compliancy with normative definitions:

a) the sampling technique
b) the calculation formula
c) the conversion of STAR_ICMi values into quality judgement (i.e. class boundaries setting).

a) the sampling technique

The data used for the Intercalibration exercise were collected by sampling along 1-2 transects across the river, depending on river type, and collecting invertebrates from all major micro-habitats occurring. A preliminary check of taxa found is done in the field, so that the possible absence of taxa which are expected in unaltered conditions can be verified with integrative sampling. A minimum number or specimens, different for each taxon, must be collected to consider valid the taxon for the computation. When a taxon is accepted, three abundance classes are usually reported for the collected taxa: Present, Abundant and Dominant. Such classes are generally interpreted in terms of relative abundance. Alternatively, an AQEM-like approach, as described in Buffagni et al. (2004) can be adopted. Before sampling, a depositional-transport sequence is identified at each site, which roughly corresponds to what is usually referred to as a pool-riffle sequence.The method for the macroinvertebrate collection was then a 'multi habitat sampling' procedure. Ten individual samples are distributed according to microhabitats occurrence in the riffle unit, taken and merged into a sample. A second merged sample is always obtained, following the same criteria, from the pool area for each site.Two taxa lists can thus be attained for each site, for the depositional and transport units respectively. An open Surber sampler is used to collect macroinvertebrates (area 0.05 m^2; mesh size 0.5 mm). All samples are collected in a quantitative way i.e. all specimens for relevant taxa are picked up and brought to the lab for identification. In some cases for particularly abundant taxa, sub-sampling in the field can be used.

b) the calculation formula

The STAR_ICMi is a multi-metric index and is composed of six metrics, which account for the main aspects present in the WFD Normative definitions (see below): ASPT, Log_{10}(sel_EPTD+1), 1-GOLD, N-taxa, EPT and Shannon-Weiner diversity. The ICMi value is calculated by the sum of all the ICMs, after attributing a weight to each metric. Hereafter, the list and category of each metric is provided (Table 1). After their normalization, the metrics are combined into the ICM index. Metrics are grouped into three groups, providing information on three major response areas: Tolerance, Abundance/Habitat and Richness/Diversity. A different weight is attributed to the metrics within each group, giving greater importance to the metrics based on the whole community (Buffagni et al., 2004). To obtain the final multimetric score, the same weight is attributed to each of the three metric groups (0.333).

Table 1. Intercalibration Common Metrics (ICMs) used in the STAR ICMi

Information type	Metric type	Metric name	Taxa considered in the metric	Literature reference	weight
Tolerance	Index	ASPT	Whole community (Family level)	e.g. Armitage et al., 1983	0.333
Abundance/ Habitat	Abundance	Log_{10} (Sel_EPTD +1)	Log(sum of Heptageniidae, Ephemeridae, Leptophlebiidae, Brachycentridae, Goeridae, Polycentropodidae, Limnephilidae, Odontoceridae, Dolichopodidae, Stratyomidae, Dixidae, Empididae, Athericidae & Nemouridae)	Buffagni et al., 2004; Buffagni & Erba, 2004	0.266
	Abundance	1-GOLD	1 - (relative abundance of Gastropoda, Oligochaeta and Diptera)	Pinto et al., 2004	0.067
Richness and Diversity	Taxa number	Total number of Families	Sum of all Families present at the site	e.g. Ofenböck et al., 2004	0.167
	Taxa number	number of EPT Families	Sum of Ephemeroptera, Plecoptera and Trichoptera taxa	e.g. Ofenbock et al., 2004; Böhmer et al., 2004.	0.083
	Diversity index	Shannon-Wiener diversity index	$D_{S-W} = -\sum_{i=1}^{s}\left(\frac{n_i}{A}\right)\cdot\ln\left(\frac{n_i}{A}\right)$	e.g. Hering et al., 2004; Böhmer et al., 2004.	0.083

c) Accordingly to the WFD requirements, the STAR ICMi class boundaries here presented for High/Good and Good/Moderate status are dedicated to R-C1 rivers i.e. they are type-specific. See next paragraphs for details on technical options used to set class boundaries.

Principles of the classification

The used approach and thus the proposed values satisfy the requirements of the WFD: type-specific adaptation of reference conditions, use of an EQR scale, REFCOND approach for setting class boundaries. As far as normative definitions in terms of kind of information provided for invertebrates are concerned, i.e. ratio sensitive/insensitive taxa, diversity and abundance, the compliancy is guaranteed by the STAR_ICM index, which directly fulfils such obligations (Buffagni et al., 2005; Wasson & Buffagni, 2005). Also, the level of biological alteration evaluated by the STAR_ICMi values complies with normative definitions in terms of: "slight deviation" of taxa richness from reference conditions; presence of sensitive taxa; presence of major taxonomic groups. In Appendix, the relationship between the quality classes based on the proposed STAR_ICMi values and each of the componing biological metrics are shown for Reference and High to Bad quality classes (Box&Whiskers plots).
The same principles were applied to most IC river types in Italy.

Definition of Reference values and dataset used

The invertebrate and pressure data used for R-C1 were provided by Lombardy EPA. Invertebrate samples were collected for standard monitoring purposes at the province scale and cover the full degradation gradient observed in the area (Milan Province, Northern Italy) i.e. from Bad status to Reference sites. Few additional data collected in an adjacent area (Novara Province, Piedmont) were provided by CNR-IRSA. In total, 365 samples collected over a period of ca 10 years ('96 to '06) are being used, including 32 samples from 5 Reference sites. Reference sites were selected on the basis of the REFCOND and CB_GIG criteria.

Boundary setting approach

Definition of High /Good boundary

The High/Good class boundary was set accordingly to a 3-step procedure.

1) First, a possible value for the H/G boundary is set to correspond to the 25th percentile of STAR_ICMi values observed at reference sites, which is considered to be a minimal and simple approach in line with WFD requirements → REFCOND approach
2) A second potential value for the boundary is calculated after testing against an independent, benchmark dataset, the AQEM/STAR Benchmark dataset (as described in Buffagni et al., 2005; 2006; Buffagni & Erba, 2006). The value obtained according to this approach should guarantee the similarity to scientifically set (and thus ecologically sound) boundaries.
3) If needed, an intermediate value comprised between the two possible boundary values defined in 1) and 2) is selected as the final H/G boundary, taking care of: a) the percentage of classification of samples from reference sites into classes lower than High (i.e. aiming at lowering this percentage); b) the pre-WFD boundary should at least be maintained (i.e. lower river protection not allowed); c) a balanced positioning between values as defined in 1) and 2) should be preferred.

Definition of Good/Moderate boundary

The Good/Moderate class boundary was set accordingly to a 3-step procedure as well.

1) The G/M boundary is set to correspond to the H/G boundary (see above) multiplied by 0.75. I.e., the range covered by STAR_ICMi values comprised between 0 and the 25th percentile of STAR_ICMi observed at reference sites was partitioned into 4 equally spaced classes, Good status being the highest in terms of STAR_ICMi. A 25% deviation from reference sites value is assumed to be, in general terms, a slight deviation → REFCOND approach
2) The boundary is refined after testing against an independent, benchmark dataset, the AQEM/STAR Benchmark dataset (as described in Buffagni et al., 2005; 2006; Buffagni & Erba, 2006) (same as for H/G boundary).
3) If needed, an intermediate value comprised between the two possible boundary values defined in 1) and 2) is selected as the final G/M boundary (same as for H/G boundary).

Point 1) for preliminary H/G and G/M boundary setting does not need any further explanation, because it simply makes use of a percentile and of an arithmetical separation into 4 equal classes. Nonetheless, such a simple option risks generating numerical results (i.e. boundaries) that are not ecologically sound. That's why point 2) and 3) cheking was adopted (see also Buffagni & Furse, 2006). The values obtained in point 1) ('REFCOND approach') represent one limit (for each of the two boundaries) for the range to be considered to finally set the class boundaries. Point 2) does require some more clarification, which is reported hereafter. The ICMi as accepted by European MSs for a standard use in the CB_GIG for comparison and harmonization, i.e. the STAR ICMi (Buffagni et al., 2005; 2006), was used in statistical testing for deriving a second value, for each boundary (G/M and H/G).
Firstly, the ICMi value for each invertebrate sample was calculated. A statistical comparison was then executed between the ICMi values found in the AQEM/STAR benchmark dataset - which is assumed to enclose WFD compliant classifications (Buffagni & Erba, 2006; Buffagni et al., 2006) - and the same observed in R-C1 dataset for the High status class as defined by using point 1) approach. The procedure is explained in details in Buffagni et al., 2005; 2006. Because R-C1 samples did not significantly differ from benchmark ones i.e showed values not lower than benchmark data, the H/G boundary obtained in point 1) was confirmed. Point 3) adjustment is thus not needed. and the boundary for STAR_ICMi for H/G class is thus finally set.
In the same way as for the adjustment of the H/G boundary, the boundary G/M was considered and the procedure of statistical comparison between Good status classes, as it was carried out for High status, was repeated. Because no statistically significant differences were observed with values obtained at Point 1), the boundary value there set was adopted.

In Table 1, the calculated values by 1), 2) and final boundaries (STAR_ICMi) for all classes are provided.

Table1. EQR scale. Boundary values reported are the lowest limit of each quality class.

Class boundary	STAR_ICMi values according to REFCOND approach	STAR_ICMi values after testing with benchmark	**Final STAR_ICMi boundaries**	final decision criteria
High/Good	0.959	0.959	**0.959**	REFCOND and benchmark testing
Good/Moderate	0.719	0.719	**0.719**	REFCOND and benchmark testing
Moderate/Poor	0.479	nc	**0.479**	REFCOND
Poor/Bad	0.240	nc	**0.240**	REFCOND

The ecological soundness of the two selected boundaries, H/G and G/M, was then validated, in terms of Normative definitions, by looking at the distribution of WFD compliant metrics values as a function of the final classification. A few examples are reported below, which support the adherence of the classification to Normative definitions for invertebrates.

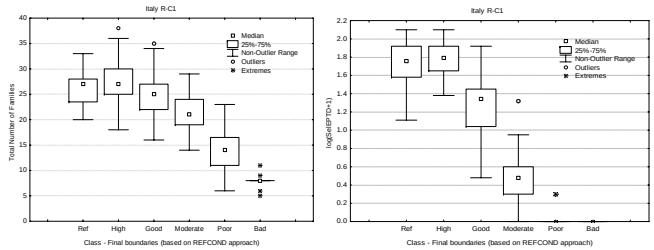
	Future revision of the classification A new list of reference sites is being derived in Italy, for all GIGs and types, to be further sampled during 2006 and 2007. This will possibly supply a larger set of invertebrate samples and pressure data. Thus, also in relation to the results of the ongoing Intercalibration activity for other Italian stream types (i.e. from the Mediterranean and Alpine GIGs) and to the definition of an improved typology for Italian rivers, some refinements of boundaries might be required and/or desirable. Essential bibliography APAT-IRSA/CNR, 2004. 'Indice Biotico Esteso (I.B.E)'. In: APAT, Manuali e linee guida 29/2003. APAT-IRSA/CNR, Metodi analitici per il controllo della qualità delle acque, Roma 3: 1115-1136. Buffagni, A., S. Erba, M. Cazzola & J. L. Kemp, 2004. The AQEM multimetric system for the southern Italian Apennines: assessing the impact of water quality and habitat degradation on pool macroinvertebrates in Mediterranean rivers. Hydrobiologia 516: 313-329. Buffagni A. & Erba S. 2006. Contribution to the CentralGIG Intercalibration activities. Notes on a possible Benchmark /Independent dataset for the IC process. CB_GIG Document, 36pp. Buffagni A. & M. T. Furse. 2006. 'Intercalibration and comparison – major results and conclusions from the STAR project'. Hydrobiologia 566: 357-364. Buffagni A., Erba S., Birk S., Cazzola M., Feld C., Ofenböck T., Murray-Bligh J., Furse M. T., Clarke R., Hering D., Soszka H. & W. van de Bund, 2005. 'Towards European Inter-calibration for the Water Framework Directive: Procedures and examples for different river types from the E.C. project STAR'. 11th STAR deliverable. STAR Contract No: EVK1-CT 2001-00089. *Quad. Ist. Ric. Acque* 123, Rome (Italy), IRSA, 468 pp. Buffagni A., Erba S., Cazzola M., Murray-Bligh J., Soszka H. & Genoni P. 2006. 'The STAR Common Metrics approach to the WFD Intercalibration Process: full application across Europe for small, lowland rivers'. Hydrobiologia 566: 379-399. European Commission, 2005. Guidance on the Intercalibration Process 2004-2006. Water Framework Directive Common Implementation Strategy (200/60/EC), Guidance Document n° 14, 2005, 31pp. Ghetti, P. F. 1997. Indice Biotico Esteso (IBE). I macroinvertebrati nel controllo della qualita` degli ambienti di acque correnti. Provincia Autonoma di Trento, 222 pp. REFCOND Guidance - Wallin, M., Wiederholm, T. & 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. 2003-03-05, 93 pp. Wasson J.G. & Buffagni A., 2005. Does the ICMi approach ensures the consistency with the WFD normative definitions? River Intercalibration - Discussion paper for the Central/Baltic GIG. Steering group & GIGs coordinators meeting, Lyon, 18-19th May 2005, 8pp.

Criteria for Boundary Setting	**High/Good boundary**	**Good/Moderate boundary**
Taxonomic composition and abundance **(see Appendix below)**	The Total Number of taxa, the number of EPT taxa, 1-GOLD, and Sel EPTD_taxa show - in High status samples - values that correspond totally or nearly totally to those observed at reference sites.	For the same metrics, the deviation from reference sites values is slight. **(see Appendix below)**
Ratio of disturbance sensitive to insensitive taxa	The sensitive to insensitive taxa ratio is reflected by the ASPT metric, by the presence/absence of indicator taxa (Sel EPTD_taxa and 1-GOLD) and by the number of EPT taxa. In	For the same metrics, the deviation from reference sites values is slight

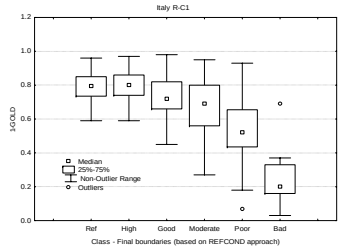
(see Appendix below)	High status samples, they show values that correspond totally or nearly totally to those observed at reference sites.	**(see Appendix below)**
Level of diversity **(see Appendix below)**	The diversity is reflected by the Total Number of taxa, number of EPT taxa and by the Shannon-Wiener metric. In High status samples, they show values that correspond to those observed at reference sites (they are even higher).	For the same metrics, the deviation from reference sites values is slight **(see Appendix below)**
STAR ICMi, in general	The High/Good boundary was set according to the procedure described in the text above, which is performed on the index values after the combination of the composing metrics. Nonetheless, even the relationship of single metrics included in the index show a good disciminatory power among quality classes (see Appendix). The REFCOND approach was used to set class boundary (25th %ile value of REF samples) and it was further validated by comparing to totally independent, benchmarking system (i.e. AQEM/STAR European dataset)	The Good/Moderate boundary was set according to the procedure described in the text above, which is performed on the index values after the combination of the composing metrics. Nonetheless, even the relationship of single metrics included in the index show a good disciminatory power among quality classes (see Appendix). The REFCOND approach was used to set class boundary (equal classes repartition starting from the 25th %ile value of REF samples) and it was further validated by comparing to totally independent, benchmarking system (i.e. AQEM/STAR European dataset)

Appendix (IT)
The response of the individual biological metrics included in the STAR_ICMi, which fulfill the WFD definitions for aquatic invertebrates in rivers to the final classification (i.e. boundaries) obtained for R-C1 type, is reported hereafter, according to the main definition categories in the WFD. The distribution of values for each metric in the 5 Ecological Status classes based on the STAR_ICMi boundaries and at Reference sites is shown in the form of Box&Whiskers plots.

Taxonomic composition and abundance
The Total Number of taxa (here Families) found in a sample can be considered one of the major indicators for taxonomic composition (see figure below, left). The shift from 27 (REF) to 25 (Good status) in the Number of Families (median value) is considered a slight change in the composition of the invertebrate community.
The EPT taxa metric as well contributes to taxonomic composition of the community (see graph in 'Ratio of disturbance sensitive to insensitive taxa').

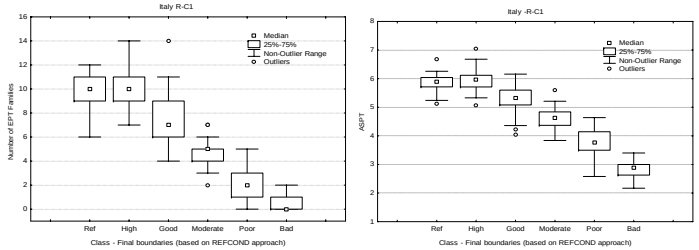


The abundance-based metric included in the STAR_ICMi (= CB_GIG ICMi) i.e. Sel EPTD_taxa (Buffagni et al., 2005; 2006) accounts for invertebrate abundance in R-C1 (see above, right). The shift from 1.75 (REF) to 1.35 (Good status) in Sel EPTD_taxa is considered a slight change in the composition and abundance of the invertebrate community. In fact, the taxa enumerated in this metric are absent in Bad and Poor status samples (i.e they are sensitive taxa, expected to disappear at altered sites), present in only 50% Moderate status samples (usually with a value below 0.5). In Good status samples, these combined taxa usually reach an abundance corresponding to a metric value around 0.75 that observed at REF sites. The overall trend of the 1-GOLD metric, which is also abundance based, is shown below.



Abundance (in terms of abundance classes) is also used in Shannon-Wiener index calculation (see 'Level of diversity').

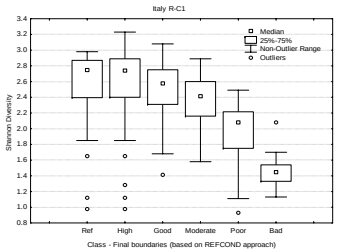
Ratio of disturbance sensitive to insensitive taxa (and missing major taxonomic groups)
EPT-Taxa can be seen as the most sensitive taxonomic groups in CB rivers types. A positive value of this metric indicates that at least one of the three most sensitive major taxonomic groups is present in the community.
In the Italian R-C1datasets, a minimum of 4 EPT Families is encountered at Good status, with 75% of samples bearing 6 or more EPT Families (below, left). The disappearance of EPT Families is only experienced in few cases in Poor status, while it becomes quite common (50% of samples) in Bad status. It appears than clearly how the most sensitive major taxonomic groups are always present in Good status samples according to the proposed boundaries.



The ASPT metric is also shown, which undoubtedly accounts for the Ratio of disturbance sensitive to insensitive taxa (figure above, right). For the Good status class, it shows a slight deviation from the level observed at Reference sites (far less than 1 unit of variation).

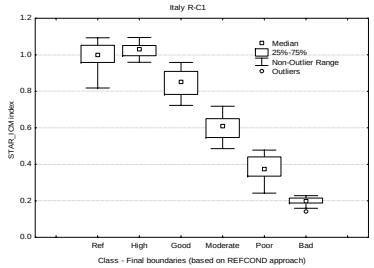
Level of diversity

The level of diversity of invertebrate taxa is present in the STAR_ICMi in terms of Shannon-Wiener diversity index (see graph below). The values of this metric in Good status samples show only slight signs of alteration from reference sites levels (Median of REF: 2.77; Median of Good status: 2.59, with a range going to values lower than 1.4).



Diversity, in terms of Richness of the community, is presented in 'Taxonomic composition and abundance' (Total Number of Families).

Overall trend of STAR_ICMi classes



The distribution of the values of the STAR_ICM index (which incorporates all the metrics shown above) in the 5 Ecological Status classes based on the boundary setting procedure explained in the text and at Reference sites is shown to describe the general trend of values into quality classes.

Country	Lithuania
Classification System:	Trent Biotic Index (BI) and Danish Stream fauna Index (DSFI)
General Description	2 macroinvertebrate classification systems were used to characterize reference conditions and different class boundaries for each intercalibration type - Trent Biotic Index (BI) and Danish Stream fauna Index (DSFI). The first one has been traditionally used for some time in Lithuania. The DSFI on the other hand is a new index for Lithuania that have been started to be applied in evaluations recently. In both cases, however, the sampling and index calculation methodologies are entrenched in national legal systems, but no official classification systems for such indexes are set in the legislation. Traditionally used quality boundaries are applied at the moment. BI is derived after the analysis of the communities of invertebrates that live in a river, taking into account the taxonomic units sensitive to pollution. The method is therefore based on: • the evaluation of the number of systematic units; • the recognition of fauna groups more or less tolerant to pollution. According to increasing degree of tolerance to pollution 7 fauna groups have been identified. To assign the sample to one of these groups, we examine the composition of the sample and search for the macroinvertebrate organism least tolerant to pollution. This one determines the fauna group which the sample belongs to. • determination of the biotic index. The chart for the determination of biotic indices has 2 entries 1. in the columns total are progressive numeric ranges of systematic units (S.U.); 2. in the rows are reported the fauna groups. ▪ Identification of pollution class. Six pollution classes have been identified, the first one (class I) corresponding to lowest level of pollution and the last one (class VI) to the highest level. The biotic index is converted in pollution class by using the following table.

Biotic index	Water quality class	Description of status
10	1	Pristine
7-9	2	Clean
5-6	3	Polluted
4	4	Moderately polluted
2-3	5	Highly polluted
0-1	6	Very much polluted

The actual values for reference conditions and class boundaries were derived referring to the table above with some slight deviations for certain types. This as due to the fact that the reference streams have been chosen on the basis of pressure criteria (agricultural landuse %, point sources, habitat hydromorphology etc) first of all. Therefore, the corresponding reference BI values happened to deviate from table values. We are of the opinion that the deviation is justifiable in the light of GIG recommendations.

DSFI is an index based both on sensitivity of macroinvertebrate taxa and diversity. DSFI is also classification system and seven quality classes (fauna-classes) can be calculated. An index value of 1 characterises a strongly impacted macroinvertebrate community. An index value of 7 characterises an unimpacted or only slightly impacted macroinvertebrate community. DSFI was not originally developed to fulfil the requirements of the WFD. However, it is almost fully in compliance with WFD, except that abundance is not used for all taxa when calculating the index value. According to Danish understanding and knowledge, DSFI is used in the following way in Denmark:

High quality = DSFI index value 7,
Good quality = DSFI index values 6 and 5
Moderate quality = DSFI index value 4

Poor quality = DSFI index value 3
Bad quality = DSFI index values 2 and 1

In Lithuania the values above also have been taken into account in the process of water status evaluation and new classification system development. However, the proposed reference and class boundary values slightly differ in our types due to different environmental conditions in Lithuania. Therefore, the DSFI that is currently being used should undergo modification procedures to fit it more to the Lithuanian conditions if it will be decided to use the DSFI in the future. Reference streams have been chosen on the basis of pressure criteria (agricultural landuse %, point sources, habitat hydromorphology etc) first of all, therefore the corresponding reference DSFI values happened to deviate from Danish classification slightly. We are of the opinion that the deviation is justifiable in the light of GIG recommendations.

Criteria for Boundary Setting	**High/Good boundary**	**Good/Moderate boundary**
Taxonomic composition and abundance	No or minor deviations of taxonomic composition and abundance from pristine state, expressed in BI (no less than 9) and DSFI (no less than 6-5)	Typically, most major taxonomic groups (orders) are found. But several families especially in important groups like Ephemeroptera, Plecoptera and case-bearing Trichoptera (EPT) are missing. Abundance of some insensitive taxa could increase. In terms of BI and DSFI values they should not be below 7-6 and 5-4 accordingly
Ratio of disturbance sensitive to insensitive taxa	No or only a very minor loss in sensitive taxa, expressed in BI (no less score than 9) and DSFI (no less than 6-5)	The number of species and individuals of sensitive taxa of Plecoptera (genus-level), Ephemeroptera (family-level) and other sensitive groups are significantly reduced at the good/moderate boundary. As a consequence the proportion of insensitive taxa becomes higher compared to the reference state. In terms of BI and DSFI values they should not be below 7-6 and 5-4 accordingly
Level of diversity	No or only a very minor loss in diversity expressed	Loss in species diversity has been estimated for

	in BI (no less than 9). In terms of DSFI, the number of "positive" taxa groups should not deviate or deviate only slightly from natural conditions (no less than 6-5)	the EPT families (see above). As a mean only about 45% of the EPT **taxa** present at high quality will be found at the good/moderate boundary. At the **family** level (all families) as a mean about 70% of the families can be found at the Good/moderate boundary. In terms of BI and DSFI scores they should not be below 7-6 and 5-4 accordingly

Country	*Luxembourg*	
Classification System:	Indice Biologique Global Normalisé (IBGN)	
General Description	The IBGN is a semi-quantitative method. The IBGN score is obtained by crossing two metrics: (1) the Faunistic Indicator Group (GFI) whose values range from 1 to 9 according to an increasing gradient of pollution sensitivity from group 1 to group 9; and (2) the number of families collected (Taxonomic Variety, VT). The IBGN is not WFD compliant since, among the four metrics recommended by the DCE (taxonomic composition, abundance, proportion of sensitive taxa and diversity), only the last two are partially taken into account by this method: the GFI corresponds to the presence of sensitive taxa and the VT to the diversity but abundances are not included in the calculation of these metrics. A predictive model (relatively similar to the RIVPACS model) is under development and aims to fulfil most recommendations of the DCE. In a first step, reference sites were selected in the monitoring network according to the criteria provided by GIG (land use and chemical reference values) and following the C/B rivers GIG screening for reference criteria. In a second step, biological indices according to the national method (IBGN) are calculated for all samples of the reference dataset, and the outliers are checked. In accordance with the WFD normative definitions, only a slight deviation of taxa richness (VT), sensitive taxa (GFI) and the disappearance of major taxonomic groups was allowed for the pool of reference sites. For each type, a reference value was derived as the median of observed values in the reference sites for the IGBN index and its separate metrics (GFI and VT). All the IBGN values were then transformed in EQR-IBGN. The minimum IBGN value is set at IBGN=1. The High/Good boundary is based on the 25th percentile of the IBGN values (combination of the VT and GFI metrics values) observed in reference sites. The "good status" lower limit was defined as the high status lower limit multiplied by 0.75.	
Criteria for Boundary Setting	**High/Good boundary**	**Good/Moderate boundary**
Taxonomic composition and abundance	The IBGN score was used as indicator of the element "taxonomic composition and abundance" (combination of the two metrics (VT and GFI). Calculated threshold values for high/good boundaries are 16, 15 and 13 for R-C3, R-C4 and R-C6, respectively (Tab. 1).	Calculated threshold values for good/moderate boundaries are 12, 11 and 10 for R-C3, R-C4 and R-C6, respectively (Tab. 1).
Ratio of disturbance sensitive to insensitive taxa	The GFI (Faunistic Indicator Group) was used as indicator of the element "ratio of sensitive to insensitive taxa". Calculated threshold values for high/good boundaries are 9, 7 and 7 for R-C3, R-C4, R-C5 and R-C6, respectively (Tab. 1).	Calculated threshold values for good/moderate boundaries are 7, 5 and 5 for R-C3, R-C4 and R-C6, respectively (Tab. 1).
Level of diversity	The VT (Taxonomic Variety) was used as indicator of the element "level of diversity". Calculated threshold values for high/good boundaries are 8, 8 and 8 for R-C3, R-C4 and R-C6, respectively (Tab. 1).	Calculated threshold values for good/moderate boundaries are 7, 5 and 5 for R-C3, R-C4, R-C5 and R-C6, respectively (Tab. 1).

Appendix (Luxembourg)

IC Types	Description	Reference values			High status			Good status		
		IBGN	GFI	VT	IBGN	GFI	VT	IBGN	GFI	VT
R-C3	Small mid-altitude siliceous	>= 16	>= 9	>= 8	>= 16	>= 9	>= 8	>= 12	>= 7	>= 6
R-C4	Medium lowland* mixed	>= 16	>= 8	>= 9	>= 15	>= 7	>= 8	>= 11	>= 5	>= 6
R-C6	Small lowland* calcareous	>= 14	>= 8	>= 7	>= 13	>= 7	>= 8	>= 10	>= 5	>= 6

Tab. 1: threshold IBGN, GFI and VT reference and ecological status values for R-C3, R-C4 and R-C6 common intercalibration rivers in Luxembourg. GFI : Faunistic Indicator Group; VT : Taxonomic Variety.
* Concerning the common intercalibration typology, it must be specified that:
- most sites in the R-C4 type have a 180-300 m altitudinal range
- most sites in the R-C6 type have a 160-300 m altitudinal range

Country	The Netherlands	
Classification System:	Dutch KRW-maatlatten (Dutch WFD-index)	
General Description	A multi-metric WFD-index, determining quality classes for Dutch running waters has been developed based on species composition and abundances. Expert judgement determined which macroinvertebrate communities characterized the different classes of ecological quality (from 'bad' to 'high'), and listed macroinvertebrate species indicative for ecological quality. Three lists of indicator species were created: one consisting of critical species, characteristic for a water type, one containing dominant positive species and one containing dominant negative species. From these lists, three metrics calculating the WFD have been developed: %KM, %(DP+KM) and %DN focussing on the relative number of sensitive species, relative abundance of dominant positive species and relative abundance of negative dominant species respectively. These metrics have been integrated into a formula for calculating WFD-index: EQR = {200*(%KM/KMmax) + 2*(100-%DN) + %(DP+KM)}/500 % KM = relative number of sensitive species in a sample KMmax = maximum achievable number of sensitive species under reference conditions % DN = relative abundance of dominant negative species % (DP+KM) = sum of relative abundances of dominant positive species and sensitive species High quality = Maatlat index value 0.8-1.0, Good quality = Maatlat index values 0.6-0.8 Moderate quality = Maatlat index value 0.4-0.6 Poor quality = Maatlat index value 0.2-0.4 Bad quality = Maatlat index values 0.0-0.2	
Criteria for Boundary Setting	**High/Good boundary**	**Good/Moderate boundary**
Taxonomic composition and abundance	A pool of macroinvertebrate samples was pre-classified by expert judgement in combination with multivariate gradient analysis. The combination of metrics that fitted the pre-classification best were selected and transformed in a formula that calculates the EQR as described above. The high/good boundary was set at an EQR-score of 0.8, whereas the good/moderate boundary was set at an EQR-score of 0.6	

Ratio of disturbance sensitive to insensitive taxa	No or only a very minor loss in sensitive taxa	The number of sensitive species and their abundances are highly reduced at the good/moderate boundary, whereas the proportion of insensitive taxa increases
Level of diversity	Minor loss in diversity; dominant positive species become less abundant	Loss in sensitive species diversity

Development of the Dutch WFD-index ('KRW-maatlat') for macroinvertebrates in running waters

M.S. van den Berg, T. Ruigrok, R. Knoben, M.C. van Riel, dd. 24-05-2006
Draft, submitted to the Central-Baltic GIG Rivers Intercalibration Workshop dated June 14th in Lithuania

Summary
A multi-metric WFD-index, determining quality classes for Dutch running waters has been developed based on species composition and abundances. Expert judgement determined which macroinvertebrate communities characterized the different classes of ecological quality (from 'bad' to'high'), and listed macroinvertebrate species indicative for ecological quality. Three lists of indicator species were created: one consisting of critical species, characteristic for a water type, one containing dominant positive species and one containing dominant negative species. From these lists, three metrics calculating the WFD have been developed: %KM, %(DP+KM) and %DN focussing on the relative number of critical species, relative abundance of dominant positive species and relative abundance of negative dominant species respectively. These metrics have been integrated into a formula for calculating WFD-index:

EQR = {200*(%KM/KMmax) + 2*(100-%DN) + %(DP+KM)}/500

% KM = relative number of critical species in a sample
KMmax = maximum achievable number of critical species under reference conditions
% DN = relative abundance of dominant negative species
% (DP+KM) = sum of relative abundances of dominant positive species and critical species

Selecting indicator species
Indicator species have been selected from indicator species lists (Verdonschot et al. 1992; Verdonschot 2000; Verdonschot & Janssen 2000; Verdonschot, 1990; Verdonschot et al., 2000; Verdonschot et al., 1999; Janssen et al., 1998; Verdonschot & Nijboer, 2004). The presence of indicator species and their abundances under reference conditions determine the typology of surface waters. Therefore, abundance classes of the indicator species selected have been added to the species list. These abundance classes are based on abundances of indicator species in natural stream types. Indicator species with abundance classes greater than 5 are referred to as dominant positive species, whereas species with lower abundances are referred to as critical species. Species that occur in high densities in impacted or polluted water systems are referred to as dominant negative species. The list containing dominant positive species was extended with species that are commonly present (i.e. >90 individuals) in natural reference waters situated in Poland, Germany, Denmark and Ukraine. Species from reference waters abroad that had already been incorporated in the list of critical species or have never been found in Dutch waters were not included in the list of dominant positive species. The lists composed were lastly judged by experts and were adjusted according to their suggestions.

Development of the Dutch WFD-metrics (%KM, %DN and %(DP+KM))

WFD-parameters were developed on the basis of medium sized lowland rivers, corresponding with intercalibration water type RC-1. This water type is well documented and biological as well as chemical variables are monitored frequently.

WFD-metrics were developed from RC-1 macroinvertebrate samples, which had been classified to EQR by expert judgement supported by multivariate data analysis (CANOCO 4.0). According to expert judgement, sufficient samples represented the classes 'bad', 'poor', 'moderate' and 'good', assuming that no Dutch streams could be qualified as 'high' due to hydromorphological alterations. Subsequently, the EQR from the classified macroinvertebrate samples was plotted to the abundances of dominant negative species, dominant positive species, critical species and rare species. Pearson coefficient analysis calculated which groups of species determined EQR and therefore had to be implemented in the WFD-index parameters. Using relative abundances instead of absolute abundances of the selected species groups improved correlation with EQR. Using relative instead of absolute abundances furthermore diminishes side effects from using different sample methods on the calculation of EQR. Critical species (KM), dominant negative species (DN) and dominant positive species (DP) determined EQR most and were therefore selected as parameters determining the WFD-index. These parameters were selected using a score system for analysing the role of the selected species groups (KM, DN and DP) on the EQR. The abundance of dominant positive species appeared to be related to the abundance of critical species as the number of critical species and the number of positive dominant species cannot be high at the same time. Positive dominant species en critical species have therefore been combined to one parameter**:** %(DP+KM). Subsequently, the scores of the parameters %KM, %DN and %(DP+KM) were translated to a normalized EQR-qualification, ranging from 0 to 1. Within this range, the class boundaries were equally distributed, with 0.6 being the most important boundary as it distinguishes between the classes 'moderate'and 'good'. As macroinvertebrate composition highly depends on water type, translations have been made according to relative abundances of DP, DN and KM in reference communities for each water type individually. Table 1 shows an example of this translation for the water type RC-1.

Table 1: translation from parameter scores to EQR for water type RC-1

Parameter	Abundance score	EQR-score
%DN (abundance)	≥ 41	0.1
	< 41	0.2
%KM (number of taxa)	≤ 7	0.1
	7 < %KM < 18	0.2
	18 ≤ %KM < 33	0.3
	≥ 33	0.5
%(DP+KM) (abundance)	< 5	0.1
	5 ≤ %KM < 25	0.2
	≥ 25	0.3

Figure 1 shows the divergent influences of the three WDF-parameters. The parameter %KM is most determining in distinguishing EQR-classes. Lower EQR-classes are characterized by high values of %DN and low values of %KM. The increase of %KM and the decrease of %DN

result in higher quality classes. From the class 'good', the role of the parameter %(DP+KM) becomes more important and determines whether the class 'high' can be achieved.

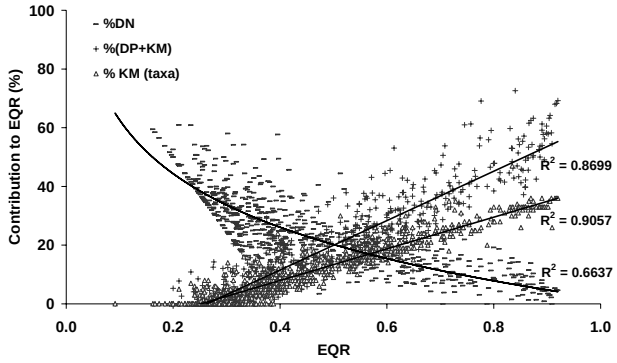


Figure 1: The influence of WFD-parameters %KM (taxa), %DN (abundance) and %(DP+KM) (abundance) on EQR

Finally, the parameters were integrated into a formula that calculates the EQR of a sample based on species composition:

EQR = {200*(%KM/KMmax) + 2*(100-%DN) + %(DP+KM)}/500

% KM = relative number of critical species in a sample
KMmax = maximum achievable number of critical species under reference conditions
% DN = relative abundance of dominant negative species
% (DP+KM) = sum of relative abundances of dominant positive species and critical species

Calibration of the Dutch WFD-index ('KRW-maatlat')

Calibration of the Dutch WFD-index to expert judgement showed equal EQR-classification for 81% of the samples tested. Expert judgement was performed according to the 'Delphi-method' by sending 30 species list of anonimised sampling sites to 10 macroinvertebrate experts distributed over the country.

Setting EQR-class boundaries

A pool of macroinvertebrate samples was pre-classified by expert judgement in combination with multivariate gradient analysis. The combination of metrics that fitted the pre-classification best were selected and transformed in a formula that calculates the EQR. The selection of these metrics and development of the formula are described above.

Selecting type-specific reference sites

Reference sites have been identified according to the criteria defined by Wasson (April, 2006). Most of the Dutch waters could not meet these requirements as most of them have been hydromorphologically altered and do not correspond with the conditions set for nitrogen and Phosphate. The criteria for the Dutch reference waters enhanced that the stream must be qualified as nature for at least 50%, may contain no more than 4% urban area, may not exceed 15 kilogram nitrogen per hectare nor 1 kilogram phosphate per hectare. Furthermore, a reference site is may not contain point sources and may not be hydromorphologically altered. Recreation or biomanipulation must be restricted to a minimum.

Table 1: Reference criteria

Parameter	RC-1	RC-4	Remark
BOD5	2.4 mg/l	2.4 mg/l	Yearly average
BOD5	3.6 mg/l	3.6 mg/l	90 percentile
O2 saturation	95-105	95-105	Yearly average
O2 saturation	85-115	85-115	10-90 percentile
N-NH4	0.1 mg N/l	0.1 mg N/l	Yearly average
N-NH4	0.25 mg N/l	0.25 mg N/l	90 percentile
P-PO4	0.04 mg P/l	0.04 mg P/l	Yearly average
N-NO3	6 mg N/l	6 mg N/l	Yearly average
N-NO3	2-4 mg N/l	2-4 mg N/l	Yearly average

The stream Hierdensche beek is the only stream that meets all criteria for a reference site.

References:

Janssen S.N., Verdonschot P.F.M.& Arts, G.H.P. (1998). Typologie van zoete duinwateren gebaseerd op macrofauna, macrofyten, diatomeeën en milieuvariabelen. IBN rapport 390: 200 pp.

Verdonschot P.F.M. (1990). Ecologische karakterisering van oppervlaktewateren in Overijssel. Het netwerk van cenotypen als instrument voor ecologisch beheer, inrichting en beoordeling van oppervlaktewateren. Provincie Overijssel, Zwolle; Rijksinstituut voor natuurbeheer, Leersum, 300 pp.

Verdonschot P.F.M., Higler L.W.G., Van der Hoek, W.F. & Cuppen J.G.M. (1992). A list of macroinvertebrates in dutch water types: a first step towards an ecological classification of surface waters based on key factors. Hydrobiol. Bull. 25:241-259.

Verdonschot P.F.M.& Janssen S.N. (2000). Natuurlijke levensgemeenschappen van de Nederlandse binnenwateren. Deel 2: Beken. Attachment in 'Handboek Natuurdoeltypen in Nederland' Document AS-12, EC-LNV, Alterra, Wageningen, 78 pp.

Verdonschot P.F.M., Nijboer, R.C., Janssen S.N. & Van der Hoorn, M.W. (2000). Waterstreefbeelden en regionale watersysteemverkenningen Limburg. Ecologische typologie, ontwikkelingsreeksen en waterstreefbeelden. Deelproject IIb: Cenotypenbeschrijvingen. Aterra, Wageningen. 171.3: 1-235.

Verdonschot P.F.M., Koopmans M. & Gerritsen R.C. (1999). Ecologisch maatweb stromende wateren. Waterschap Veluwe en Waterschap Vallei en Eem, 142 pp.

Verdonschot P.F.M. & Nijboer, R.C.(2004). Macrofauna en vegetatie van de Nederlandse beken. Een aanzet tot beoordeling van de ecologische toestand. Wageningen, Alterra, Research Instituut voor de groene ruimte. Alterra-rapport 756: 326 pp.

Country	PL	
Classification System:	Adapted to Polish conditions Biological Monitoring Working Party score (BMWP-PL) verified by the modified Margalef Diversity index	
General Description	In Poland river assessment method based on macroinvertebrates compliant with WFD requirements has been still under development. Therefore for the intercalibration purposes, preliminary reference values of both indices, H/G and G/M boundaries have been established. Proposed values has not gained an acceptance of relevant authorities yet and should has been treated just as an proposal. Reference values of both indices constitute the median values from the reference sites while high/good boundary - 25th of percentile. The range from H/G boundary to the worst case is splitted into 4 equal intervals according to REFCOND guidance. The assignment into one of five quality classes is based on the worst case.	
Criteria for Boundary Setting	**High/Good boundary**	**Good/Moderate boundary**
Taxonomic composition and abundance	Major Taxonomic Groups (MTG) defined mostly at the level of order (e.g. Plecoptera, Trichoptera) or class (e.g. Oligochaeta, Gastropoda). Mean number of MTG constitutes 12 and is comparable with the number of MTG expected in the reference state. It is observed an increase in abundance.	Similarly to reference state mean number of MTG, constitutes 12. However groups sensitive to pollution become less frequent or rare. Abundance moderately increases comparing with the abundance in reference conditions.
Ratio of disturbance sensitive to insensitive taxa	Preliminary H/G boundary is set as 0,9 of EQR_BMWP-PL. Mean value of BMWP-PL similar to the value expected in the reference state. On average only one taxa sensitive to pollution (e.g. EPT taxa) is absent. Proportion of sensitive and insensitive taxa is balanced.	0,7 is set as the G/M boundary. It corresponds to the decreasing number of sensitive taxa and increasing abundance of insensitive taxa.
Level of diversity	Preliminary boundary values of EQR_Margalef diversity index is set at the level 0,9. Average value of Margalef diversity in high state is slightly higher than the values expected in the reference state. Inconsistent results suggest that a replacement with more relevant metric in the final, WFD compliant, method of assessment is needed.	Preliminary boundary values for EQR_Margalef diversity index is set at the level 0,7. Mean value of diversity index slightly decreases compared with the values expected at reference conditions.

Country	Spain	
Classification System:	INV8 F1 a multimetric index for R-C2 stream type	
General Description	A multimetric index built as a simple average combination of metrics, having discriminatory efficiency between reference sites and multiple pressure influenced sites. Indices used: number of families, number of EPT families, number of sensitive families, Bray-Curtis similarity index, Log (Trichoptera and Plecoptera abundance + 1), % 3 dominant families, % Oligochaeta and % sensitive families, based on family level invertebrate data. EQR multimetric values are normalised by the median value of the reference population to reflect the WFD concept of reference condition.	
Criteria for Boundary Setting	**High/Good boundary**	**Good/Moderate boundary**
Taxonomic composition and abundance	EQR of *number of families, number of EPT families,* and *abundance of Trichoptera and Plecoptera* are plotted against the multimetric EQR (Fig 1). The EQR value of 0.930 is the H/G boundary.	All EQRs for richness and abundance metrics cross the line of the *% of 3 dominant families*, this crossing representing the center of the moderate class. The lowest limit of the G/M boundary results in an EQR of 0.730.
Ratio of disturbance sensitive to insensitive taxa	EQR of *number of sensitive families* and values of *% 3 dominant families, % Oligochaeta* and *% sensitive families* are plotted against the multimetric EQR (Fig 1). At high status the values of the metrics indicate a reference community represented equally by sensitive taxa of the type specific community and by the tolerant dominant taxa. The initial crossing of sensitive *(% sensitive families*) vs tolerant taxa *(% 3 dominant families*), results in an EQR of 1.03 indicating the center of the high class. The EQR value of 0.930 (H/G boundary), represents approximately the 84% of the *number of sensitive taxa* value at the centre of the high class.	With increasing pressure there is a decrease of the *number of sensitivity taxa* and an increase of the dominance of the pressure-favored families *(% 3 dominant families*), the EQR = 0.730, were this 2 metrics cross represents the lowest limit of the G/M boundary. The cross of tendencies followed by *% Oligochaeta* and *% sensitive taxa*, also indicates the center of the moderate class. The EQR value of 0.730 (G/M boundary), represents approximately a 69% of the *number of sensitive taxa* value at the centre of the high class.
Level of diversity	The *Bray-Curtis similarity index* is plotted against the multimetric EQR (Fig 1). The high class over an EQR of 0.930 represents the highest level of similarity between samples (>60%), as *Bray-Curtis index* represents the beta diversity measurement for the reference type-specific community.	The crossing of the *Bray-Curtis* metric with the *% of 3 dominant families* indicates the center of the good class, with a community able to quickly recover towards the reference state. At an EQR of 0.730, this boundary represents a *Bray-Curtis* similarity of 52%.

Country	Spain	
Classification System:	INV6 F3 a multimetric index for R-C3 stream type	
General Description	A multimetric index built as a simple average combination of metrics which showed discriminatory efficiency between reference sites and multiple pressure sites. Indices used: number of EPT families, number of sensitive families, % of sensitive families, Bray-Curtis similarity index, % Abundance classes EPT, % 3 dominant families, based on family level invertebrate data. Using the multimetric index value the EQR can be calculated. EQR multimetric values are normalised by the median value of the reference population to reflect the WFD concept of reference condition.	
Criteria for Boundary Setting	**High/Good boundary**	**Good/Moderate boundary**
Taxonomic composition and abundance	EQR of *number of EPT families*, and the *% abundance classes of EPT* are plotted against the multimetric EQR (Fig 2). The EQR value of 0.930 is the H/G boundary.	The *number of EPT families* crosses the regression line of the *% 3 dominant families*, this crossing representing the center of the moderate class. The lowest limit of the G/M boundary results in an EQR of 0.730.
Ratio of disturbance sensitive to insensitive taxa	EQR of *number of sensitive families* and values of *% 3 dominant families* and *% sensitive families* are plotted against the multimetric EQR (Fig 2). At high status the values of the metrics indicate a reference community represented equally by sensitive taxa of the type specific community and by the tolerant dominant taxa. The initial crossing of sensitive *(% sensitive families*) vs tolerant taxa *(% 3 dominant families*), results in an EQR of 1.1 indicating the center of the high class. The EQR value of 0.930 (H/G boundary), represents approximately a 83% of the *number of sensitive taxa* value at the centre of the high class.	With increasing pressure there is an important decrease of the number of sensitivity taxa and an increase of the dominance of the pressure-favored families *(% 3 dominant families*), where this 2 metrics cross, it represents the lower limit of the G/M boundary (EQR of 0.730). The EQR value of 0.750 (G/M boundary), represents approximately a 61% of the *number of sensitive taxa* value at the centre of the high class.
Level of diversity	The *Bray-Curtis* similarity index is plotted against the multimetric EQR (Fig 2). The high class over an EQR of 0.930 represents the highest level of similarity between samples (>60%), as *Bray-Curtis index* represents the beta diversity measurement for the reference type-specific community.	The crossing of the *Bray-Curtis* metric with the percentage of the *% 3 dominant families* indicates the center of the good class, with a community able to quickly recover towards the reference state. At an EQR of 0.730 this boundary represents a *Bray-Curtis* similarity of 54%.

Appendix (Spain)
Figure 1: R-C2 type. Normalised (Number of families, number of EPT families, number of sensitive families, Log (Trichoptera and Plecoptera abundance + 1)), and Bray-Curtis similarity index, % 3 dominant families, % Oligochaeta and % sensitive families, based on family level invertebrate data Vs Normalised INV8 F1 EQR (Blue line: H/G boundary; Green line: G/M boundary)

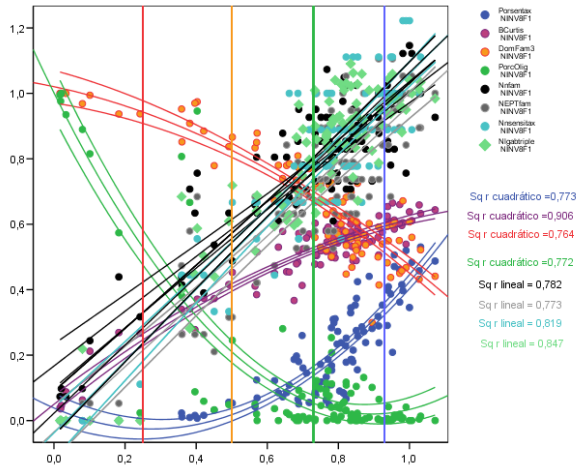
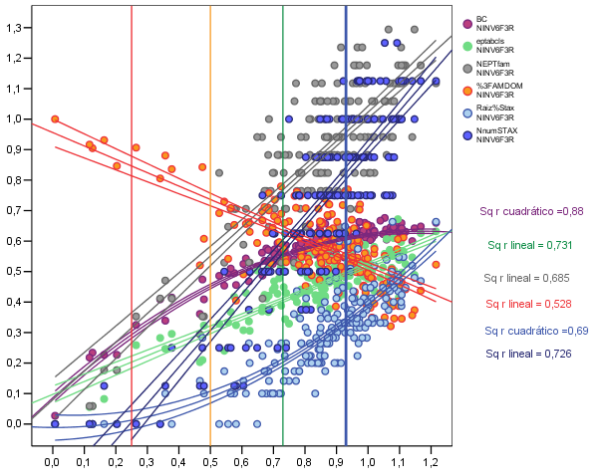


Figure 2: R-C3 type. Normalised (Number of EPT families, number of sensitive families), and % abundance of classes EPT, and Bray-Curtis similarity index, % 3 dominant families, and % sensitive families, based on family level invertebrate data Vs Normalised INV6 F3 EQR (Blue line: H/G boundary; Green line: G/M boundary)



Country	**Sweden**	
Classification System:	General Quality Assessment Classification	
General Description	The DJ-index (Dahl & Johnson 2005) is a multimetric index consisting of measures of diversity (number of EPT taxa), composition (% Crustacea and % EPT taxa) and tolerance ((ASPT and Sabrobic Index; Zelinka & Marvan 1961), is used to assess effects of nutrient enrichment	
Criteria for Boundary Setting	**High/Good boundary**	**Good/Moderate boundary**
Taxonomic composition and abundance	Defined as the 25th-percentile of EQRs of "reference" populations (see above)	Defined using either percentile distribution or linear regression. No ecological thresholds have been established. A 25% decrease in EQR from the high-good boundary, which equates to a 44% deviation in EQR values from the average (median) reference EQRs was used.
Ratio of disturbance sensitive to insensitive taxa	Defined as the 25th-percentile of EQRs of "reference" populations (see above)	Defined using either percentile distribution or linear regression. No ecological thresholds have been established. A 25% decrease in EQR from the high-good boundary, which equates to a 44% deviation in EQR values from the average (median) reference EQRs was used.
Level of diversity	Defined as the 25th-percentile of EQRs of "reference" populations (see above)	Defined using either percentile distribution or linear regression. No ecological thresholds have been established. A 25% decrease in EQR from the high-good boundary, which equates to a 44% deviation in EQR values from the average (median) reference EQRs was used.

Country	UK	
Classification System:	General Quality Assessment Classification (RIVPACS)	
General Description	RIVPACS used to estimate "reference" condition of the values of the indices used in the current UK classification scheme. Indices used: Average Score per Taxon (ASPT), Number of Taxa (NTAXA), based on family level invertebrate data. Using observed values of ASPT and NTAXA for a site and the expected values obtained from RIVPACS, the Ecological Quality Index (EQI) can be calculated for each of the indices. EQI values are converted to EQRs using the median value at reference sites that comply with REFCOND guidance and the specific criteria and values agreed by the GIGs.	
Criteria for Boundary Setting	**High/Good boundary**	**Good/Moderate boundary**
Taxonomic composition and abundance	No missing groups.	Major taxonomic groups were defined at the level of order. The average reference condition value for orders is circa 10 orders. The point where typically one order (one major group) is missing occurs at an ASPT value of 0.9: boundary for G/M (Figure 1).
Ratio of disturbance sensitive to insensitive taxa	EQI proportion sensitive and EQI proportion insensitive are plotted against ASPT EQI. The point at which on average the proportions of sensitive and insensitive taxa are both exactly as expected occurs at an ASPT EQI of approximately 1.055. This represents the middle of 'high' (Figure 2). Below this ASPT EQI point it is unlikely that the proportion of expected sensitive taxa will exceed the proportion of expected insensitive taxa. An ASPT EQI value of 1.0 (H/G boundary) represents a location in the 'crossover zone' where the proportion of sensitive is typically lower than the expected middle of "high".	An ASPT EQI of 0.9 represents a valid G/M boundary (Figure 3); at this point the proportions of sensitive and insensitive taxa are consistently different. At this ASPT EQI value the proportion of sensitive taxa is almost always lower than expected at reference condition and the proportion of insensitive taxa is almost always higher than expected at reference condition. At the boundary the ratio of sensitive taxa to insensitive taxa is typically half of that expected at reference condition. The actual number of sensitive taxa is typically just over half that expected on average at reference condition.
Level of diversity	Relationship between NTAXA EQI and normative definitions has not been considered in detail yet. Initial analyses support use of an NTAXA EQI of 1.0 as the H/G.	Initial analyses support use of an NTAXA EQI of circa 0.8 as the G/M boundary.

Appendix (UK)

Figure 1: Missing Major Groups Vs Normalised ASPT EQI (Blue line: H/G boundary; Green line: G/M boundary)

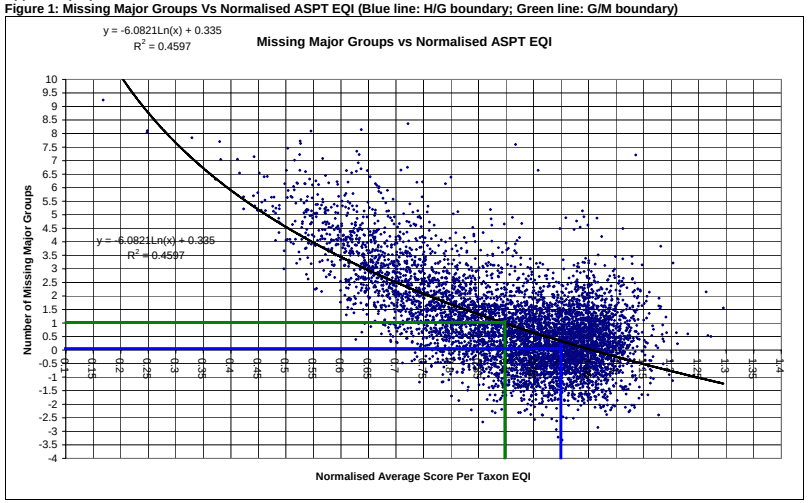


Figure 2: EQI Proportion "Sensitive" and "Insensitive" Vs ASPT EQI (Blue line: H/G boundary; Green line: G/M boundary)