Collaborative Project (large-scale integrating project)
Grant Agreement 226273
Duration: March 1st, 2009 – February 29th, 2012

Water bodies in Europe: Integrative Systems to assess Ecological status and Recovery

DELIVERABLE

# Deliverable 3.4-4: Fish indicators for ecological status assessment of lakes affected by eutrophication and hydromorphological pressures

*- Provisional report -*

**Lead contractor:** CEMAGREF Aix-en-Provence
**Contributors:** Stéphanie PEDRON, Julien De BORTOLI, Christine ARGILLIER

**Due date of deliverable:** Month 30
**Actual submission date:** Month 20

| Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)<br>Dissemination Level | | |
|---|---|---|
| PU | Public | X |
| PP | Restricted to other programme participants (including the Commission Services) | |
| RE | Restricted to a group specified by the consortium (including the Commission Services) | |
| CO | Confidential, only for members of the consortium (including the Commission Services) | |

# CONTENT

# INTRODUCTION

One of the objectives of work package 3.4 of WISER is to develop a fish-based ecological status indicator for European lakes exposed to hydromorphological and eutrophication pressures, including uncertainty assessment. This indicator has to follow the requirement of the Water Framework Directive (WFD; 2000/60/EC) i.e. the status of the fish fauna should be assessed with the following criteria: species composition, abundance and age structure (Annex V 1.2.1 of this directive). This European index is dedicated to the help of the intercalibration exercise achievement. We present here the method implemented to address this issue.

In Lakes, some studies have already assessed the response of individual fish metrics to human stresses such as acidification (Appelberg *et al*. 2000), eutrophication (Jennings *et al.* 1999) or land use (Drake and Pereira 2002) but only at a regional scale. In these studies, natural parameters influencing environment variability are considered as negligible, therefore variability of fish communities (through metrics) is only considered as a response to pressures. Moreover, in most of these studies, the reference is more or less considered as the "best condition" observed in the dataset and this reference is seldom defined.
From a general point of view, these approaches raise two questions that have to be solved before starting with metric selection at the European scale in the framework of the WFD: which environmental parameters are influencing fish communities at such a large scale? And how to define the reference conditions?

We will present below the result of a literature review on fish based metrics already used in bioassessment of lakes or reservoirs quality. This review allows us to perform a list of potential metrics for an European lake fish index. We will then present the available data and method developed to select metrics responding to eutrophication at the European scale before to conclude on required improvement and perspectives for the next months.

## 1 Review of metrics used for lakes IBI development

Biological integrity was described as ``the capability of supporting and maintaining a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of natural habitats of the region'' (Karr and Dudley 1981). A widely used standardized method for measuring the ecological status of aquatic ecosystems is the index of biotic integrity (IBI) define par James R. Karr (1981). He suggested to monitor water resources using fish to assess "biotic integrity" and emphasized that fish communities respond to human alterations in a predictable and quantifiable manner. IBI, as a multimetric indicator helps the quantification and reflects the overall biological condition of a water body (Barbour *et al.* 1995).
A large amount of publications all over the world followed his first version of IBI developed for Midwestern streams and number of scientists have tested and/or adapted the concept of multimetric approach to lentic systems of their own countries and regional features. Thereby lake's IBI and RFAI (Reservoir Fish Assemblage Index) were developed in Tennessee (Mc Donough and Hickman 1995, Minns, Cairns *et al*. 1994; Hickman and McDonough 1996; Belpaire, Smolders *et al.* 2000; Lyons, Gutierrez-Hernandez *et al*. 2000), in North East lakes of United States (Hughes *et al*., 1992; Whittier, 1999), in Wisconsin (Jennings *et al.* 1999), in Florida (Schulz *et al.* 1999) and in Minnesota (Drake & Pereira 2002; Drake & Valley 2005) but also in Mexico (Lyons *et al.* 2000) and in Europe (Belpaire *et al.* 2000, Holmgren *et al.* 2007). Applying the IBI to lentic systems seems more difficult than for rivers because lakes exhibit large scale variation regionally in physical and biological characteristics (Jackson, Peres-Neto *et al*. 2001).

## 1.1 Worldwide indices

In each new publication, the list of metrics changes more or less with the region, country, and lake type where the index was applied, but most IBI use several components of fish communities recommended by the WFD.
We summarize below all metrics calculated by Central and North American authors (Table.1).

**Table.1 List of metrics used in studies aiming to develop a fish-based assessment system adapted to different lake types of USA and Mexico.**

| CATEGORY | METRICS | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *SPECIES COMPOSITION* | | | | | | | | | | | | | |
| | Number of species | ■ | | ■ | ■ | ■ | ■ | | | ■ | | | |
| | Number of native species | | ■ | | | | | ■ | | ■ | ■ | ■ | ■ |
| | Number of common native species | | | | | | | | | | ■ | | |
| | Number of native Chirostoma | | | | | | | | | | ■ | | |
| | Number of native Goodeidae | | | | | | | | | | ■ | | |
| | Number of big native species | | | | | | | | ■ | | | | |
| | Number of non native species | | | | | | | | ■ | | | | |
| | Number of little Cyprinids and Darters | ■ | | | | | | | | | | | |
| | Number of Cyprinids species | | | | | | | | | | | ■ | ■ |
| | Number of centrarchids species | | | | | | | ■ | | | | | |
| | Number of Sucker | ■ | | ■ | ■ | | ■ | | | | | | |
| | Number of native Cyprinids | | ■ | | | | | ■ | | | | | |
| | Number of centrarchids species | | ■ | | | | | | | | | | |
| | Number of non indigeneous species | | ■ | | | | | | | | | | |
| | Number of Lepomis species | | | ■ | ■ | ■ | ■ | | | ■ | | | |
| | Number of benthic species | | | | | ■ | | ■ | | | | ■ | ■ |
| | Number of little benthic species in littoral zone | | | | | | | | | | | ■ | |
| | Maximum standard length of native species | | | | | | | | | | ■ | | |
| Tolerance/Intol. | | | | | | | | | | | | | |
| | Number of intolerant species in littoral zone | | | | | | | | | | | ■ | |
| | Number of intolerant species | ■ | ■ | ■ | ■ | ■ | ■ | | ■ | ■ | ■ | ■ | ■ |
| | Number of tolerant species | | | | | | | | ■ | ■ | | | ■ |
| | % individuals intolerant species | | | | | | | ■ | | | | | ■ |
| | % biomass tolerant species | | | | | | | | | | ■ | | ■ |
| | % biomass intolerant species | | | | | | | | | | | | ■ |
| | % individuals tolerant species | ■ | | ■ | ■ | ■ | ■ | ■ | | | | | |
| *TROPHIC COMPOSITION* | | | | | | | | | | | | | |
| | % individuals of benthic insectivores | ■ | | | | | | | | | | | |
| | % individuals omnivores (without shad & bluegill) | ■ | | | | | | | | | | | |
| | % individuals omnivores | | | ■ | ■ | ■ | ■ | ■ | ■ | | | | |
| | % biomass omnivores | | | | | | | | | | | | ■ |
| | % individuals piscivores (without shad & bluegill) | ■ | | | | | | | | | | | |
| | % individuals top carnivores | | | | | ■ | | ■ | ■ | | ■ | | |
| | Proportion biomass top carnivores | | | | | | | | | | | | ■ |
| | Proportion biomass insectivores | | | | | | | | | | | | ■ |
| | % individuals invertivores | | | ■ | ■ | | ■ | | ■ | | | | |
| | % individuals shad & bluegill adults | ■ | | | | | | | | | | | |
| | % individuals shad & bluegill YOY | ■ | | | | | | | | | | | |
| | Number of piscivore species | | | | ■ | | ■ | | | ■ | | | |
| | Number of generalist species | | | | | | | | | ■ | | ■ | ■ |
| | Number of invertivore species | | | | | | | | | ■ | | | |
| | Number of insectivore species | | | | | | | | | | | ■ | ■ |
| | Number of top carnivore species | | | | | | | | | | | ■ | |

| CATEGORY | METRICS | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***ABUNDANCE*** | | | | | | | | | | | | | |
| | Number of individuals | ■ | | ■ | ■ | | ■ | | | | | | |
| | Proportion of little benthic individuals | | | | | | | | | | | | ■ |
| | Number of individuals (without gizzard shad) | | | | | ■ | | | | | | | |
| | % non indigeneous biomass | | ■ | | | | | | | | | | |
| | % non indigeneous individuals | | ■ | | | ■ | | ■ | | | | | |
| | Number of native individuals | | ■ | | | | | | | | | | |
| | Biomass of native individuals | | ■ | | | | | | | | | | |
| | % individuals associated to the lake | | | | | ■ | | | | | | | |
| | % individuals of dominant species | | | | ■ | | ■ | | | | | | |
| | % individuals non native species | | | | | | | | ■ | | | | |
| | % biomass exotic species | | | | | | | | | | ■ | | |
| | % native individuals benthic | | | | | | | ■ | | | | | |
| ***CONDITION & REPRODUCTION*** | | | | | | | | | | | | | |
| Health | | | | | | | | | | | | | |
| | Proportion with disease, tumors, anomalies... | | | ■ | ■ | ■ | ■ | | | | | | |
| | Health indice | | | ■ | | | | | | | | | |
| Ecological niche | | | | | | | | | | | | | |
| | % individuals lithophilic | | | | | | | ■ | | | | | |
| | % individuals phythophilic | | | | | | | ■ | | | | | ■ |
| | Number phythophilic species in littoral zone | | | | | | | | | | | ■ | |
| | Number lithophilic species | | | ■ | ■ | | ■ | | | | | | |
| | Number phythophilic species | | | | | ■ | | ■ | | | | ■ | ■ |
| Migration | | | | | | | | | | | | | |
| | Number of migrating species | | | ■ | | | | | | | | | |

**1** : Karr & Dionne, 1991, **2** : Minns *& al.,* 1994, **3** : Jennings Fore & Karr, 1995, **4** : Hickman G & Mc Donough, 1996, **5** : Thoma, 1999, **6** : Mc Donough & Hickman, 1999, **7** : Jennings *& al.*, 1999, **8** : Whittier, 1999, **9** : Schulz *& al.*, 1999, **10** : Lyons *& al.,* 2000, **11** : Drake & Pereira, 2002, **12** : Drake & Valley, 2005.

## 1.2 European indices

In Europe, most countries have not yet included fish in their routine ecological assessment tools. Nevertheless, currently, two Scandinavian countries (Sweden, Finland) of the northern GIG (Geographical intercalibration Group) have finalised their IBI development. In the other GIGs few countries are well advanced: Flanders, Austria… (Rask, Olin & Ruuhijärvi 2009; Belpaire, Smolders *et al*. 2000; Gassner, Tischler *et al*. 2003; Jaarsma, Klinge & Pot 2007). Others European countries are also working on the development of fish based index (Germany, France, Ireland…) but these indices have not been published yet. All metrics used in application or in the last steps of development of an IBI in these European countries are sum up in table.2.

**<u>Table.2</u> List of metrics used in studies aiming to develop a fish-based assessment system adapted to different lake types of Europe and ichtyofauna.**

| CATEGORY | METRICS | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| ***SPECIES COMPOSITION*** | | | | | | | | |
| | Number of species | X | | | | | X | |
| | Number of native species | | X | X | | X | | |
| | Number of native species with an abundance index of 2 or 3 | | | | X | | | |
| | Species type | X | | | | | | |
| | Relative biomass of non native species | X | X | X | | | | |
| | Lost in native species | | | | X | | | |
| | % native species | | | | | X | | |
| | Indices stock/density for the dominant specie of a type | | | | X | | | |
| | Occurrence of indicator species | | | | | | | X |
| Tolerance/Intol. | Number of sensitive species to acidity | | X | X | | | | |
| | Biomass proportion of tolerant species to O2 decreased/ Low oxygen levels | | X | X | | | X | |
| | Lost of intolerant species | | | | X | | | |
| ***DIVERSITY*** | | | | | | | | |
| | Shannon-Weaver in number | | X | X | | X | | |
| | Shannon-Wheaver in biomass | | | | | X | | |
| | Simpson in number | | | | | X | | |
| | Simpson in biomass | | | | | X | | |
| ***TROPHIC COMPOSITION*** | | | | | | | | |
| | ratio in biomass of piscivorous/ non piscivorous | X | | | | | | |
| | % biomass of piscivore percids | | X | X | | X | | |
| ***ABUNDANCE*** | | | | | | | | |
| | Relative biomass of rudd | X | | | | | | |
| | Relative biomass of roach | X | | | | | | |
| | Relative biomass of bream | X | | | | | X | |
| | Relative biomass of perch + roach | | | | | | X | |
| | Recruitment and biomass of tench | X | | | | | | |
| | Recruitment and biomass of pike | X | | | | | | |
| | Biomass of native species | | | | | X | | |
| | Average weight of individuals | | | | | X | | |
| | Increased in biomass of native species | | | | X | | | |
| | Decreased in biomass of native species | | | | X | | | |
| | Total biomass | X | | | | | | |
| | CPUE number of native species | | X | X | | | | |
| | CPUE biomass of native species | | X | X | | | | |
| | BPUE | | | | | | | X |
| | NPUE | | | | | | | X |
| | Ration Perch/ Cyprinids in biomass | | | | | | | |
| | Proportion of Cyprinids in biomass | | X | X | | | | X |
| ***CONDITION & REPRODUCTION*** | | | | | | | | |
| | Number of reproductive native species | | | | X | | | |
| Ecological niche | % biomass of phytophilic species | | | | | | X | |

**1**: Belpaire, Smolders *et al*. 2000 (Belgium), **2**: Appelberg, Bergquist *et al.* 2000 (Sweden), **3**: Tammi, Lappalainen *et al.* 2001 (Finland), **4**: Gassner, Tischler *et al*. 2003 (Austria), **5**: Holmgren, Kinnerbäck *et al.* 2007 (Sweden), **6**: Jaarsma, Klinge & Pot (eds) 2007 (Netherlands), **7**: Rask, Olin & Ruuhijärvi 2009 (Finland).

# 2 Materiel and methods

## 2.1 Candidate metrics in the frame of the WISER project

The definition of a metric is described as a measurable variable or process that represents an aspect of the biological structure, function, or other component of the fish community and changes in value along a gradient of human influence.
In the frame of the WISER project, metrics tested are related to composition and abundance of fish communities. No metrics based on age structure and sensitive species have been studied. The assignment of each species to functional guilds is given in Annex 1.

The aim in this study was to have a wide choice of candidate metrics with several modes of calculation i.e. measured in different ways but assessing the same aspect of functional community. In most studies, few metrics are retained *a priori* from expert knowledge and scarce are the authors who conduct a rigorous procedure step by step with objective criteria and statistical procedures for selection (Hughes *et al.* 1998).
Here, once a previous exhaustive list of metrics performed, a selection has been done based on ecological knowledge, recommendations of the guidance and data available and then on statistical results. Based on these statements, the explanations of our choice in excluding the irrelevant metrics are described below.

### *2.1.1 Ecological knowledge (Trends of variation)*

Depending on the underlying biological hypotheses, a candidate metric should be proposed in relation to the expected variation with human disturbances (based on previous studies), and this would help for selecting the most relevant metrics. The list of metrics with the expected trends with different kind of degradation is presented in table.1.

### *2.1.2 Guidance requirements*

We examined (1) the distribution of values taken by each metric and (2) species composition on which each metric was calculated.

First, we proceed to the identification and exclusion of numerically unsuitable measures following the recommendation of Hering, Feld *et al. (*2006). Metrics with a narrow range of values or many outliers and extreme values were deleted (Figure.1). The native and lithophile abundance metrics show a low variability, which can be simply revealed by box-plots (Figure.2).

Secondly, species composition reveals some more irrelevant metrics, based on trophic guild and family. For the herbivore trait, only few individuals are present in only one natural lake, and for the Goodeid and Athenid (*Atherina boyeri* (Risso, 1810)) families, no species were identified in the natural lakes of the database.

All metrics considered after these statements are presented in grey in Table.1.

**Table.1 List of metrics to be tested on WISER database, and expected variation with degradation. In grey, the metrics we do not select *a priori* because of the irrelevance on our dataset and lack of data.**

| Metrics | Present In available Assessment system | shore line degradation | Eutrophication* | Water level regulation** | Answer to degradation |
|---|---|---|---|---|---|
| **SPECIES COMPOSITION** | | | | | |
| Total number of species | | - | - | - | ↓ |
| Total number of native species | EQR8 | - | ↑ | - | ↓ / ↑ |
| Number of cyprinids species | | - | - | - | ↓ |
| Number of native atherinids species | | - | - | - | -- |
| Number of native goodeids species | | - | - | - | -- |
| Number of native cyprinids species | | - | - | - | ↓ |
| Number of salmonids species | | - | - | - | ↓ |
| **DIVERSITY/ ABONDANCE** | | | | | |
| Relative biomass of native species* | EQR8 | - | ↑ | - | ↓ |
| Total biomass of native species | | - | - | - | ↓ |
| Relative number of native fish species* | EQR8 | - | ↑ | - | ↓ |
| Shannon-Weaver (numbers) | | - | - | - | ↓ |
| Simpson's Dn (numbers) | EQR8 | - | ↓ | ↓ | ↓ |
| Simpson's Dw (biomass) | EQR8 | - | ↑ | ↓ | ↓ / ↑ |
| Equitability index | | - | - | - | ↓ |
| Total biomass | | - | - | - | ↑ |
| Relative number of cyprinids | | - | - | - | ↓ |
| Relative biomass of cyprinids | EQR4 | - | - | - | ↑ |
| Ratio Perch/Cyprinids (biomass) | EQR8 | - | ↓ | - | ↓ |
| Relative number of salmonids (& biomass) | | - | - | - | ↓ / - |
| Relative number of percids (& biomass) | | - | - | - | -- |
| Total number of individuals | | - | - | - | ↓ |

| | | | | | |
|---|---|---|---|---|---|
| BPUE | EQR4 | - | - | - | ↓ / ↑ |
| CPUE | EQR4 | - | - | - | ↓ |
| Relative biomass of roach (& abundance) | | ↑ | - | - | ↑ |
| Relative biomass of rudd (& abundance) | | ↑ | ↓ | ↓ | ↓ |
| Relative biomass of bream (& abundance) | | - | ↑ | - | ↑ |
| Mean mass (from total catch) | | - | - | - | ↑ |
| Relative biomass of non native species | | - | - | - | ↑ |
| **TROPHIC GUILD** | | | | | |
| Relative biomass of piscivore percids | EQR8 | - | - | - | ↓ |
| Number of invertivore species | | - | - | - | ↓ |
| Number of omnivore species | | - | - | - | ↑ |
| Number of planctivore species | | - | - | - | -- |
| Number of strict piscivore species | | - | - | - | ↓ |
| Number of herbivore species | | - | - | - | -- |
| Relative number of omnivore (& biomass) | | - | - | - | ↑ |
| Relative number of invertivore (& biomass) | | - | - | - | ↓ |
| Relative number of piscivore | | - | - | - | ↓ |
| **REPRODUCTIVE GUILD** | | | | | |
| Number of phytophile species | | ↓ | ↓ | ↓ | ↓ |
| Relative number of phytophile | | ↓ | ↓ | ↓ | ↓ |
| Relative number of lithophile | | - | - | - | -- |
| Relative biomass of strict lithophile | | - | - | - | ↑ |
| Relative biomass of strict phytophile | | ↓ | ↓ | ↓ | ↓ |

* Increased of algal growth, reduced water clarity & loss of submerged vegetation

** Loss of inundated areas & emergent vegetation

* Total biomass (g) and total number of individuals of all native species, divided by the number of nets.

*(Source: Overview report of biological assessment methods used in national WFD monitoring programmes. FIRST DRAFT. Methods for lakes, exported from Waterview2-Database on assessment method for lakes, rivers, coastal and transitional waters in Europe and WISER work package 2.2- http://www.wiser.eu. Birk Sebastian, 2010.)*

It was decided to not integrate unknown species and hybrids in the calculation of metrics for functional guilds (*Abramis sp*., *Coregonus sp*., *Cottus sp*., *Mugilidae unknown*, *Cyprinidae unknown*, *Liza aurata* and *Liza ramada)* because the traits could be different from one species to another, even in the same family; Nevertheless they were kept for the calculation of species richness.
The metrics based on functional traits shared by less than three species were omitted. It was the case of the metrics: number of lithophile and number of phytophile species where 94 and 93% of the campaigns respectively were composed by only 1, 2 or 3 species.
This definitive set of metrics was expected to adequately reflect community richness and functioning.

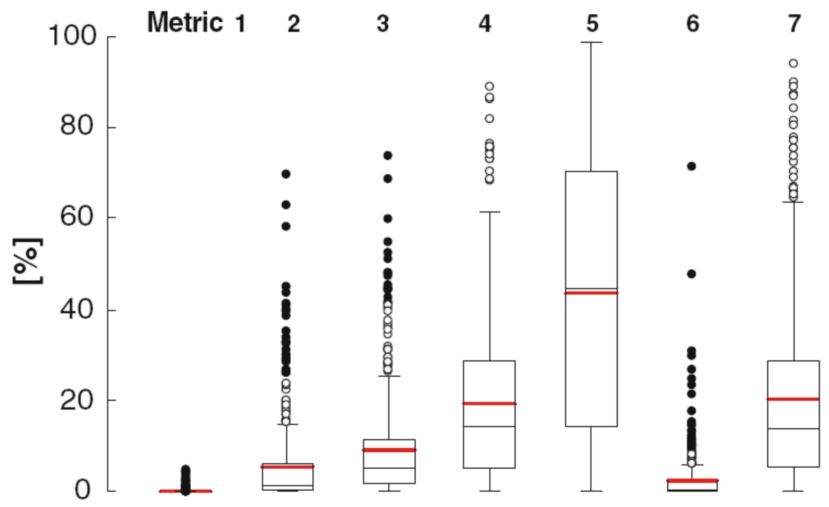



**Figure.1 Boxplots of numerically unsuitable metrics (1–3, 6) and suitable metrics (Metrics 4–5, 7). Circles indicate outliers (○) and extremes (●).** (*From Hering, Feld et al. 2006*)

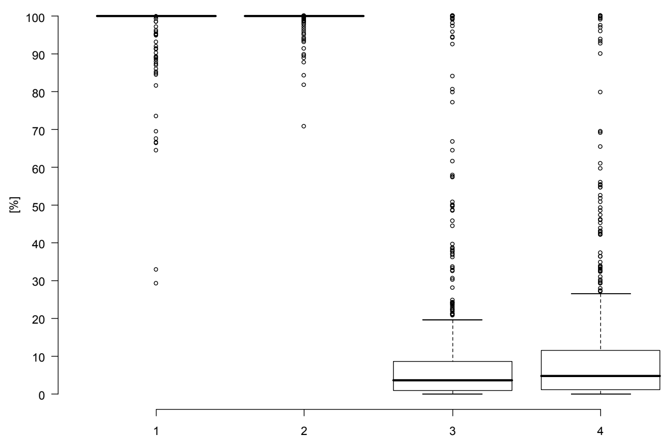



**Figure.2 Distribution of (1) Biomass of native species, (2) Number of native species, (3) Number of lithophile species and (4) Biomass of lithophile species in percentage for natural lakes sampled with CEN benthic standard.**

## 2.2 Study sites

### 2.2.1 The initial database

The entire European database created during the intercalibration process is composed of 2107 lakes; 'lakes' is the generic term used for both types of lentic ecosystems. They are divided in 1833 water bodies with natural origin, called "natural lakes" and 274 systems created artificially by damming, called "reservoirs". All these sites were sampled from 1993 to 2008 with 31 natural lakes added in 2009 in the frame of the WISER project. All fish data were asked in association to the respective environmental characteristics, climate variables and anthropogenic catchment-scale pressures available.
A large amount of sampling method are available in Europe, but to get a comparable dataset, we considered exclusively the lakes sampled with the CEN benthic multimesh gillnets (C.E.N 2005), which decreased the dataset to 1840 lakes: 1760 natural lakes and 80 reservoirs.
The primary uses of impounded waters (hydropower, flood control and issuance of drinking water) produce unnatural variation in water levels that impose to biota a stress additive to the environmental one. The fluctuations in flow rates are not available at European scale; therefore these reservoirs were not considered in our approach.

### 2.2.2 The final dataset used for analyses

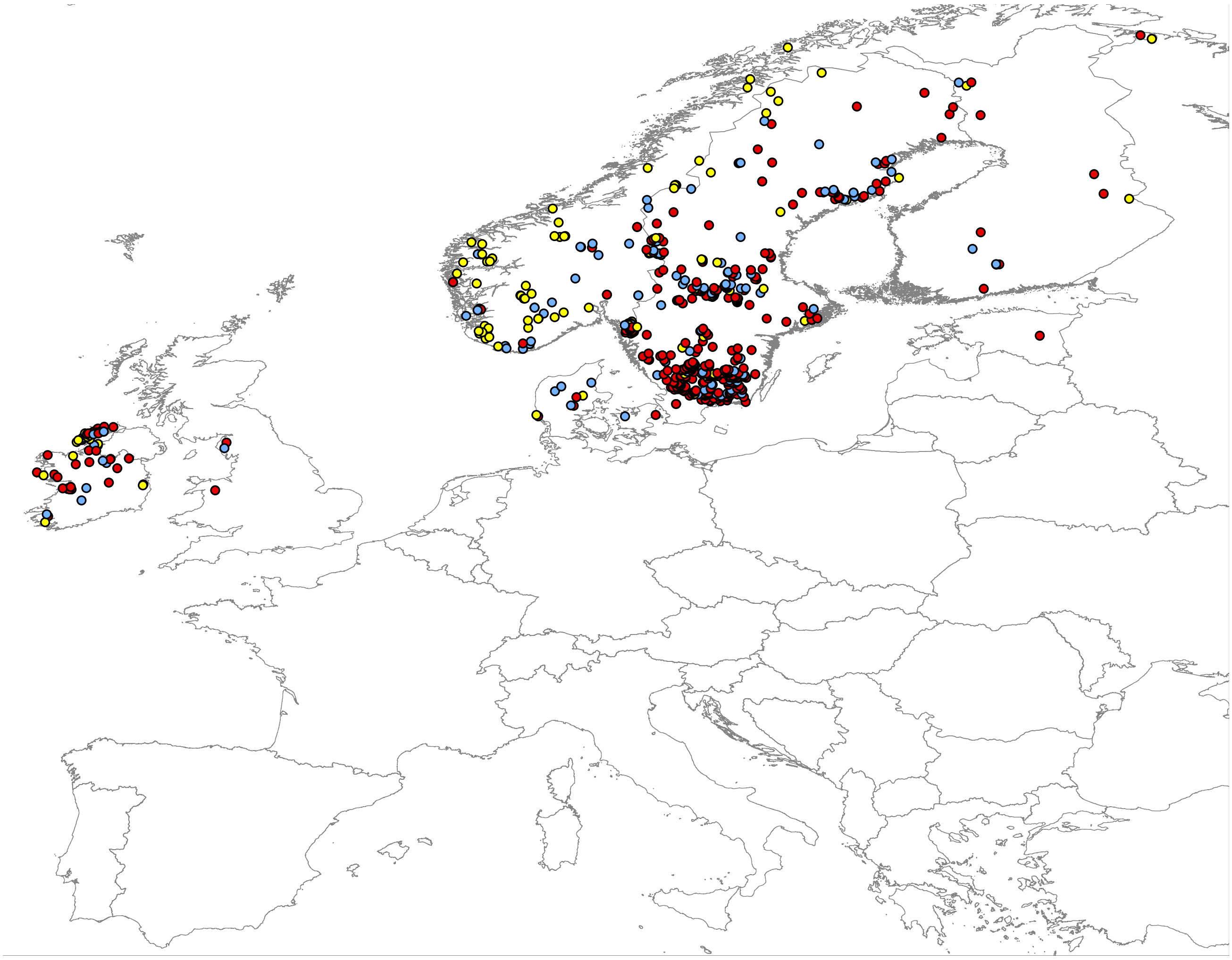
One campaign per lake was kept in the global dataset because of all repeated samplings located in Scandinavian countries (SE and NO) and also because of one year environmental data. Therefore, the last campaign of the time series sampled with CEN benthic multimesh gillnets was selected.
It is well known that species diversity is rather low in Europe (except in Danube basin) compared to the diversity of North America. Based on our dataset, 36.9% of the sampled lakes contain low diversity, i.e. less than 3 species (Table.2). These poor species lakes are all part of the Nordic GIG and mainly located in Scandinavian countries (Figure.3).

**Table.2 Distribution of the 1760 natural lakes related to species richness (RS).**

| RS | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Nb lakes | 139 | 190 | 321 | 310 | 209 | 188 | 122 | 84 | 67 | 58 | 31 | 20 | 12 | 3 | 2 | 2 | 1 | 1 |

In such conditions, as the efficiency of an index based on fish community structure with low species richness is obviously low, it was decided not considering lakes with less than 3 species (Schmedtje *et al*. 2009). A total of 1097 natural lakes emerged at European scale (Table.3a).

In this study, among these 1097 natural lakes, 419 were selected based on the availability of environmental parameters and pressures for these sites (Table.3b).
These lakes are mainly located in the Nordic GIG (Figure.4). At a GIG level, only the Nordic (NO) and the Central-Baltic (CB) ones are relevant for any statistical analyses, as only one and 12 lakes are present in the Mediterranean (MED) and alpine (AL) GIG respectively (Table.4).



**<u>Figure.3</u> Distribution map of natural lakes sampled with multi-mesh CEN standard method and respective species richness (RS) <= 3.**

**Table.3** Total number of lakes by Member State (MS) present in the global dataset (a) and in the dataset with all environmental parameters available (b).

| MS | Nb of lakes |
|---|---|
| Denmark | 73 |
| Estonia | 21 |
| Finland | 87 |
| France | 32 |
| Germany | 75 |
| Ireland | 41 |
| Italy | 4 |
| Norway | 7 |
| ROI/ NI | 4 |
| Slovenia | 2 |
| Sweden | 748 |
| UK | 3 |
| Total | 1097 |

| MS | nb lakes |
|---|---|
| Denmark | 41 |
| Estonia | 21 |
| Finland | 77 |
| France | 27 |
| Germany | 69 |
| Ireland | 34 |
| Italy | 1 |
| Norway | 0 |
| ROI/ NI | 3 |
| Slovenia | 0 |
| Sweden | 146 |
| UK | 0 |

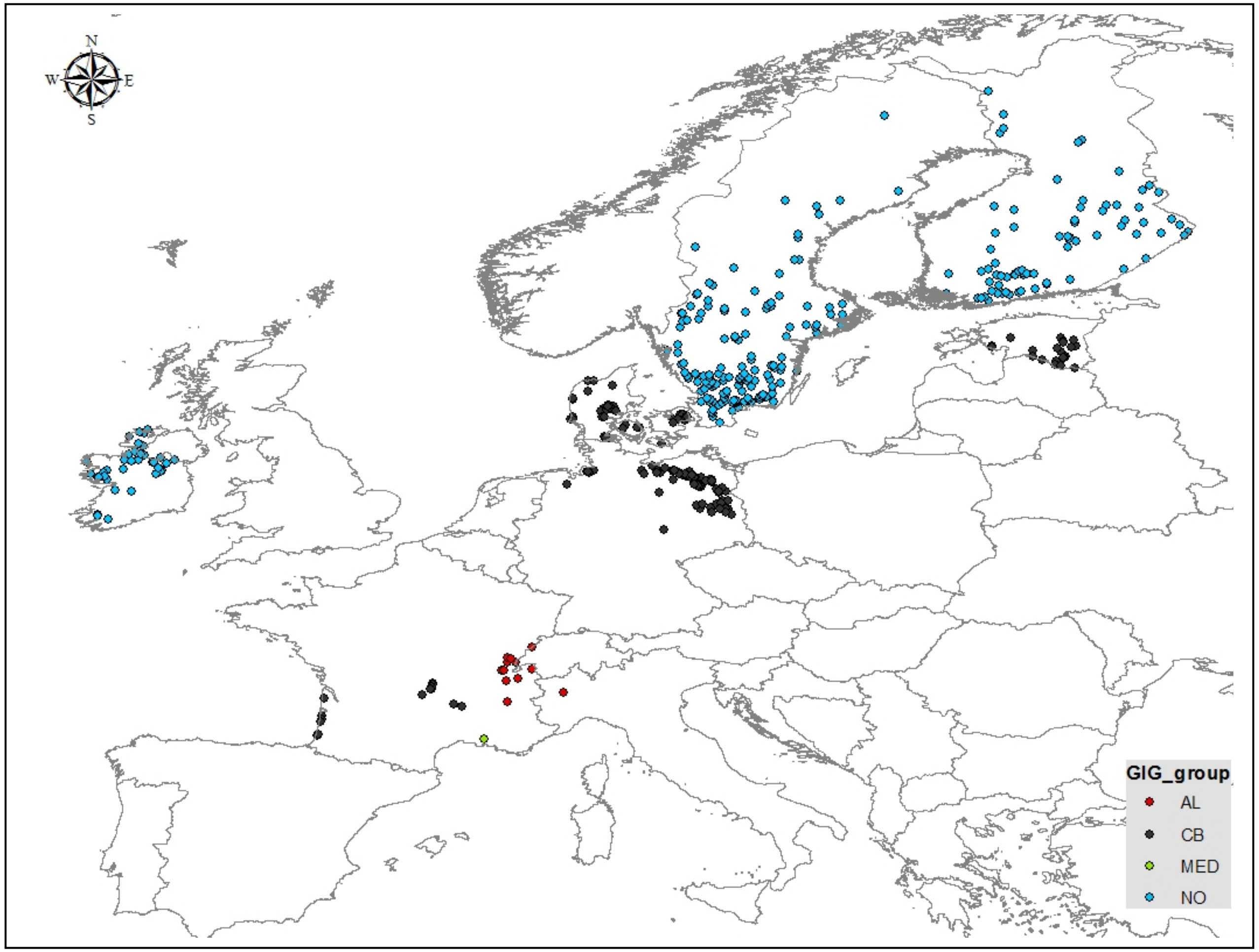



**Figure.4** Distribution of the 419 lakes used for the creation of the models among European map. GIGs (Alpine (AL), Central- Baltic (CB), Mediterranean (MED) and Nordic (NO)) are represented in colour.

**Table.4** Distribution by GIG (Alpine (AL), Central-Baltic (CB), Mediterranean (MED) and Nordic (NO) of the 419 lakes with all environmental parameters available.

| GIG_group | Nb lakes |
|---|---|
| AL | 12 |
| CB | 146 |
| MED | 1 |
| NO | 260 |
| Total | 419 |

## 2.3 Environmental parameters

Because we did not perform type-specific but site- specific analyses, impact assessment at broad spatial scales requires the consideration of environmental variables that are not modified by human activities and could well describe community structure. An efficient control of natural ecological patterns known to determine the variability of fish communities across sites is then needed. Hence, the variables given in Table 5 were included in the models. Maximum depth ($Z_{max}$) and Lake Area ($L_A$) are strong drivers of fish species richness (Barbour and Brown 1974; Eadie *et al.* 1986). Altitude ($Alt$) parameter can be related to isolation and climatic data (Godinho *et al.* 1998; Hinch *et al.* 1991; Magnuson *et al.* 1998; Tonn *et al.* 1990). No mountain lakes above 1500m were included because species richness is generally low; moreover, in these lakes, fish communities are generally strongly influenced by human introductions (Argillier *et al.* 2002) and fish is not considered as a relevant bioindicator to assess ecological status (Ministère de l'Ecologie et du Développement Durable, 2006).

Catchment area ($A_{DB}$) can be considered as a surrogate for habitat diversity upstream from the lake (Irz, Argillier *et al.* 2004). January to December mean yearly air temperatures ($T_{January}$ & $T_{December}$) were obtained from the climate CRU model (New *et al.* 2002). January and July mean temperature allowed to derive the following independent variable related to temperature requirements of living organisms (Daufresne and Boet 2007; Irz *et al.* 2007; Mason *et al.* 2008; Rathert, White *et al.* 1999).

(i) $AveT = (T_{January} \_ T_{December})/12$

(ii) $AmpT = T_{July} - T_{January}$

The geology (*G)* represents one of the ground characteristics of the Lake catchment area and is defined as calcareous or siliceous. Therefore it relates to the water chemistry and buffering capacity (Brousseau, Baccante *et al*. 1985; Alpay, Veillette *et al.* 2006). Consequently, this parameter is influencing lakes' productivity, i.e low specific richness in lakes with low pH (Koskenniemi *et al.* 1990; Matuszek & Beggs 1988 ; Rago & Wiener 1986). At a local scale, this parameter could also influence the nature of the species present (Rahel & Magnuson 1983).

These parameters were also retained because of their availability (for example mean depth was excluded because of too many unknown values). The square parameters of all environmental variables were also added in the model because of the polynomial character of the response.

Maximum depth, lake Area and catchment area were log-transformed for graphical display and analyses. A correlation between all natural parameters was performed to check their independence.

**Table.5 List of environmental parameters, with units, mean and range included in the models.**

| Parameters | Definition | units | mean | range |
|---|---|---|---|---|
| $Z_{max}$ | *Max depth* | *Meters (m)* | *17.07* | *0.17 - 110* |
| $L_A$ | *Lake area* | *Square Kilometers (km2)* | *5.77* | *0.05 - 116.50* |
| $A_{DB}$ | *Catchment area* | *Square Kilometers (km2)* | *139.01* | *0.05 - 10628.89* |
| *Alt* | *Altitude* | *Meters (m)* | *128.74* | *-1.00 - 1200* |
| *AveT* | *Average temperature* | *Degree Celsius (°C)* | *6.19* | *-2.15 - 14.04* |
| *AmpT* | *Amplitude temperature* | *Degree Celsius (°C)* | *19.18* | *8.5 - 30* |
| *G* | *Geology* | *Siliceous or Calcareous (Si & Ca)* | -- | -- |

## 2.4 Anthropogenic pressures

Three catchment-scale pressures and two local-scale variables were collected for the overall dataset (N = 419) but only two of them were included in the models. Within each lake catchment, we estimated (i) natural land cover percentages and (ii) acidification pressure. The latter pressure is mainly based on a direct measure of the pH on the lake and on an expert opinion if the pH was below 6 to know if the acidification was a pressure or not. Both parameters conduct to a yes/no assessment.
We assumed that these two variables derive from GIS/ experts opinions or direct measures on the lake reflecting the anthropogenic pressures undergone by lakes at the catchment scale. The percentage of natural on the catchments areas is considered as the reverse of the pressure and was arcsine-square-root transformed.

## 2.5 Statistical approach

### 2.5.1 Reference conditions

The first step to build a fish-based index is the agreement on reference conditions (RC). Following the Guidance: "*High status or Reference Conditions should reflect a state in the present or in the past corresponding to very low pressure, without the effects of major industrialisation, urbanisation and intensification of agriculture, and with only very minor modification of physico-chemistry, hydromorphology and biology*». They could be determined from existing sites, from models, from paleolimnological reconstructions, from expert judgement or from some combination of these (WFD; 2000/60/EC).

In this study, reference conditions were obtained from two different ways:
- Reference sites (88 natural lakes) identified during the intercalibration process, with general reference thresholds established on the level of anthropogenic pressure and proved by expert judgement (depending on their relevance for the lake ecosystem), and
- Hindasting approach where reference conditions are set to establish a "natural trophic state" by modelling.

*Approach based on reference sites*
During the intercalibration process, a set of reference criteria and thresholds with no or minor human impact on the environment was performed (Table.6). A total of 88 sites appeared, distributed among member state as following (table 7). Once the reference model build (on

reference sites), it is applied on all sites (reference+ disturbed) to get the reference conditions (Figure.5a), i.e. values that disturbed sites should get if they were in reference.

**Table.6 List of criteria and reference conditions established in Ranco by all member states in Europe.**

| | Criteria | Thresholds |
|---|---|---|
| **Eutrophication** | % land use « natural » | >80% or class 1<br>(Rejection threshold = 70%) |
| | Population density | 10 hab.km$^{-2}$ or class 1<br>(Rejection threshold at 30hab/km²) |
| | Ptot (µg/l) | 20<br>(Rejection threshold at 50µg/l) |
| **Acidification** | pH | > 6 & if <6 : expert jugement |
| **Hydromorphology** | Impoundment upstream | Expert judgment |
| | Loss of connectivity downstream | Expert judgment |
| | Water level fluctuation | Expert judgment |
| | Shoreline Bank modification | Expert judgment |
| **Activity on the lake** | Urbans/industrials discharge | Expert judgment |
| | Stocking | Expert judgment |
| | Biological or chemical manipulation | Expert judgment |
| | Fishing activities | Expert judgment |
| | Others activities | Expert judgment |

**Table.7** Number of reference lakes among the 419 lakes of the dataset.

| MS | Reference IC |
|---|---|
| Estonia | 6 |
| Finland | 27 |
| Ireland | 1 |
| Sweden | 38 |
| France | 5 |
| Germany | 11 |

This method was truly criticized because of the threshold settled, which were considered too high by some GIGs. Consequently, the "hindcasting" method was also developed.

*Reference condition from the hindcasting model method*

The Hindcasting method removes the need to select and classify reference sites, eliminating a potential bias in Lake bioassessment. This approach is different by the point that the model includes anthropogenic factors as predictor variables in addition to environmental parameters.

Some pressures are injected in the model during his creation and then set to zero to get site-specific values expectations in the absence of anthropogenic pressures or reference conditions (Figure.5b). In our study, to predict reference conditions, no acidification pressure was taken into account and the CLC natural was set to 90%.

If the model is used to recalculate fish community metrics, once the pressures set to zero, the model output represents expected fish community in that lake in the absence of pressure (Baker, Wehrly *et al.* 2005).

One hypothesis of the Hindcasting method is the assumption to have a dataset covering a large scale of pressure acting on lakes, meaning that when new lakes will be added *a posteriori*, results won't be supposed to change.

### 2.5.2 Variable selection

First, classic monotonic transformations of the metrics were used to meet the requirements of the linear model (normality, linearity): count (abundance, richness) and biomass metrics were log-transformed; proportion metrics were arcsine-square root transformed, whereas diversity indices were kept raw. The abundance metrics were computed two different ways: (i) total number/biomass of individuals sharing a trait divided by the total number/biomass of individuals and (ii) number of individuals sharing a trait caught by unit effort.

All candidate metrics were then scaled and each metric was regressed using a stepwise linear multiple regression analyses based on the Akaike Information criterion. The three sets of predictors involved in the generalised linear regression model (GLM) were: the seven environmental variables and their squared value (for non linear response), the natural land use (NLU) and the acidification pressure (pressacid2O). Selection of predictors was from the complete models below, based on the 88 reference sites (1) or on the 419 sites for the hindcasting procedure (2):

(1) Observed metric ~ $L_A + A_{DB} + Z_{max}$ + Alt + AveT + AmpT + G

(2) Observed metric ~ natural environment + pressures
Observed metric ~ ($L_A + A_{DB} + Z_{max}$ + Alt + AveT + AmpT + G) + NLU + pressacid2O

Where $L_A$ *(Lake area)*, $A_{DB}$ *(Drainage basin area)*, $Z_{max}$ *(Max depth)*, Alt *(Altitude)*, *AveT (Average temperature)*, AmpT *(Amplitude temperature)*, G (*Geology*), NLU (*Natural land cover*) and pressacid2O (*Acidification pressure*),

With this method, only the relevant environmental parameters explaining the model based on reference sites will be kept whereas both environmental parameters and the pressures will be integrated in the hindcasting model.
To know the real participation of each variable included in the model, a hierarchical partitioning is usually used (Chevan *et al. 1991*). The variance part of each variable explaining the model was then given.

### 2.5.3 Metric normalisation

Once the reference conditions values obtained by the hindcasting method, they were compared to the observed ones, present in the dataset. For each metric, the difference between the observed and the predicted values, corresponding to the residuals of the models and here called "Metric_result" was calculated (Figure.5).

The WFD explicitly states that the purpose of expressing results as an EQR is to provide a common scale of ecological quality.
The use of EQRs is prescribed in Annex V, 1.4.1 of the WFD and in the CIS guidance on monitoring. It is defined as follows: "Ecological Quality Ratio (EQR) - The ration between the value of the observed biological parameter for a given surface water body and the expected value under reference conditions. The ration shall be expressed as a numerical value between 0 and 1…"

For metric's normalisation here, as explained in the WISER guidelines, the upper and lower anchors which mark the indicative range of a metric are empirically set and defined as "1" (upper anchor) and "0" (lower anchor), respectively.
The upper anchor corresponds to the upper limit of the metric's value under reference conditions. The lower anchor corresponds to the lower limit of the metric's value under the worst attainable conditions (minimum observed metric value).
Each metric result was translated into a value between 0 and 1 (Ecological Quality Ratio) from the "Metric_result" first obtained, using the following formula:

$$\text{Value} = \frac{\text{Metric\_result} - \text{Lower\_Anchor}}{\text{Upper\_Anchor} - \text{Lower\_Anchor}}$$

for metrics decreasing with increasing impairment, and

$$\text{Value} = 1 + \frac{\text{Metric\_result} - \text{Lower\_Anchor}}{\text{Upper\_Anchor} - \text{Lower\_Anchor}}$$

for metrics increasing with increasing impairment.

High ecological status is represented by values close to one and bad ecological status by values close to zero.

### 2.5.4 Metrics selection

According to the WISER guidelines for indicator development (Hering *et al.* 2009), "*An ideal metric should be responsive to stressors, have a low natural variability, provide a response that can be distinguished from natural variation, and be interpretable (Hering et al. 2006). A candidate metric's results must show a significant correlation to the stressor gradient. This correlation can be positive or negative, either across the whole stressor gradient or measured for a part thereof (e. g. only moderate to high quality sites). Metrics fulfilling this criterion are, in principal, suited to assessing the degradation of the ecosystem type and can be selected as candidate metrics.*"

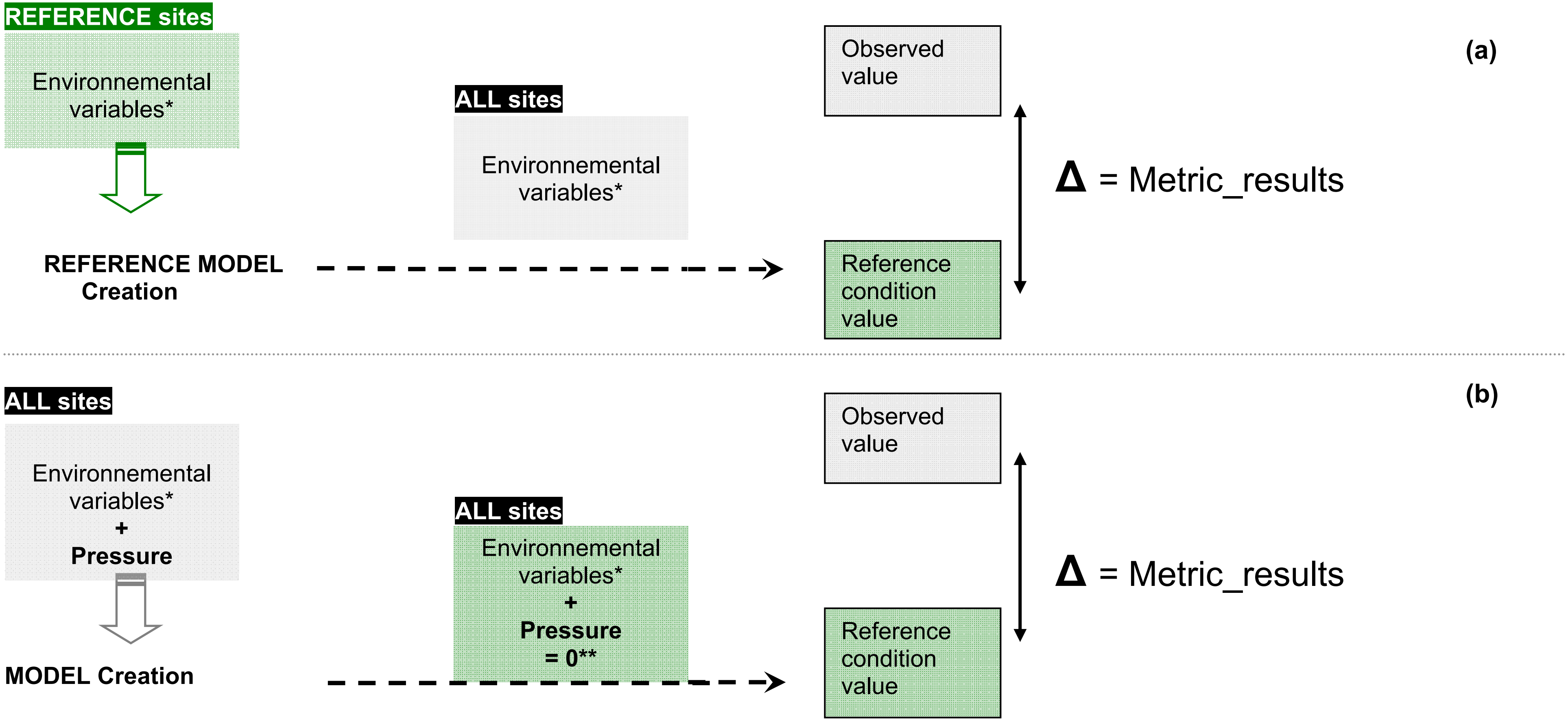
A large amount of authors described possible approaches for metric selection (e.g. Barbour *et al.* 1992, 1999; Karr and Kerans 1992; Karr and Chu 1999; Buffagni *et al.* 2004; Hering *et al.* 2004; Ofenböck *et al*. 2004; Vlek *et al.* 2004; Pont *et al.* 2006), but here, the metric was considered only if: (i) the adjusted R-squared of the resulting model was higher than 0.2 and the variation trend of the metric was conformed to the bibliography, (ii) Spearman's rank correlation analysis between metrics and natural land cover was significant (>0.2) and (iii) the metric response to the stressor gradient shows a narrow range of distribution. For validation, boxplots values on both reference and disturbed sites were checked to be statistically different by the Mann-Whitney test (P≤ 0.05). The other metrics were excluded.

Analyses were computed with R software (R Development Core Team 2007) and performed at the European scale.

### 2.5.5 Metrics compilation and boundaries setting

For integration of a metric in a multi metric index (MMI), the criterion to be met is the non redundancy of the latter. Most studies used a correlation test as an indicator of redundancy (Hugues *et al.* 2004, Mc Cormick *et al.* 2001 and Oberdorff *et al.* 2002). A spearman's rank correlation test was then performed to identify these metrics.

Compilation of metrics to build an index could be done by different ways, but here as first convenient approach, we chose a simple addition of the core metrics (expressed in EQR).
Two methods were used to set the H/G boundary: the index value on reference sites was selected and also the index response to the pressure gradient. Tests were performed with different lake percentile of index distribution: 10, 15 or 20%. For all other boundaries, 2 methods were also used: by making 4 homogeneous groups for values below H/G boundary or by clustering method (here k-means algorithm).




* max depth/ lake area/ Altitude/ average temperature/ amplitude temperature/ catch area/ geology
** Percentage CLC "Natural" = 90

**Figure.5** Concepts of reference condition set based on reference sites (a) and the hindcasting method (b).

# 3 Results

## 3.1 Study sites and natural parameters

None of the natural variables show high correlation, except average and amplitude temperature. For these variables, the correlation's coefficient is "-0.80" but they were kept both because they are ecologically relevant and provide different information (Table.8).

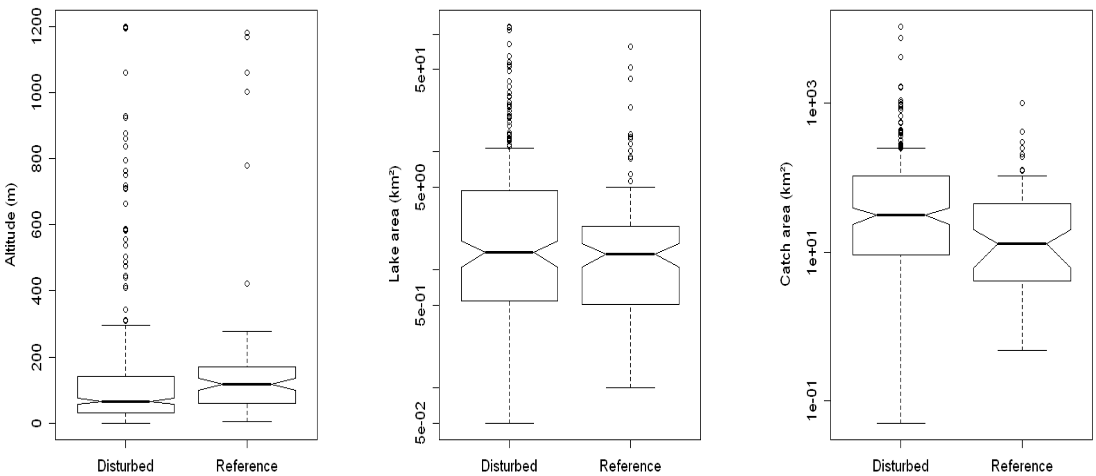
**Table.8** Correlation table of environmental parameters selected (with significance).

| | $Alt$ | $L_A$ | $Z_{max}$ | $A_{DB}$ | $AveT$ | $AmpT$ |
|---|---|---|---|---|---|---|
| Altitude ($Alt$) | 1.00 | | *** | * | *** | *** |
| lake_area ($L_A$) | -0.01 | 1.00 | *** | *** | . | *** |
| max_depth ($Z_{max}$) | 0.28 | 0.42 | 1.00 | *** | | . |
| catch_area ($A_{DB}$) | -0.11 | 0.70 | 0.21 | 1.00 | | |
| ave_temperature ($AveT$) | -0.56 | -0.08 | -0.05 | 0.03 | 1.00 | *** |
| amp_temperature ($AmpT$) | 0.41 | 0.19 | 0.08 | 0.01 | -0.80 | 1.00 |

*Signification codes: 0 '***' 0.001'**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

## 3.2 Reference sites

The 88 reference sites cover a wide range of values among the environmental parameters (Figure.6). The distribution of these sites mainly reflects the overall dataset, except for the average temperature and thermal amplitude. The median of the average temperature is lower (i.e. colder) on reference sites than on the disturbed ones, with extreme values going below 0°C. The median for thermal amplitude is higher, with a range going to 30°C.



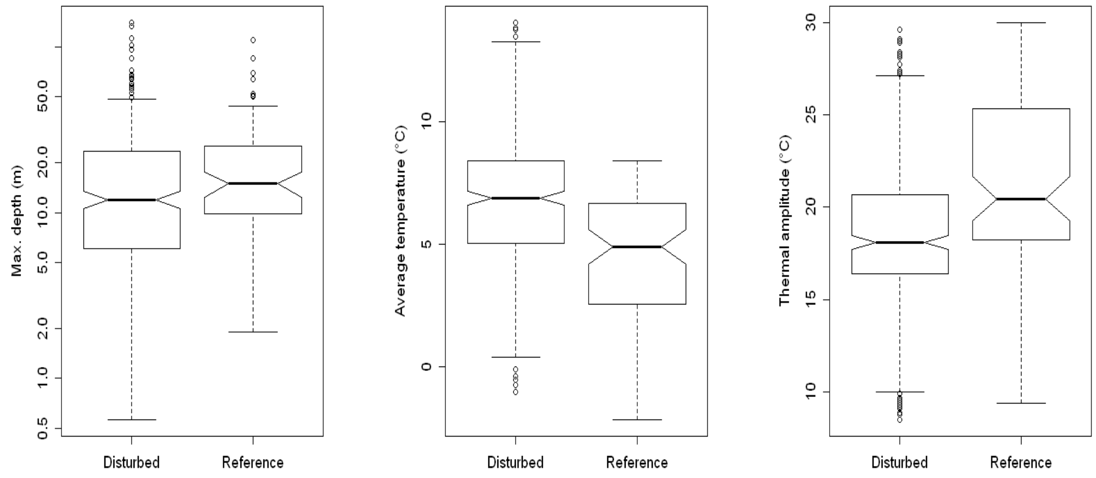



**Figure.6** Boxplots of environmental parameters (Altitude, Lake Area, Catchment area, Max depth, Amplitude temperature and Thermal amplitude) for disturbed and reference natural lakes.

## 3.4 Selection of relevant metrics

After a first selection based on the adjusted R-squared of the resulting model and the trend of variation known for the metrics (not detailed here), 10 and 12 specific fish traits (e.g. 18 and 24 metrics) displayed a significant response (>0.2) to the natural land cover (inverse of the pressure) for both respectively i.e. reference sites and hindcasting approaches (Table.9).

**Table.9** Spearman's rank correlation between natural land use (NLU) and Relative/ Absolute abundance metrics (each cell is a combination trait & calculation mode) for the model built on reference sites (a) and for the hindcasting model (b).

**a)**

| | Absolute Number (CPUE) | Relative Number | Absolute Biomass (BPUE) | Relative Biomass |
|---|---|---|---|---|
| Planctivore | -0.28 | | -0.37 | |
| Omnivore | -0.23 | | -0.35 | |
| Specialist | | | -0.27 | |
| Perch | -0.32 | -0.21 | -0.36 | |
| Roach | | | -0.31 | |
| Rudd | | 0.21 | | 0.24 |
| Cyprinidae | | | -0.37 | |
| Salmonidae | 0.29 | 0.31 | | 0.29 |
| Percidae | -0.39 | | | |
| All individuals | -0.36 | | -0.42 | |

**b)**

| | Absolute Number (CPUE) | Relative Number | Absolute Biomass (BPUE) | Relative Biomass |
|---|---|---|---|---|
| Planctivore | -0.43 | | -0.48 | -0.28 |
| Omnivore | -0.42 | | -0.49 | -0.28 |
| Invertivore | | | | 0.23 |
| Piscivore | | | -0.29 | 0.31 |
| Specialist | | | -0.23 | |
| Perch | -0.35 | | -0.31 | |
| Bream | | | -0.24 | |
| Roach | -0.32 | | -0.46 | -0.26 |
| Cyprinidae | | | -0.47 | -0.28 |
| Salmonidae | 0.22 | 0.29 | | 0.21 |
| Percidae | -0.40 | | | |
| All individuals | -0.54 | | -0.53 | |

To exclude the redundant metrics, correlations between metrics each other are needed for the reference sites and for the hindcasting models (Table 10).

The non-redundant metrics (<0.8) that show the best correlation and narrow distribution to the stressor gradient were CPUE and BPUE. These two metrics are already part of some national assessment systems developed in Europe. Consequently, they were chosen to describe the procedure below and compiled in a biological index. They were developed on both models (hindcasting and reference sites).
The other metrics show large distribution ranges of response to the pressure gradient (Figure.7). Setting class boundaries in such a relationship could contribute to the misclassification of lots of sites. Moreover, combination of some of these metrics with the CPUE and BPUE tends to decrease the final correlation score to the stressor gradient. Consequently, at that stage, no other metrics have been included in the common index. Nevertheless, one question that arises is the absolutely necessity to keep the higher correlation score to build an efficient MMI.

**Table.10** Spearman's rank correlation (>0.8) between the potential relevant metrics for the model applied on reference sites (a) and the hindcasting model (b).

**a)**

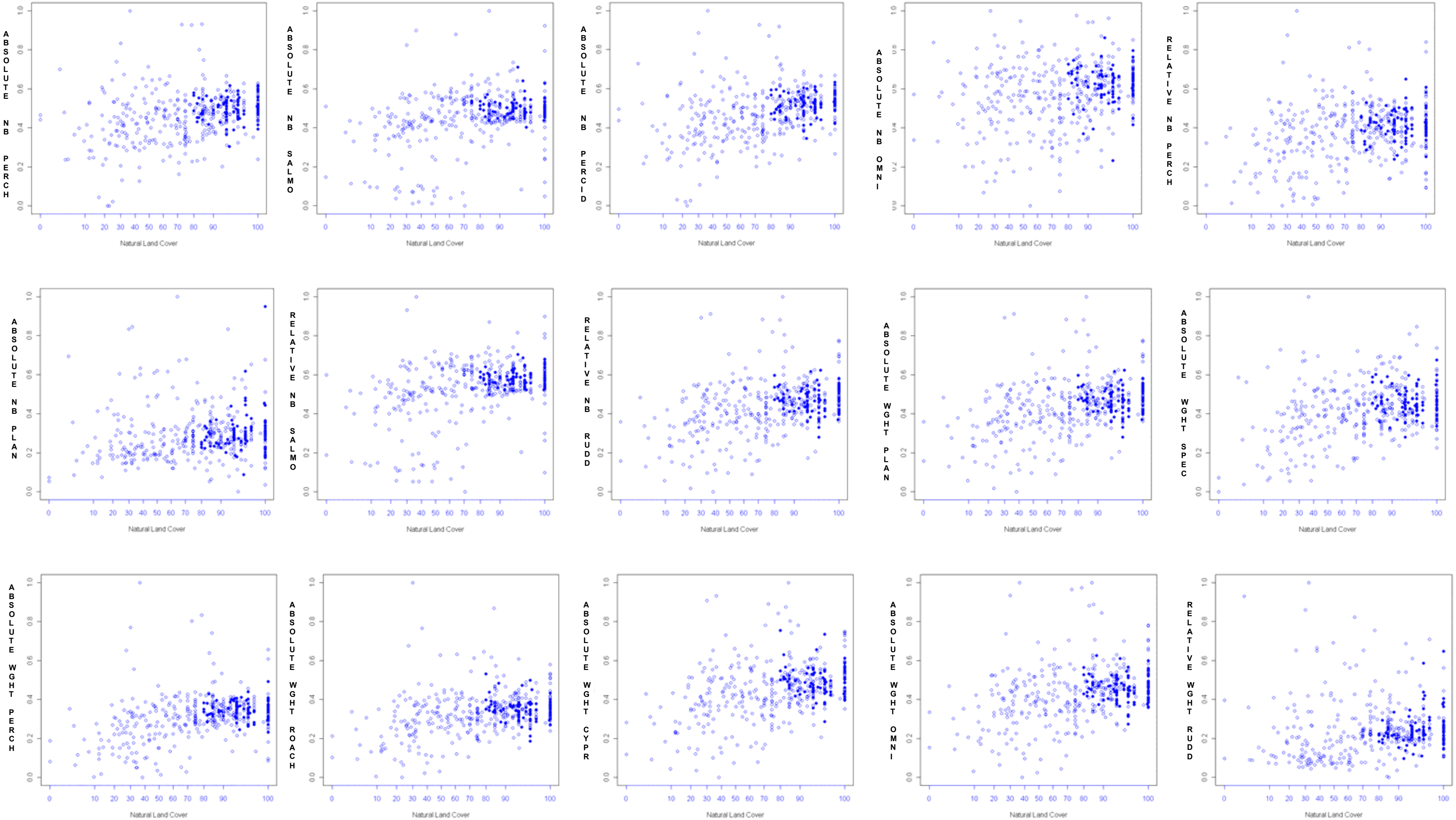
| | | Absolute number | | | |
|---|---|---|---|---|---|
| | | Planctivore | Cyprinidae | Salmonidae | Percidae |
| Absolute number | Omnivore | 0.95 | 0.9 | | |
| | Perch | | | | 0.96 |
| | CPUE | 0.81 | | | 0.82 |
| Relative number | Salmonidae | | | 0.96 | |

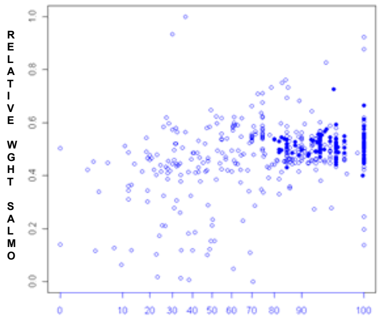
| | | Absolute number | Relative number | | Absolute biomass | | |
|---|---|---|---|---|---|---|---|
| | | Salmonidae | Rudd | Salmonidae | Roach | Omnivore | Cyprinidae |
| Absolute biomass | Omnivore | | | | | | 0.91 |
| | Roach | | | | | 0.85 | 0.83 |
| | Planctivore | | | | 0.86 | 0.91 | 0.95 |
| Relative biomass | Rudd | | 0.8 | | | | |
| | Salmonidae | 0.9 | | 0.93 | | | |

**b)**

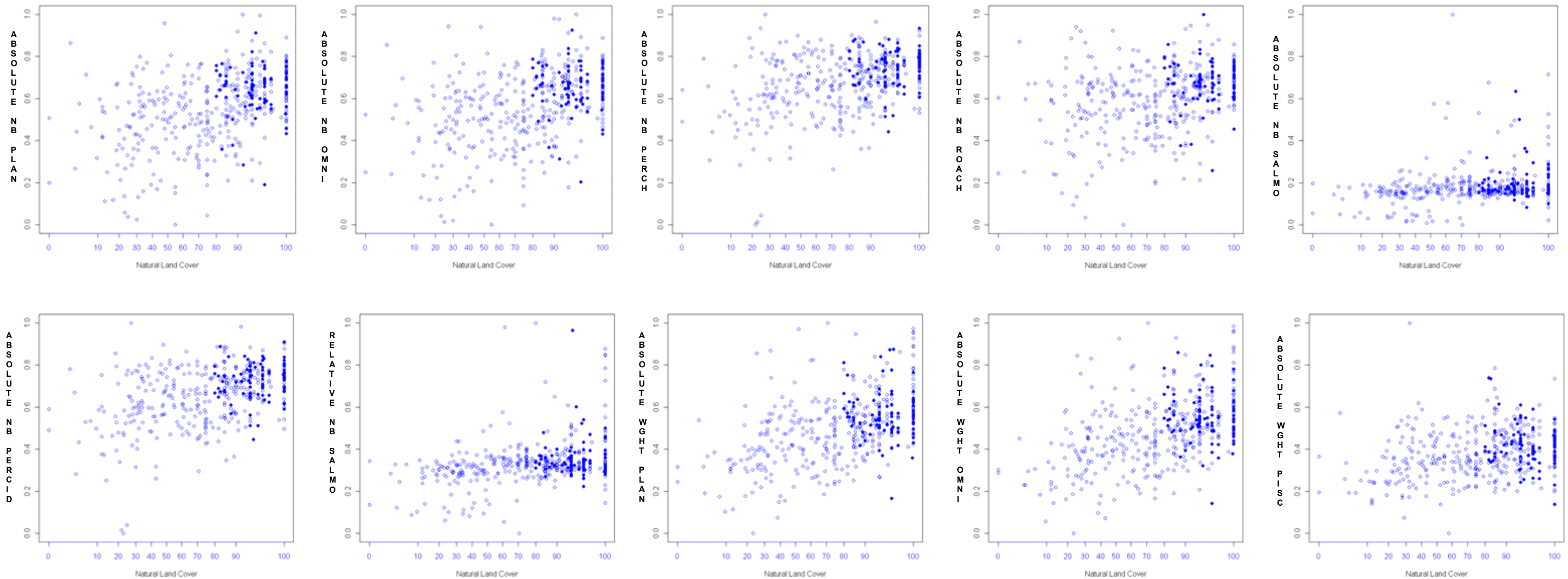
| | | Absolute number | Absolute biomass | | | Relative biomass | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Salmonidae | Planctivore | Roach | Omnivore | Piscivore | Planctivore | Roach | Cyprinidae | Salmonidae |
| Relative number | Salmonidae | 0.83 | | | | | | | | 0.91 |
| Absolute biomass | Roach | | 0.83 | | | | | 0.82 | | |
| | Cyprinidae | | 0.96 | 0.81 | 0.93 | | | | | |
| | Omnivore | | 0.93 | 0.85 | | | | | | |
| Relative biomass | Invertivore | | | | | 0.86 | 0.9 | | 0.88 | |
| | Planctivore | | | | | 0.93 | | | | |
| | Cyprinidae | | | | | 0.94 | | | | |

**(a)**


ABSOLUTE NB PERCH
ABSOLUTE NB SALMO
ABSOLUTE NB PERCID
ABSOLUTE NB OMNI
RELATIVE NB PERCH
ABSOLUTE NB PLAN
RELATIVE NB SALMO
RELATIVE NB RUDD
ABSOLUTE WGHT PLAN
ABSOLUTE WGHT SPEC
ABSOLUTE WGHT PERCH
ABSOLUTE WGHT ROACH
ABSOLUTE WGHT CYPR
ABSOLUTE WGHT OMNI
RELATIVE WGHT RUDD
Natural Land Cover


RELATIVE WGHT SALMO

(b)


ABSOLUTE NB PLAN
ABSOLUTE NB OMNI
ABSOLUTE NB PERCH
ABSOLUTE NB ROACH
ABSOLUTE NB SALMO
ABSOLUTE NB PERCID
RELATIVE NB SALMO
ABSOLUTE WGHT PLAN
ABSOLUTE WGHT OMNI
ABSOLUTE WGHT PISC
Natural Land Cover

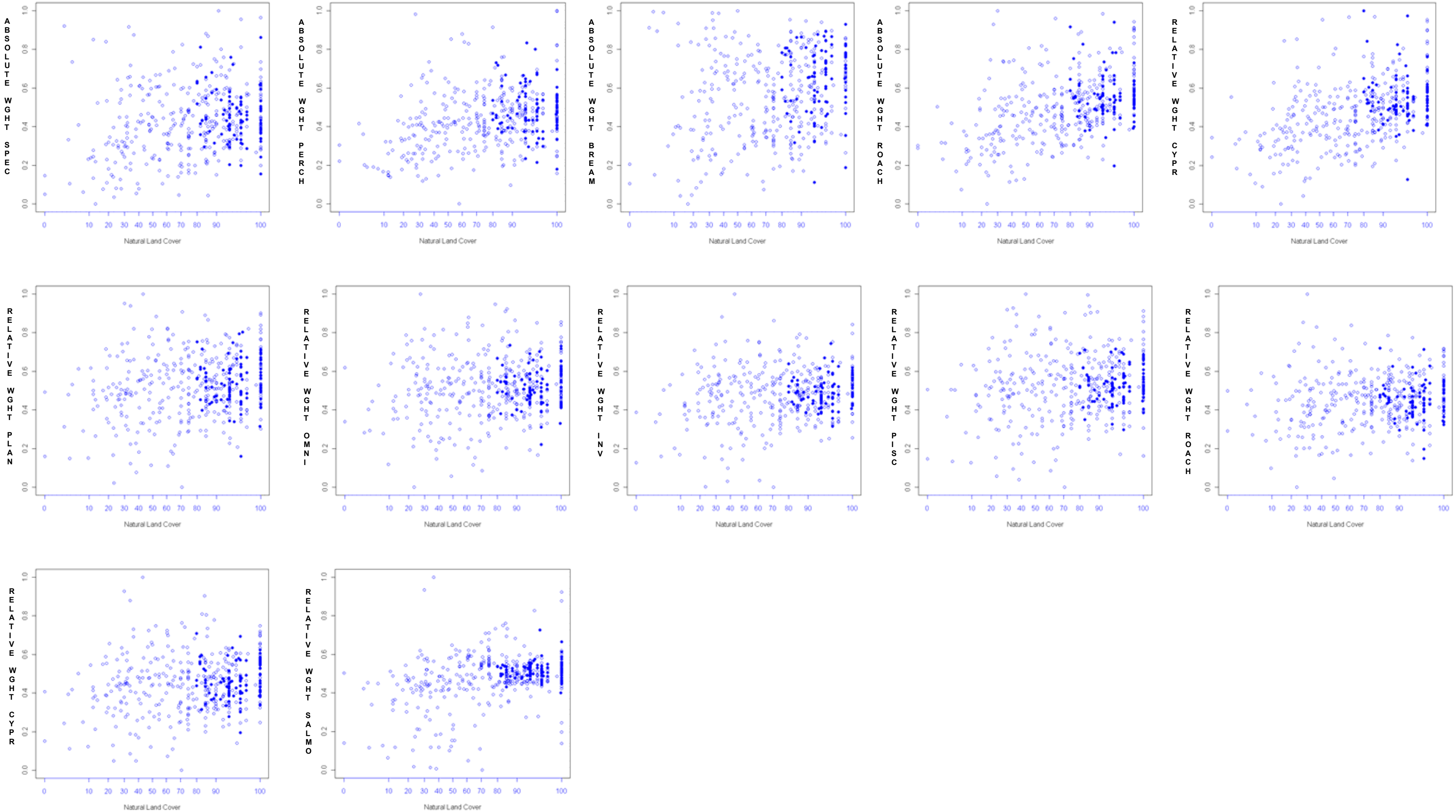



**Figure.7** Distribution range of the candidate metrics' EQR values against the natural land cover for the model applied on reference sites (a) and the hindcasting one (b).

## 3.5 Relevant metrics on reference sites

### 3.5.1 CPUE

All models included at least one significant coefficient for environmental parameters, thereby confirming that environmental patterns have to be taken into account when studying at a broad-scale relationship between fish metrics and anthropogenic pressures (Table.11). This is particularly shown here for the CPUE metric calculated on reference model, with 46.21% of the variance explained by environmental parameters such as the max depth ^2, average temperature ^2 and amplitude temperature, lake area ^2, catch area and amplitude temperature^2, catch area ^2, altitude, altitude^2 and geology in a lesser extent.

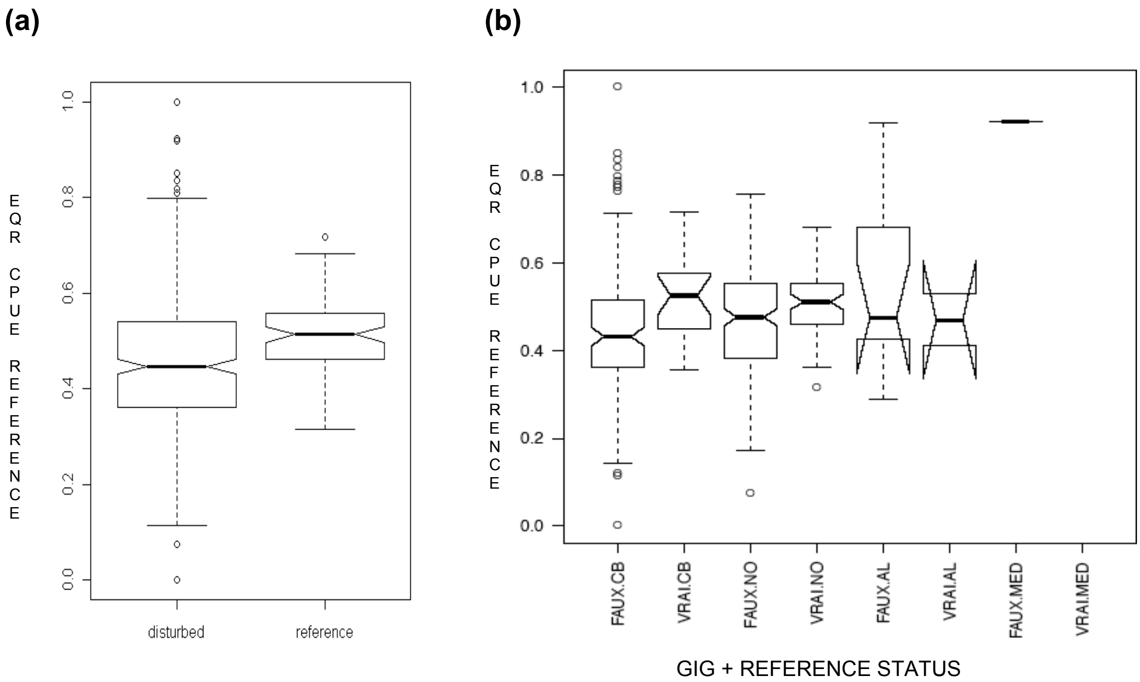
The correlation between EQR values of the CPUE metric and CLC natural defined in application of the reference sites method is less important than those determined by the hindcasting method (around 36%).

The comparison of EQR boxplots of the metric by GIG between impacted and reference sites is shown Figure.8. At the European scale, mean EQR value for disturbed sites is significantly lower than on reference sites (*p-value = 2.726e-05*). The Mann-Whitney test reveals a significant difference between the mean EQR values of the metric on impacted and reference sites for the Nordic and the Central Baltic GIGs (*p-value(NO)* = 0.01079 and *p-value(CB) = 0.004384*). This difference is not so clear in the others GIGs but only 1 and 12 lakes are present in the Mediterranean and the Alpine GIG respectively. Most of the lakes in the Nordic GIG are located in Sweden; no difference appears because most of the lakes in this region have an important natural land cover on their catchments, also the disturbed sites (Figures 9 & 10).

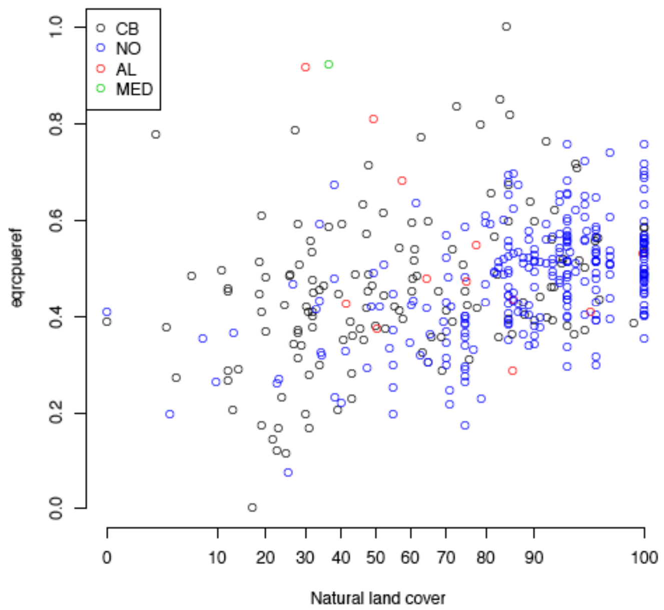
**Table.11** Results of the stepwise multiple linear regressions for the CPUE metric, done on parameters included in the "reference sites" model with the coefficient and significance associated.

| | Coefficient | Significance |
|---|---|---|
| (Intercept) | -3.78E+00 | *** |
| I(log10(max_depth)^2) | -2.38E-01 | *** |
| I(log10(lake_area)^2) | 1.22E-01 | * |
| log10(catch_area) | 2.67E-01 | * |
| I(log10(catch_area)^2) | -8.81E-02 | . |
| Altitude | -7.16E-04 | |
| I(Altitude^2) | 6.91E-07 | |
| I(ave_temperature^2) | 1.05E-02 | ** |
| amp_temperature | 1.67E-01 | ** |
| I(amp_temperature^2) | -2.73E-03 | * |
| geolsiliceous | 1.48E-01 | |

*Signification codes: 0 '***' 0.001'**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*




**<u>Figure.8</u>** Boxplots of the EQR values obtained by modelling on reference sites for the CPUE metric including all sites **(a)** and by GIG **(b)** where *Alpine (AL), Central- Baltic (CB), Mediterranean (MED) and Nordic (NO)*. Reference sites are indicated as “reference” or “VRAI” in the reference status and disturbed sites as “disturbed” or “FAUX”.




**<u>Figure.9</u>** Relation between the EQR values obtained by modelling on reference sites for the CPUE metric and the Natural land cover by GIG (*Alpine (AL), Central- Baltic (CB), Mediterranean (MED) and Nordic (NO))*.

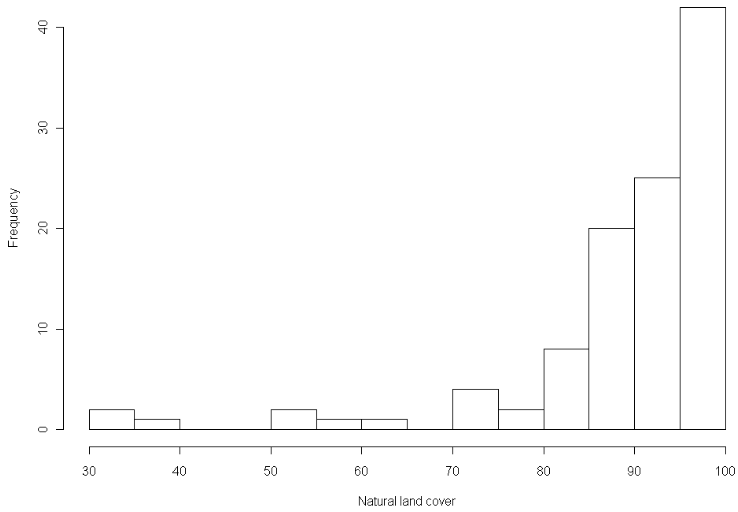



**Figure.10** Histogram of Natural land cover percentage on Swedish disturbed sites.

### *3.5.2 BPUE*

All parameters explaining 41.45% of the model are presented in the table.12 and are: max depth, lake area, average temperature and average temperature^2, catchment area and catchment area^2 to a lesser extent, with amplitude temperature, geology and amplitude temperature^2.

As for the CPUE metric, Mann-Whitney test reveals that mean EQR values for disturbed sites is significantly lower than on reference sites at the European scale (*p-value = 3.273e-05*) and at a GIG scale: NO GIG (*p-value =0.04172*) and CB GIG (*p-value = 0.0004394*) (Figure.11). No significant difference has been found for the two other GIG (MED and AL). The correlation of the EQR values of the BPUE metric and the percentage of natural land cover is 41.98% (Figure.12).

**Table.12** Results of the stepwise multiple linear regressions for the BPUE metric, done on parameters included in the "reference sites" model with the coefficient and the significance associated.

| | Coefficient | Significance |
|---|---|---|
| (Intercept) | -8.21E-03 | |
| log10(max_depth) | -4.89E-01 | *** |
| log10(lake_area) | 1.98E-01 | *** |
| I(log10(catch_area)^2) | -3.17E-02 | * |
| I(Altitude^2) | 2.54E-07 | * |
| ave_temperature | -1.31E-01 | *** |
| I(ave_temperature^2) | 1.55E-02 | *** |
| amp_temperature | 9.89E-02 | . |
| I(amp_temperature^2) | -2.67E-03 | * |
| geolsiliceous | 2.08E-01 | * |

*Signification codes: 0 '***' 0.001'**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

**(a)** **(b)**

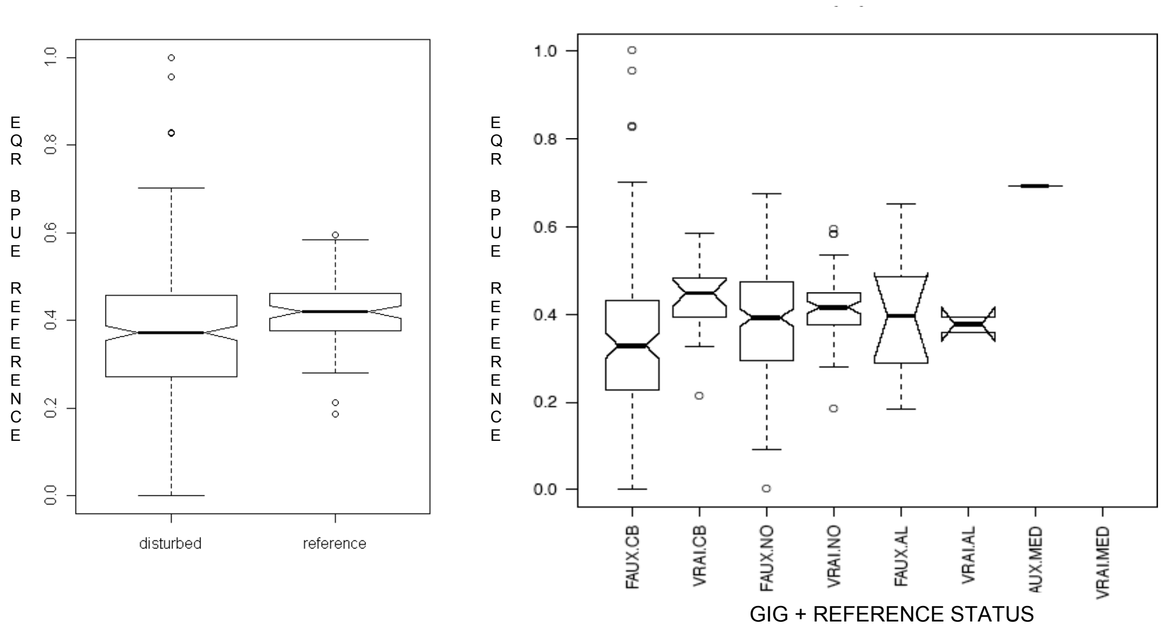



**Figure.11** Boxplots of the EQR values obtained by modelling on reference sites for the BPUE metric by reference status at the European scale **(a)** and by GIG **(b)** where *Alpine (AL), Central- Baltic (CB), Mediterranean (MED) and Nordic (NO)*. Reference sites are indicated as "reference" or "VRAI" in the reference status and disturbed sites as "disturbed" or "FAUX".

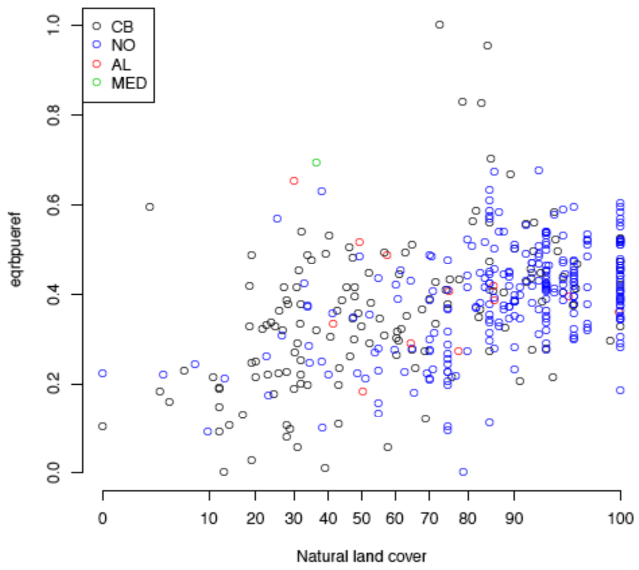



**Figure.12** Relation between the EQR values obtained by modelling on reference sites for the BPUE metric and the Natural land cover by GIG (*Alpine (AL), Central- Baltic (CB), Mediterranean (MED) and Nordic (NO))*.

## 3.6 Relevant metrics with the hindcasting model

### *3.6.1 CPUE*

Environmental parameters selected in the stepwise multiple linear regressions with the CPUE metric are max depth^2, lake area, altitude, altitude^2, amplitude temperature, amplitude temperature^2, catch area and average temperature to a lesser extent (Table.11). All these parameters explain 53.18% of the natural variability (Adjusted R-squared). The geology included in the model does not appear an explanatory variable.

The hierarchical partitioning analysis identified and confirmed variables explaining the model (Figure.13). The latter included at least one pressure, thereby confirming that they play an important role in explaining a part of the variability. Pressures included in the model are:

- acidification (*pressacid2O)* and
- eutrophication via the proxy "land use" (*asin (sqrt (CLC_percNatural/100)))*.

The EQR values for the CPUE metric presented in figure.14 show approximately the same shape as for the reference sites models at both scales (Europe and GIG). The Mann-Whitney test on all data gave a mean EQR value for disturbed sites significantly lower than on reference sites (*p-value = 5.092e-10)* and also on NO and CB GIG *(p-value = 0.000517 and p-value = 3.955e-05* respectively*).*

**Table.11** Results of the stepwise multiple linear regressions for the CPUE metric, done on parameters included in the hindcasting model with the coefficient and significance associated.

| | Coefficient | Significance |
|---|---|---|
| (Intercept) | -1.68E+00 | *** |
| I(log10(max_depth)^2) | -2.45E-01 | *** |
| log10(lake_area) | 1.53E-01 | *** |
| log10(catch_area) | -5.69E-02 | * |
| Altitude | -1.43E-03 | *** |
| I(Altitude^2) | 1.17E-06 | *** |
| I(ave_temperature^2) | 1.25E-03 | . |
| amp_temperature | 1.24E-01 | *** |
| I(amp_temperature^2) | -2.32E-03 | *** |
| asin(sqrt(CLC_percNatural/100)) | -4.41E-01 | *** |
| pressacid2O | -1.38E-01 | * |

*Signification codes: 0 '***' 0.001'**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

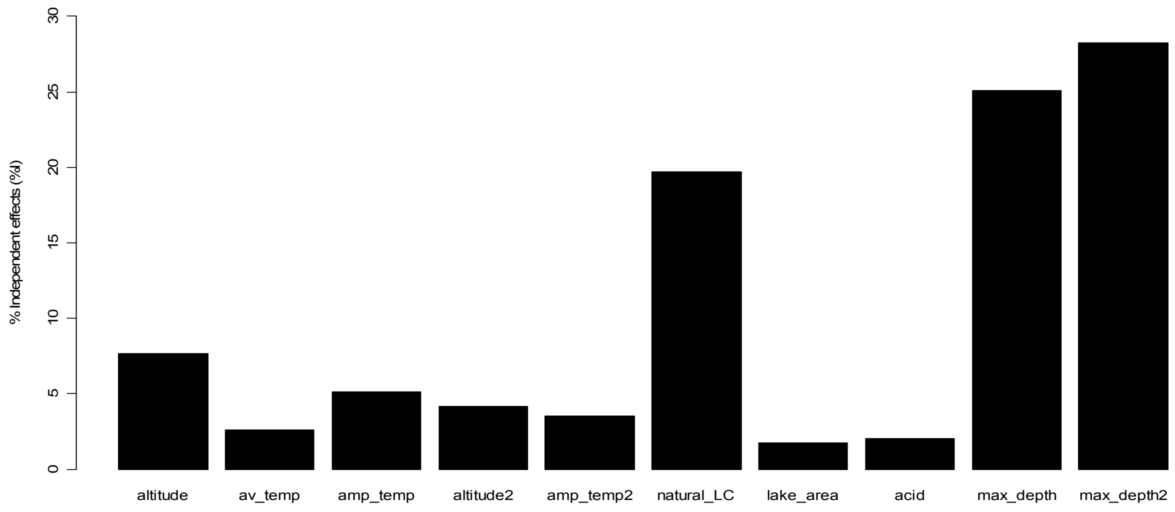



**Figure.13** Result of the hierarchical partitioning for the CPUE metric values obtained from the hindcasting model.

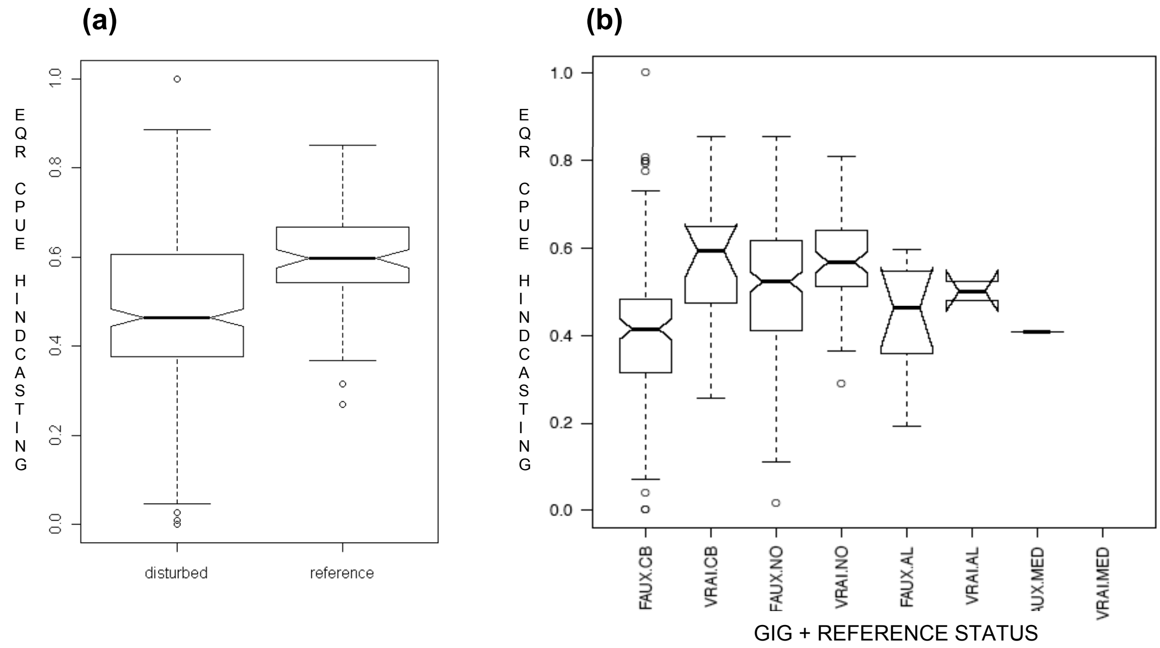



**Figure.14** Boxplots of the EQR values obtained with the hindcasting model for the CPUE metric at European scale **(a)** and by GIG **(b)** where *Alpine (AL), Central- Baltic (CB), Mediterranean (MED) and Nordic (NO)*. Reference sites are indicated as “reference” or “VRAI” in the reference status and disturbed sites as “disturbed” or “FAUX”.

A positive correlation of 54.21% was observed between the EQR values and the percentage of CLC natural on the catchment (the main pressure as we can see on Figure.13) of the lakes for all GIGs. But most of the percentage of natural land cover reaching the 100% are from the Nordic lakes and particularly Swedish ones (Figure.15). The variation scale is large and sites with for example 90% of natural land cover on their catchment show EQR between 0.3 and 0.8.

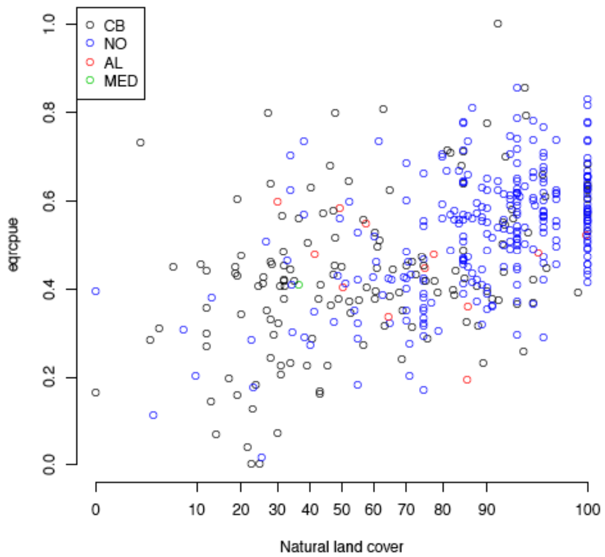



**Figure.15** Relation between the EQR values obtained with the hindcasting model for the CPUE metric and the Natural land cover by GIG *(Alpine (AL), Central- Baltic (CB), Mediterranean (MED) and Nordic (NO))*.

### *3.6.2 BPUE*

All parameters explaining 52.94% of the model for the BPUE metric are the lake area, the altitude and the altitude squared, the average temperature and the average temperature^2, the amplitude temperature and the amplitude temperature^2, the max depth and the max depth^2 and the catchment area. Both pressure (Natural land cover and acidification) are also explaining the model (Table.12) but mainly the land cover.
Mean EQR value for disturbed sites is significantly lower than on reference sites at the European scale (*p-value = 1.487e-07*) and for the Nordic and Central Baltic GIG (*p-value = 0.01679*and *p-value = 2.009e-05* respectively) (Figure.16). Correlation between EQR values of this metric and Natural land cover are 50.38% (Figure.17).

**Table.12** Results of the stepwise multiple linear regressions for the BPUE metric, done on parameters included in the hindcasting model with the coefficient and significance associated.

| | Coefficient | Significance |
|---|---|---|
| (Intercept) | 1.05E+00 | *** |
| log10(max_depth) | -1.94E-01 | * |
| I(log10(max_depth)^2) | -1.05E-01 | * |
| log10(lake_area) | 1.16E-01 | *** |
| log10(catch_area) | -3.31E-02 | |
| Altitude | -1.02E-03 | *** |
| I(Altitude^2) | 9.80E-07 | *** |
| ave_temperature | -1.14E-01 | *** |
| I(ave_temperature^2) | 6.87E-03 | *** |
| amp_temperature | 8.58E-02 | *** |
| I(amp_temperature^2) | -2.67E-03 | *** |
| asin(sqrt(CLC_percNatural/100)) | -3.58E-01 | *** |

| | |
|---|---|
| pressacid2O | -1.30E-01 * |

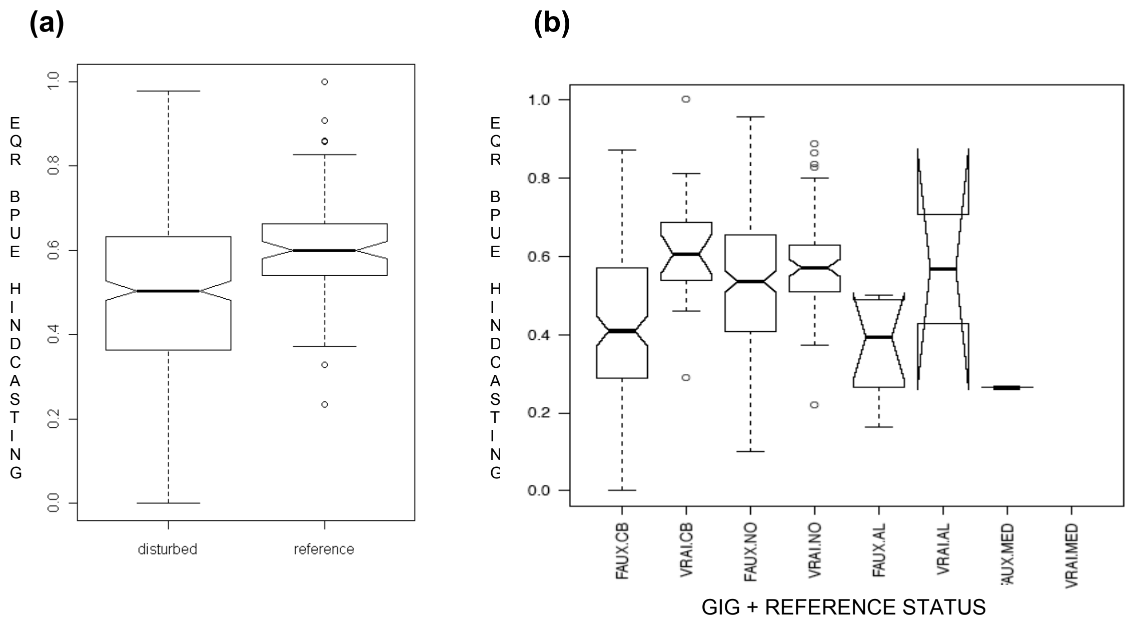



**Figure.16** Boxplots of the EQR values obtained with the hindcasting model for the BPUE at the European scale **(a)** and by GIG **(b)** where *Alpine (AL), Central- Baltic (CB), Mediterranean (MED) and Nordic (NO)*. Reference sites are indicated as “reference” or “VRAI” in the reference status and disturbed sites as “disturbed” or “FAUX”.

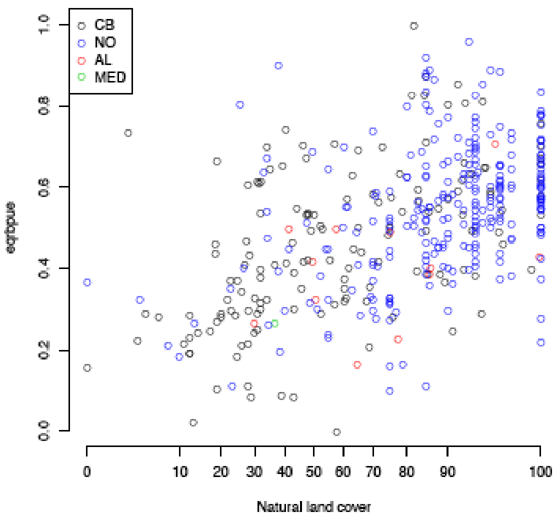



**Figure.17** Relation between the EQR values obtained with the hindcasting model for the BPUE metric and the Natural land cover by GIG (*Alpine (AL), Central- Baltic (CB), Mediterranean (MED) and Nordic (NO))*.

## 3.7 Comparison between the reference sites and hindcasting model

The comparison of the EQR values of the two different models show high correlation rate: 83.26% for the CPUE metric and 83.95 % for the BPUE one. The methods are very similar and give almost the same results: some points are out of the scatter (Figure.18). The outliers are French lakes few representative of the overall dataset (see figure 4) and more or less geographically isolated i.e. south of France (one in the Mediterranean GIG) and South West of France for the Central Baltic ones.

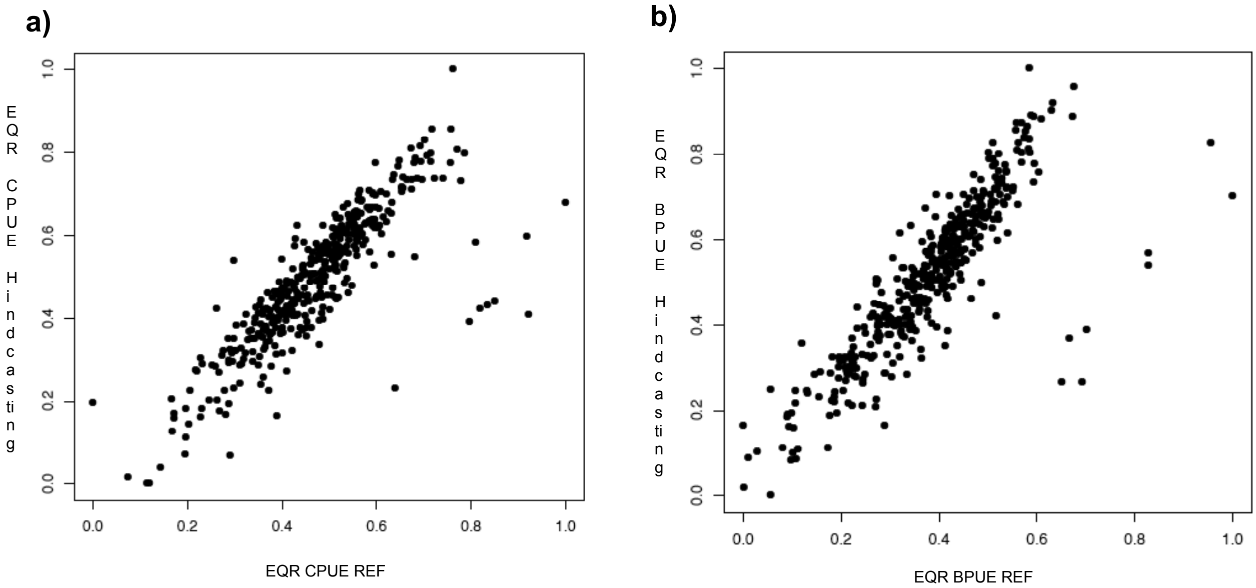



**Figure.18** Comparison of EQR values obtained by the hindcasting model and by the intercalibration's reference sites model for the CPUE metric (a) and the BPUE metric (b).

A regression line was drawn between the EQR values from the reference sites and the hindcasting models for the two metrics (CPUE & BPUE). Residuals were then analysed depending on their distance to the regression line to understand characteristics of these sites from an environmental point of view. It appears that the sites out of the scatter have an average temperature higher than the remaining ones (Figure.19).

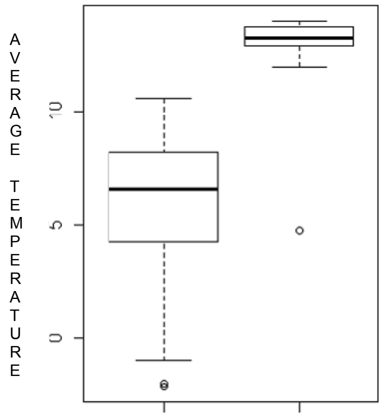



**Figure.19** Boxplots of average temperature for sites near the regression line (on the left) and the outliers (on the right).

## 3.8 IBI development and definition of class boundaries

The next step to use previous results in the intercalibration exercise is to define the ecological classes' boundaries, especially the High/Good and Good/Moderate boundaries. At present time we decided to keep only BPUE and CPUE to compute common index because this is the best metrics combination in regards to the stressor gradient.

### 3.8.1 The High/Good boundary

- **Using index values on reference sites**

Common method previously used in the intercalibration process is based on the distribution of values on reference sites into percentile. The choice of relevant percentile is related to the confidence of the reference list. For our index, the H/G boundary should be located in the lower part of the distribution (percentile 5, 10, 15 or 20%) (Table.13).

**Table.13** Distribution of reference values into percentile to the selection of H/G boundary.

| Percentile | 0% | 5% | 10% | 15% | 20% | 25% | 30% | 35% | 40% | 45% | 50% |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Index value | 0,60 | 0,84 | 0,94 | 1,01 | 1,05 | 1,08 | 1,10 | 1,14 | 1,16 | 1,17 | 1,19 |

However, some lakes with significant fish exploitation were included among reference sites (but considered as reference by experts). It was decided to exclude them before computing percentiles because fish exploitation has probably an effect on BPUE and CPUE:

| Percentile | 0% | 5% | 10% | 15% | 20% | 25% | 30% | 35% | 40% | 45% | 50% |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Index value | 0,71 | 0,88 | 1,03 | 1,05 | 1,08 | 1,10 | 1,14 | 1,16 | 1,17 | 1,18 | 1,21 |

| | 0.88 | | 1.03 | | 1.05 | | 1.08 | |
|---|---|---|---|---|---|---|---|---|
| | <H | H | <H | H | <H | H | <H | H |
| Disturbed | 40% | 60% | 57% | 43% | 58% | 42% | 60% | 40% |
| Reference | 7% | 93% | 18% | 82% | 20% | 80% | 25% | 75% |

The percentile 10 was retained as H/G boundary. This decision represents a compromise. Indeed, with lowest value (0.88), too many disturbed sites are assigned to high status. With higher value (1.08), too many sites are assigned to status below high. The figure 20 of confusion matrices also helps to choose the percentile. The percentile 30 give simultaneously the maximum of true positives and true negatives, but 35% of reference sites are not in high status. By selecting the percentile 10, only less than 20% of reference sites are not in high status and we believe more relevant promoting true positives.

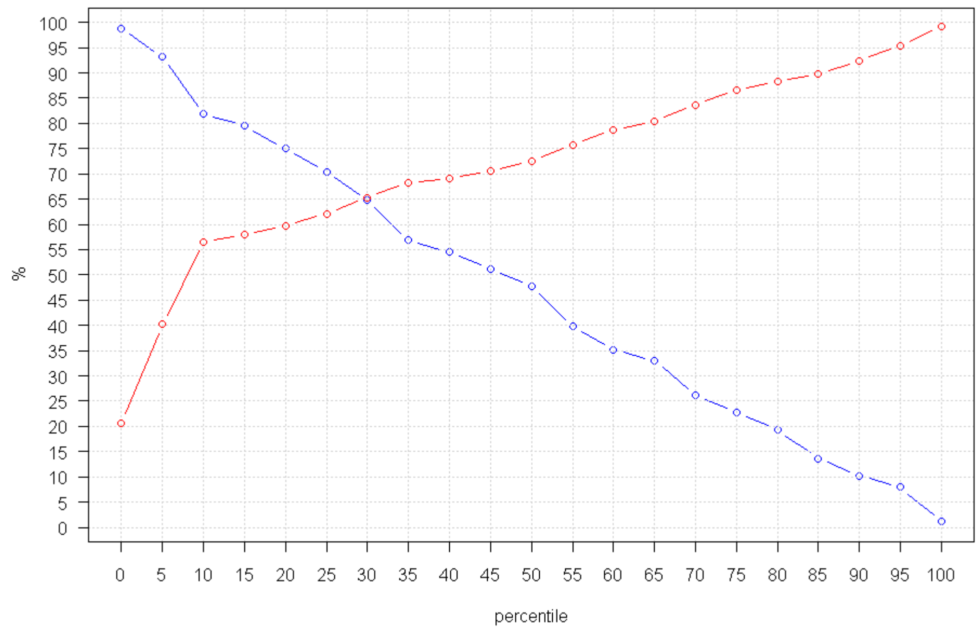
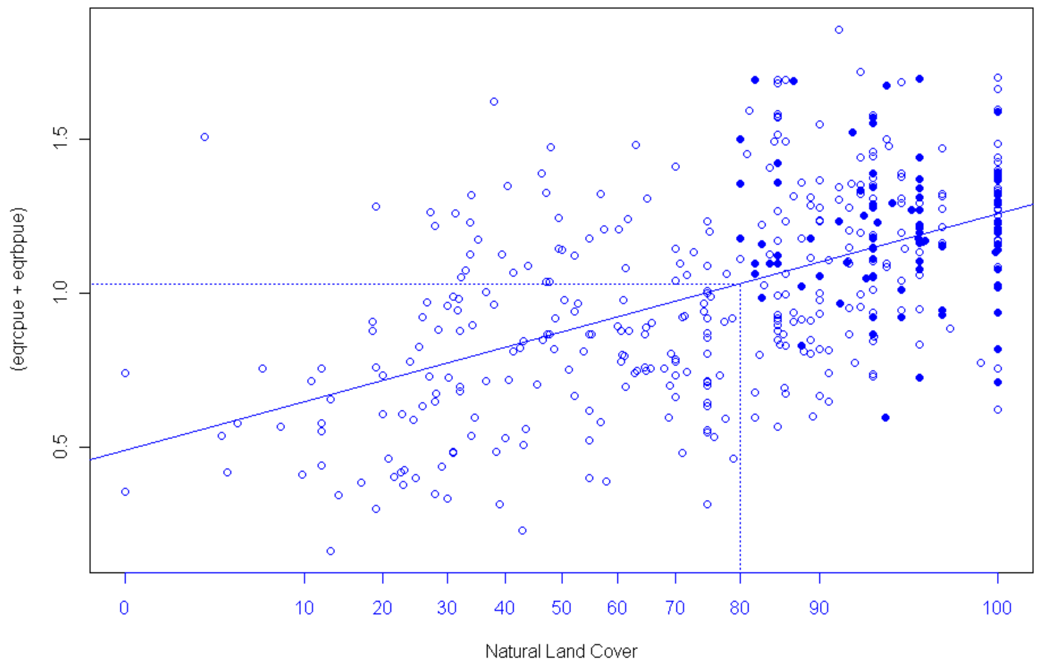



**Figure.20** Graphic from confusion matrices. In blue the percentage of reference sites in high status. In red the percentage of disturbed sites in class below high status.

- **Using index response to stressor gradient**

Here, we can consider a linear relationship between Index value and stressor gradient. Therefore, a linear model can be built to predict the expected index value for a certain amount of pressure. In the task of defining the reference sites, catchment natural land cover of the reference sites was set below 80%. So the prediction with this threshold could give H/G boundary. The corresponding Index Value is 1.03 (Figure.21).




**<u>Figure.11</u>** Relationship between Index (sum of EQRcpue an EQRbpue) and Natural land cover in the catchment. Full dots corresponding to reference sites. Open dots corresponding to disturbed sites. The blue line is the regression line.

### *3.8.2 The other boundaries (G/M, M/P and P/B)*

- **<u>using index response to stressor gradient</u>**

This approach is simply the extension of the one above. We have to define pressure values to predict index boundaries for each ecological class. We supposed a linear relationship between the index and the pressure (transformed form of land cover). Consequently, equal classes of pressure can be used. The corresponding natural land cover percentages are 54.5% (G/M), 27.6% (M/P) and finally 7.5% (P/B), applying the linear model with these predictors, the expected index values for the boundaries are: 0.895 for G/M, 0.759 for M/P and 0.624 for P/B (<u>Figure.22</u>).

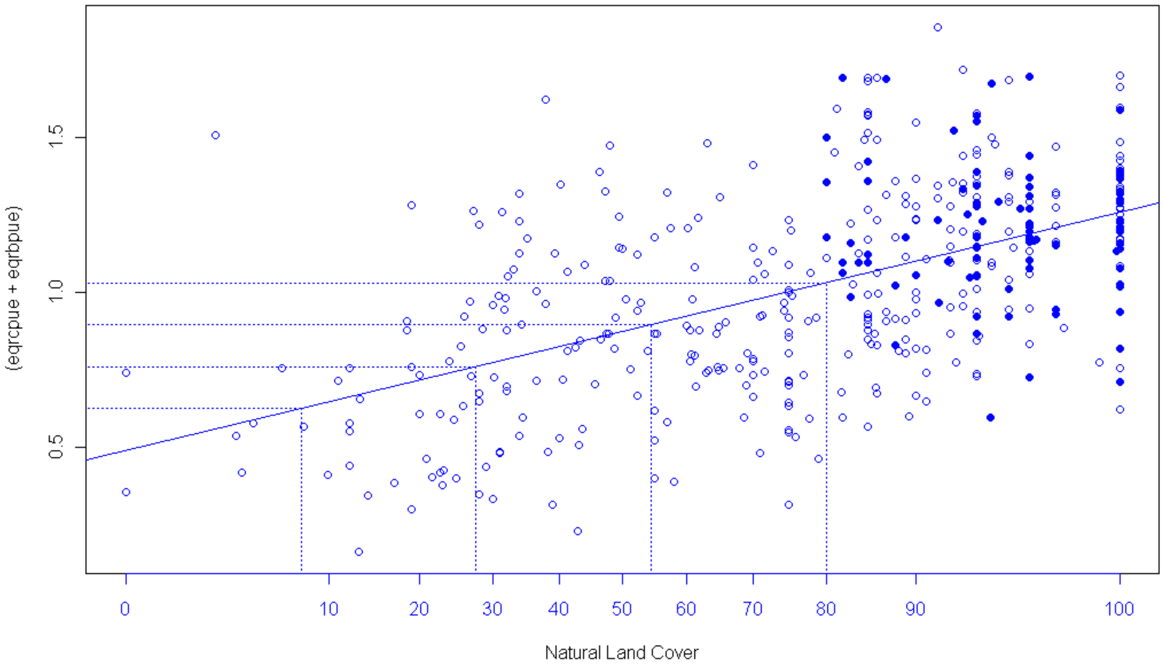



**Figure.22** Relationship between Index (sum of EQRcpue an EQRbpue) and Natural land cover in the catchment. Full dots corresponding to reference sites. Open dots corresponding to disturbed sites. The blue line is the regression line.

If we apply all the boundaries derived from the linear model, the distribution of sites into ecological classes is as follow:

| | B | P | M | G | H |
|---|---|---|---|---|---|
| Number | 48 | 43 | 49 | 58 | 221 |
| *%* | *11* | *10* | *12* | *14* | *53* |

More than 50% of sites appear in high status and the remaining sites are equally distributed in the other classes.

Find below the distribution of sites into the 5 ecological classes regarding the reference/disturbed feature. Six reference sites had been assigned to status worse than good. We did not find any parameter in the database that could explain this result. Forty five percent of disturbed sites are assigned to high status and 14.5% to good one. Others are equally distributed in the remaining degraded classes.

| | disturbed | reference |
|---|---|---|
| H | 149 | 72 |
| G | 49 | 10 |
| M | 50 | 3 |
| P | 43 | 2 |
| B | 52 | 1 |

- **<u>using index values distribution</u>**

Automatic clustering (the k-means method for example) can be used to create G/M, M/P and P/B boundaries. For index values below H/G boundary, k-means will build groups so as to get lower within groups variance. Boundaries values are then derived from minimum and maximum index values of adjacent groups: each observation is assigned to one group and the boundary corresponds to the mean between the minimum index value in a group and the maximum index value in the group below. By applying this approach we got 0.851 for G/M, 0.684 for M/P and 0.492 for P/B (Figure.23).

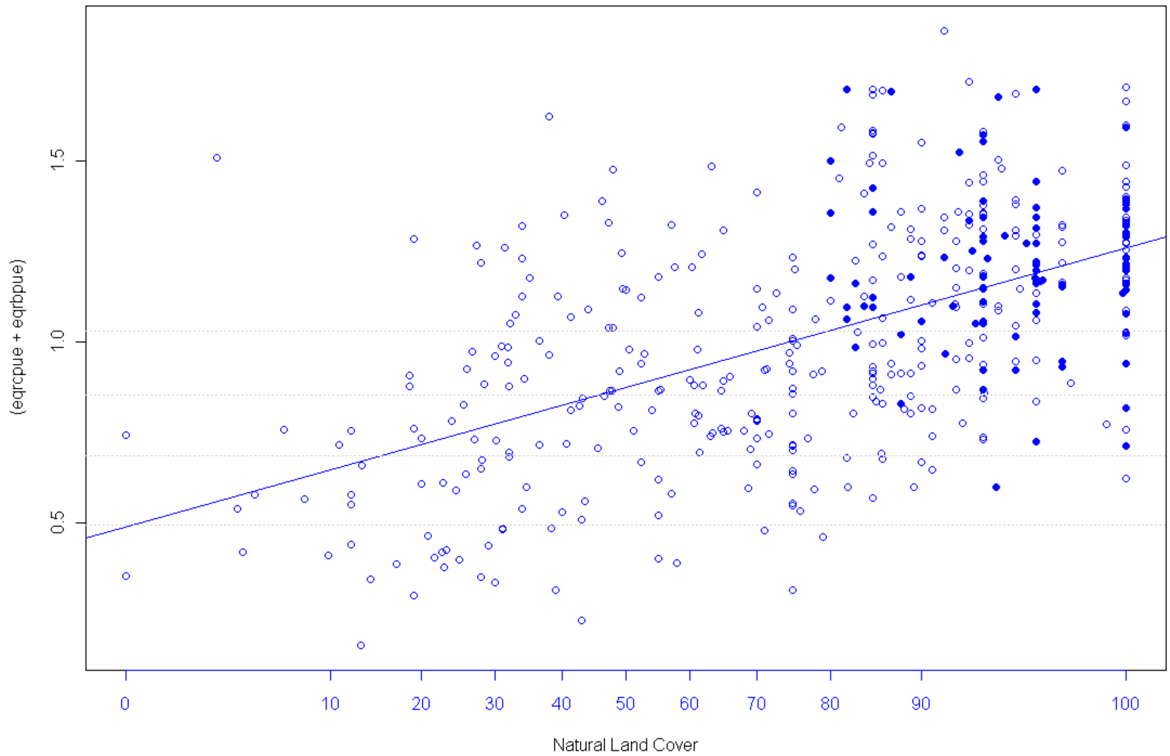



**Figure.23** Relationship between Index (sum of EQRcpue an EQRbpue) and Natural land cover in the catchment. Full dots corresponding to reference sites. Open dots corresponding to disturbed sites. The blue line is the regression line.

If we set 1.03 as H/G boundary, sites are assigned into ecological classes as follows:

| | B | P | M | G | H |
|---|---|---|---|---|---|
| Number | 26 | 33 | 63 | 76 | 221 |
| *%* | *6* | *8* | *15* | *18* | *53* |

The percentage in each category is decreasing with status degradation. With these boundaries, no reference is assigned to bad status and only one to poor status (see below). Near 65% of sites not recognized as reference (disturbed) are at least in good status.

| | disturbed | reference |
|---|---|---|
| H | 149 | 72 |
| G | 65 | 11 |
| M | 59 | 4 |
| P | 32 | 1 |
| B | 26 | 0 |

However k-means method is not stable. Indeed it is closely related to dataset used. Adding new sites will probably give other boundaries.

# 4 Discussion

The present study demonstrated how hindcasting modelling of fish-based metrics enabled to assess lakes' current conditions, even at a broader large scale than submitted by previous authors (Baker et al. 2005; Kilgour and Barton 1999). Two non-redundant metrics displaying the most significant responses to the reverse of the eutrophication pressure (natural land cover) were selected and combined into an index of biotic integrity at the European scale.

## 4.1 Targeted pressures

In this study, interest has focused on eutrophication and acidification.
The natural land cover was used as the reverse of the pressure and for a catchment scale indicator for eutrophication. Even if the total phosphorus was available for almost all lakes in our dataset, we did not keep it as a pressure in the models for three main reasons.
First, even if nutrient inputs have long been considered as major drivers of fish communities in lakes because eutrophication implies oxygen depletion, organic sediment accumulation (Carpenter *et al.* 1998; Harper 1992) and a transfer of primary productivity from macrophytes to phytoplankton (Leach *et al.* 1977), it's clear now that this view oversimplifies the lake ecosystems functioning. More complex processes interact as the food web with the top-down controls and the recycling of organic matter in the biofilm, directly related to lake morphology (Mehner *et al.* 2004) and the buffering capacity (Shaw and Kelso 1992). Human activities also produce different strains that interact with abiotic and biotic factors in shaping the structure and variation of fish communities. Based on theses statements, dissecting the natural from the human induced sources of total phosphorus can be extremely difficult.
Secondly, the stepwise multiple linear regressions integrate the "total phosphorus" in the model and delete the "maximum depth" when both were added in the selection process, which could be considered as a problem from a functional point of view. Indeed, a large amount of studies have demonstrated the relationship between the natural variations of total phosphorus in a lake with the depth (Cardoso, Solimini *et al.* 2007). For the hindcasting model, it appears that setting a unique threshold of Total phosphorus for all lakes (with a range of maximum depth from 0.2 to 110 m) to get the reference conditions was not relevant, also for the natural variability.
The last reason was the heterogeneity of the phosphorus data collected in the different countries. Sampling protocols are often different and calculation is based on the average of

two, three or four samplings a year. Only one discrete measure was integrated in the final database and was sometimes not synchronous with the fish sampling. The measure should be extrapolated at the date of the sampling event to get continuous natural variation of this parameter during all the year.
The second pressure considered i.e. acidification explains also a part of the fish community variability; this strong effect on fish communities as waters acidify has already been described (Somers & Harvey 1984). Even if anthropogenic acidification has been mainly restricted to the past two to four decades, it generally not provides sufficient time for selection or colonization by tolerant species. Natural acidification could occur in some regions due to the presence of high concentration of organic acids from adjacent wetlands, but in this case, it was not considered as a pressure.

The WFD also recommends the assessment of hydromorphological pressures. It involves detailed recording of shoreline characteristics and stressors, modifications to the hydrological regime and impact of lake uses. Such data are not available at the European scale and more detailed information should be collected to integrate properly this pressure.

## 4.2 Selected metrics for potential use in IBI

Overall, 18 and 24 candidate fish metrics selected for the reference sites and for the hindcasting models, displayed a significant response to anthropogenic pressures. All these metrics could potentially be included in a multimetric index (MMI). Both composition and abundance requirements of the WFD are covered but abundance is mainly represented by the trophic guild.
Species classification into guilds is convenient due to the functional information it gives. The classification into trophic guilds of species (Annex 1) is based on their diet or their manner of feeding but without considering size and local particularities. For species with restricted diets, it does not raise problem, however, many fishes vary their diets as they grow from a fry to an adult and the size of the fishes was not taken into account in this study ; all fishes of the same species were awarded the same trait for all countries.
None of the selected metrics derived from the reproductive guild. These metrics are probably more relevant to assess littoral habitat alteration, a pressure not considered here because of the lack of homogenous information on this type of pressure at the European scale. Native-related guild had a limited distribution and was no more considered in this study.
Other metrics of potential interest, derived from the species status (exotic and introduced), hybridization, tolerance, individuals' conditions, were not considered in this study due to the lack of data, but they could be collected for future needs.

Populations' size-class distributions also have to be included in the quality assessment tool, but this aspect will be considered later by the lake fish WISER group. .

In Europe, some countries such as Finland and Sweden already integrated the CPUE & BPUE metrics (calculated on all individuals or only on native species) in their assessment system (Appelberg *et al.* 2000, Rask, Olin & Ruuhijärvii 2009, Tammi *et al.* 2001).
By combining these metrics in an indice, only the abundance parameter required by the WFD is fulfilled.
A preliminary work (not presented here), shows that combination of some of composition metrics with the CPUE and BPUE tends to decrease the final correlation score to the stressor gradient. Hence, a perspective could be to test other types of combination to include composition without reducing the correlation score to the stressor gradient.
And, as explained in the results, is there a real necessity to keep the higher correlation score to build an efficient MMI?

## 4.3 Which approach: Reference sites or hindcasting?

For the reference approach, the main issue is probably the extrapolation beyond the range of the calibration: that means environmental gradient of reference sites must cover the whole gradient of dataset otherwise we risk rough prediction errors. It was shown here that the distribution of our reference sites reflects the overall environmental gradient considered except regarding temperature. This is a potential source of bias in the assessment system developed.
The choice of reference criteria and thresholds for reference sites selection is also a key step: ideally no significant impact on biology should be detected on reference sites so that reference models only describe natural variations. Some sites selected as in reference conditions by member state were in fact clearly impacted by local activities (information collected in parallel to the present study). This misclassification induces also a bias in the approach based on reference sites.

Regarding the hindcasting model, the main issue is extrapolation: indeed if there is no site with no or low pressures in the dataset used to build the model, the risk of prediction errors for reference conditions is high. The stressor gradient should be as wide as possible, which is the case here. Reference conditions are usually modelled by the hindcasting approach when not enough reference sites are available (Baker, Wehrly et al. 2005). Nevertheless, in the described approach, reference sites were included in the hindcasting model, to have a wide range of pressure and environmental parameters. For a better comparability between both methods, the reference sites had to be excluded. This has been recently done (results not shown) and the same results occurred.

So, for both models, almost the same metrics based on trophic and taxonomic guilds were retained. The relationships between these metrics and pressures are compliant to those found in previous studies (Table.1). So, which approach is the best?
Considering the disadvantages and uncertainties around the selection of reference sites (subjective/expert opinion, narrow environmental gradient) and the close results of both models, the hindcasting ones seems to be the best oncoming.
Whatever the approach, the estimation of uncertainty for the assessment system should be performed during the next year. The workpackage 6.1 provided the guidelines for a future application.

## 4.4 Geographical representativeness

Data distribution over Europe is heterogeneous. A lot of lakes are located in the Scandinavian region; those in the Central part are scarce nay inexistent and only one lake is in the Mediterranean part of Europe.
During comparison between both models, outliers appear to have some higher average temperature than the remaining. These sites do not belong to the environmental range of all sites. These lakes from the South and South West of France are in extreme situations in terms of climate and do not match with any other lakes in Europe. To improve this situation, increasing southern data collection on natural lakes (from Greece, Spain and/or Italia for example) could be recommended.
As explained above, also reference sites do not cover the all environmental variability and if more member state or more lakes among Europe could be included, wider the range of distribution could be. Some efforts should be done in the future to get a good overview and a better geographical representativeness of these natural lakes in Europe.

## 4.5 Definition of Class boundaries

Several approaches have been proposed to define class boundaries for the current index (based on WPUE and CPUE). Whatever the method, values for H/G boundary are very similar and consequently can be considered robust. On contrary, there is no consistency for lower boundaries. But the approach using index response to gradient stressor is probably better and comparable with that developed for chlorophyll-a in some GIGs during the first round (equal log class distribution approach, in link with relationship between chlorophyll-a and total phosphorus).

However all approaches are closely related to the criteria used for the definition of reference sites: obviously for the H/G boundary (percentile of index values on the reference sites and expected index value from reference threshold), but also for the other boundaries (clustering below H/G boundary or equal groups regarding pressure below H/G boundary). So the choice of reference sites and pressures thresholds to identify them is particularly crucial and need to be agreed.

Moreover, at this stage the fish index is composed of BPUE and CPUE. As explained above, new combination rules could be tested to include other metrics. One way could be to combine normalized EQR that implies to define boundaries for each metrics. It will be a hard work but necessary in order to take into account the difference between metrics curves responses to stressor.

## CONCLUSIONS and PERSPECTIVES

The European indicator presented here is based on two metrics representative of the fish abundance. Therefore, it follows partly the requirement of the Water Framework Directive (WFD; 2000/60/EC) since species composition and age structure were not included yet (Annex V 1.2.1 of this directive). Nevertheless, potential abundance fish metrics could also be added in a near future. To do that, more analyses are required, particularly on rules of metrics' aggregation. Similarly, responses of metrics based on size structure to environmental parameters are currently under analyses by the lake fish group.
The interest of the two metrics proposed (CPUE and BPUE) is that they could be easily calculated by the member state and permit an intercalibration at the European scale.
Another interesting point could be comparing the data obtained by hydroacoustics and those by gillnetting.

The hindcasting model has proved, in our dataset, to be a relevant method for the development of an assessment tool. This method will be used to select metrics responding to pressures on the reservoirs included in the database and on the low species richness lakes. Later, analyse of fish communities of the lakes sampled with other types of multimesh gillnets (included in the database) could also be tried in case no fish based ecological assessment methods were available in the countries using these non standardised sampling protocols.

**ANNEX 1** Assignment of the 70 fish species (present in the dataset) into reproductive, trophic and habitat guilds used to derive community traits. Two classifications were used, one with a binary code (a) and one with the name (b).

**(a)**

| | | | | Reproductive guild | | | | | | Trophic guild | | | | | | | Food habitat | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Id_taxon | Family | Genus | Species | PHYT | LITH_LIPE | PELA | OSTR | ARIAD | SPEL | INV | BENT | PISC | PLAN | HERB | PARA | DETR | BENT | WC |
| ABRABR | Cyprinidae | *Abramis* | *brama* | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| ALBUBI | Cyprinidae | *Alburnoides* | *bipunctatus* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| ALBUAL | Cyprinidae | *Alburnus* | *alburnus* | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| ALBUME | Cyprinidae | *Alburnus* | *mento* | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| ALOSFA | Clupeidae | *Alosa* | *fallax* | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| AMEIME | Ictaluridae | *Ameiurus* | *melas* | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| ANGUAN | Anguillidae | *Anguilla* | *anguilla* | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| ASPIAS | Cyprinidae | *Aspius* | *aspius* | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| BALLBA | Cyprinidae | *Ballerus* | *ballerus* | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| BARBBR | Cobitidae | *Barbatula* | *barbatula* | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| BARBBA | Cyprinidae | *Barbus* | *barbus* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| BLICBJ | Cyprinidae | *Blicca* | *bjoerkna* | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| CARAAU | Cyprinidae | *Carassius* | *auratus* | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 |
| CARACA | Cyprinidae | *Carassius* | *carassius* | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 |
| CARAGI | Cyprinidae | *Carassius* | *gibelio* | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 |
| CLUPSP | Clupeidae | *Clupea* | *sprattus* | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| COBITA | Cobitidae | *Cobitis* | *taenia* | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| COREAL | Salmonidae | *Coregonus* | *albula* | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| COREAU | Salmonidae | *Coregonus* | *autumnalis* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| CORELA | Salmonidae | *Coregonus* | *lavaretus* | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| COREPE | Salmonidae | *Coregonus* | *peled* | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| COTTGO | Cottidae | *Cottus* | *gobio* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| COTTPO | Cottidae | *Cottus* | *poecilopus* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 |
| CYPRCA | Cyprinidae | *Cyprinus* | *carpio* | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| ESOXLU | Esocidae | *Esox* | *lucius* | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| GASTAC | Gasterosteidae | *Gasterosteus* | *aculeatus* | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| GOBIGO | Cyprinidae | *Gobio* | *gobio* | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |

| | | | | Reproductive guild | | | | | | Trophic guild | | | | | | | Food habitat | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Id_taxon | Family | Genus | Species | PHYT | LITH_LIPE | PELA | OSTR | ARIAD | SPEL | INV | BENT | PISC | PLAN | HERB | PARA | DETR | BENT | WC |
| GYMNCE | Percidae | *Gymnocephalus* | *cernuus* | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| HYPOMO | Cyprinidae | *Hypophthalmichthys* | *molitrix* | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| HYPONO | Cyprinidae | *Hypophthalmichthys* | *nobilis* | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| LEPOGI | Centrarchidae | *Lepomis* | *gibbosus* | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| LEUCDE | Cyprinidae | *Leucaspius* | *delineatus* | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| LEUCID | Cyprinidae | *Leuciscus* | *idus* | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| LEUCLE | Cyprinidae | *Leuciscus* | *leuciscus* | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| LOTALO | Lotidae | *Lota* | *lota* | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| MICRSA | Centrarchidae | *Micropterus* | *salmoides* | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| MISGFO | Cobitidae | *Misgurnus* | *fossilis* | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| NEOGME | Gobidae | *Neogobius* | *melanostomus* | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| ONCOMY | Salmonidae | *Oncorhynchus* | *mykiss* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| OSMEEP | Osmeridae | *Osmerus* | *eperlanus* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| PERCFL | Percidae | *Perca* | *fluviatilis* | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| PHOXPH | Cyprinidae | *Phoxinus* | *phoxinus* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| PLATFL | Pleuronectidae | *Platichthys* | *flesus* | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| POMAMI | Gobiidae | *Pomatoschistus* | *minutus* | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| PSEUPA | Cyprinidae | *Pseudorasbora* | *parva* | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 |
| PUNGPU | Gasterosteidae | *Pungitius* | *pungitius* | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| RHODAM | Cyprinidae | *Rhodeus* | *amarus* | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| RUTIAU | Cyprinidae | *Rutilus* | *aula* | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 |
| RUTIME | Cyprinidae | *Rutilus* | *meidingeri* | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| RUTIRU | Cyprinidae | *Rutilus* | *rutilus* | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 |
| SALAFL | Blenniidae | *Salaria* | *fluviatilis* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| SALMFE | Salmonidae | *Salmo* | *ferox* | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| SALMNI | Salmonidae | *Salmo* | *nigripinnis* | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| SALMSA | Salmonidae | *Salmo* | *salar* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| SALMST | Salmonidae | *Salmo* | *stomachicus* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| SALMTR | Salmonidae | *Salmo* | *trutta* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| SALMTF | Salmonidae | *Salmo* | *trutta fario* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |
| SALMTT | Salmonidae | *Salmo* | *trutta trutta* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| SALVFO | Salmonidae | *Salvelinus* | *fontinalis* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |

| | | | | Reproductive guild | | | | | | Trophic guild | | | | | | | Food habitat | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Id_taxon | Family | Genus | Species | PHYT | LITH_LIPE | PELA | OSTR | ARIAD | SPEL | INV | BENT | PISC | PLAN | HERB | PARA | DETR | BENT | WC |
| SALVNA | Salmonidae | *Salvelinus* | *namaycush* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| SALVUM | Salmonidae | *Salvelinus* | *umbla* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| SANDLU | Percidae | *Sander* | *lucioperca* | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| SCARER | Cyprinidae | *Scardinius* | *erythrophthalmus* | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 |
| SILUGL | Siluridae | *Silurus* | *glanis* | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| SQUACE | Cyprinidae | *Squalius* | *cephalus* | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| TELESO | Cyprinidae | *Telestes* | *souffia* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| THYMTH | Salmonidae | *Thymallus* | *thymallus* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| TINCTI | Cyprinidae | *Tinca* | *tinca* | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 |
| TRIGQU | Cottidae | *Triglopsis* | *quadricornis* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| VIMBVI | Cyprinidae | *Vimba* | *vimba* | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |

**(b)**

| Id_taxon | Family | Genus | Species | Reproductive guild | Trohic guild | Food habitat |
|---|---|---|---|---|---|---|
| ABRABR | Cyprinidae | *Abramis* | *brama* | PHLI | PLAN | BENT |
| ALBUBI | Cyprinidae | *Alburnoides* | *bipunctatus* | LITH | INV | WC |
| ALBUAL | Cyprinidae | *Alburnus* | *alburnus* | PHLI | PLAN | WC |
| ALBUME | Cyprinidae | *Alburnus* | *mento* | LITH | PLAN | WC |
| ALOSFA | Clupeidae | *Alosa* | *fallax* | LITH | PLAN | BENT |
| AMEIME | Ictaluridae | *Ameiurus* | *melas* | LITH | OMNI | BENT |
| ANGUAN | Anguillidae | *Anguilla* | *anguilla* | PELA | INV/PISC | WC |
| ASPIAS | Cyprinidae | *Aspius* | *aspius* | LITH | PISC | BENT |
| BALLBA | Cyprinidae | *Ballerus* | *ballerus* | PHYT | PLAN | WC |
| BARBBR | Cobitidae | *Barbatula* | *barbatula* | PHLI | INV | BENT |
| BARBBA | Cyprinidae | *Barbus* | *barbus* | LITH | INV | BENT |
| BLICBJ | Cyprinidae | *Blicca* | *bjoerkna* | PHYT | OMNI | BENT |
| CARAAU | Cyprinidae | *Carassius* | *auratus* | PHYT | OMNI | BENT |
| CARACA | Cyprinidae | *Carassius* | *carassius* | PHYT | OMNI | BENT |
| CARAGI | Cyprinidae | *Carassius* | *gibelio* | PHYT | OMNI | BENT |
| CLUPSP | Clupeidae | *Clupea* | *sprattus* | PELA | PLAN | WC |

| Id_taxon | Family | Genus | Species | Reproductive guild | Trohic guild | Food habitat |
|---|---|---|---|---|---|---|
| COBITA | Cobitidae | *Cobitis* | *taenia* | PHYT | BENT | BENT |
| COREAL | Salmonidae | *Coregonus* | *albula* | LITH | PLAN | WC |
| COREAU | Salmonidae | *Coregonus* | *autumnalis* | LITH | INV/PLAN | WC |
| CORELA | Salmonidae | *Coregonus* | *lavaretus* | LITH | INV | WC |
| COREPE | Salmonidae | *Coregonus* | *peled* | LITH | PLAN | WC |
| COTTGO | Cottidae | *Cottus* | *gobio* | LITH | INV | BENT |
| COTTPO | Cottidae | *Cottus* | *poecilopus* | LITH | OMNI | BENT |
| CYPRCA | Cyprinidae | *Cyprinus* | *carpio* | PHYT | OMNI | BENT |
| ESOXLU | Esocidae | *Esox* | *lucius* | PHYT | PISC | WC |
| GASTAC | Gasterosteidae | *Gasterosteus* | *aculeatus* | ARIAD | INV | BENT |
| GOBIGO | Cyprinidae | *Gobio* | *gobio* | PHLI | INV | BENT |
| GYMNCE | Percidae | *Gymnocephalus* | *cernuus* | PHLI | OMNI | BENT |
| HYPOMO | Cyprinidae | *Hypophthalmichthys* | *molitrix* | PELA | PLAN | WC |
| HYPONO | Cyprinidae | *Hypophthalmichthys* | *nobilis* | PELA | PLAN | BENT |
| LEPOGI | Centrarchidae | *Lepomis* | *gibbosus* | LITH | INV | WC |
| LEUCDE | Cyprinidae | *Leucaspius* | *delineatus* | PHYT | OMNI | WC |
| LEUCID | Cyprinidae | *Leuciscus* | *idus* | PHLI | INV/PISC | WC |
| LEUCLE | Cyprinidae | *Leuciscus* | *leuciscus* | LITH | OMNI | WC |
| LOTALO | Lotidae | *Lota* | *lota* | LITH | PISC | WC |
| MICRSA | Centrarchidae | *Micropterus* | *salmoides* | ARIAD | PISC | WC |
| MISGFO | Cobitidae | *Misgurnus* | *fossilis* | PHYT | BENT | BENT |
| NEOGME | Gobidae | *Neogobius* | *melanostomus* | SPEL | INV | BENT |
| ONCOMY | Salmonidae | *Oncorhynchus* | *mykiss* | LITH | INV/PISC | WC |
| OSMEEP | Osmeridae | *Osmerus* | *eperlanus* | LITH | INV/PISC | WC |
| PERCFL | Percidae | *Perca* | *fluviatilis* | PHLI | INV/PISC | WC |
| PHOXPH | Cyprinidae | *Phoxinus* | *phoxinus* | LITH | INV | WC |
| PLATFL | Pleuronectidae | *Platichthys* | *flesus* | PELA | INV/PISC | BENT |
| POMAMI | Gobiidae | *Pomatoschistus* | *minutus* | OSTR | INV/PISC | BENT |
| PSEUPA | Cyprinidae | *Pseudorasbora* | *parva* | PHLI | OMNI | WC |
| PUNGPU | Gasterosteidae | *Pungitius* | *pungitius* | PHYT | INV | BENT |
| RHODAM | Cyprinidae | *Rhodeus* | *amarus* | OSTR | OMNI | WC |
| RUTIAU | Cyprinidae | *Rutilus* | *aula* | PHYT | OMNI | BENT |
| RUTIME | Cyprinidae | *Rutilus* | *meidingeri* | PHLI | INV | BENT |
| RUTIRU | Cyprinidae | *Rutilus* | *rutilus* | PHLI | OMNI | WC |

| Id_taxon | Family | Genus | Species | Reproductive guild | Trohic guild | Food habitat |
|---|---|---|---|---|---|---|
| SALAFL | Blenniidae | *Salaria* | *fluviatilis* | LITH | INV | BENT |
| SALMFE | Salmonidae | *Salmo* | *ferox* | LITH | PISC | WC |
| SALMNI | Salmonidae | *Salmo* | *nigripinnis* | LITH | PLAN | WC |
| SALMSA | Salmonidae | *Salmo* | *salar* | LITH | INV/PISC | WC |
| SALMST | Salmonidae | *Salmo* | *stomachicus* | LITH | INV | BENT |
| SALMTR | Salmonidae | *Salmo* | *trutta* | LITH | INV/PISC | WC |
| SALMTF | Salmonidae | *Salmo* | *trutta fario* | LITH | INV/PISC | WC |
| SALMTT | Salmonidae | *Salmo* | *trutta trutta* | LITH | INV/PISC | WC |
| SALVFO | Salmonidae | *Salvelinus* | *fontinalis* | LITH | INV/PISC | WC |
| SALVNA | Salmonidae | *Salvelinus* | *namaycush* | LITH | INV/PISC | WC |
| SALVUM | Salmonidae | *Salvelinus* | *umbla* | LITH | INV/PISC | WC |
| SANDLU | Percidae | *Sander* | *lucioperca* | PHLI | INV/PISC | WC |
| SCARER | Cyprinidae | *Scardinius* | *erythrophthalmus* | PHYT | OMNI | WC |
| SILUGL | Siluridae | *Silurus* | *glanis* | PHYT | PISC | WC |
| SQUACE | Cyprinidae | *Squalius* | *cephalus* | PHLI | OMNI | WC |
| TELESO | Cyprinidae | *Telestes* | *souffia* | LITH | INV | WC |
| THYMTH | Salmonidae | *Thymallus* | *thymallus* | LITH | INV | WC |
| TINCTI | Cyprinidae | *Tinca* | *tinca* | PHYT | OMNI | BENT |
| TRIGQU | Cottidae | *Triglopsis* | *quadricornis* | LITH | INV | BENT |
| VIMBVI | Cyprinidae | *Vimba* | *vimba* | LITH | INV | BENT |

**REFERENCES :**


Alpay, S., J. J. Veillette, A. S. Dixit, et al. (2006). Regional and historical distributions of lake-water pH within a 100-km radius of the Horne smelter in Rouyn-Noranda, Quebec, Canada**:** 179-186.

Appelberg, M., B. C. Bergquist, et al. (2000). "Using fish to assess environmental disturbance of Swedish lakes and streams - a preliminary approach." Verhandlungen der Internationalen Vereinigung fuer Limnologie **27**: 311-315.

Appelberg, M. (2000). Swedish standard methods for sampling freshwater fish with multi-mesh gillnets. Drottningholm (Sweden), Institute of freshwater research**:** 28.

Argillier, C., O. Pronier and P. Irz (2002). "Approche typologique des peuplements piscicoles lacustres Français. I. Les communautés des plans d'eau d'altitude supérieure à 1500 m." Bulletin Français de Pêche et de Pisciculture **365/366**: 373-387.

Baker, E. A., K. E. Wehrly, et al. (2005). "A multimetric assessment of stream condition in the northern lakes and forests ecoregion using spatially explicit statistical Modeling and regional normalization." Transactions of the American Fisheries Society **134**(3): 697-710.

Barbour, C. D. and J. H. Brown (1974). "Fish species diversity in lakes." American Naturalist **108**: 473-489.

Barbour, M. T., J. L. Plafkin, B. P. Bradley, et al. (1992). "Evaluation of Epas Rapid Bioassessment Benthic Metrics - Metric Redundancy and Variability among Reference Stream Sites." Environmental Toxicology and Chemistry **11**(4): 437-449.

Barbour M.T., J.B. Stribling and J.R. Karr (1995). Multimetric approach for establishing biocriteria and measuring biological condition. Pages 63-77 in W.S Davis and T.P. Simon, editors. Biological assessment and criteria: tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida.

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling. 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish, Second Edition. EPA 841-B-99-002. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

Belpaire, C., R. Smolders, et al. (2000). "An Index of Biotic Integrity characterizing fish populations and the ecological quality of Flandrian water bodies." Hydrobiologia **434**(1-3): 17-33.

Brousseau, C. S., D. Baccante and L. W. Maki (1985). "Role of Bedrock and Surficial Geology in Determining the Sensitivity of Thunder-Bay Area Lakes to Acidification." Journal of Great Lakes Research **11**(4): 501-507.

Buffagni, A., S. Erba, M. Cazzola, et al. (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**(1): 313-329.

C.E.N., 2005. Water quality - Guidance on the scope and selection of fish sampling methods. prEN 14962 european standard draft. 24 pp.

Cardoso, A. C., A. Solimini, et al. (2007). "Phosphorus reference concentrations in European lakes." Hydrobiologia 584: 3-12.

Carpenter, S. R., N. F. Caraco, D. L. Correll, et al. (1998). "Nonpoint pollution of surface waters with phosphorus and nitrogen." Ecological Applications **8**(3): 559-568.

Chevan, A. and M. Sutherland (1991). “Hierarchical partitioning”. The American Statistician **45**:90-96.

Clarke, R. 2000. Uncertainty in estimates of biological quality based on RIVPACS. *in* J. F. Wright, D. W. Sutcliffe, and M. T. Furse, editors. Assessing the biological quality of freshwaters. Rivpacs and other techniques. Freshwater Biological Association, Ambleside, Cumbria, UK.

Daufresne, M and P. Boet (2007). “Climate change impacts on structure and diversity of fish communities in rivers.” Global Change Biology **13**(12): 2467-2478.

Drake, M. T. and D. L. Pereira (2002). "Development of a fish-based index of biotic integrity for small inland lakes in central Minnesota." North American Journal of Fisheries Management **22**(4): 1105-1123.

Drake, M. T. and R. D. Valley (2005). "Validation and application of a fish-based index of biotic integrity for small central Minnesota lakes." North American Journal of Fisheries Management **25**(3): 1095-1111.

Eadie, J. M., T. A. Hurly, R. D. Montgomerie, et al. (1986). "Lakes and rivers as islands: species-area relationships in the fish faunas of Ontario." Environmental Biology of Fishes **15**: 81–89.

European Commission (2000). Directive 2000/60/EC of the European Council and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. Official Journal of the European Communities, L327: 1-72.

Gassner, H., G. Tischler, et al. (2003). "Ecological integrity assessment of lakes using fish communities - suggestions of new metrics developed in two Austrian prealpine lakes." International Review of Hydrobiology **88**(6): 635-652.

Godinho, F. N., M. T. Ferreira and M. I. Portugal e Castro (1998). "Fish assemblage composition in relation to environmental gradients in Portuguese reservoirs." Aquatic Living Resources **11**(5): 325-334.

Harper, D. (1992). Eutrophication of freshwaters. Principles, problems and restoration. New York.

Hering, D., J. Bohmer, P. Haase, et al. (2004). "New methods for assessing freshwaters in Germany." Limnologica **34**(4): 281-282.

Hering, D., C. K. Feld, et al. (2006). "Cook book for the development of a Multimetric Index for biological condition of aquatic ecosystems: Experiences from the European AQEM and STAR projects and related initiatives." Hydrobiologia **566**: 311-324.

Hering, D., S. BIrk, A.L. Solheim, L. Carvalho et al. (2009). “Deliverable 2 2-2: Guidelines for indicator development”. 22p.

Hickman, G. D. and T. A. McDonough (1996). Assessing the reservoir fish assemblage index: A potential measure of reservoir quality. Multidimensional approaches to reservoir fisheries management. L. E. Miranda and D. R. DeVries. Chattanooga, Tenessee, American Fisheries Society. **16:** 85-97.

Hinch, S. G., N. C. Collins and H. H. Harvey (1991). "Relative abundance of littoral-zone fishes: Biotic interactions, abiotic factors, and postglacial colonization." Ecology **72**(4): 1314-1324.

Holmgren, K., A. Kinnerbäck, et al. (2007). "Bedömningsgrunder för fiskfaunans status i sjöar." Fiskeriverket Informerar **3**: 54.

Hughes, R. M., T. R. Whittier, S. A. Thiele, et al. (1992). Lake and stream indicators for the Unites States Environmental Protection Agency's environmental monitoring and assessment program. Ecological indicators - Vol. 1. D. H. McKenzie, D. E. Hyatt and V. J. McDonald. London, Elsevier Applied Science**:** 305-335.

Hughes R.M., P.R. Kaufmann, A.T. Herlihy, T.M. Kincaid, L. Reynolds & D.P. Larsen (1998). A process for developing and evaluating indices of fish assemblage integrity. Canadian Journal of Fisheries and Aquatic Sciences: **55**, 1618-1631.

Hughes, R. M., S. Howlin and P. R. Kaufmann (2004). "A biointegrity index (IBI) for coldwater streams of Western Oregon and Washington." Transactions of the American Fisheries Society **133**(6): 1497-1515.

Irz, P., J. De Bortoli, et al. (2007). "Controlling for natural variability in assessing the response of fish metrics to anthropogenic pressures for Northeast U.S.A. lakes." Aquatic Conservation: Marine and Freshwater Ecosystems **18**(5): 633-646.

Irz, P., C. Argillier and T. Oberdorff (2004). "Native and introduced fish species richness in French lakes: local and regional influences." Global Ecology and Biogeography **13**: 335-344.

Jaarsma, N., M. Klinge & R. Pot (eds), 2007. Achtergronddocument Referenties en Maatlatten Vissen ten behoeve van de Kaderrichtlijn Water. STOWA, Utrecht.

Jackson, D. A., P. R. Peres-Neto, et al. (2001). "What controls who is where in freshwater fish communities - the roles of biotic, abiotic, and spatial factors." Canadian Journal of Fisheries and Aquatic Sciences **58**(1): 157-170.

Jennings, M. J., L. S. Fore, et al. (1995). "Biological monitoring of fish assemblages in Tennessee Valley reservoirs." Regulated Rivers: Research & Management **11**(3-4): 263-274.

Jennings, M. J., J. Lyons, et al. (1999). Toward the development of an index of biotic integrity for inland lakes in Wisconsin. Assessing the sustainability and biological integrity of water resource quality using fish communities. T. P. Simon. Boca Raton, Florida, CRC Press**:** 541-562.

Karr J.R. (1981). Assessment of biotic integrity using fish communities. Fisheries **6**: 21-27.

Karr J.R. and DR. Dudley (1981). Ecological perspective on water quality goals. Environmental management **5**: 55-68.

Karr, J. R. and M. Dionne (1991). Designing surveys to assess biological integrity in lakes and reservoirs. Proceedings of symposium - Biological criteria: research and regulation, Washington, DC, Office of Water, U.S. Environmental Protection Agency.

Karr, J. R. and B. L. Kerans. 1992. Components of biological integrity: their definition and use in development of an invertebrate IBI. Pages 1-16 in W. S. Davis and T. P. Simon (eds.), Proceedings of the 1991 Midwest Pollution Control Biologists meeting. U. S. Environmental Protection Agency, Chicago, Illinois. EPA-905/R-92/003.

Karr J.R.and Chu E.W. 1999. Restoring Life in Running Waters: Better Biological Monitoring. Island Press, Washington, DC. 206 pp

Kilgour, B. W., and D. R. Barton. 1999. Associations between stream fish and benthos across environmental gradients in southern Ontario, Canada. Freshwater Biology 41:553-566.

Koskenniemi, E., E. K. Ranta, R. Palomaki, et al. (1988). On the natural and introduced fish fauna in Finnish reservoirs. Management of Freshwater Fisheries. Proceedings of a Symposium Organized By the European Inland Fisheries Advisory Commission, Goeteborg, Sweden, 31 May 3 Jun 1988., Goetboerg, PUDOC.

Leach, J. H., M. G. Johnson, J. R. M. Kelso, et al. (1977). "Responses of percid fishes and their habitats to eutrophication." Journal of the Fisheries Research Board of Canada **34**(10): 1964-1971.

Lyons, J., A. Gutierrez-Hernandez, et al. (2000). "Development of a preliminary index of biotic integrity (IBI) based on fish assemblages to assess ecosystem condition in the lakes of central Mexico." Hydrobiologia **418**: 57-72.

Magnuson, J. J., W. M. Tonn, A. Banerjee, et al. (1998). "Isolation vs. extinction in the assembly of fishes in small northern lakes." Ecology **79**(8): 2941-2956.

Mason N.W.H., P. Irz, et al. (2008). Evidence that niche specialization explains species-energy relationships in lake fish communities. Journal of animal ecology, **77**: 285-296.

Matuszek, J. E. and G. L. Beggs (1988). "Fish species richness in relation to lake area, pH, and other abiotic factors in Ontario lakes." Canadian Journal of Fisheries and Aquatic Sciences **45**: 1931 - 1941.

Mc Cormick F.H., R.M. Hugues and P.R. Kaufmann, D.V. Peck, J.L. Stoddard and A.T. Herlihy (2001). Development of an index of biotic integrity for the Mid-Atlantic Highlands region. Transactions of the American Fisheries Society 130, 857-877.

McDonough TA, Hickman GD. 1999. Reservoir Fishery Assessment Index development: a tool for assessing ecological health in Tennessee Valley Authority impoundments. In Assessing the Sustainability and Biological Integrity of Water Resource Quality Using Fish Communities, Simon TP (ed.). CRC Press: Boca Raton, FL; 523–540.

McDonough, T. A. and G. D. Hickman (1999). Reservoir Fishery Assessment Index development: a tool for assessing ecological health in Tennessee Valley Authority impoundments. Assessing the sustainability and biological integrity of water resource quality using fish communities. T. P. Simon. Boca Raton, Florida, CRC Press: 523-540.

Mehner, T., R. Arlinghaus, S. Berg, et al. (2004). "How to link biomanipulation and sustainable fisheries management: a step-by-step guideline for lakes of the European temperate zone." Fisheries Management and Ecology **11**(3-4): 261-275.

Ministère de l'Ecologie et du Développement Durable (2006). Circulaire DCE 2006/16 relative à la constitution et la mise en œuvre du programme de surveillance (contrôle de surveillance, contrôles opérationnels, contrôles d'enquête et contrôles additionnels) pour les eaux douces de surface (cours d'eau, canaux et plans d'eau) en application de la directive 2000/60/DCE du 23 octobre 2000 du Parlement et du Conseil établissant un cadre pour une politique communautaire dans le domaine de l'eau. Bulletin Officiel**:** 14.

Minns, C. K., V. W. Cairns, et al. (1994). "An Index of Biotic Integrity (IBI) for fish assemblages in the littoral zone of Great Lakes areas of concern." Canadian Journal of Fisheries and Aquatic Sciences **51**(8): 1804-1822.

New M., D. Lister, M. Hulme and I. Makin (2002). A high resolution data set of surface climate over global land areas. Climate Research, **21**: 1-25.

Oberdorff T., D. Pont, B. Hugueny & J.P. Porcher (2002). Development and validation of a fish based index for the assessment of "river health" in France. Freshwater Biology 47, 1720-1734

Ofenbock, T., O. Moog, J. Gerritsen, et al. (2004). "A stressor specific multimetric approach for monitoring running waters in Austria using benthic macro-invertebrates." Hydrobiologia **516**(1): 251-268.

Pont, D., B. Hugueny, U. Beier, et al. (2006). "Assessing the biotic integrity of rivers at the continental scale: a European approch." Journal of Applied Ecology **43**: 70-80.

R Development Core Team (2007). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rago, P. J. and J. G. Wiener (1986). "Does pH affect fish species richness when lake area is considered ?" Transactions of the American Fisheries Society **115**: 438-447.

Rahel, F. J. and J. J. Magnuson (1983). "Low pH and the absence of fish species in naturally acidic Wisconsin lakes: inferences for cultural acidification." Canadian Journal of Fisheries and Aquatic Sciences **40**: 3-9.

Rask M., M. Olin & J. Ruuhijärvi (2009). Fish based assessment of ecological status of Finnish lakes loaded by diffuse nutrient pollution from agriculture. Fisheries management and and Ecology (pre print).

Rathert, D., D. White, et al. (1999). "Environmental correlates of species richness for native freshwater fish in Oregon, USA." Journal of Biogeography **26**(2): 257-273.

Schmedtje, 2009. WG ECOSTAT, Guidance on the intercalibration process. 2008-2011 [Version 6.0]. 47p.

Schulz, E. J., M. V. Hoyer, et al. (1999). "An Index of Biotic Integrity: A test with limnological and fish data from sixty Florida lakes." Transactions of the American Fisheries Society **128**(4): 564-577.

Shaw, M. A. and J. R. M. Kelso (1992). "Environment factors influencing zooplankton species composition of lakes in north-central Ontario, Canada." Hydrobiologia **241**(3): 141-154.

Somers, K. M. and H. H. Harvey (1984). "Alteration of lake fish communities in response to acid precipitation and heavy-metal loadning near Wawa, Ontario." Canadian Journal of Fisheries and Aquatic Sciences **41**: 20-29.

Tammi, J., A. Lappalainen and M. Rask (2001). Using Swedish fish index FIX in assessing degradation of Finnish eutrophic lakes - what does fish community data tell about them? Classification of ecological status of lakes and rivers, ThemaNord. **584:** 37-39.

Thoma, F. (1999). Biological monitoring and an index of biotic integrity for lake Erie's nearshore waters. Assessing the sustainability and biological integrity of water resource quality using fish communities. T. P. Simon. Boca Raton, Florida, CRC Press: 417-461.

Tonn, W. M., J. J. Magnuson, M. Rask, et al. (1990). "Intercontinental comparison of small-lake fish assemblages: The balance between local and regional processes." American Naturalist **136**(3): 345-375.

Vlek, H. E., P. F. M. Verdonschot and R. C. Nijboer (2004). "Towards a multimetric index for the assessment of Dutch streams using benthic macroinvertebrates." Hydrobiologia **516**(1): 173-189.

Weigel, B. M., L. Z. Wang, P. W. Rasmussen, et al. (2003). "Relative influence of variables at multiple spatial scales on stream macroinvertebrates in the Northern Lakes and Forest ecoregion, USA." Freshwater Biology **48**(8): 1440-1461.

Whittier, T. R. (1999). Development of IBI metrics for lakes in Southern New England. Assessing the sustainability and biological integrity of water resource quality using fish communities. T. P. Simon. Boca Raton, Florida, CRC Press: 563-584.