Fisheries Management and Ecology, 2007, **14**, 441–452

Spatially based methods to assess the ecological status of riverine fish assemblages in European ecoregions

S. SCHMUTZ, A. MELCHER, C. FRANGEZ & G. HAIDVOGL

Institute for Hydrobiology and Aquatic Ecosystem Management, University of Natural Resources and Applied Life Sciences, Vienna, Austria

U. BEIER

National Board of Fisheries, Institute of Freshwater Research, Drottningholm, Sweden

J. BÖHMER

Institute for Zoology, University Hohenheim, Stuttgart, Germany

J. BREINE & I. SIMOENS

Institute for Forestry and Game Management, Groenendaal-Hoeilaart, Belgium

N. CAIOLA & A. DE SOSTOA

Department of Animal Biology, Faculty of Biology, University of Barcelona, Barcelona, Spain

M. T. FERREIRA & J. OLIVEIRA

Forest Department, Instituto Superior de Agronomia, Lisbon Technical University, Tapada da Ajuda, Lisbon, Portugal

G. GRENOUILLET & D. GOFFAUX

Facultés Universitaires N.D. de la Paix (University of Namur), Namur, Belgium

J. J. DE LEEUW

Netherlands Institute for Fisheries Research (RIVO), IJmuiden, The Netherlands

R. A. A. NOBLE

University of Hull, International Fisheries Institute, Hull, UK

N. ROSET

Office National de l'Eau et des Milieux Aquatiques-Délégation Régionale de Lyon, Bron, France

T. VIRBICKAS

Department of Aquatic Ecosystems, Institute of Ecology of Vilnius University, Vilnius, Lithuania

 doi: 10.1111/j.1365-2400.2007.00582.x

Abstract The objective was to develop spatially based (type-specific) methods to assess the ecological status of European rivers according to the EU Water Framework Directive. Some 15 000 samples from about 8000 sites were pre-classified within a five-tiered classification system based on hydromorphological and physico-chemical pressures. The pre-classification was used to identify reference conditions and to calibrate the assessment methods. Clustering reference sites based on relative species composition resulted in 60 fish assemblage types within 11 of the ecoregions under study. Discriminant function analyses (DFAs) were employed to identify environmental parameters characterising fish assemblage types; altitude, river slope, wetted width, mean air temperature and distance from source were the principal predictors. These environmental parameters were used to assign impacted sites with altered fish assemblage composition to the reference fish assemblage type. Metrics (fish assemblage descriptors) responding to human pressures were selected based on correlation and DFAs. Assessment methods were developed for 43 fish assemblage types. Metrics based on individual sentinel species were more often used in type-specific methods than metrics related to reproduction, habitat and feeding. Metrics based on long-distance migrants and potamodromous species were more sensitive to human pressures than overall composition metrics, e.g. total number of species. Only some of the tested metrics showed pressure-specific responses, i.e. reacted to one type of pressure but not to others. Insectivorous, intolerant and lithophilic species exclusively responded (decreased) to chemical and hydromorphological pressures in 14–19%. Omnivorous species was the only metric type that showed a consistent reaction (increase) to continuum disruptions in 25% of the cases. Accuracy of methods based on cross-validation with pre-classification varied between 47% and 98% (mean 81%) when contrasting calibration data set (class 1 and 2) with degraded sites (class 3, 4 and 5).

KEYWORDS: ecological status assessment, fish, index of biotic integrity, methods, Water Framework Directive.

Correspondence: Stefan Schmutz, Institute for Hydrobiology and Aquatic Ecosystem Management, University of Natural Resources and Applied Life Sciences, Max Emanuel-Str. 17, 1180 Vienna, Austria (e-mail: stefan.schmutz@boku.ac.at)

Introduction

The development of fish-based methods to assess human-induced impacts on aquatic ecosystems has been strongly influenced by the index of biotic integrity (IBI; Karr 1981). IBIs are based on the assumption that within a given entity the variability of fish communities is low enough to be able to distinguish between natural and human-induced variability. Although IBIs have been developed worldwide (Roset, Grenouillet, Goffaux, Pont & Kestemont 2007), appropriate delineation of regions or river types is a major issue in developing regional IBIs (Strange 1998) because of large-scale natural variability in fish communities. IBIs developed in North America account for natural variability by regionalisation and metric adjustments to longitudinal gradients (e.g. Hughes, Howlin & Kaufmann 2004).

The spatially based approach is the underlying methodological principle of the European Union's Water Framework Directive (WFD; 2000/60/EEC) for assessing the ecological status of running waters. The concept is that rivers can be classified into units (e.g. river segments) with homogenous characteristics. The WFD offers two options: System A and System B; both use only abiotic criteria for river typology. However, as the hypothesis is that the less the biotic heterogeneity within identified types the higher the accuracy of the employed IBI (Fausch, Lyons, Karr & Angermeier 1990; Smogor & Angermeier 1998, 2001), biotic typologies should be used. In principle, two spatial dimensions structure fish assemblages at the large scale: (1) the zoogeographic pattern reflecting mainly climatic gradients across the continental scale and (2) the longitudinal pattern within each river at the catchment scale caused by changes in environmental parameters along the river course (Beier, Degerman, Melcher, Rogers & Wirlöf 2007). Ecoregions are supposed to provide a spatial framework for ecosystem assessment at the large scale (Omernik 1987, 1995). Illies (1967) introduced a European classification system, dividing the continent into 25 ecoregions. Although Illies' ecoregion system is the only widely used classification, and has been adopted by the WFD, it has never been evaluated for its ability to discriminate among fish communities at a continental scale (Reyjol, Hugueny, Pont, Bianco, Beier, Caiola, Casals, Cowx, Economou, Ferreira, Haidvogl, Noble, de Sostoa, Vigneron & Verbickas 2007). The longitudinal zonation of rivers with a sequence of distinct fish communities was developed more than a century ago (Fritsch 1872). Several key parameters, e.g. slope/width ratio (Huet 1949), have been used to explain longitudinal fish community patterns in specific regions, but no attempt has been made to analyse longitudinal patterns at the ecoregion and continental scale. In the WFD, the classification of surface water bodies is based on the assumption that an abiotic river typology is adequate to stratify fish communities, but no efforts to validate this assumption at a European scale have been undertaken.

According to the IBI concept, assessment of ecological status is based on the comparison between observed and expected values of a set of defined metrics (fish assemblage descriptors) (Noble, Cowx, Goffaux & Kestemont 2007) combined into a single index (Karr 1981). Expected conditions can be defined in different ways (Hughes 1995), one is to use best available (least-disturbed) conditions resulting in unequal thresholds for less and more impacted fish assemblage types. However, the WFD requires standardised reference conditions showing no, or only minor, anthropogenic alterations. Therefore, most of the existing methods are not compliant with the WFD.

Information on pressures is necessary to distinguish between reference and impacted sites and for calibrating or scoring of metrics. However, compared with other aspects of IBI development, less emphasis has been dedicated to quantifying precisely the level of degradation.

Spatially based assessment methods (SBM) were developed for the WFD based on data from ecoregion 1 (Iberian Peninsula), 2 (Pyrenees) 4 (Alps) 8 & 13 (Western Highlands & Western Plains), 9 (Central Highlands), 14 (Central Plains), 15 (Baltic Province), 18 (Great Britain) 20 & 22 (Borealic Upland & Fenno-Scandinavian Shield). Detailed descriptions of ecoregional SBM developments are found in project reports (http://fame.boku.ac.at) and in examples presented in this volume on ecoregion 1 (Ferreira *et al.* 2007), ecoregions 8 & 13 (Grenouillet *et al.* 2007), ecoregion 15 (Virbickas & Kesminas 2007) and ecoregion 18 (Noble, Cowx & Starkie 2007). The objective of this paper was to summarise and to compare methodologies and results in these 11 ecoregions (Fig. 1) with respect to: (1) standardised identification and characterisation of fish assemblage types; (2) abiotic characteristics which discriminate among fish assemblage types; (3) metrics response to human pressures; (4) metrics aggregation and status classification; and (5) accuracy of assessment methods.

Figure 1. Illies' (1967) ecoregions. Ecoregions analysed in this paper are in bold.

Methods

All data were extracted from an extensive database (FIDES; Beier *et al.* 2007). For the spatially based approach, data from 14 789 fishing occasions at 7977 sites collected in 11 ecoregions and 11 countries were available. Fisheries data were based on single-pass electric fishing methodology only (Beier *et al.* 2007). Species were classified as alien or native (Noble *et al.* 2007).

Pre-classifications of sites based on potential pressures (Degerman, Beier, Breine, Melcher, Quataert, Rogers, Roset & Simoens 2007) were used to select unimpacted or weakly impacted sites for fish assemblage typology development and to calibrate the developed assessment methods against a five-tiered pressure gradient. The five classes of pressure status were defined following the normative classification of biological quality elements of the WFD: a pressure level that is supposed to result in minor impact on fish is equivalent to class 1, slight = class 2, moderate = class 3, strong = class 4 and severe = class 5.

Fish-based river typologies (i.e. fish assemblage types) for ecoregions were developed by cluster analyses and other techniques (see results) using only unimpacted or weakly impacted sites. To allocate new (impacted) monitoring sites to fish assemblage types (hereafter fish types) several environmental descriptors (Beier *et al.* 2007) were tested for their power to discriminate among fish types using stepwise discriminant analysis (SPSS 12.0®). Discriminant function analysis (DFA) generates functions with linear combinations of variables that maximises the probability of correctly assigning observations to their pre-determined groups. DFA can also be used to classify new observations into one of the groups (Quinn & Keough 2002). DFA has been used for the same and similar purposes in RIVPACS (Moss 2000), a system developed for invertebrates in the UK, and for predictive fish models (Joy & Death 2002).

More than 400 potential metrics (Noble *et al.* 2007) were analysed for their capacity to respond to human pressures. Candidate metrics, e.g. metrics with clear dose–response relationship, were tested for redundancy. Pressure-specific response of metrics was tested as follows. A metric was classified as pressure specific if it responded (Spearman's $r > |0.6|$) within a fish type to one type of pressure only, i.e. chemical (mean of nutrients/organic input and toxics/acidification/O_2), physical (mean of morphology and hydrology) or connectivity but not to another pressure type. Different options to aggregate final metrics and to set class boundaries among ecological status classes were evaluated (e.g. box plots and multivariate statistics). Error of spatially based methods was assessed by cross-validation between pre-classification of pressures and ecological status classification. Accuracy differences were tested among ecoregions and fish zones by ANOVA using the robust Welch-statistic for heterogeneous variance (SPSS 12.0®). Huet's classification (Huet 1949) was used to analyse differences along the longitudinal zonation.

Results

Data available for the analyses were originally collected for different purposes (e.g. fish species inventory, impact assessment studies and conservation monitoring; Beier *et al.* 2007). Hence, available fish data were unevenly distributed across Europe and countries and covered mainly western (Northern Portugal, France, England, Belgium and the Netherlands) and central Europe (Austria) (Beier *et al.* 2007; Reyjol *et al.* 2007).

Reference conditions and pressure status

Five core pressure variables (morphology, hydrology, nutrients/organic input, toxics/acidification/O_2 and connectivity) were combined to a mean global pressure index and used to select weakly impacted sites in all ecoregions. Other pressure variables (land use, urbanisation, riparian zone, sediment load, upstream dam, salinity, impact of stocking and introduction of fish) were added to account for regional peculiarities, where data were available (Table 1). In general, weakly impacted sites (calibration sites) encompass only pressure status classes 1 and 2 (minor and slight impact expected). However, because of the low number of non-impacted sites in ecoregions 1, 14, 15 and 18 some pressure variables also included sites with

Table 1. Pressure variables used to distinguish between calibration and impacted data set in different ecoregions (for definition of pressure variables see Degerman *et al.* 2007)

Human pressures	Iberian Peninsula 1	Pyrenees 2	Alps 4	Western Highlands & Western Plains 8 & 13	Central Highlands 9	Central Plains 14	Baltic Province 15	Great Britain 18	Borealic Upland & Fenno-Scandian Shield 20 & 22
Morphology	×	×	×	×	×	×	×	×	×
Hydrology	×	×	×	×	×	×	×	×	×
Toxics/acidification/O_2	×	×	×	×	×	×	×	×	×
Nutrients/organic input	×	×	×	×	×	×	×	×	×
Connectivity_segment	×	×		×	×		×	×	×
Land_use	×								
Urbanisation	×								
Riparian_zone	×								
Sediment_load	×								
Upstream_dam	×								
Salinity	×								
Impact_of_stocking									×
Introduction_of_fish									×

pressure status class 3, resulting in a contribution of moderately impacted sites of 28, 54, 26 and 39 percent in these four ecoregions. In total, 1920 calibration sites were separated from 4554 'impacted sites', encompassing 2200 rivers and 102 fish species.

Fish types

Ecoregions 20 and 22 as well as 8 and 13 were combined as preliminary analyses did not reveal significant differences in fish-type patterns based on cluster analyses. As a rule, relative abundance (catch data of the first run) and hierarchical cluster analyses (Ward's method) were used to group calibration samples with similar species composition except in ER2 and ER8/13 (absolute abundance), ER18 (absolute abundance standardised per species) and ER 20/22 (expert judgement process) (Table 2). In total, 60 fish types were identified ranging from two (Pyrenees) to eight types per ecoregion (median: 6.6 types per ecoregion). Eight types were added based on historical data or expert judgement but are not included in further analyses. Following Huet's classification, the highest proportion of types (31 types) corresponded to the trout, 13 to the grayling, 14 to the barbel and two to the bream zone. The mean number of fish species per type and sample ranged from 0.9 to 17.2 (median 5) with the lowest values in the trout region (0.9–5.9). *Salmo trutta fario* L. was the dominating species in 20 types followed by *Phoxinus phoxinus* (L.) (10 types), *Rutilus rutilus* (L.) and *Salmo salar* L. (four types each). When only the five most abundant species per fish type were considered, the total diversity comprised only 47 species.

Discriminating fish types

In general, DFAs were used to identify environmental descriptors predicting fish types and assigning degraded sites to fish types. In ecoregion 20/22, an expert decision system was employed (Table 2). Correct assignment of sites to fish types using environmental descriptors ranged from 18% (ER 14) to 100% (ER 1). Between 2 and 6 environmental descriptors (median 5) were used to discriminate fish types. Altitude, river slope, wetted width, mean air temperature (annual and/or July) and distance from source were the prevailing environmental descriptors.

Metrics response

In all ecoregions, the mean of the five human pressure variables was used as an index for human pressure. Spearman's rank correlation (nine ecoregions) and discriminant analysis (four ecoregions) were the primary methods to analyse metrics response to human

Table 2. Methods used for spatially based assessment method development

Method	Iberian Peninsula 1	Pyrenees 2	Alps 4	Western Highlands & Western Plains 8 & 13	Central Highlands 9	Central Plains 14	Baltic Province 15	Great Britain 18	Borealic Upland & Fenno-Scandian Shield 20 & 22
Fish types									
Cluster analyses	×	×	×	×	×	×	×	×	–
Expert judgement	–	–	–	–	–	–	–	–	×
Discrimination between types									
Discriminant analyses	×	×	×	×	×	×	×	×	–
Selection of metrics									
Spearman rank correlation	×	×	×	×	×	×	×	×	–
Test of metrics redundancy									
Spearman rank correlation	×	×	×	–	×	–	×	×	–
Class boundaries									
Discriminant analyses	–	–	×	×	×	–	–	×	–
Box plots	×	×	–	–	–	×	×	–	–
Mann-Whitney *U*-test	×	–	–	–	–	–	–	–	–
Kruskal–Wallis	–	×	–	–	–	–	–	–	–
Metrics combination									
Discriminant analyses	–	–	×	×	×	–	–	×	–
Multimetric index	×	×	–	–	×	×	×	–	–
Method validation									
Independent data set	–	–	–	×	–	–	–	×	–

pressures (Table 1). For a more detailed description of how discriminant analysis was used for ecological status allocation see Melcher, Schmutz, Haidvogl & Moder (2007).

Only responsive metrics (Spearman's rank correlation $P < 0.05$ or $r > |0.4|$) were selected for further analyses. Redundancy among metrics was tested by Spearman's rank correlation and metrics with $r > |0.80|$ were excluded. In ecoregions 1, 2 and 14, each metric was scored from 1 to 5 by plotting metric values against mean pressure (box plots). The final index was calculated as the average metric score. In ecoregion 9, metrics were normalised from 0 to 1 and the average of the metrics was scored by box plots (index vs mean pressure). Scores and class boundaries were set either for each single metric (ecoregions 9, 14 and 15) or by discriminant analysis (ecoregions 4, 8/13, 15 and 18), as it was done for metrics selection using the pre-classification of pressures for calibration (Table 1).

Spatially based methods were developed for 43 fish types in eight ecoregions (total number of samples: 10 120). The number of available samples per fish type ranged from 16 to 1120 (median 235). For the remaining fish types ($n = 16$) the number of samples or their distribution along pressure gradient was not adequate for method development. All fish types in ER 20/22 were omitted because of lack of sufficient impacted sites. The Swedish database used for this approach covered only small streams that are less likely to be impacted.

In total, 130 different metrics (of 451 metrics tested) were selected with a median of 9.2 metrics per fish type, with percentage of insectivorous individuals, percentage of lithophilic individuals and percentage of lithophilic species selected most often. The 10 most often used metrics represent about one-third of selected metrics. All of the 10 most often used metrics represent relative measurements (nine of 10) or presence/absence information (Fig. 2).

Similar metrics, based on grouping metrics into different types (i.e. functional metrics and sentinel metrics) and variants (i.e. same units) were used for all fish zones, i.e. trout, grayling and barbel zone (in the bream region only two fish types were identified). Metrics based on individual sentinel species were most commonly selected followed by metrics related to reproduction, habitat and feeding. About one-third of all sentinel metrics were based on *S. trutta fario*. Metrics based on long-distance migrants and potamodromous species were more important than overall composition metrics (e.g. total number of species) (Table 3). With respect to metric variants, density metrics (ind ha^{-1} and % ind ha^{-1}) were used as often as all other metrics such as number of species, biomass or 0+ fish (Table 4).

Only some of the tested metrics showed pressure-specific response, i.e. reacted to one type of pressure but not to another. Insectivorous, intolerant and lithophilic metrics exclusively responded (decreased) to chemical and hydromorphological pressures in only 14–39% of the cases. However, in some cases, these metrics showed no or an opposite response (increase) to connectivity disruptions. Only omnivorous species (metrics) showed a consistent reaction (increase) to disruptions of the continuum, in 25% of the cases (Table 5).

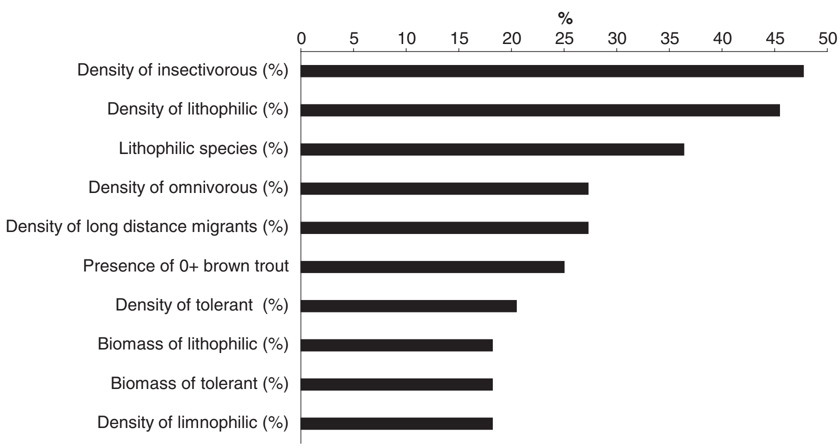

Figure 2. Ten metrics most often used within 43 fish type-specific methods of eight ecoregions (calculated as the percentage of use within the 43 fish types, (%) represent relative measurements).

Table 3. Relative importance of metric types (% of all metrics used) in different fish zones

	Total	Trout	Grayling	Barbel	Bream
Number of types	44	31	13	14	2
Sentinel	23.4	18.4	21.2	31.8	5.0
Reproduction	17.6	17.1	18.2	15.6	35.0
Habitat	15.1	17.7	21.2	10.4	10.0
Feeding	14.9	15.2	16.7	12.3	25.0
Tolerance	10.1	7.6	7.6	12.3	20.0
Migration	7.3	6.3	12.1	6.5	5.0
Overall composition	6.0	9.5	1.5	5.2	0.0
Longevity	2.8	3.2	0.0	3.9	0.0
Historical metrics	1.5	3.8	0.0	0.0	0.0
0+	1.3	1.3	1.5	1.3	0.0

Sentinel: based on individual sentinel species; overall composition: species diversity, native species, exotic species, etc.; historical metrics: based on absence/presence of historically documented species; 0+: occurrence, number or relative number of YOY individuals of sentinel species; for further details on metrics definition see Noble *et al.* (2007).

Table 4. Relative importance of metric variations (% of all metrics used) in different fish zones

	Total	Trout	Grayling	Barbel	Bream
No. of fish types	44	31	13	14	2
% density	31.2	20.9	45.5	32.0	60.0
Density	20.4	17.7	15.2	28.1	0.0
% species	16.9	20.9	13.6	13.7	20.0
Biomass	9.6	9.5	10.6	10.5	0.0
No. of species	9.6	16.5	6.1	5.2	0.0
% biomass	6.8	6.3	3.0	7.2	20.0
Presence 0+	4.0	5.7	4.5	2.6	0.0
% presence 0+	1.5	2.5	1.5	0.7	0.0

% density/% biomass: relative density/biomass; density/biomass: number of individuals per hectare; % species: proportion of species compared with all species; (%) presence 0+: (%) number of sentinel species with YOY fish; all metrics are based on single-pass electric fishing samples; for further details on metrics definition see Noble *et al.* (2007).

Method accuracy

Accuracy of methods based on cross-validation with pre-classification varied between 47% and 98% (mean 81%) when contrasting calibration data set (class 1 and 2) with degraded sites (classes 3, 4 and 5). Method accuracy was different among ecoregions (Welch test = 7.914, d.f. = 7, $P = 0.002$) but homogeneous across fish zones (Welch test = 0.018, d.f. = 3, $P = 0.996$). The lowest accuracy was achieved in the Alps ecoregion (58%), and the highest in Western Highlands/Western Plains (87%). Mean type I error in classification, i.e. a site is classified by the method as impacted, although the pre-classification indicated no severe pressures, was 8.9% (range 0–47%) and similar to type II error, i.e. a site is classified by the method as unimpacted although the pre-classification indicates severe pressures, ranging from 0% to 33% (mean 10.2 %). Sample size had a significant, but marginal, influence on the method accuracy ($r^2 = 0.13$, $F = 7.4$, $P < 0.01$). Methods were validated by independent data sets in only two ecoregions (Table 2).

Table 5. Pressure-specific response of functional metrics calculated as exclusive response (Spearman's $r_s > |0.6|$) to one of the pressure groups (in percent); − indicates a decrease; + an increase of the metrics (only metrics that show pressure-specific response in at least 25% of the 43 methods in one category are shown; sentinel species metrics are not considered)

	Pressure					
	Chemical		Hydro-morpho-logical		Connectivity	
Metric	−	+	−	+	−	+
Insectivorous	39	0	20	0	0	9
Intolerant	39	0	16	0	14	5
Lithophilic	25	0	14	0	0	2
Rheophilic	27	0	9	0	0	0
Omnivorous	0	7	0	7	0	25

Discussion

This paper represents one of the first attempts to develop data-driven assessment methods for running waters at the European scale. Based on a large European database (FIDES), sampling sites were pre-classified according to human pressures to distinguish between calibration and impacted sites. Calibration sites were used to develop fish typologies of running waters within 11 ecoregions. Environmental descriptors were identified for discriminating among fish types and for allocating impacted sites to fish types. Type-specific metrics responding to pressures were used to discriminate between ecological status classes allowing the allocation of new sites to ecological status classes (Fig. 3).

Fish types

The data covered a wide range of different ecoregions and rivers types representative of western, central and northern Europe. However, there was a paucity of data for Mediterranean and eastern countries.

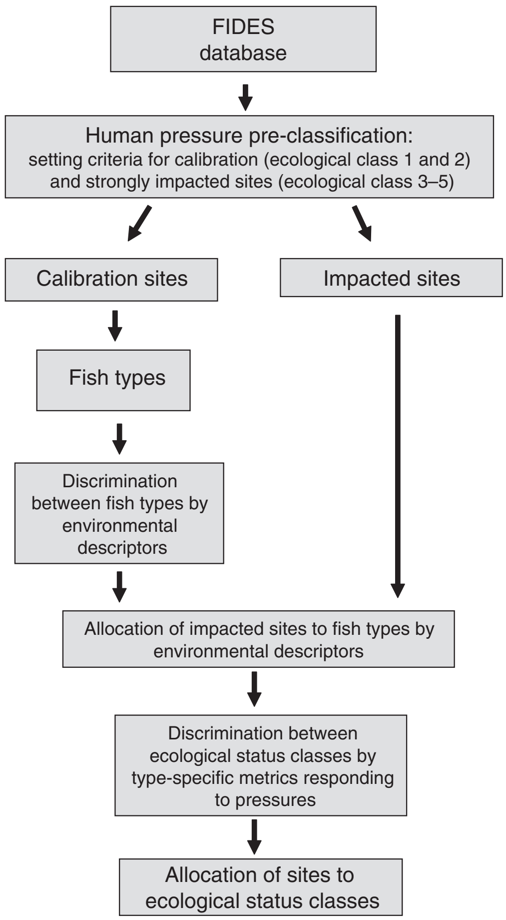

Figure 3. Flow chart for the development of spatially based assessment methods.

One of the major challenges of this pan-European approach was the standardised selection of reference sites according to the WFD criteria. Strictly following the WFD criteria resulted in a lack of reference sites in some ecoregions (ecoregions 1, 14, 15 and 18) and sites with pressure status 3 had to be included for the development of the fish types. Such an approach seems not to violate the WFD rules as calibration data were only used for fish type identification but not for discriminating among the different levels of deterioration within the final assessment methods. However, there still remains the question of how much deterioration at calibration sites affects fish typology.

Most IBIs were developed for specific regions or countries. Within-region variability of reference sites has received little attention. In the proposed approach, reference sites were clustered using relative species composition into fish types and two to eight distinct fish types were identified in each of the ecoregions. As the level of discrimination among groups within cluster analyses is arbitrary and relies on expert judgement, other thresholds would have resulted in different groupings. However, there is a trade-off between defining groups as small and hence homogeneous as possible and the availability of data. As data availability varied among ecoregions and some types by nature are scarce the criteria for defining thresholds in the examination of the cluster dendograms also differed among ecoregions. It was assumed that these differences did not affect the accuracy of assessment methods developed as Melcher *et al.* (2007) showed more broadly defined fish assemblage types produced reliable results.

The spatio-temporal pattern of fish communities is a result of a plethora of factors acting across several scales. Both top-down factors, e.g. zoogeography, geomorphology, and bottom-up factors, e.g. habitat availability and competition, are known to determine the structure of fish assemblages. Smogor & Angermeier (1998) used physiographical regions for delineating IBI regions to account for upland and lowland river section differences. To account for within-region variability, Fausch, Karr & Yant (1984) suggested using stream order or catchment area. Some authors have tested stream-size dependence of IBI scores or metrics (e.g. Mundahl & Simon 1998) or adjusted metrics to watershed size (e.g. Hughes *et al.* 2004). However, in most IBI developments no adjustments have been made (Smogor & Angermeier 1998). The DFAs showed that not only stream size (wetted width), but also altitude, slope, mean air temperature and distance from source have effects on species composition, parameters that correlate only partly with stream size. A median number of five environmental descriptors were used to discriminate among fish types. Most identified environmental parameters structuring fish assemblages describe the longitudinal gradient of streams and rivers.

In the USA, separate IBIs have been developed for cold (e.g. Hughes *et al.* 2004) and warm water streams (e.g. Karr 1981). The present results, however, show that fish communities not only differ considerably both among biogeographical regions and among cold and warm waters but also within cold and warm waters as a result of the longitudinal zonation. For example, in the Alps ecoregion, six different fish types were found for cold water streams (trout and grayling fish zone) showing different proportions of brown trout or even dominance of cyprinids (*Leuciscus souffia* Risso). However, developing a high number of spatially homogeneous units accounting mainly for longitudinal gradients resulted in a deficit of number of sites left per individual fish type preventing IBI development in some of the types.

The diversity of fish fauna in European rivers is low compared with North America. The mean expectation of number of fish species based on the database was four, showing an increase from the trout to bream zones. Forty-seven species accounted for the five most abundant species expected at a given site in all the fish types identified. Although Mediterranean rivers are under-represented in the database, about one-quarter of all species are endemic to the Mediterranean region underlining the distinctness of southern European fish fauna and by comparison, the homogeneous fish communities in western, central and northern Europe.

Human pressure pre-classification

Hydromorphological alterations, water pollution and continuum disruptions combined into a mean pressure index, functioned as a calibration index for metric selection and class boundary setting. Unlike other IBI developments, a five-tiered pre-classification system was used. This had several advantages: (1) identification of fish types where insufficient numbers of reference sites or impacted sites with different levels of degradation were available – these could be eliminated before proceeding with IBI development; (2) pre-classification was used to distinguish between reference and impacted sites and to calibrate the fish-based index; and (3) comparing IBI scores with pressure pre-classification enabled cross-validation of the methods developed.

Metrics

Although more than 400 metrics were tested, few were eventually employed. The most commonly selected metrics were percentage of insectivorous individuals, percentage of lithophilous individuals and percentage of lithophilous species. Insectivorous (alternately invertivorous; Oberdorff, Pont, Hugueny & Chessel 2001; Belpaire, Smolders, Vanden Auweele, Ercken, Breine, Van Thuyne & Ollevier 2000; Breine, Simoens, Goethals, Quataert, Ercken, Van Liefferinghe & Belpaire 2004) and lithophilous (Oberdorff & Hughes 1992) were used in other European IBIs.

Grouping metrics into metric types and exploring their relative importance showed that sentinel species were the predominant type of metric used in the 43 methods (Table 3). The spatially based approach minimised natural variability within fish types. This favoured the selection of species-specific (sentinel) metrics over other metrics, as the clustering of fish types was based on relative species composition. Species were classified as sentinel species if they were typical of the distinct fish communities (e.g. fish zones), sensitive to human disturbances and sufficiently abundant and well distributed under undisturbed conditions (Noble *et al.* 2007). About one-third of all sentinel metrics were based on brown trout, indicating that in species poor fish types sentinel species metrics were more important than functional metrics (Hughes *et al.* 2004).

In addition to species-specific metrics, functional metrics referring to reproduction, habitat and feeding were used in all types and fish zones. Newly introduced metrics such as long-distance migrants and potamodromous species were less often used, but are more frequently employed (7.1%) than metrics such as total number of species (6.1%) that were used in earlier IBIs (Karr 1981). This might arise because species diversity can react to human pressures both in terms of a decrease (e.g. loss of intolerant species) and an increase (e.g. additional tolerant species). An increase of species number was observed, for example, in impounded rivers (Martinez, Chart, Trammell, Wullschleger & Bergersen 1994). Historical metrics (e.g. % of original species) were rarely used because few countries (Austria, Germany and the Netherlands) provided full data sets. However, historical metrics are the only option to develop assessment methods in regions and fish types, e.g. ecoregions 14 and large rivers, where appropriate reference sites are no longer available (de Leeuw, Buijse, Haidvogl, Lapinska, Noble, Repecka, Virbickas, Wisniewolski & Wolter 2007).

With respect to metric variants, density metrics (ind ha^{-1} and % ind ha^{-1}) were used as often as all other metric variants combined, i.e. number of species, biomass and 0+ fish. This confirms the general trend in IBI developments that at least semi-quantitative information on fish communities is necessary to detect human impacts. Biomass and 0+ data were supplied by only a few countries; therefore, those metrics were under-represented in the database, but proved to be effective where data were available (e.g. ecoregion Alps and fish-type hyporhithral, see Melcher *et al.* 2007).

Only some of the tested metrics showed pressure-specific response, i.e. reacted to one type of pressure but not to another. Insectivorous, intolerant and lithophilic species decreased exclusively to chemical and hydromorphological pressures in 9–39% of the cases. However, in some cases, those species show no or an opposite response (increase) to connectivity disruptions. Omnivore metrics were the only ones that showed a consistent reaction (increase) to continuum disruptions in 25% of the cases.

After identifying impacted sites, the next step of water management is to identify appropriate restoration measures. For that, the index or individual metrics should be able to infer which type of restoration measures would be appropriate. More detailed investigations on pressure-specific response of fish are necessary to fulfil this management requirement.

Multivariate analyses

Discriminant function analyses were used to develop assessment methods elsewhere. Wright, Armitage, Furse & Moss (1984) used DFA to predict reference conditions based on environmental parameters and Joy & Death (2002) used DFA to find discriminant functions of landscape variables to best differentiate among IBI classes. However, no method appears to have used DFA for metric selection (stepwise procedure) and class assignment, i.e. index computation. The advantage of this approach is that metric combinations are selected that have the highest probability of explaining biological responses to human pressures. Metrics are weighted in the discriminant function according to their individual contribution to the overall pressure–index relationship. In all IBIs developed so far, each metric receives equal weight by summing or averaging individual metric scores. However, the sensitivity of metrics may differ considerably depending on the type of metric. For example, the principle of the Multi Level Fish-based Assessment method (MuLFA; Schmutz, Kaufmann, Vogel, Jungwirth & Muhar 2000) is based on the assumption that metrics derived from higher biological organisational level (e.g. species occurrence) might react only to high doses, while metrics of lower levels (e.g. population age structure) are sensitive to lower doses of human pressures. Using stepwise procedures of multivariate statistics might solve this problem in selecting the most influential metrics. However, care is needed because there are several potential pitfalls (Quinn & Keough 2002). Stepwise procedures are dependent on defining thresholds (e.g. significance level < 0.10) for selecting variables to be included in the model, but these are always arbitrary. Different stepwise techniques can produce very different final models even from the same data, impeding meaningful interpretation. In addition, the problem of collinearity among predictors in general increases with the number of variables. Therefore, uni-variate testing, pre-selection of candidate metrics and elimination of redundant metrics before applying DFA, as carried out in the proposed approach, is recommended. Automated fitting of all subsets (i.e. potential variable combinations) would be an alternative to stepwise procedures and is becoming more feasible with increasing computer power (Melcher *et al.* 2007).

Method accuracy

Accuracy of methods based on cross-validation with pre-classification varied between 47% and 98% (mean 81%) when contrasting calibration data set with degraded sites. Method accuracy did not vary among fish zones indicating consistency along the river continuum although human pressures tended to increase in an upstream–downstream continuum. Reduced availability of reference sites in lowland rivers did not hamper the development of robust methods for the barbel zone. However, only two fish types were assigned to the bream zone indicating a lack of available data in those fish types. The same method accuracy in fish types was achieved with low and high species diversity demonstrating that by selecting adequate metrics low species diversity does not undermine the development of IBIs.

Method accuracy varied considerably among ecoregions. Lowest accuracy was achieved in the ecoregion Alps. More than 90% of Alpine data were from Austria where rivers are generally only impacted by physical pressures but rarely by chemical pressures (BMLFUW 2002). Averaging physical and chemical pressures – as for the global pressure index of the pre-classification – leads to an underestimation of human pressure for Austrian rivers. Consequently, the mismatch between the pre-classification and the observed fish status is higher. Alternative approaches for defining a global pressure index for rivers impacted by a single pressure should be developed in future.

Based on a large data set, a consistent procedure for developing type-specific assessment methods could be applied to eight European ecoregions. However, splitting the database into 60 fish types impeded the method development or lowered accuracy in some of the types because of an insufficient number of remaining samples and hampered method validation with independent data. To overcome limitations of the spatially based approach at the ecoregional level merging fish types across ecoregions is proposed (Melcher *et al.* 2007).

Acknowledgments

We would like to thank the European Commission for funding the FAME project (EVR1-CT-2001-00094). Special thanks are dedicated to the many institutions in Europe providing data for the FAME project.

References

Beier U., Degerman E., Melcher A., Rogers C. & Wirlöf H. (2007) Processes of collating a European fisheries database to meet the objectives of the European Union Water Framework Directive. *Fisheries Management and Ecology* **14,** 407–416.

Belpaire C., Smolders R., Vanden Auweele I., Ercken D., Breine J., Van Thuyne G. & Ollevier F. (2000) An index of biotic integrity characterizing fish populations and the ecological quality of Flandrian water bodies. *Hydrobiologia* **434,** 17–33.

BMLFUW (2002) *Gewässerschutzbericht 2002*. Wein: Bundesministerium für Landwirtschaft, Forstwirtschaft, Umwelt und Wasserwirtschaft, 212 pp.

Breine J., Simoens I., Goethals P., Quataert P., Ercken D., Van Liefferinghe C. & Belpaire C. (2004) A fish-based index of biotic integrity for upstream brooks in Flanders (Belgium). *Hydrobiologia* **522,** 133–148.

Degerman E., Beier U., Breine J., Melcher A., Quataert P., Rogers C., Roset N. & Simoens I. (2007) Classification and assessment of degradation in European running waters. *Fisheries Management and Ecology* **14,** 417–426.

Fausch K.D., Karr J.R. & Yant P.R. (1984) Regional application of an index of biotic integrity based on stream fish communities. *Transactions of the American Fisheries Society* **113,** 39–55.

Fausch K.D., Lyons J., Karr J.R. & Angermeier P.L. (1990) Fish communities as indicators of environmental degradation. In: S.M. Adams (ed.) *Biological Indicators of Stress in Fish*, Vol. 8. Bethesda, MD: American Fisheries Society Symposium, pp. 123–144.

Ferreira T., Caiola N., Casals F., Oliveira J.M. & de Sostoa A. (2007) Assessing perturbation of river fish communities in the Iberian Ecoregion. *Fisheries Management and Ecology* **14,** 519–530.

Fritsch A.J. (1872) Die Wirbeltiere Böhmens. Ein Verzeichnis aller bisher in Böhmen beobachteten Säugetiere, Vögel, Amphibien und Fische. *Archiv für naturwissenschaftliche Landesdurchforschung von Böhmen* **2,** 1–152.

Grenouillet G., Roset N., Goffaux D., Breine J., Simoens I., de Leuuw J.J. & Kestemont P. (2007) Fish assemblages in European Western Highlands and Western Plains: a type-specific approach to assess ecological quality of running waters. *Fisheries Management and Ecology* **14,** 509–517.

Huet M. (1949) Apercu dés relations entre la pente et les populations des eaux courantes. *Schweizerische Zeitschrift fur Hydrologie* **11,** 333–351.

Hughes R.M. (1995) Defining acceptable biological status by comparing with reference conditions. In: W.S. Davies & T.P. Simon (eds) *Biological Assessment and Criteria; Tools for Water Resource Planning and Decision Making*. Boca Raton, FL: Lewis Publishers, pp. 31–47.

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

Illies J. (1967) *Limnofauna europaea. Eine Zusammenstellung aller die europäischen Binnengewässer bewohnenden mehrzelligen Tierarten mit Angaben über ihre Verbreitung und Ökologie*. Stuttgart: Gustav Fischer Verlag, 454 pp (in German.).

Joy M.K. & Death R.G. (2002) Predictive modelling of freshwater fish as a biomonitoring tool in New Zealand. *Freshwater Biology* **47,** 2261–2275.

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

de Leeuw J.J., Buijse A.D., Haidvogl G., Lapinska M., Noble R.A.A., Repecka R., Virbickas T., Wiśniewolski W. & Wolter C. (2007) Challenges in developing fish-based ecological assessment methods for large floodplain rivers. *Fisheries Management and Ecology* **14,** 483–494.

Martinez P.J., Chart T.E., Trammell M.A., Wullschleger J.G. & Bergersen E.P. (1994) Fish species composition before and after construction of a main stem reservoir on the White River, Colorado. *Environmental Biology of Fishes* **40,** 227–239.

Melcher A., Schmutz S., Haidvogl G. & Moder K. (2007) Spatially based methods to assess the ecological status of European fish assemblage types. *Fisheries Management and Ecology* **14,** 453–463.

Moss D. (2000) Evolution of statistical methods in RIVPACS. In: J.F. Wright, D.W. Sutcliffe & M.T. Furse (eds) *Assessing the Biological Quality of Fresh Waters: RIVPACS and other Techniques*. Ambleside, UK: Freshwater Biological Association, pp. 25–38.

Mundahl N.D. & Simon T.P. (1998) Development and Application of an index of biotic integrity for Coldwater Streams of the Upper Midwestern United States. In: T.S. Simon (ed.) *Assessing the Sustainability and Biological Integrity of Water Resources Using Fish Communities*. Boca Raton, FL: CRC Press, pp. 383–416.

Noble R.A.A., Cowx I.G., Goffaux D. & Kestemont P. (2007) Assessing the health of European rivers using functional ecological guilds of fish communities: standardising species classification and approaches to metric selection. *Fisheries Management and Ecology* **14,** 381–392.

Noble R.A.A., Cowx I.G. & Starkie A. (2007) Development of fish-based methods for the assessment of ecological status in English and Welsh rivers. *Fisheries Management and Ecology* **14,** 495–508.

Oberdorff T. & Hughes R.M. (1992) Modification of an index of biotic integrity based on fish assemblages to characterize rivers of the Seine Basin, France. *Hydrobiologia* **228,** 117–130.

Oberdorff T., Pont D., Hugueny B. & Chessel D. (2001) A probabilistic model characterizing fish assemblages of French rivers: a framework for environmental assessment. *Freshwater Biology* **46,** 399–415.

Omernik J.M. (1987) Ecoregions of the conterminous United States. *Annals of the Association of American Geographers* **77,** 118–125.

Omernik J.M. (1995) Ecoregions: a spatial framework for environmental management. In: W.S. Davies & T.P. Simon (ed.) *Biological Assessment and Criteria; Tools for Water Resource Planning and Decision Making*. Boca Raton, FL: Lewis Publishers, pp. 49–62.

Quinn G.P. & Keough M.J. (2002) *Experimental Design and Data Analysis for Biologists*. Cambridge: Cambridge University Press, 537 pp.

Reyjol Y., Hugueny B., Pont D., Bianco P.G., Beier U., Caiola N., Casals F., Cowx I., Economou A., Ferreira T., Haidvogl G., Noble R., De Sostoa A., Vigneron T. & Virbickas T. (2007) Patterns in species richness and endemism of European freshwater fish. *Global Ecology and Biogeography* **16,** 65–75.

Roset N., Grenouillet G., Goffaux D., Pont D. & Kestemont P. (2007) A review of existing fish assemblage indicators and methodologies. *Fisheries Management and Ecology* **14,** 393–405.

Schmutz S., Kaufmann M., Vogel B., Jungwirth M. & Muhar S. (2000) A multi-level concept for fish-based, river-type-specific assessment of ecological integrity. In: M. Jungwirth, S. Muhar & S. Schmutz (eds) *Assessing the Ecological Integrity of Running Waters*. Dordrecht: Kluwer Academic Publishers, pp. 279–289.

Smogor R.A. & Angermeier P.L. (1998) Relations between fish metrics and measures of anthropogenic disturbance in three IBI regions in Virginia. In: T.P. Simon (ed.) *Assessment Approaches for Estimating Biological Integrity using Fish Assemblages*. Boca Raton, FL: Lewis Publishers, pp. 585–610.

Smogor R.A. & Angermeier P.L. (2001) Determining a regional framework for assessing biotic integrity of Virginia streams. *Transactions of the American Fisheries Society* **130,** 18–35.

Strange R.M. (1998) Historical biogeography, ecology and fish distribution: conceptual issues for establishing IBI criteria. In: T.P. Simon (ed.) *Assessment Approaches for Estimating Biological Integrity using Fish Assemblages*. Boca Raton, FL: Lewis Publishers, pp. 65–78.

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

Wright J.F., Armitage P.D., Furse M.T. & Moss D. (1984) The classification of sites on British rivers using macroinvertebrates. *Verhandlungen der Internationalen Vereinigung für Theoretische und Angewandte Limnologie* **22,** 1939–1943.