Establishing demographic reference condition procedures for Italian river fish bioindicator

Giovanni Rossi, Andrea Marchi, Salvatore De Bonis, Gianluca Zuffi, Gaetano Caricato, Stefano Goffredo, Cristina Martone, Stefano Macchio, Gian Luigi Rossi & Stefania Balzamo

View supplementary material

Published online: 20 Jul 2021.

Submit your article to this journal

Article views: 434

View related articles

View Crossmark data

Taylor & Francis
Taylor & Francis Group

Establishing demographic reference condition procedures for Italian river fish bioindicator

Giovanni Rossi[a], Andrea Marchi[b], Salvatore De Bonis[c], Gianluca Zuffi[b], Gaetano Caricato[d], Stefano Goffredo[e], Cristina Martone[f], Stefano Macchio[f], Gian Luigi Rossi[g] and Stefania Balzamo[f]

[a]Arpae - Regional Agency for Prevention, Environment and Energy of Emilia-Romagna, Bologna, Italy; [b]Hydrosynergy SC - Spin-off of the University of Bologna, San Lazzaro di Savena, Bologna, Italy; [c]Arpa Lazio - Regional Agency for Environmental Protection of Lazio, Rome, Italy; [d]Arpa Basilicata - Regional Agency for Environmental Protection of Basilicata, Matera, Italy; [e]BiGeA - Department of Biological, Geological and Environmental Sciences, University of Bologna, Bologna, Italy; [f]ISPRA - Italian Institute for Environmental Protection and Research, Rome, Italy; [g]ENEA - Italian National Agency for New Technologies, Energy and Sustainable Economic Development - Division Protection and Enhancement of the Natural Capital - Laboratory Biodiversity and Ecosystem Services, Saluggia Research Center, Saluggia, Vercelli, Italy

ABSTRACT

The Water Framework Directive 2000/60/EC (WFD) requires Member States of the European Union (EU) to use a pool of bioindicators, that is, Biological Quality Elements (BQEs), including fish species and each Member State shall, in fact, certify the compliance with the ecological quality goals of the WFD. The New Italian Index of the Ecological State of Fish Communities (NISECI) is the WFD-compliant index for the analysis of the fish species BQE in Italian river water bodies, both wadable and non-wadable, that are not (using WFD definitions) heavily modified or artificial. The WFD requires each index to relate the conditions found in the field to reference conditions corresponding to the absence of anthropogenic alterations or to minor anthropogenic alterations. The present study defines the logical basis on the model of refinement for the reference conditions of the metrics and submetrics of NISECI relating to the biological condition of indigenous and non-indigenous species. The model was tested on an empirical basis using data from 284 ichthyological samplings taken from river basins throughout Italy. In order to clean random errors and stochastic noise due to possible different sampling methods, the data was only collected using protocols that guarantee adequately homogeneous and reliable data of abundance (overall and by size). The developed approach provided comparable and homogeneous results throughout Italian watersheds; this allowed a standardized and robust refinement of the threshold values of the demographic parameters of NISECI. The procedure used, consistent with the spatially based reference condition and

CONTACT Giovanni Rossi ✉ grossi@arpae.it 🗨 Arpae - Regional Agency for Prevention, Environment and Energy of Emilia-Romagna, Via F. Rocchi 19, 40138 Bologna, Italy.

Supplemental data for this article is available online at https://doi.org/10.1080/02705060.2021.1946180.

expert judgement reference condition approaches envisaged by the WFD, improves the analytical performance of the NISECI by increasing its power of resolution and is applicable in every Italian zoogeographic and ecological context.

1. Introduction

1.1. The NISECI in European and Italian legislation

The Water Framework Directive 2000/60/EC (WFD) is the operational instrument of the European Water Policy. This Directive aims to prevent the deterioration of the status of water bodies in the European Union (EU) and possibly to improve their condition to achieve rivers, lakes, transitional waters, marine-coastal waters, groundwater that are in 'good status'. Regarding perennial surface waters, this Directive requires each Member State to use a pool of bioindicators, called Biological Quality Elements (BQEs), including fish species. In compliance with this Directive, the Italian legislation (particularly Legislative Decree 152/2006, as amended and supplemented by Ministerial Decree 260/2010) had initially identified the Index of the Ecological State of Fish Communities (ISECI) by Zerunian et al. (2009) as the index for the analysis of the fish fauna BQE in river water bodies. The ISECI has undergone a national validation process that has led to the development of the New Italian Index of the Ecological State of the NISECI Fish Communities (Macchio et al. 2017). The latter, as envisaged by the implementation process of Directive 2000/60/EC (Common Implementation Strategy, CIS) has, therefore, passed the European intercalibration exercise. The legislative replacement of the ISECI by the NISECI was ratified by Commission Decision (EU) 2018/229 of 12 February 2018 establishing, pursuant to Directive 2000/60/EC of the European Parliament and of the Council, the values of the Member States' monitoring system classifications resulting from the intercalibration exercise. In compliance with this Decision, the NISECI is the index to be used for the analysis of the fish species EQR in river water bodies, both wadable and non-wadable, which are not heavily modified or artificial. The NISECI sampling protocol for the collection of ichthyological data in the wadable water bodies (Macchio and Rossi 2014) has been formalized by the Italian Institute for Environmental Protection and Research (ISPRA) and included in the Manual 'Biological Methods for inland surface waters'. Known as ISPRA 2040, this sampling protocol is essentially a removal method (Zippin 1958) which, in successive capture steps, allows the estimation of the catchability and numerical abundance of individual populations.

1.2. The NISECI structure and functionality

The main criteria used by the NISECI to assess the ecological status of a waterway are the naturalness of the fish community (understood as the completeness of the composition in indigenous species expected in relation to the zoogeographic and ecological framework) and the biological condition of the populations present (quantified positively for the expected indigenous species and negatively for the non-indigenous ones), in terms of abundance and population structure such as to guarantee the ability to reproduce and to have normal ecological-evolutionary dynamics. These criteria are linked to the requirements of the Water Framework Directive, 2000/60/EC, reiterated in the corresponding rules of transposition at national level (Legislative Decree 152/06 and subsequent

Table 1. Threshold values of EQR$_{NISECI}$ CLASSES for Alpine and Mediterranean areas of Italian territory.

EQR$_{NISECI}$ CLASSES	Alpine area	Mediterranean area
I	$0.80 \leq$ EQR$_{NISECI}$	$0.80 \leq$ EQR$_{NISECI}$
II	$0.52 \leq$ EQR$_{NISECI}$ <0.80	$0.60 \leq$ EQR$_{NISECI}$ <0.80
III	$0.40 \leq$ EQR$_{NISECI}$ <0.52	$0.40 \leq$ EQR$_{NISECI}$ <0.60
IV	$0.20 \leq$ EQR$_{NISECI}$ <0.40	$0.20 \leq$ EQR$_{NISECI}$ <0.40
V	EQR$_{NISECI}$ <0.20	EQR$_{NISECI}$ <0.20

Abbreviations: EQR$_{NISECI}$ = ecological state value before refinement are shown in black, the values after refinement are shown in grey.

amendments and additions), which envisage that, for the definition of the ecological status of river water bodies, the fish fauna BQE must be considered, assessing specific composition, abundance and age structure.

The NISECI is made up of three main indicators:

X_1 Presence/absence of species;

X_2 Biological condition of the populations of expected indigenous species;

X_3 Presence of non-indigenous or hybrid species, structure of the relative populations and numeric ratio (in terms of number of species and number of individuals) compared to the expected indigenous species detected.

The multimetric formulation of the index is given by:

$$\text{NISECI} = 0.1 \times X_1^{0.5} + 0.1 \times X_2^{0.5} + 0.8 \times (X_1 \times X_2) - 0.1 \times (1 - X_3)$$
$$\times \left(0.1 \times X_1^{0.5} + 0.1 \times X_2^{0.5} + 0.8 \times (X_1 \times X_2)\right)$$

Value of NISECI, as well as indicators and submetrics values, range from 0 to 1 (included).

Ecological state values are then expressed as EQRs after the equation

$$\text{EQR}_{\text{NISECI}} = (\log \text{NISECI} + 1.1283)/1.0603$$

Since WFD ecological indexes have to describe ecological status in 5 ordinal classes, EQR$_{NISECI}$ values are then grouped in five EQR$_{NISECI}$ CLASSES on the basis of threshold values of EQR$_{NISECI}$. Thresholds values were defined after the intercalibration exercise required by WFD in order to standardize the performances of the pool of river fish bioindicators available in the European Community.

After subdivision of Italian territory in Alpine area (northern of the Po-Tanaro rivers line) and Mediterranean area (the remaining Italian territory), threshold values are defined as in the following Table 1.

1.2.1. Indicator X_1

It compares the specific composition of the indigenous fish community observed with that expected. The species belonging to Salmonidae sensu Nelson (including *Thymallus thymallus*), Esocidae and Percidae are defined as species of major ecological and functional importance. The value of the indicator is given by the function:

$X_1 = (1.2 \times n_i + 0.8 \times n_a)/(1.2 \times m_i + 0.8 \times m_a)$

where:

n_i = number of expected and sampled indigenous species of major ecological and functional importance;

n_a = number of other expected and sampled indigenous species;

m_i = number of expected indigenous species of greater ecological-functional importance;

m_a = number of other expected indigenous species.

Table 2. Population structure and demographic consistency of indigenous species.

Population structure	Score	Demographic consistency	Score
Well structured	1	As expected	1
Moderately structured	0.5	Intermediate	0.5
Destructured	0	Scarce	0

Values and corresponding levels of judgement.

Table 3. Size groups of fish species and associated size classes.

Very small fish	Small fish	Medium fish	Large fish
CL1 < 3.5 cm	CL1 < 4.5 cm	CL1 < 8.0 cm	CL1 < 25.0 cm
3.5 cm ≤ CL2 < 4.5 cm	4.5 cm ≤ CL2 < 9.0 cm	8.0 cm ≤ CL2 < 17.0 cm	25.0 cm ≤ CL2 < 45.0 cm
4.5 cm ≤ CL3 < 6.0 cm	9.0 cm ≤ CL3 < 13.0 cm	17.0 cm ≤ CL3 < 21.0 cm	45.0 cm ≤ CL3 < 65.0 cm
6.0 cm ≤ CL4 < 8.0 cm	13.0 cm ≤ CL4 < 15.0 cm	21.0 cm ≤ CL4 < 30.0 cm	65.0 cm ≤ CL4 < 80.0 cm
CL5 ≥ 8.0 cm	CL5 ≥ 15.0 cm	CL5 ≥ 30.0 cm	CL5 ≥ 80.0 cm

Abbreviations: CL = size class.

1.2.2. Indicator X_2

The biological condition of each of the indigenous species present is given by the integration of population structure and demographic consistency (numerical abundance) that can assume the scores 0, 0.5 and 1 corresponding to three levels (Table 2). The three levels are assigned with reference to threshold values that have been experimentally identified on Italian nationwide datasets.

The final value of the indicator is given by the function:

$$X_2 = \frac{\sum_{i=1}^{n}(0.6 \ \times \ X_{2,a,i} \ + 0.4 \ \times \ X_{2,b,i})}{n}$$

Where:
n = number of expected indigenous species sampled;
i = single indigenous species sampled;
$X_{2,a}$ = submetric related to the population structure in age groups;
$X_{2,b}$ = submetric related to numerical abundance.

1.2.2.1. Submetric $X_{2,a}$: population structure. The assessment of the age of the individuals sampled is not carried out directly (analysis of scales or otoliths), but using an indirect method, based on the relationship between age and length, considering the second as a proxy for the first. All species of fish have been assigned to one of four size groups based on the typical size of the species; five different size classes (CLz, with z = {1, 2, 3, 4, 5}) were identified within each size group, to which the individuals sampled are assigned on the basis of body length (Table 3).

Once the individuals sampled have been assigned to the size classes, the population structure in age groups is assessed by integrating two criteria (Table 4), each of which expresses a response based on three scores (1, 2, 3) expressed on the basis of pre-established objective rules:

- *Criterion A – Distribution by size classes:* the score is assigned in function of the number of length classes populated in the sample;
- *Criterion B – Adults/juveniles numerical ratio:* the score is assigned in function of the ratio between the number of adults AD (CL4 + CL5) and the number of juveniles JUV (CL2 + CL3). The criterion identifies four threshold values and assigns the optimal values in the central range of distribution.

Table 4. Method of assignment of scores and judgements for submetric $X_{2,a}$.

Criterion A			
Score		**Distribution by size classes**	
1		at least four out of five classes are populated	
2		three out of five classes are populated	
3		no more than two out of five classes are populated	
Criterion B			
Score		**Adults/juveniles numerical ratio**	
1		$0.67 \leq AD/JUV \leq 1.5$	
2		$0.5 \leq AD/JUV < 0.67$	
		$1.5 < AD/JUV \leq 2$	
3		$AD/JUV > 2$	
		$AD/JUV < 0.5$	
Score			
Criterion A	**Criterion B**	**Final judgement**	**Final score**
1	1	well structured	1
1	2	well structured	1
1	3	moderately structured	0.5
2	1	moderately structured	0.5
2	2	moderately structured	0.5
2	3	destructured	0
3	1	destructured	0
3	2	destructured	0
3	3	destructured	0

Abbreviations: AD = adults; JUV = juveniles.

Table 5. Method of assignment of scores and judgements for submetric $X_{2,b}$.

Criterion	Final judgement	Final score
density observed $\geq$ 2nd tertile	as expected	1
1st tertile $\leq$ density observed < 2nd tertile	intermediate	0.5
density observed < 1st tertile	Scarce	0

1.2.2.2. Submetric $X_{2,b}$: demographic consistency. The assessment of demographic consistency is carried out by assigning the values observed to one of three categories of abundance. These categories have been defined in Macchio et al. (2017) on the basis of threshold values for the distribution of frequency of density values obtained from regional fish maps. The threshold values used to separate the three categories of abundance were (Table 5) the 1st tertile of the distribution of frequency (cumulative percentage of the sample = 33%) and the 2nd tertile (cumulative percentage of the sample = 66%).

1.2.3. Indicator X_3

Non-indigenous species are divided into three groups according to their harmfulness, defined on the basis of the decreasing level of impact on the indigenous fish species listed in three lists (List 1: Species with high harmfulness; List 2: Species with medium harmfulness; List 3: Species with moderate harmfulness). The lists of species belonging to the three different groups are defined on the basis of the assessments carried out by Zerunian et al. (2009). The value of X_3 is determined as follows:

State 1) absence of non-indigenous species $X_3 = 1$

State 2) Presence of species belonging to list 1, with at least one well-structured population $X_3 = 0$

State 3) Total number of non-indigenous fish $\geq$ total number of indigenous fish (belonging to expected species) $X_3 = 0$

State 4) In all other cases, the following formula is calculated:

$X_3 = 0,5 \, (a_{min} + b)$

where:

a_{min} = the lowest value of 'a' found in the sample observed:

- Presence of species belonging to List 1 with a population that is not well structured: a = 0.5
- Total number of non-indigenous species belonging to List 2 $\geq$ total number of indigenous species: a = 0.5
- Total number of non-indigenous species belonging to List 2 < total number of indigenous species: a = 0.75
- Total number of non-indigenous species belonging to List 3 $\geq$ total number of indigenous species: a = 0.75
- Total number of non-indigenous species belonging to List 3 < total number of indigenous species: a = 0.85

b = i ii iii
where:

- i = Proportion of non-indigenous species with a well-structured population compared to the total number of non-indigenous species present, multiplied by 0;
- ii = Proportion of non-indigenous species with a moderately-structured population compared to the total number of non-indigenous species present, multiplied by 0.5;
- iii = Proportion of non-indigenous species with a destructured population compared to the total number of non-indigenous species present, multiplied by 1.

1.3. Refining the NISECI calculation parameters

The original publication of the NISECI defines calculation parameters and application modalities identified on a historical-bibliographic basis and experimentally validated at national and district level. As reported in the Ministerial Decree 260/2010, in the original publications (Zerunian et al. 2009; Macchio et al. 2017) and in the dedicated literature (Rossi et al. 2015, 2016, 2017; De Bonis et al. 2017), the NISECI is an index designed to be applicable in all Italian hydrographic basins through a process of calibration of the calculation parameters depending on the different local, ecological and zoogeographical characteristics. This procedure, made possible by the explicit logical and mathematical structure of the index and by the partially 'open source' nature of its calculation parameters, allows the refinement of the functionality of the NISECI without compromising its legislative validity. Thanks to the fact that the formulation of the NISECI is constructed using a deductive approach (Top-Down), while the values of the coefficients of the calculation parameters are processed according to an inductive approach (Bottom-Up), it is possible to model variables of different scales (at Community level and at Population level, respectively), dramatically reducing the amount of experimental data needed to obtain a model using the inductive approach alone. For each detailed zoogeographical-ecological zone, therefore, the theoretical fish communities, described in the original publication, must be refined through ecological observations on the actual or potential habitats and historical-bibliographic analysis of the knowledge of the fish species. As stated in Ministerial Decree 260/2010: 'The Regions that, following the aforementioned surveys, have refined the expected fish communities, shall forward the results of the surveys carried out and the corresponding information, together with the supporting scientific documentation, to the Ministry of the Environment and Protection of the Territory and Sea'. Similarly, the reference conditions for the biological condition of individual

species can be refined in relation to the geographical and ecological context of the water body under consideration (*spatially based reference condition*). A key requirement of Directive 2000/60/EC is that the reference conditions must correspond to the absence of anthropogenic changes or minor anthropogenic changes. These conditions can be defined through existing situations, when available; through a modelling approach, when allowed by the quantity and quality of the data; through expert judgement (pending the consistency of the data) only when the previous options are not possible. Before application of the NISECI for monitoring purposes, the following must be defined for each individual fish species in each individual water body: (1) whether the species is autochthonous, that is, indigenous, to its zoogeographic context; (2) whether it should be expected in the local ecological context; (3) the threshold values for associating quality judgments with the observed values of the metrics and submetrics relating to the biological condition of indigenous and non-indigenous species. Starting from experimental case studies, the present study defines and generalizes the modalities for refining the parameters used to calculate the NISECI, in the various Italian geographical contexts, in relation to: (1) population structure of the indigenous species (Criterion A and Criterion B of submetric $X_{2,a}$); (2) demographic consistency of the indigenous species (submetric $X_{2,b}$); (3) population structure of the non-indigenous species (State 2 and State 4 of indicator X_3). Integrating the analysis of experimental data and the scientific synaecological and autoecological knowledge of fish species, said refinement methods allow the definition of the *spatially based* and *expert judgement reference conditions* for every single Italian ecological and zoogeographical context. The aim of this study is to provide those responsible for refining the NISECI with a unique and standardized methodology capable of simplifying the refining process and delivering comparable and homogeneous results throughout the country.

2. Materials and methods

2.1. Area of study

The methods of refinement (i.e., logical model of refinement) of the calculation parameters (i.e., threshold values of submetrics $X_{2,a}$ and $X_{2,b}$) have been elaborated on a rational basis, taking into account the need not to create logical distortions when the biological characteristics of the different species vary (e.g., annual or multi-year biological cycle, early or late sexual maturation, type of body size). The calculation parameters were then calculated and tested on an empirical basis using data from 284 ichthyological samplings from the hydrographic basins of Lombardy, Veneto, Emilia-Romagna, Tuscany, Marche, Basilicata and Calabria either embodied in the databases of the Department of Biological, Geological and Environmental Sciences of University of Bologna, of the Regional Agency for the Protection of the Environment (ARPA) of Basilicata, of the academic Spin-Off Hydrosynergy, of one of the in Authors (De Bonis et al. 2017) or available in literature (Confortini et al. 2008).

2.2. Sampling procedures

Since many the ISPRA 2040 protocol, not all the disposable data were obtained from ichthyological samples caught with the sampling procedure formalized by ISPRA. In order to contain random errors and stochastic noise due to the different sampling methods, only data collected using protocols that guarantee data of abundance, overall and by size, adequately homogeneous and reliable for the refinement of the biological condition of the

Table 6. (A) Information for the refinement of the NISECI parameters obtained from data collected using the ISPRA 2040 protocol, (B and C) minimum information for the refinement of individual NISECI metrics required for the use of bibliographic data collected using verse protocols from ISPRA 2040.

Parameter		(A) Information on ISPRA 2014	(B) Minimum information for refinement of community of reference and abundance thresholds	(C) Minimum information for refinement of population structure
Presence/absence	rare	YES	NO	NO
of species	common	YES	YES	YES
Demographic consistency		YES	YES	YES
Population structure	length	YES	YES	YES
in classes of	age	YES	NO	YES
ratio of adults - juveniles		YES	NO	YES

populations, have been considered. In addition to the data collected with the ISPRA 2040 protocol, other data found in the bibliography and collected using different protocols were also used, as follows:

- sampled stretches of at least 50 m in length for average wetland widths of less than 5 m and of at least 100 m for average widths of more than 5 m;
- censuses by electrofishing with removal methods (Moran 1951; Zippin 1956, 1958);
- registration of the single body size with a precision of 1 mm.

These conditions were considered sufficient to ensure performances similar to the ISPRA 2040 protocol (see quantitative section) exclusively for the analysis of demographic consistency and population structure in classes of length. Moreover, given the incomplete homogeneity of the representation of data from different sources, not all censuses were used to refine each of the parameters considered. With regard to the refinement of the reference community and the abundance thresholds, the works carried out with protocols that included all the information B of Table 6 were considered sufficient while, for aspects linked to the population structure in age groups, a stricter criterion was applied (only data from protocols contemplating the information C of Table 6). The information on the population structure in age groups by scalimetric survey or analysis of the growth rate (e.g., Von Bertalanffy Growth Model), allowed the estimation of the relationship between body length and age groups and the analysis of the relationship between the abundance of juvenile and adult age groups. The workflow for the dataset preparation is supplied as flowchart diagrams in Supplementary material (Supplementary Figures 3–4).

2.3. Experimental design

The two main variables that can influence the threshold values of the Total Lengths between different age groups are the geographical origin of the populations and the local ecological context. The populations of each species have therefore been subdivided according to the geographical area of origin and to the environmental and geomorphological characteristics of the wet river channel of the sampling stations (channel-reach scale morphology) as defined by Montgomery and Buffington (1997) and described in the System for Stream Hydromorphological Assessment, Analysis and Monitoring (IDRAIM) methodology (Rinaldi et al. 2016). To eliminate self-referentiality in the test phase, the data available were divided into a training dataset (74% of the data available), with which the values of the calculation parameters were obtained, and a test dataset (26% of the data

available), with which the validity of the model was tested. The latter dataset is fully compliant with the ISPRA 2040 sampling protocol.

2.4. Calculation

2.4.1. General criteria for the estimation of demographic consistencies

The estimation of the demographic consistencies was carried out by a classical Maximum Likelihood models on the basis of data collected by the removal method (Zippin 1958) at $s = \{2, 3, 4\}$ successive trapping steps. These models were chosen, despite the existence of more modern statistical alternatives, because they can be applied to large datasets using simple spreadsheets and without the need for time-consuming calculations.

The general formula for estimating numerical abundance (N) according to the Maximum Likelihood estimates of Moran (1951) and Zippin (1956, 1958) is:

$$N = C_s/(1 - q^s)$$

(1) where:

$s =$ number of trapping steps;

$C_s = \sum_{i=1}^{s} C_i =$ total number of individuals caught in s trapping steps;

$q = 1\text{-}p$

$p =$ likelihood of being caught in each of the s trapping steps.

For two and three-sample methods, mathematical functions with explicit solutions are available in literature.:

Two-sample method (Seber and Le Cren 1967)

$$N = C_1^2/(C_1 - C_2)$$
$$p = 1 - C_2/C_1$$

Three-sample method (Junge and Libosvarsky 1965)

$$N = \frac{6x^2 - 3xy - y^2 + y \times \left(y^2 + 6xy - 3x^2\right)^{1/2}}{18(x - y)}$$

$$p = \frac{3x - y - \left(y^2 + 6xy - 3x^2\right)^{1/2}}{2x}$$

Where:

$$x = 2C_1 + C_2$$
$$y = C_1 + C_2 + C_3$$

For the four-sample method, as no explicit mathematical functions are available, the general model of the Maximum Likelihood estimates of Moran (1951) and Zippin (1956, 1958) was used, providing implicit solutions (iterative algorithms or graphic procedures) in the joint determination of N and p starting from the equations:

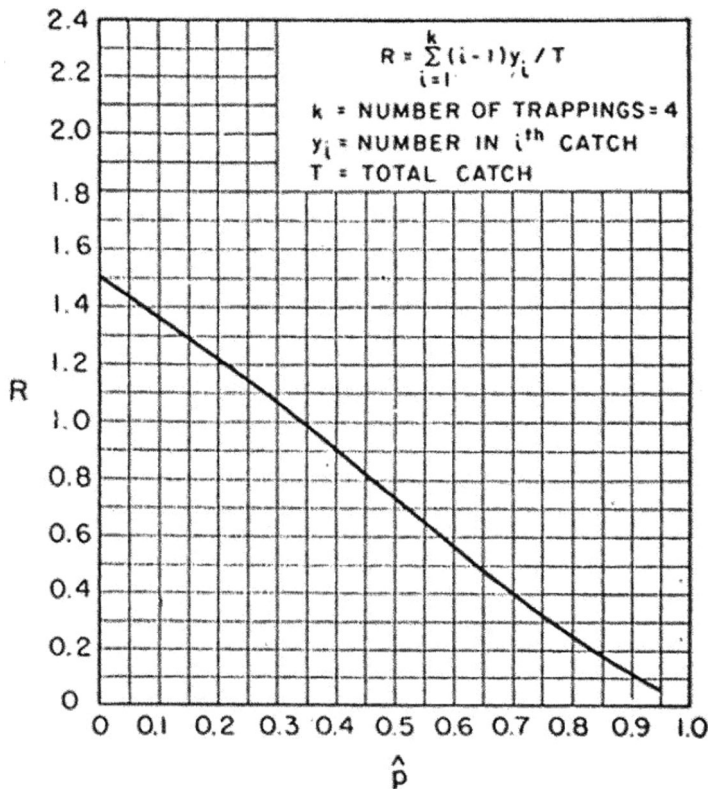$$N = C_S/\left(1 - q^S\right)$$

and

Figure 1. Graph for determination of the catch parameter (p) in the Maximum Likelihood estimates of Moran (1951) and Zippin (1956, 1958) starting from the calculated values of the variable R for the four-sample method ($s = 4$); from Zippin (1956).

$$q/p - s \times q^s/(1 - q^s) = \left(\sum_{i=1}^{s} (i - 1) \times C_i \right)/C_s$$

Zippin (1956), with

$$R = q/p - s \times q^s/(1 - q^s)$$

and in the specific case of $s = 4$

$$R = \frac{\sum_{i=1}^{s=4}(i - 1) \times C_i}{C_s} = \frac{C_2 + 2C_3 + 3C_4}{C_1 + C_2 + C_3 + C_4}$$

provides a graphic procedure for the estimation of p (and, therefore, of N) through the correlation with the variable R; this procedure involves the calculation of R, the determination on the graph shown in Figure 1 of the corresponding value of p and, therefore, the estimation of N.

To improve the calculation of p and obtain numerical solutions on large datasets using simple spreadsheets, the graph for $s = 4$ has been reconstructed by calculating the values of R corresponding to 21 different catch values perfectly constant in the four trapping steps $p_{j=1}^{21} = \{0.1, p_{i=1}^{19}/20, 0.99\}$ and the relationship between the variables p and R obtained by polynomial interpolation. To calculate R, the number of specimens captured in each ith step for each jth perfectly constant catch value was obtained as:
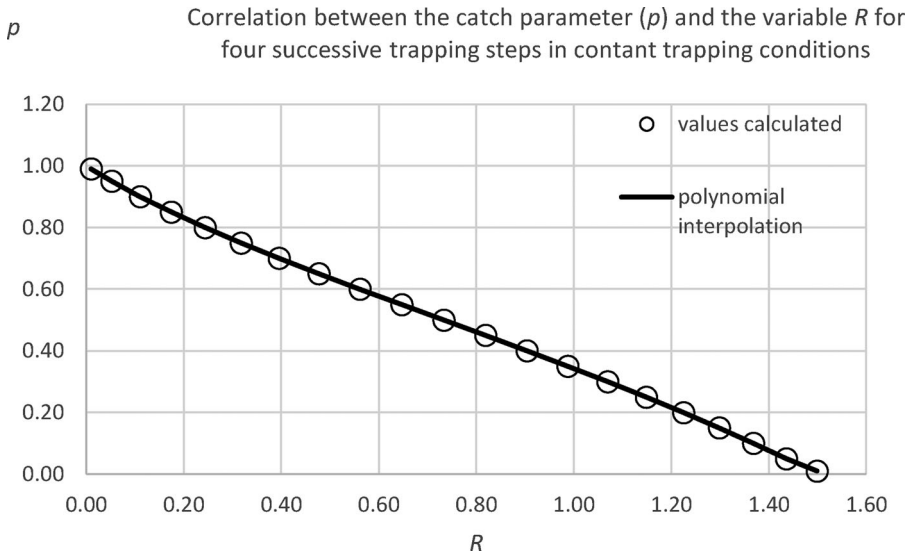
Figure 2. Correlation between the catch parameter (p) and the variable R in the Maximum Likelihood estimates of Moran ([1951](#)) and Zippin ([1956](#), [1958](#)) for 4-step catch sampling.

$$C_{i=1}^{s=4} = p_{j=1}^{21} \times (N - C_{i-1})$$

With the variable R depending on the numerical ratios between catches in the various steps, but independent of the individual catch values and of N, $N = 100$ was chosen arbitrarily. The dataset was therefore interpolated with polynomial regression of the tenth degree (having a correlation coefficient of $R^2 = 0.999$) elaborated using the 'epr' Package of software 'R' ([Figure 2](#)):

$$p = 2.384647585 \times R^{10} - 16.64495935 \times R^9 + 49.37677198 \times R^8 - 81.21779764 \times R^7$$
$$+ 81.15187585 \times R^6 - 50.97199729 \times R^5 + 20.47908891 \times R^4 - 5.70232908 \times R^3$$
$$+ 1.511123163 \times R^2 - 1.023973119 \times R + 1.000296192$$

The same approach can easily be extended for $s > 4$.

The formula for calculating the variance of the variables N and p is given in Seber ([1982](#)).

For $s > 2$:

$$V[N] = \frac{N \times (1 - q^s) \times q^s}{(1 - q^s)^2 - (p \times s)^2 \times q^{s-1}}$$

$$V[p] = \frac{N \times (1 - q^s) \times q^s}{N \times [q \times (1 - q^s)^2 - (p \times s)^2 \times q^s]}$$

For $s = 2$:

$$V[N] = \frac{C_1^2 \times C_2^2 \times (C_1 + C_2)}{(C_1 - C_2)^4}$$

$$V[p] = \frac{C_2 \times (C_1 + C_2)}{C_1^3}$$

Assuming a normal error distribution, the 95% confidence intervals can be calculated as:

$$N \pm 1.96 \times V[N]^{1/2}$$
$$p \pm 1.96 \times V[p]^{1/2}$$

It must, however, be considered that, when p and N are too small, the distribution of N is asymmetrical and $V[N]$ too narrow, and so the confidence intervals calculated as above are closer to 90% than to 95% (Zippin 1956; Robson and Regier 1968).

All the models used require trapping to be constant or for all the trapping steps to be comparable. Slight deviations from this requirement, due to stochastic factors, are tolerable while a net change in electrofishing efficiency between the various trapping steps leads to unrepresentative estimates. This situation can be found, for example, with small benthonic species; at the first step with electrofishing, passive swimming induced by the electric current can have such a direction that many of the specimens remain in the den; these specimens can then come out of the shelter passively and slowly as a result of the water current or by fleeing voluntarily and quickly when their motor skills are fully restored. At the second step, these animals are, therefore, 'bewildered and in the open field' and it is common to observe a greater number of them than at the first step. To assess the applicability and reliability of these models, it is therefore necessary to consider some mathematical, logical and statistical aspects. These models are not mathematically applicable when (Seber and Whale 1970):

$R > s/2$

meaning:

$C_2 > C_1$ (for $s = 2$);
$C_3 > C_1$ (for $s = 3$);
$C_3 + 3C_4 > 3\,C_1 + C_2$ (for $s = 4$).

For $R = s/2$, the models are logically inapplicable, because they render a zero-catchability value. For $s > 2$, it is possible to test the validity of the assumption of constant catchability through a goodness of fit test. White et al. (1982) describe the procedure by means of a simple chi-square test considering the expected values, resulting from the estimation of abundance and catchability. As pointed out in Lockwood and Schneider (2000), the chi-square test must be performed with $s-2$ degrees of freedom because N and p are estimated quantities; and for this reason, the test is not applicable when $s = 2$. If the chi-square test is significant, the difference between expected and observed data is of such magnitude that the model cannot be applied. In cases where a model on s trapping steps is not applicable, an attempt may be made to achieve an estimate by using only the partial sample obtained with fewer trapping steps. In fact, providing at least two trapping steps, it is mathematically possible to get the estimate excluding data from the last trapping steps. It is also possible, to obtain an estimate excluding data from the first trapping steps. At the end, the number of individuals actually caught in the steps excluded from the application of the model must be added to the expected number of individuals obtained by applying the model with the partial sample. Conversely it is not correct to exclude data obtained from intermediate trapping steps. The applicability of the model must obviously be tested also for each attempt with partial samples. For the purpose of refining the threshold values using the procedures indicated, all cases in which the conditions of

applicability are not met or the goodness of fit test is significant must therefore be excluded. It is also necessary to exclude cases that lead to excessively wide confidence intervals; as a practical rule, cases in which the coefficient of variation of the estimate (cv) is greater than 30% should be excluded (Zippin 1956):

$$cv = \frac{V[N]^{1/2}}{N} > 0,3$$

Lastly, given the reduced efficiency of the calculation of the variances when p and N are small, as reported in Seber (1982), it is necessary to exclude cases where:

$$N \times p^3 < 16 \times q^2 \times (1+q)$$

2.4.2. Submetric $X_{2,a}$ – definition of Size Classes (CL) and refinement of threshold values between Size Classes (CL)

Submetric $X_{2,a}$ envisages the subdivision of each sample in five different size classes (CL) according to the body length of the specimens caught. These size classes must be functional to the representation of the population structure in age groups for which they are proxies. This subdivision (data binning) was deemed necessary to make the submetric as resistant as possible to seasonal variations of individual dimensions.

The breakdown of the population structure into a finite number of size classes entails a likelihood (variable depending on the species and the geographical and ecological context) that more than one age group can fall, at least partially, into a single length class and vice versa.

The association between size classes and age groups must also be made for taking into account the definition of Criterion B of Submetric $X_{2,a}$ (ration of adults – juveniles) . According to this criterion, the larger size classes (CL4 and CL5) are to be associated with adult forms (AD), while it is assumed that the immediate smaller size classes (CL2 and CL3) are mainly populated by juvenile forms (JUV). According to this scheme, all or most of the individuals included in CL4 should have reached their first sexual maturity. In the calculation of the juvenile/adult ratio using the total length classes, the smaller size class (CL1) corresponding to the fry (AV) is excluded; this exclusion is motivated by the extreme inter and intra-annual numerical variability (which is also usually affected by high inaccuracy in the sample estimate) of the cohort 0^+ (individuals under one year of age), a component that is assumed to dominate CL1 (Thompson and Rahel 1996; Dolan and Miranda 2003; Miranda and Dolan 2003; Korman et al. 2009; Millar et al. 2016; Hedger et al. 2018). The age groups (CE) above 0^+ have been divided into juveniles and adults on the basis of the years required to reach sexual maturity, information gathered from the main bibliographic texts (Holcik 1986; Gandolfi et al. 1991; Hoestlandt 1991; Bănărescu 1999; Bănărescu and Paepke 2001; Bănărescu and Bogutskaya 2003; Miller 2003; Zerunian 2004) complemented by evidence of reproductive maturity recorded in the field (nuptial tubercles, gamete emission, etc.). If first sexual maturity (CE_{sex}) is reached at different ages in the two sexes, the age at which both sexes are able to reproduce with each other for the first time has been taken as a precautionary reference. For species in which both sexes are mature from the fourth year (CE 3^+) onwards, the CL were assigned according to the following criteria (scheme for merging age groups-length classes):

$CE0^+ = CL1$

$CE1^+ \leq CL2 \leq \lfloor CE_{sex-1}/2 \rfloor$

Length/age frequency distribution of *Telestes muticellus* Bonaparte, 1837

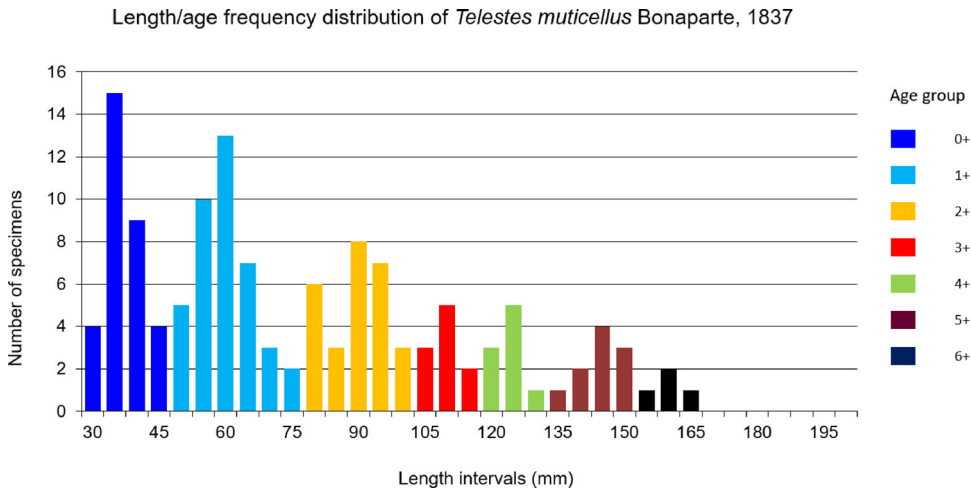**Figure 3.** Example of population structure by size intervals of 0.5 cm with determination of age groups by scalimetric analysis.

$\lfloor CE_{sex-1}/2 \rfloor < CL3 \le CE_{sex-1}$

$CE_{sex} \le CL4 \le \lfloor (CE_{sex} + CE_{max})/2 \rfloor$

$\lfloor (CE_{sex} + CE_{max})/2 \rfloor < CL5$

Where:

$CE0^+$ = age group of fish born in the year (first year of age);

$CE1^+$ = age group of fish born the previous year (second year of age);

CE_{sex} = age group in which both sexes are sexually mature for the first time;

CE_{sex-1} = age group before reaching sexual maturity in both sexes;

CE_{max} = maximum age group still fertile (reproductive life);

$\lfloor CE_{sex-1}/2 \rfloor$ = intermediate age group between $CE0^+$ and CE_{sex-1};

$\lfloor (CE_{sex} + CE_{max})/2 \rfloor$ = intermediate age group between CE_{sex} and CE_{max};

$\lfloor \ \rfloor$ = Rounding down.

The workflow for the definition of Size Classes (CL) is supplied as flowchart diagram in Supplementary material (Supplementary Figure 5).

The refinement of the length threshold values between the size classes (CL) was carried out on the basis of the distributions of frequency of the total body lengths (LT) for narrower class intervals than those used in the original draft of the NISECI, i.e., in intervals of size (InT) of 0.5 cm for very small and small species and 1 cm for medium and large species; it follows that each CL of NISECI is divided into several InT. The smallest InT ($InT_{z,min}$) and biggest ($InT_{z,max}$) values were identified for each CLz of each population; then the median values of the distributions of $InT_{z,min}$ (Me_$InT_{z,min}$) and $InT_{z,max}$ (Me_$InT_{z,max}$) for all the available populations were identified. The species-specific threshold value between two adjacent size classes CLz and $CLz + 1$ was then calculated as: (Me_$InT_{z,max}$ + Me_$InT_{z+1,min}$)/2. Only populations for which the length/age ratio was available were considered for this purpose (by scalimetric survey and/or analysis of the growth rate), with a sample of specimens corresponding to at least one intermediate

population size in indicator $X_{2,b}$ and characterized by a continuous distribution of the sizes (Figure 3).

In species in which sexual maturity is reached at or before the 3rd year of age ($CE_{sex} \leq 2^+$), age differences of less than one year are considered significant for the purpose of identifying threshold values for separating length classes.

The following criteria are adopted:

A. If first sexual maturity is reached at the 3rd year of age ($CE_{sex} = 2^+$), juvenile individuals (JUV) consist exclusively of $CE1^+$, which in this case would combine CL2 and CL3. Individuals of $CE1^+$ were divided between the two classes using the median value of LT distribution of $CE1^+$ as the discriminating value.
B. If first sexual maturity is reached at the 2nd year of age ($CE_{sex} = 1^+$) all the individuals belonging to age groups above $CE0^+$ were divided between CL2, CL3, CL4 and CL5, in this case, the first three quartiles of the Lt distribution of individuals over the age of $CE0^+$ were used as the discriminant threshold LT.
C. If sexual maturity is reached at the 1st year of age ($CE_{sex} = 0^+$) all the individuals in the population were divided between CL1, CL2, CL3, CL4 and CL5 using the first four quintiles of the Lt distribution as the threshold LT. The synoptic framework of the breakdown of the age groups (CE) into size classes (CL) is shown in Table 7.

The workflow for the refinement of threshold values between Size Classes is supplied as flowchart diagrams in Supplementary material (Supplementary Figures 6–7).

2.4.3. Submetric $X_{2,a}$ – Criterion B refinement of the threshold values of the AD/ JUV ratio

For each species, the threshold values of the ratio between the number of adult and juvenile individuals were defined on the basis of the sextiles of the frequency distribution of this parameter in the training dataset, assigning the optimal values to the central sextiles. This choice was made so as to comply with Criterion B of Submetric X_2, which involves four threshold values, and in order to evenly distribute the cases available in six symmetrical scoring ranks with respect to the optimal value. The extreme cases corresponding to the complete absence of one of the two categories were considered as outliers and therefore excluded from the analysis (AD = 0 or JUV = 0). The scores were then assigned as in Table 8.

Cumulative catches have been used for the definition of sextiles, regardless of the number of subsequent trapping steps.

In conditions of constant catchability between successive steps and assuming negligible differences in the likelihood of capture p between the age groups above 0^+, that is:

$p_{CL2} \approx p_{CL3} \approx p_{CL4} \approx p_{CL5}$

using (1) the general formula for estimating numerical abundance (N)

and indicating as

C_{sCLz}, the total catches in the s electrofishing steps carried out for the zth size class, we obtain the following formulation for the AD/JUV ratio:

$$\frac{\mathbf{AD}}{\mathbf{JUV}} = \frac{C_{sCL4} + C_{sCL5}}{C_{sCL2} + C_{sCL3}}$$

The workflow for the refinement of the threshold values of the AD/JUV ratio is supplied as flowchart diagram in Supplementary material (Supplementary Figure 8)

Table 7. Synoptic framework of the scheme of merging age group-length classes.

CE_{sex}	CL1	CL2	CL3	CL4	CL5
e	0^+	$\{1^+, \ldots, \lfloor(e-1)/2\rfloor\}$	$\{\lfloor(e-1)/2\rfloor+1, \ldots, e-1\}$	$\{e, \ldots, \lfloor(e+CE_{max})/2\rfloor\}$	$> \lfloor(e+CE_{max})/2\rfloor$
$\ldots$					
8^+	0^+	$\{1^+, 2^+, 3^+\}$	$\{4^+, 5^+, 6^+, 7^+\}$	$\{8^+, \ldots, \lfloor(8^+ + CE_{max})/2\rfloor\}$	$> \lfloor(8^+ + CE_{max})/2\rfloor$
7^+	0^+	$\{1^+, 2^+, 3^+\}$	$\{4^+, 5^+, 6^+\}$	$\{7^+, \ldots, \lfloor(7^+ + CE_{max})/2\rfloor\}$	$> \lfloor(7^+ + CE_{max})/2\rfloor$
6^+	0^+	$\{1^+, 2^+\}$	$\{3^+, 4^+, 5^+\}$	$\{6^+, \ldots, \lfloor(6^+ + CE_{max})/2\rfloor\}$	$> \lfloor(6^+ + CE_{max})/2\rfloor$
5^+	0^+	$\{1^+, 2^+\}$	$\{3^+, 4^+\}$	$\{5^+, \ldots, \lfloor(5^+ + CE_{max})/2\rfloor\}$	$> \lfloor(5^+ + CE_{max})/2\rfloor$
4^+	0^+	1^+	$\{2^+, 3^+\}$	$\{4^+, \ldots, \lfloor(4^+ + CE_{max})/2\rfloor\}$	$> \lfloor(4^+ + CE_{max})/2\rfloor$
3^+	0^+	1^+	2^+	$\{3^+, \ldots, \lfloor(3^+ + CE_{max})/2\rfloor\}$	$> \lfloor(3^+ + CE_{max})/2\rfloor$
2^+	0^+	1^+	1^+	$\{2^+, \ldots, \lfloor(3^+ + CE_{max})/2\rfloor\}$	$> \lfloor(2^+ + CE_{max})/2\rfloor$
1^+	0^+	$LT_{1+} < 50^{th}\ \%_{ile}$ $> 0^+$ $LT_{>0+} < 25^{th}\ \%_{ile}$	$LT_{1+} \geq 50^{th}\ \%_{ile}$ $> 0^+$ $25^{th}\ \%_{ile} \leq LT_{>0+} < 50^{th}\ \%_{ile}$	$> 0^+$ $50^{th}\ \%_{ile} \leq LT_{>0+} < 75^{th}\ \%_{ile}$	$> 0^+$ $LT_{>0+} \geq 75^{th}\ \%_{ile}$
0^+	$LT < 20^{th}\ \%_{ile}$	$20^{th}\ \%_{ile} \leq LT < 40^{th}\ \%_{ile}$	$40^{th}\ \%_{ile} \leq LT < 60^{th}\ \%_{ile}$	$60^{th}\ \%_{ile} \leq LT < 80^{th}\ \%_{ile}$	$LT \geq 80^{th}\ \%_{ile}$

Abbreviations: CL = size class; CE_{sex} = age group in which both sexes are sexually mature for the first time; CE_{max} = maximum age group typical of the species; LT = total individual length; LT_{1+} = total length of individuals in age group 1+; $LT_{>0+}$ = total length of individuals in age group greater than 0+; $\%_{ile}$ = percentile; e = age group greater than 8+.

Table 8. How to refine threshold values and assign scores for criterion B of submetric $X_{2,a}$.

Score	Criterion
3	AD/JUV < 1st sextile
2	1st sextile ≤ AD/JUV < 2nd sextile
1	2nd sextile ≤ AD/JUV ≤ 4th sextile
2	4th sextile < AD/JUV ≤ 5th sextile
3	5th sextile < AD/JUV

Abbreviations: AD = adults; JUV = juveniles.

Table 9. Refinement of threshold values and assignment of scores and judgments for submetric $X_{2,b}$.

Criterion	Final judgement	Final score
Estimated density ≥ 2nd tertile of the estimates	as expected	1
1st tertile of the estimates ≤ estimated density < 2nd tertile of the estimates	intermediate	0.5
Estimated density < 1st tertile of the estimates	scarce	0

Abbreviations: Stat = statistic value; Sig = significance; df = degrees of freedom.

2.4.4. Submetric $X_{2,b}$ – refinement of the demographic consistency threshold values

The density threshold values were calculated according to the procedure identified by Macchio et al. (2017). In order to standardize the data obtained from sampling carried out under different conditions of catchability and possibly with a different number of trapping steps, estimated density values (Table 9), from demographic consistencies calculated as reported in subsection 2.1.1., were used instead of the densities observed.

The workflow for the refinement of the demographic consistency threshold values is supplied as flowchart diagram in Supplementary material (Supplementary Figure 9).

3. Results and discussion

Quantitative data available, mostly collected in wadable watercourses, have made it possible to refine the threshold values of populations distributed among the ecological zones of salmonids and rheophilic cyprinids. The test of the efficiency of the proposed refinement methodology was then carried out on 74 population data not used for the processing of the model. For this purpose, only data collected using the ISPRA 2040 protocol and referring to the ecological zones of salmonids and rheophilic cyprinids were considered. To assess the efficiency of the refinement method proposed here, the frequency distributions before and after refinement of the variables were compared: values of submetric $X_{2,a}$, values of the NISECI, values of the EQR_{NISECI}, classes of the ecological state of the EQR_{NISECI} (Figure 4).

The efficiency of the methodology for refining submetric $X_{2,b}$ with respect to standard threshold values has not been tested here; the original publication of the NISECI does not in fact provide standard threshold values to be refined but assumes that the refining for this submetric must be compulsory.

The following have been considered as positive evaluation criteria:

1. the obtaining of a more symmetrical and less flattened distribution for indicator X_2, with a median value closer to 0.5 and with greater dispersion of values; this condition was considered to be consistent with the spatially based reference condition approach in which the best possible conditions found in the geographical and ecological context of the water body in question are adopted as reference condition;
2. a statistically significant difference between the before and after refinement condition in the distribution of values of all the metrics considered;

Frequency distribution of test variables before and after refinement
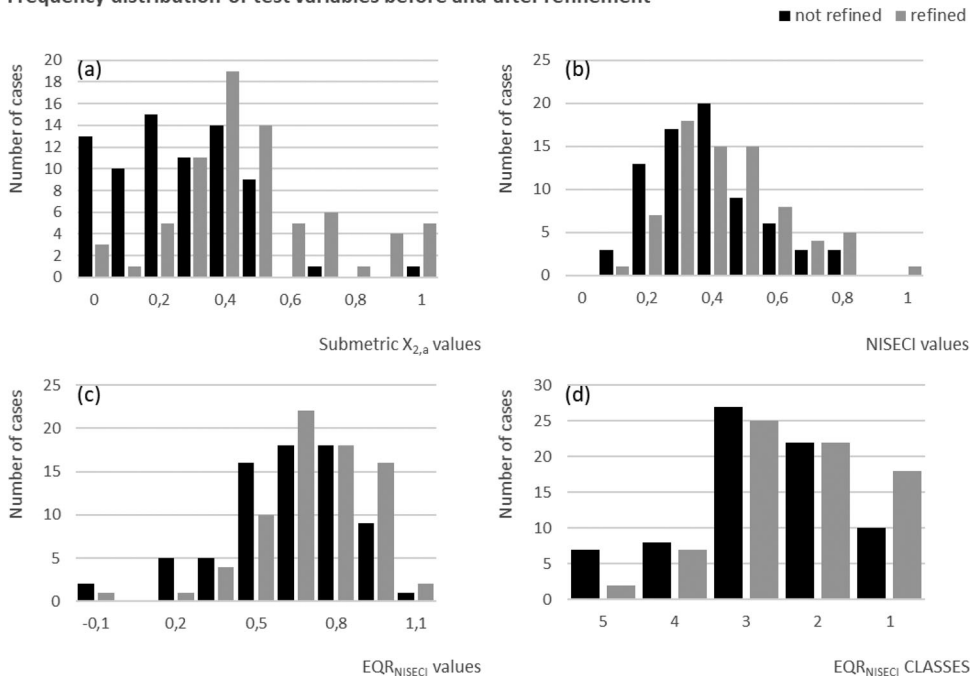
■ not refined ■ refined

Figure 4. Frequency distribution of the test variables; (a) values of submetric $X_{2,a}$; (b) values of the NISECI; (c) values of the EQR_{NISECI}; (d) EQR_{NISECI} CLASSES. The values before refinement are shown in black, the values after refinement are shown in grey.

Table 10. Descriptive statistics of submetric $X_{2,a}$ in before and after refinement conditions.

	Submetric $X_{2,a}$	
Descriptive statistics	Not refined	Refined
Mean	0.24	0.46
Median	0.20	0.40
Variance	0.04	0.06
Std. Deviation	0.19	0.25
Minimum	0.00	0.00
Maximum	1.00	1.00
Total Range	1.00	1.00
Interquartile Range	0.28	0.29
Adjusted Skewness	1.05	0.61
Excess Kurtosis	2.17	0.03

3. a numerically significant difference between the before and after refinement condition in the allocation of ecological quality ratio scores.

3.1. Evaluation criterion 1

Table 10 shows the descriptive statistics of the frequency distribution of submetric $X_{2,a}$ in conditions before and after refinement.

Submetric $X_{2,a}$ after refinement shows a distribution of values with a greater dispersion (Variance = 0.06; Interquartile Range = 0.29), more symmetrical (Adjusted Skewness = 0.61, median value 0.4) and more bell-shaped (Excess Kurtosis = 0.03). The form acquired by the values of submetric $X_{2,a}$ after refinement indicates a greater

Table 11. Normality test of variables for values X_{2a}, NISECI values, EQR_{NISECI} values, EQR_{NISECI} classes before and after refinement.

Tests of normality		Kolmogorov-Smirnov		Shapiro-Wilk		
		Stat	Sig.	Stat	Sig.	Df
values X_{2a}	Before	0.11	0.02	0.91	0.00	74
	After	0.15	0.00	0.94	0.00	74
NISECI values	Before	0.11	0.04	0.97	0.04	74
	After	0.11	0.02	0.96	0.04	74
EQR_{NISECI} values	Before	0.09	0.19	0.95	0.01	74
	After	0.08	0.20	0.94	0.00	74
EQR_{NISECI} classes	Before	0.20	0.00	0.90	0.00	74
	After	0.19	0.00	0.89	0.00	74

Table 12. Summarized results of the 2-paired-samples non-parametric tests applied to the variables of interest before and after refinement.

	Significance		
	Sign Test	Stuart-Maxwell test	Bhapkar test
$X_{2,a}$ values	p = 0.00	NA	
NISECI values	p = 0.00		
EQR_{NISECI} values	p = 0.00		
EQR_{NISECI} classes	p = 0.00	p = 0.00	p = 0.00

Abbreviations: NA = test not applicable.

discriminatory power between similar but not identical situations, that is, a better descriptive capacity of the different possible cases. After refinement, submetric $X_{2,a}$ is able to describe both optimal and critical situations, as opposed to before refinement, when almost exclusively poor and intermediate situations are described.

3.2. Evaluation criterion 2

The variables of interest for each individual sample have a before and after condition with non normal data distribution (Table 11).

To assess the statistical significance of the effect of refinement on the results of the variables of interest, 2-paired-samples non-parametric tests were used. In particular, given the non-symmetrical distribution of the before and after differences (data not shown) for all the variables, both continuous and ordinal, the paired-samples sign test was applied (Sprent 1993). For the 'classes of the ecological quality of the EQR_{NISECI}' variable, being ordinal, the *marginal homogeneity test* of Stuart-Maxwell (Stuart 1955; Maxwell 1970; Everitt 1977) and Bhapkar (1966), were also used as confirmation. The first two tests were carried out using the PASW Statistic 17 software, the last one using the Bhapkar test.xlsx spreadsheet (Stikker 2017).

The results of the tests are summarized in Table 12.

For all the variables, the refinement of the threshold values was found to have a highly significant effect. In other words, as a result of the refinement process, the variables considered acquired new and more efficient capacities for assessing fish populations.

3.3. Evaluation criterion 3

The results of the assignment of quality scores (classes) of the EQR_{NISECI} before and after refinement are shown in comparative form as a table with two entries (Table 13).

Table 13. Assignment of EQR$_{NISECI}$ CLASSES; comparison between conditions before and after refinement.

EQR$_{NISECI}$ CLASSES			After refinement CLASS					
			I	II	III	IV	V	Total
Before	CLASS	I	10	–	–	–	–	10
Refinement		II	7	15	–	–	–	22
		III	1	7	17	2	–	27
		IV	–	–	6	2	–	8
		V	–	–	2	3	2	7
Total			18	22	25	7	2	74

The refinement of the threshold values has a numerically significant effect on the assignment of quality scores. As a result of the refinement, the changes in the score led to an improvement of a quality class in 35% of the stations and to the deterioration of a class in 3% of the stations.

4. Conclusions

The present study has allowed the identification of methods for the refinement of the threshold values of the demographic parameters of indigenous and non-indigenous fish species necessary for the calculation of the NISECI and the assignment of quality scores (classes) of the EQR$_{NISECI}$. The procedure used, consistent with the spatially based reference condition and expert judgement reference condition approaches of the WFD, is applicable in every zoogeographic and ecological Italian context and on the basis of the three assessment criteria used, improves the analytical performance of the NISECI by increasing its power of resolution. In other words, the NISECI acquires a greater ability to describe and classify fish populations by identifying differences in greater detail. The procedure used is designed to refine the descriptive capacity of the ecological quality of the fish species with the NISECI with further data, available in the bibliography (e.g., fish maps or atlas). The bibliographic examination of the data available at Italian nationwide level highlighted products for which, for example, demographic consistency values collected in a sufficiently robust way were not coupled with structural analysis in appropriate age groups and vice versa. The criteria for judging whether and which bibliographic data can be used to refine individual demographic parameters have therefore been defined here. This procedure maximizes the possibility of having a sufficient critical mass of data to correctly represent the demographic variability of individual territorial contexts. This study also defines the minimum level of detail that must be guaranteed to allow the refinement of the demographic threshold values for classification purposes. It should be noted that the model presented here is designed to be used and implemented in different Italian contexts in operational mode, without prejudice to the need for ichthyological and ecological skills. The summary of the procedures identified, broken down into subroutines, is shown in the form of a flowchart as Supplementary material (Supplementary Figures 1–9). The approach used here represents a first method for refining the NISECI; in the near future, further models will be tested in terms of cost benefits (e.g., amount of data needed to ensure sufficient reliability) and technical applicability. For the latter, it is necessary to envisage that complex models can be used as an alternative to those proposed here, as long as they are developed by staff with statistical, ichthyological and ecological skills that guarantee the validation of the results.

Acknowledgements

Disclosure statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

Authors contributions

GR: Conceptualization, Methodology, Formal Analysis, Investigation, Resources, Data Curation, Writing – Original Draft, Visualization. AM: Methodology, Investigation, Resources, Data Curation. SDB: Conceptualization, Methodology, Investigation, Resources, Data Curation, Writing – Review & Editing. GZ: Methodology, Investigation, Resources, Data Curation. GC: Investigation, Resources. SG: Supervision. CM: Conceptualization, Writing – Review & Editing, Project Administration. SM: Conceptualization, Validation, Writing – Review & Editing. GLR: Conceptualization, Validation, Writing – Review & Editing. SB: Conceptualization, Writing – Review & Editing, Project Administration.

Notes on contributors

Giovanni Rossi: Biologist with specialization in freshwater ichthyology at the Arpae - Regional Agency for Prevention, Environment and Energy of Emilia-Romagna. Trained his scientific background thanks to more than 20 years of academic research programs and freelance activities regarding biodiversity conservation, phylogeography, ecological status assessment, biological monitoring and environmental impact assessment.

Andrea Marchi: Biologist with 10 years of experience in environmental monitoring and applied ecology as a technical adviser. In his activities, he has developed expertise in the study of fish communities in relation to environmental alteration and in the application of biological quality index for inland surface waters.

Salvatore De Bonis: Biologist of aquatic systems for more than ten years, is involved in biological monitoring in inland waters. He is currently regional referent for the 'fish fauna' biological quality element for the environmental protection agency in the Lazio region of Italy, and a member of working groups for the assessment of the ecological status of inland waters.

Gianluca Zuffi: graduated in Biological Sciences, over the years, he has implemented his ecological skills with studies and research in the field of molecular biology. It cooperates with administrators, technicians, fishing associations, sector volunteers, to share information and optimize collaboration relationships for the management and conservation of freshwater fish fauna and habitats.

Gaetano Caricato: Zootechnician with specialization in Aquaculture at University of Studies of Basilicata. Currently, he is senior technician at the 'Ecosystems, Biodiversity and Land Use Unit' of the Science Technical Committee of ARPA (Regional Agency for Environmental Protection) of Basilicata.

Stefano Goffredo: Associate Professor of Ecology at the Department of Biological, Geological, and Environmental Sciences, Alma Mater Studiorum - Università di Bologna. Areas of scientific interest: coral ecology and biology – freshwater ecology – citizen Science. In his research, he has developed original aspects in the interpretation of growth and population data, population dynamics, reproductive biology and conservation monitoring.

Cristina Martone: Naturalist, with specialization in freshwater ecology, analysis of diatomic communities, ecotoxicology, ichthyoculture. Engaged for 15 years in laboratory and field activities, she now works as a researcher at the Laboratory Department of ISPRA.

Stefano Macchio: Biologist with specialization in applied ecology and environmental monitoring through vertebrate bioindicators. Engaged for more than 20 years in field activities and data analysis in ISPRA.

Gian Luigi Rossi: Senior Researcher at Research Center ENEA Saluggia, in the Laboratory of Biodiversity and Ecosystem Services of the Department for Sustainability. His work fields are river ecology, environmental quality assessment through the use of bioindication and synthetic biotic indices, planning and management of protected areas, management of wetland environments.

Stefania Balzamo: Chemist with more than 20 years of experience in the field of environmental chemistry and sampling; she has worked for the Environmental Ministry to implement the WFD in Italy; she attends to the WG of ECOSTAT and now she is the head of Laboratory Department of ISPRA.

Data availability statement

The Authors do not have the permission to share the main part of their data because was collected for third parties. The freely shareable data that support the findings of this study are available from the corresponding author, [GR], upon reasonable request.

References

Bănărescu PM. 1999. The freshwater fishes of Europe. Volume 5/I. Cyprinidae 2. Part I: Rhodeus to Capoeta. Wiebelsheim: Aula-Verlag.

Bănărescu PM, Bogutskaya NG. 2003. The freshwater fishes of Europe. Volume 5/II. Cyprinidae 2. Part II. Barbus. Aula-Verlag, Wiebelsheim.

Bănărescu PM, Paepke H-J. 2001. The freshwater fishes of Europe. Volume 5/III. Cyprinidae 2. Part III: Carassius to Cyprinus. Gasterosteidae. Wiebelsheim: Aula-Verlag.

Bhapkar VP. 1966. A note on the equivalence of two test criteria for hypotheses in categorical data. J Am Stat Assoc. 61(313):228–235. https://www.tandfonline.com/doi/abs/10.1080/01621459.1966.10502021.

Confortini I, Turin P, Salviati S, Zanetti M, Tuzzato B, Bilò MF, Maio G, Marconato E, Busatto T. 2008. Carta Ittica della provincia di Verona. Provincia di Verona, Verona. https://portale.provincia.vr.it/uffici/uffici/10017/611/documenti/pesca-e-gestione-delle-acque/carta-ittica-della-provincia-di-verona-2/at_download/file.

De Bonis S, Rossi G, Giorgio A, Bianco PG. 2017. Monitoraggio dell'ittiofauna del Fiume Crati. Affinamento della comunità ittica di riferimento per il calcolo dell'ISECI e proposta di una zonazione ittica per la gestione alieutica. Biologia Ambientale. 31(2017):105–109. http://www.cisba.eu/images/rivista/biologia_ambientale/Ba2017/P06_DeBonis_ISECI_Crati.pdf.

Dolan CR, Miranda LE. 2003. Immobilization thresholds of electrofishing relative to fish size. Trans Am Fish Soc. 132(5):969–976.

Everitt BS. 1977. The analysis of contingency tables. London: Chapman & Hall.

Gandolfi G, Zerunian S, Torricelli P, Marconato A. 1991. I Pesci delle acque interne italiane. Roma: Istituto Poligrafico e Zecca dello Stato. XVI + 617 pp.

Hedger RD, Diserud OH, Sandlund OT, Saksgård L, Ugedal O, Bremset G. 2018. Bias in estimates of electrofishing capture probability of juvenile Atlantic salmon. Fish Res. 208:286–295.

Hoestlandt H. 1991. The freshwater fishes of Europe. Volume 2. Cluepeidae, Anguillidae. Wiebelsheim: Aula-Verlag.

Holcik J. 1986. The freshwater fishes of Europe. Volume 1/I. Petromyzontiformes. Wiebelsheim: Aula-Verlag.

Junge CO, Libosvarsky J. 1965. Effects of size selectivity on population estimates based on successive removals with electrical fishing gear. Zool Listy. 14:171–178.

Korman J, Yard M, Walters C, Coggins LG. 2009. Effects of fish size, habitat, flow, and density on capture probabilities of age-0 rainbow trout estimated from electrofishing at discrete sites in a large river. Trans Am Fish Soc. 138(1):58–75.

Lockwood RN, Schneider JC. 2000. Stream fish population estimates by mark-and-recapture and depletion methods. In: Schneider JC, editor. Manual of fisheries survey methods II: with periodic updates. Michigan Department of Natural Resources, Fisheries Special Report, Chapter 7. Vol. 25. Ann Arbor (MI): Michigan Department of Natural Resources. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.214.5797.

Macchio S, Rossi GL. 2014. 2040 - Protocollo di campionamento e analisi della fauna ittica dei sistemi lotici guadabili. In: Mazzella D, editor. Metodi Biologici per le acque dolci superficiali. ISPRA, Manuali e Linee Guida 111/2014. Roma: ISPRA.

Macchio S, Rossi GL, Rossi G, De Bonis S, Balzamo S, Martone C. (2017). Nuovo Indice dello Stato Ecologico delle Comunità Ittiche (NISECI). ISPRA, Manuali e Linee Guida 159/2017. Roma: ISPRA.

Maxwell AE. 1970. Comparing the classification of subjects by two independent judges. Br J Psychiatry. 116(535):651–655.

Millar CP, Fryer RJ, Millidine KJ, Malcolm IA. 2016. Modelling capture probability of Atlantic salmon (*Salmo salar*) from a diverse national electrofishing dataset: implications for the estimation of abundance. Fish Res. 177:1–12.

Miller PJ. 2003. The freshwater fishes of Europe. Volume 8/I. Mugilidae, Atherinidae, Atherinopsidae, Blenniidae, Odontobutidae, Gobiidae 1. Wiebelsheim: Aula-Verlag.

Miranda LE, Dolan CR. 2003. Test of a power transfer model for standardized electrofishing. Trans Am Fish Soc. 132(6):1179–1185.

Montgomery DR, Buffington JM. 1997. Channel-reach morphology in mountain drainage basins. Geol Soc Am Bull. 109(5):596–611. https://doi.org/10.1130/0016-7606(1997)109 < 0596:CRMIMD > 2.3.CO;2.

Moran PAP. 1951. A mathematical theory of animal trapping. Biometrika. 38(3-4):307–311.

Rinaldi M, Surian N, Comiti F, Bussettini M. 2016. IDRAIM – Sistema di valutazione idromorfologica, analisi e monitoraggio dei corsi d'acqua – Versione aggiornata 2016 – ISPRA – Manuali e Linee Guida 131/2016. Roma: ISPRA.

Robson DS, Regier HA. 1968. Estimation of population number and mortality rates. In: Ricker WE, editor. Methods for assessment of fish production in fresh waters. IBP Handbook. Vol. 3. Oxford: Blackwell. p. 124–158.

Rossi G, Falconi R, Capostagno S, Marchi A, Zuffi G, Ferri D, Zaccanti F. 2015. Applicazione ed analisi critica dell'ISECI nelle acque montane, pedemontane ed altoplaniziali dell'Emilia-Romagna (Distretti del Po e dell'Appennino settentrionale). It J Fresh Ichthyol. 2015(2):76–80. http://www.aiiad.it/ijfi/index.php/ijfi/article/view/110.

Rossi G, Marchi A, Valli M, Zuffi G, Ferri D, Zaccanti F, Falconi R. 2016. Analisi della funzionalità dell'ISECI nella rappresentazione di scenari teorici e sperimentali: proposte di calibrazione di parametri di calcolo. It J Fresh Ichthyol. 2016(3):55–62. http://www.aiiad.it/ijfi/index.php/ijfi/article/view/148.

Rossi G, Zuffi GL, Marchi A, Caricato G. 2017. Problematiche nell'affinamento delle comunità ittiche di riferimento per l'applicazione dell'ISECI in aree con carenze informative e quadro zoogeografico incerto. Quali i rischi applicativi? Il caso di studio del bacino del Basento in Basilicata. It J Fresh Ichthyol. 2017(4):125–133. http://www.aiiad.it/ijfi/index.php/ijfi/article/view/102.

Seber GAF. 1982. The estimation of animal abundance and related parameters. 2nd ed. London: C. Griffin & Co., Ltd; p. 654.

Seber GAF, Le Cren ED. 1967. Estimating population parameters from catches large relative to the population. J Anim Ecol. 36(3):631–643. https://www.jstor.org/stable/2818.

Seber GAF, Whale JF. 1970. The removal method for two and three samples. Biometrics. 26(3):393–400. https://www.jstor.org/stable/2529096.

Sprent P. 1993. Applied nonparametric statistical methods. 2nd ed. London: Chapman and Hall.

Stikker PH. 2017. Bhapkar test [Excel spreadsheet]. [accessed 2019 July 26]. https://drive.google.com/file/d/0B3VsxBCbBYAKZ091M0RQRndpd3c/view.

Stuart AA. 1955. A test for homogeneity of the marginal distributions in a two-way classification. Biometrika. 42(3-4):412–416. https://www.jstor.org/stable/2333387.

Thompson PD, Rahel FJ. 1996. Evaluation of depletion-removal electrofishing of brook trout in small Rocky Mountain streams. N Am J Fish Manage. 16(2):332–339. https://doi.org/10.1577/1548-8675(1996)016 < 0332:EODREO > 2.3.CO;2.

White GC, Anderson DR, Burnham KP, Otis DL. 1982. Capture-recapture and removal methods for sampling closed populations. Los Alamos: Los Alamos National Laboratory.

Zerunian S, Goltara A, Schipani I, Boz B. 2009. Adeguamento dell'Indice dello Stato Ecologico delle Comunità Ittiche alla Direttiva Quadro sulle Acque 2000/60/CE. Biologia Ambientale. 23 (2):15–30. http://www.cisba.eu/images/rivista/biologia_ambientale/Ba2009-2/Zerunian-Indice_Ittico_ISECI-WFD.pdf.

Zerunian S. 2004. Pesci delle acque interne d'Italia. Quad. Cons. Natura, 20, Min. Ambiente-Ist. Naz. Fauna Selvatica; p. 257. https://www.minambiente.it/sites/default/files/archivio/biblioteca/qcn_20.pdf.

Zippin C. 1958. The removal method of population estimation. J Wildl Manage. 22(1):82–90. https://www.jstor.org/stable/379730. 1.

Zippin C. 1956. An evaluation of the removal method of estimating animal populations. Biometrics. 12(2):163–189. https://www.jstor.org/stable/3001759.