Ecological Indicators 96 (2019) 52–58

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Original Articles

Monte-Carlo methods to assess the uncertainty related to the use of predictive multimetric indices

Maxime Logez*, Anthony Maire[1], Christine Argillier

Irstea, UR RECOVER, Pôle R&D AFB-Irstea Hydroécologie des Plans d'eau, Centre d'Aix-en-Provence, 3275 Route Cézanne, F-13182 Aix-en-Provence, France

ARTICLE INFO

Keywords:
Ecological assessment
Multimetric index
Bioindication
Lakes
WFD
Fish

ABSTRACT

The publication of the Water Framework Directive by the European commission in 2000 has promoted the development of many multimetric biological indices to assess the ecological status of European waterbodies. These ecological assessments are based on the measurement of deviations between a metric's (characteristic of assemblages) observed values (obtained by sampling) and a metric's expected values in the absence of anthropogenic stressors (reference conditions). In addition, the confidence in the ecological status evaluation provided by the different biological indices is required. Numerous sources of uncertainty due to sampling variability or operator bias, for example, are often considered on observed metric values, whereas uncertainty associated with expected metric values are seldom discussed. In this study, we developed a methodology based on Monte-Carlo methods to assess the uncertainty associated with the establishment of reference values for multimetric predictive indices. This was done by randomly generating reference values and propagating the uncertainty throughout the computation of the index. This methodology can be applied to a wide variety of indices as long as it is possible to make assumptions about the statistical distributions of some of the index's numerical components (e.g. coefficients of the statistical models, metric values). The European Lake Fish Index was used to illustrate the methodology and show how this method can provide valuable information on the confidence in the ecological status defined by the index. These results also revealed that the degree of uncertainty varied between the ecological classes, which were highest for the "Moderate" class and lowest for the "Poor" and "High" classes for the ELFI.

1. Introduction

The Water Framework Directive (WFD; European Union, 2000) quickened the development of biological indices to assess the ecological status of European waterbodies (Hering et al., 2006). It also provided a strong framework for the development of these indices. Indeed, to be WFD-compatible, the indices have to fulfil a panel of criteria such as considering various aspects of biological communities (e.g. abundance, diversity, composition, tolerance, age classes), which are also called metrics. In addition, these multimetric indices have to attribute one ecological class to each waterbody, among the five of the WFD: Bad, Poor, Moderate, Good and High (European Union, 2000).

One of the most important requirements of the WFD is the assessment of the ecological status of a given waterbody by comparing the observed characteristics of the biological communities to the characteristics that would have been observed in reference conditions. The concept of reference conditions has been widely debated, and various definitions have been used in practice (Stoddard et al., 2006). To define the baselines to which current communities should be compared, authors used historical records (Muxika et al., 2007) or defined theoretical features that a community should present (Oberdorff and Hughes, 1992; Karr, 1981; Karr et al., 1986; Maire et al., 2015; Hughes et al., 1998) or used community features observed in minimally disturbed conditions (Oberdorff et al., 2002; Pont et al., 2007; Marzin et al., 2014; Mondy et al., 2012).

Once environmental gradients influence community features (e.g. Logez et al., 2013), two main methodologies were used to control these variations when establishing reference condition values (Roset et al., 2007). The 'type-specific' approach defines reference conditions for groups of water bodies that are generally close from a hydro-morphological point of view and located in geographical areas encompassing similar and stable environmental conditions (e.g. Mondy et al., 2012). The 'site-specific' approach consists in defining reference conditions for a given site depending on its environmental conditions, generally

* Corresponding author.
E-mail address: maxime.logez@irstea.fr (M. Logez).
[1] Present address: EDF R&D, LNHE (Laboratoire National d'Hydraulique et Environnement), 6 quai Watier, F-78401 Chatou, France.

https://doi.org/10.1016/j.ecolind.2018.08.051
Received 16 May 2018; Received in revised form 20 August 2018; Accepted 23 August 2018
1470-160X/

through statistical models relating metric values to environmental variables (Logez and Pont, 2011; Oberdorff et al., 2002; Pont et al., 2007; Marzin et al., 2014). These two approaches tend to control the environmental effect on metric variability by defining homogenous regions (e.g. 'ecoregions'; Wasson et al., 2002) or by explicitly taking into account the heterogeneity of environmental conditions when defining reference values. Despite their differences, the metric value assigned to the reference condition is generally the average or expected value of this metric in a given environment or region, regardless of the approach considered. To be consistently representative of the reference conditions, the expected value of a metric has to be consistently estimated.

The WFD also recommends estimating the confidence related to the ecological assessment of waterbodies. This is of prime importance since forthcoming environmental policies will be based on this assessment, with all the economic consequences induced. To estimate the confidence associated with the ecological assessment, comprehensive information about all the methods involved are required, especially regarding the property of the different biological quality elements (BQEs). Until now the uncertainty associated with the assessment of the ecological status has been mainly addressed with regard to the sampling variability of the BQE (Clarke et al., 2002; Clarke, 2000; Clarke and Hering, 2006; Wiederkehr et al., 2016), the operator bias (Wiederkehr et al., 2015) or the role played by combining the rules of the different BQEs (Moe et al., 2015).

Clarke (2010) has developed the WISERBUGS software, which takes into account the uncertainty associated with the metric's observed values (e.g. sampling variability), once the various sources of uncertainty have been estimated (i.e. their variance) (Thackeray et al., 2013). Using this method, uncertainty assessments have been carried out for numerous metrics that are sensitive to eutrophication in lakes for the different BQEs acknowledged by the WFD (Lyche-Solheim et al., 2013). However, whereas the ecological assessment of a given water body is explicitly based on the average value observed in reference conditions, the variability associated with establishing the reference condition has not been sufficiently studied (but see Marzin et al., 2014).

To quantify this source of uncertainty, we can consider that each metric should not be summarized solely by an average value, but that it would instead be necessary to consider its dispersion. Assessing the potential range of values that a metric could display under reference conditions in a given environment, rather than a single, average value, would allow us to evaluate the uncertainty of the ecological assessment due to the definition of reference values. As defined by the WFD, the ecological assessment of waterbodies is based on comparing the observed and expected values of each metric. Our idea was to extend this comparison from a single reference value (i.e. the average value) to a distribution of reference values and thus to account for the uncertainty associated with the establishment of reference conditions (Logez, 2010; Marzin et al., 2014).

In this study, our main purpose was to provide a methodology based on Monte-Carlo methods (Manly, 1997; Clarke, 2010) to assess the uncertainty associated with the estimation of reference condition values. Compared to Marzin et al. (2014), we were not seeking to integrate the uncertainty associated with reference conditions into the computation of a given index but to develop a methodology that would account for this source of uncertainty for already existing indices.

As a study case, this methodology was developed for the European Lake Fish Index (ELFI), a multimetric predictive index that was not formerly designed to account for uncertainty (Argillier et al., 2013). The ELFI uses the hindcasting approach to predict site-specific reference condition values (Launois et al., 2011; Kilgour and Stanfield, 2006). We present herein the methodological framework before illustrating the outcomes of this methodology using the ELFI as an example of uncertainty assessment associated with a predictive index. Finally, we assess how the uncertainty spread among the five ecological classes to determine whether the same confidence is maintained in the ecological assessment whatever the assigned ecological status.

2. Material and methods

2.1. European Lake fish index

The ELFI is a multimetric index developed to fulfil WFD needs and to evaluate the ecological status of European natural lakes. It is the official method used in France for the assessment of lake ecological status based on fish communities. This index was developed at the European scale on the basis of a data set consisting of natural lakes located in 10 European countries (Argillier et al., 2013). The method involves statistical models to relate metric variability to environmental and pressure variables. This multimetric index was built to respond to eutrophication, which was measured through two stressor variables: total phosphorus (TP) and the proportion of non-natural land uses in the lake drainage area (CLCNONAT; for further detail see Argillier et al., 2013).

The ELFI is a predictive index based on three metrics: total catches per unit effort (CPUE), total biomass per unit effort (BPUE) and catches per unit effort of omnivorous species ($CPUE_{omni}$) (Argillier et al., 2013). To limit skewness and heteroscedasticity, these metrics were log-transformed and were assumed to be normally distributed. The reference values of each metric are predicted using multiple linear regression models (MLRs) and the hindcasting approach (Kilgour and Stanfield, 2006; Baker et al., 2005). Unlike most predictive multimetric indices (Bady et al., 2009; Pont et al., 2007; Marzin et al., 2014) that have used only reference sites to calibrate the statistical models (Logez and Pont, 2011; Oberdorff et al., 2001), the ELFI integrated reference and disturbed sites into the calibration data set. The variability of metric values was modelled using both environmental variables (e.g. lake area) and stressors (e.g. non-natural land cover) (Table 1). To predict reference values for each lake, it was necessary to assign low values to the stressors corresponding to a theoretical low level of impairment. These reference values of stressors were the same for all lakes (Table 1).

Metric observed values were compared to metric reference values and transformed into an ecological quality ratio (EQR; discrepancy between the observed value and the expected value in reference conditions) following the computation process of the ELFI (Ritterbusch et al., 2017; Argillier et al., 2013). The EQR of the three metrics were then averaged and rescaled between 0 and 1 to get a single index value, which was finally derived into ecological classes according to the class boundary values that have been standardized between European countries (Ritterbusch et al., 2017).

2.2. Uncertainty associated with reference condition values

The methodology developed in this study was designed for predictive multimetric indices such as the European fish-based indices for rivers (EFI and EFI+; Pont et al., 2007; Bady et al., 2009) or lakes

Table 1
Environmental variables used as independent variables to predict reference values. The numerical values displayed here correspond to the values observed for Lake Bourget, provided as an example.

Independent variables	Code	Values used	Units
Environment			
Lake area	SLAC	43.72	km^2
Drainage basin area	SBV	589	km^2
Maximum depth	PMAX	145	m
Elevation	ALTI	231.5	m
Pressure			
Total phosphorous	TP	5	mg/L
Non-natural land cover	CLCNONAT	10	%