A procedure to update quality indices based on species abundances: an example using the EPI-L diatom index

Aldo Marchetto,[1,2] Simona Musazzi[1*]

[1]National Research Council, Water Research Institute (CNR-IRSA), Verbania-Pallanza; [2]National Biodiversity Future Center (NBFC), Palermo, Italy

ABSTRACT

We propose a procedure to update ecological quality indices using species autoecology estimated by gradient analysis when new species are found, or taxonomy is updated. When updating an index, the new values must be comparable to the old ones to avoid recalibration and quality class boundary changes. As gradient length influences the values of the species optima, we propose to avoid adding new species to existing lists and we suggest recalibrating the index with a new calibration dataset and rescaling species optima. This final step reduces index updating-induced quality classification changes. An example is shown using the EPI-L diatom index, a quality index for lakes affected by trophic pressure.

INTRODUCTION

Bioassessment of the ecological quality of water bodies is an important tool for their protection and management. Under the impulsion of the Water Framework Directive (European Commission, 2000), a number of quality indices were developed to evaluate the ecological quality status of river, lakes and transitional and coastal water (Birks *et al*., 2012).

A large majority of the indices used in Europe are based on the presence or on the abundance of selected species (Birks *et al*., 2012), as single metrics or as components of multi-metric indices. Their definition includes developing lists of "species scores" or "indicator values", to be used to calculate quality metrics. Species lists cannot be exhaustive, and they may need to be updated to follow the finding of new species and/or changes in taxonomy. A procedure for adding or replacing species in these lists is then necessary to ensure their use over long periods. However, when updating a species list, attention should be paid not to modify the evaluation of the ecological quality of waterbodies based on that specific quality index. Following the WFD, waterbodies are classified in "quality classes", based on the values of some quality indices compared to "class boundaries" defined through complex intercalibration procedures (Birks, 2013) and reported in national and European legislation (European Commission, 2018). If a new index is introduced, a new intercalibration procedure must be performed and class boundaries must be defined.

Species scores can be obtained by assigning conventional values based on expert judgment or literature reviews (Kelly and Whitton 1995; Salmaso *et al*., 2006) or may be obtained estimating from an environmental gradient a single list of species optima using for example abundance-weighted averaging regression (Sgro *et al*., 2007), or multiple indicator values (Potapova *et al*., 2004). In the case of weighted averaging, calculated species optima are unbiased estimates of actual species optima if the samples are equally spaced over the environmental gradient and they are closely spaced with respect to species tolerances (ter Braak and Looman, 1986). However, at the extremes of the gradient, these assumptions cannot be verified as the frequency distributions of taxa are truncated, causing overestimation of the optima at the lower side of the gradient and their underestimation at the higher side.

This effect causes a shrinkage of the species optima toward the centre of the gradient (ter Braak and van Dam, 1989), which depend on the specific data set used for the

Corresponding author: simona.musazzi@cnr.it

Key words: benthic diatoms; ecological quality index.

Citation: Marchetto A, Musazzi S. A procedure to update quality indices based on species abundances: an example using the EPI-L diatom index. *J Limnol 2023;82:2151*.

Edited by: Michela Rogora, *National Research Council, Water Research Institute (CNR-IRSA), Verbania Pallanza, Italy.*

Received: 16 May 2023.
Accepted: 30 September 2023.

J. Limnol., 2023; 82:2151
DOI: 10.4081/jlimnol.2023.2151

calibration. As a consequence, it is not possible to add new data to an existing data set using a new set of samples, as these samples will not be part of the original calibration data set, but of a different data set including new samples and new counts and then the optima will not be comparable.

Nevertheless, when it arises the need to calculate species optima for taxa not included in the original species list, such as new discovered species never detected before in a geographic region, or new classifications of species previously misidentified or unclassified, this can be done using new counts, but it is necessary to calibrate again the index as a whole using an updated data set in order to assure that all species optima are obtained using the same calibration data set.

However, a new data set covers a different gradient than the original one, so the calculated trophic weights would be different. This would make differences in the classification results obtained with the old and the new version of the index. A procedure is then necessary to make this updated index comparable to the previous version, in order to avoid altering the classification results.

The EPI-L index

The original EPI-L index is a quality index based on benthic (periphytic and epilithic) diatoms to be used for lakes affected by a trophic pressure. It was developed by Marchetto *et al*. (2013) on the basis of 108 samples from 80 lakes in Italy. 109 species showing abundance higher than 3% in at least 1 lake and higher than 1% in at least 3 lakes were considered.

For each *i*-th species, a trophic weight (p_i) was obtained by the average of the logarithm of the epilimnetic total phosphorus concentration (TP, in μg L^{-1}), weighted by the relative abundance (as number of counted frustules) of that species in all the lakes.

The indicator value (v_i) was obtained as the inverse of the average of the squared differences between the trophic weight of the species and the epilimnetic total phosphorus concentration in each lake, weighted by the abundance of that species in the lake itself. Indicator values higher than 30 were replaced with 30.

The calibrated EPI-L index for each *j*-th lake was then calculated on the basis of the relative abundance ($a_{i,j}$) of each species using a linear function of the Zelinka and Marvan (1961) formula:

$$EPI-L = 4 - 2\frac{\sum_{i=1}^{n} a_i p_i v_i}{\sum_{i=1}^{n} a_i v_i}$$

The index value was not calculated for those lakes for which the sum of the relative abundance of the species included in the score list was lower than 70%.

Updating procedure

A new data set including 167 diatom species found in 185 epilithic and epiphytic samples from 97 lakes was used, including not more than 3 samples for each lake. The new data set included 58 counts from the previous version of EPI-L and 35 new counts obtained from samples used in the previous version to include newly defined species. The remaining samples of the old data set were excluded, as taxonomic verification was not possible, because we could not obtain the original diatom slides. In order to include new species not present in the old data set, further 50 samples from 22 new lakes were also included, together with 41 new samples from lakes used in the old data set (*Tab. S1*). The updated version of EPI-L comprises recently described species such as *Achnanthidium neomicrocephalum, Brachysira neglectissima, Cocconeis pseudolineata, Fragilaria perdelicatissima, Gomphonema elegantissimum* and *Sellaphora nigri* and taxa assigned to the new genera *Halamphora, Pantocsekiella, Delicatophycus* and *Gogoreieva*.

Trophic weights and indicator values were obtained using the same procedure as for the original EPI-L and are reported in *Tab. S2*.

As the new data set covers a different gradient than the original one, it was expected that calculated trophic weights would be different for the same species in the two data sets. In fact, the range of the updated optima was slightly smaller than that of the original optima (Fig. 1). The dependence of the trophic weights from the dataset used is a very critical point when updating a quality index, as it is crucial to avoid that the redefinition of the index could lead to changes in the index values. To ensure that the estimation of the quality status of the lakes would be the same using the old and the updated version of index, species optima were rescaled. In the case of the revised EPI-L, an ordinary least square linear regression was used, but non-linear relationships can also be used (Marchetto, 1994).

A regression of the new trophic weights against the old ones is an important and necessary procedure to ensure that the index values obtained using the new definition of the index will be as close as possible to the values obtained with the previous definition. Without the rescaling procedure, recomputing a new updated index either using only a new dataset, either using the old and new data set together, would result in systematic differences between the quality index calculated for the same lakes using the previous and the updated version of the index. This would require performing a new calibration process for the quality index and may entail changes in the boundaries between quality classes.

After rescaling, the two sets of trophic weights are comparable, and it would be possible to use the rescaled weights of the new species together with the original weights of the original dataset. However, we recommend to use the rescaled values obtained from the new dataset for the ecological classifications, because some species included in the data base used for the previous version of the index have been split, or merged or transferred in different genera in the new data set. For all the samples used for the new calibration of the EPI-L index, the values of the original and updated indices were calculated, and they are compared in Fig. 2 for all the samples of the new data set for which the sum of the relative abundances of the species included in both original and updated species lists was at least 70%.

In spite of more marked individual differences between the two versions of the index in some lakes, the median difference between them is low (0.06 units) before rescaling, and it dropped to -0.03 units after rescaling species optima. Moreover, rescaling the optima caused a

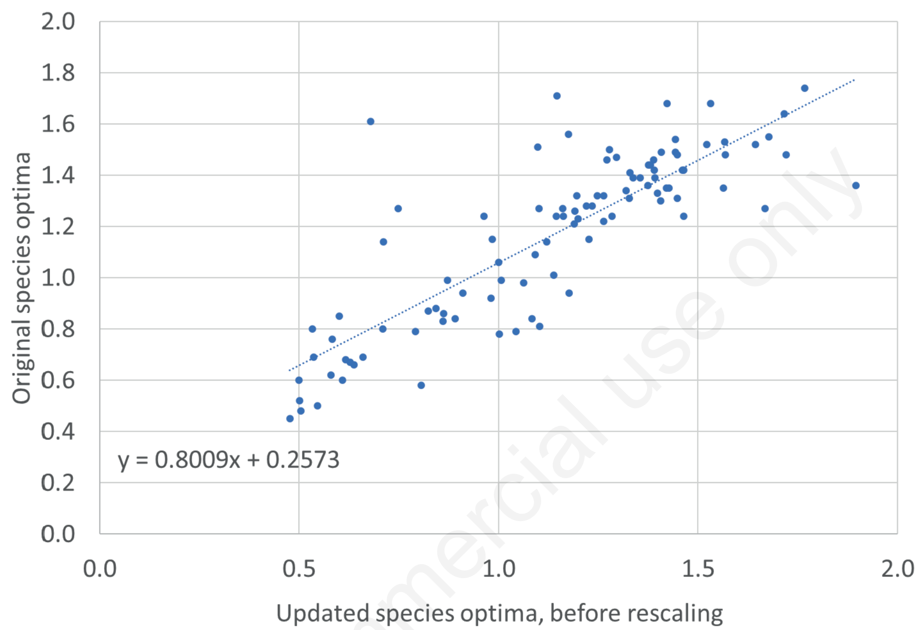

Fig. 1. Comparison between original and revised species optima before rescaling.

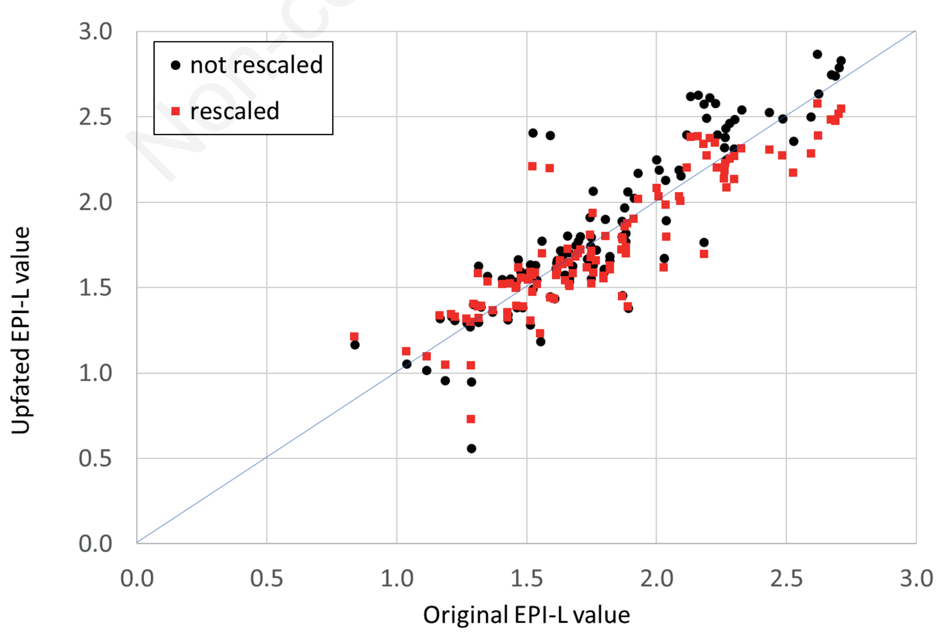

Fig. 2. Comparison between original and revised EPI-L values for the samples for which the sum of the relative abundances of the species included in both original and updated species lists was at least 70%. Black, not rescaled; red, rescaled. The 1:1 line is also shown.

reduction of the maximum absolute difference between original and updated index from 0.88 to 0.69 units.

Even if in this case the rescaling procedure only leaded to small changes in the index values, we strongly recommend rescaling species optima when updating a quality index based on species lists and gradient analysis, in order to limit the possible changes in the quality classification of water bodies.

ACKNOWLEDGEMENTS

We gratefully thank the following colleagues, who provided diatom counting for the definition of the EPI-L index and for the present update: Renate Alber, Berta Thaler (Provincial Environmental Protection Agency of Bolzano), Arianna Macor, Raffaella Zorza (Regional Environmental Protection Agency of Friuli Venezia Giulia), Laura Mancini, Stefania Marcheggiani, Camilla Puccinelli, Claudia Vendetti (National Institute of Health), Chiara Agostinelli, Elena Arnaud, Andrea Beghi, Fabio Buzzi, Elisa Carena, Marco Fioravanti, Riccardo Formenti, Matteo Galbiati, Filippo Galimberti (Regional Environmental Protection Agency of Lombardia), Pierluigi Fogliati, Arianna Nicola (Regional Environmental Protection Agency of Piemonte), Silvia Costaraoss, Sabrina Pozzi, Paola Testa (Provincial Environmental Protection Agency of Trento), Susanna Cavalieri, Federica Cimoli, Giovanna Marchi, Chiara Rustighi (Regional Environmental Protection Agency of Toscana), Stefania Bracchi, Isa Crescentini, Valentina Della Bella, Margerita Di Brizio, Rosalba Padula, Egiziana Rinaldi (Regional Environmental Protection Agency of Umbria).

The authors acknowledge the support of the National Biodiversity Future Center to Consiglio Nazionale delle Ricerche, funded by the Italian Ministry of University and Research, PNRR, Missione 4, Componente 2, "Dalla ricerca all'impresa", Investimento 1.4 Project CN00000033.

The useful comments of two anonymous reviewers helped to improve the quality of this manuscript.

REFERENCES

Birk S, Bonne W, Borja A, Brucet S, Courrat A, Poikane S, et al., 2012. Three hundred ways to assess Europe's surface waters: An almost complete overview of biological methods to implement the Water Framework Directive. Ecol Indic 18:31-41.

Birk S, Willby NJ, Kelly MG, Bonne W, Borja A, Poikane S, van de Bund W, 2013. Intercalibrating classifications of ecological status: Europe's quest for common management objectives for aquatic ecosystems. Sci Total Environt 454-455:490-499.

European Commission, 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. OJ L 327, 22.12.2000, p. 1-73. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32000L0060

European Commission, 2018. 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 State monitoring system classifications as a result of the intercalibration exercise and repealing Commission Decision 2013/480/EU. OJ L 47, 20.2.2018, p. 1-91. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32018D0229

Kelly MG, Whitton BA, 1995. The Trophic Diatom Index: a new index for monitoring eutrophication in rivers. J Appl Phycol 7:433-444.

Marchetto A, 1994. Rescaling species optima estimated by weighted averaging. J Paleolimnol 12:155-162.

Marchetto A, Bracchi S, Buzzi F, Carena E, Cavalieri S, Cimoli F, et al., 2013. [Indice per valutazione della qualità delle acque lacustri italiane a partire dalle diatomee epifitiche ed epilitiche (EPI-L)], p. 75-92. In: A. Marchetto (ed.), [Indici per la valutazione della qualità ecologica dei laghi]. [Report in Italian]. Report CNR-ISE 02-13. Available from: http://www.vb.irsa.cnr.it/it/wfd

Potapova MG, Charles DF, Ponader HC, Winter DM, 2004. Quantifying species indicator values for trophic diatom indices: a comparison of approaches. Hydrobiologia 517:25-41.

Salmaso N, Morabito G, Buzzi F, Garibaldi L, Simona M, Mosello RT, 2006. Phytoplankton as an indicator of the water quality of the deep lakes south of the Alps. Hydrobiologia 563:167-187.

Sgro GV, Reavie ED, Kingston C, Kireta AM, Ferguson MJ, Danz NP, Johansen JR, 2007. A diatom quality index from a diatom-based total phosphorus inference model. Environ Bioindicat 2:15-34.

ter Braak CJF, Looman CWN, 1986. Weighted averaging, logistic regression and the Gaussian response model. Vegetatio 65:3-11.

ter Braak CJF, van Dam H, 1989. Inferring pH from diatoms: a comparison of old and new calibration methods. Hydrobiologia 178:209-223.

Zelinka M, Marvan P, 1961. [Zur Präzisierung der biologischen Klassifikation der Reinheit fließender Gewässer].[Article in German]. Arch Hydrobiol 57:389-407.

Online supplementary material:

Tab. S1. Percent relative abundances of diatoms used for updating EPI-L TP: average epilimnetic or surface water total P concentration ($\mu g\ L^{-1}$). Sample code: N new sample from new lake; P new sample from a lake included in the data set of the previous version of EPI-L; PR sample included in the data set of the previous version of EPI-L, not revised; PRr sample included in the data set of the previous version of EPI-L, revised. The full name of each species is reported in Tab. S2.

Tab. S2. Updated and original trophic weights and indicator values. Values in bold italics refer to species synonyms.