

Biomonitoring with macroinvertebrate communities in Italy: What happened to our past and what is the future?

Tiziano BO,¹ Alberto DORETTO,² Alex LAINI,³ Francesca BONA,² Stefano FENOGLIO^{1*}

¹Department of Science and Technological Innovation, University of Piemonte Orientale, Viale T. Michel 25, 15121 Alessandria;

²Department of Life Sciences and Systems Biology, University of Turin, Via Accademia Albertina 13, 10123 Turin; ³Life Sciences Department, University of Parma, Parco Area delle Scienze 11/A, 43124 Parma, Italy

*Corresponding author: stefano.fenoglio@uniupo.it

ABSTRACT

This paper reviews the history and development of biological water quality assessment using macroinvertebrates in Italy. Italy was one of the first European countries to officially adopt a biomonitoring system based on benthic invertebrates, the “Indice Biotico Esteso” (IBE). After the European Water Framework Directive (WFD) 2000/60/EC, this method was replaced by the “Standardisation of River Classifications Intercalibration Common Metrics” (STAR_ICM) index, which met the new requirements. As this method has been employed for some years, it could be useful to take a provisional stock and to provide some suggestions to ameliorate the current biomonitoring approach, also trying to minimize the break with past practices and better harmonize the history of biomonitoring in Italy. One of the most evident difference between past and current approach is related to the amount of time and effort required in the application of the two methods. STAR_ICM is a scientifically rigorous and modern method, but much more time-consuming and challenging in both field and laboratory efforts. This fact has various disturbing practical repercussions, *i.e.*, the environmental agencies have generally reduced the number of sampling stations routinely monitored during the year. The aim of our work is to propose some operational changes that would help to simplify and expedite the monitoring process. In particular, regarding fieldwork, we focus on the time and effort required for macroinvertebrate collection, while for laboratory activity we suggest a reshaping of the requested taxonomic detail. Moreover, in this way the data provided by the new approach could be compared with the long time series available from the previous application of IBE.

Key word: IBE; STAR_ICMi; WFD 2000/60; environmental quality assessment; running waters.

INTRODUCTION

A brief history of biological monitoring with macroinvertebrates

In the last decades, biological monitoring of running water systems has become increasingly important as an indispensable complement to traditional chemical-physical techniques in the evaluation of anthropic impacts (Barbour *et al.*, 1999; Birk *et al.*, 2012; Friberg, 2014). A wide range of techniques blossomed throughout the last century (Hellowell, 1986) and, whilst a variety of biological groups continued to be considered (*e.g.*, bacteria, benthic algae, fish), the use of benthic macroinvertebrates became by far the most common method (Metcalf, 1989). Nowadays, benthic macroinvertebrates represent the most widely used group of organisms in freshwater biomonitoring, due to their different sensitivity to changes in both chemical characteristics of the water column and physical properties of habitats (Rossaro *et al.*, 2011; Szivák, and Csabai, 2012). Macroinvertebrates are a heterogeneous group of ubiquitous and abundant organisms, relatively easy to collect, identify and enumerate (Bonada *et al.*, 2006). Furthermore, the relatively long length of life cycles of many species and

their constant presence in the same locality make the analysis of their community structure an effective tool to detect the occurrence of human pressures over long time periods (Allan and Castillo, 2007).

Benthic macroinvertebrates have a long history as key component of biomonitoring tools, dating back to the beginning of the 20th century (Cairns and Pratt, 1993). In fact, the idea of using macroinvertebrates as biological indicators began in Europe with the studies of Kolkwitz and Marsson (1908). Their Saprobien system relied on the fact that some organisms could be used as indicators of specific environmental conditions, and was essentially aimed at relating the organic load to the presence and distribution of benthic invertebrates in rivers. This system is still widely adopted in Central Europe, mainly in countries with German influence, while it has found minor acceptance in other areas. Its main biases are the specific and geographically restricted taxonomic approach, and limited applicability in the detection of pollution other than organic load. For these reasons other indices were developed, combining the indicator value concept with biodiversity and relative abundance of different selected groups. The Trent Biotic Index was a pioneering and seminal approach, designed to assess the water quality status

of the Trent River (Woodiwiss, 1964). This method took two main aspects of benthic communities: the biological richness, *i.e.* the number of collected taxa, and the presence of some key groups, characterized by different levels of tolerance to environmental alteration. The index ranges from zero (polluted condition) to ten (clean waters). In those years other methods arose, and among them specific scores were attributed to different selected taxa according to their abundance (Chandler, 1970) or tolerance. These approaches developed into true biotic score indices, among which the Biological Monitoring Working Party (BMWP; Hellowell, 1986) assumed a particular importance. In this method, taxonomic identification is easier, because performed at family level. Each family is associated to a specific score, depending on its sensitivity to environment alteration; for example, Heptageniidae are scored ten, while Chironomidae two.

Biomonitoring methods that use macroinvertebrates to assess lotic ecosystem quality are nowadays a key topic in freshwater ecology (Guareschi *et al.*, 2017) and they are employed in many countries, such as United States of America, United Kingdom, Australia, New Zealand, Canada, European Union. Furthermore, their diffusion as bioindicators is also growing in South America, East Asia, Africa and other areas (Buss *et al.*, 2015).

The Italian situation before the Water Framework Directive (2000/60/EC)

Italy was among the first European countries to adopt a biomonitoring system based on benthic macroinvertebrates. In fact, since 1975 many field studies were conducted in our country with the aim of obtaining and calibrating a reliable method. For example, in a study realized in the Parma River (Emilia Romagna), Ghetti and Bonazzi (1977) compared the reliability of several indices, while Casellato and collaborators (1980) investigated the applicability of the French Verneaux and Tuffery's Biotic Index to the Brenta River (Trentino). After some adaptations to the Italian environmental conditions and comparisons with other European methods, the "Indice Biotico Esteso" (EBI; Ghetti, 1986) later renamed IBE (Ghetti, 1997) was calibrated and adopted for running water biomonitoring in Italy. According to this procedure, benthic invertebrates were collected with kick-nets (21 mesh cm⁻¹). In wadeable environments, transects were realized kick-sampling with the net from one bank to another, and samples had to be accurately collected in all microhabitats, in order to include the entire local biodiversity. Collected macroinvertebrates were field sorted and then identified at the taxonomic level required for each group (*e.g.*, genus for Plecoptera and Ephemeroptera, family for Trichoptera and Diptera). The taxonomic list was then transformed into a numeric value, using a double entry table. This table considered the tax-

onomic richness in columns, varying from poor (0-1 taxa) to very rich and biodiverse environments (>35 taxa). The horizontal entry took into consideration the sensitivity of different benthic taxa: the highest row was represented by Plecoptera, whilst the lowest by Oligochaeta and Chironomidae. Combining taxonomic richness and presence of these selected groups, the final index ranged from 0 to 14. These values were finally subdivided into 5 quality classes (see Ghetti, 1997 for further details).

The Water Framework Directive (2000/60/EC) and its consequences on water biomonitoring in Europe

The Water Framework Directive (WFD) represents today the main normative reference for all EU member states in the field of water monitoring and conservation (Collins and Anthony, 2008). Published in 2000, the WFD has profoundly changed management practices placing greater emphasis on ecosystem integrity rather than on the simple detection of pollution (Hering *et al.*, 2010). In fact, this was the most noteworthy and innovative aspect, because water quality assessment shifted from a merely chemical to an ecological approach (Nôges *et al.*, 2009). In this context, the conformity of the whole community in comparison to unaffected conditions must be considered rather than the individual taxon-stressor relationship (Birk and Hering, 2006). Following this holistic approach, some innovative elements were introduced. First, several biological components, named "Biological Quality Elements" (BQEs), are simultaneously taken into consideration to provide a comprehensive evaluation of the environmental condition. For lotic environments, BQEs include phytoplankton, phytobenthos, macrophytic flora, benthic macroinvertebrates and fish (Hering *et al.*, 2003). Moreover, the WFD requires a "type-specific" approach (Hering *et al.*, 2006, Verdonschot and Nijboer, 2004), including six different categories of aquatic ecosystems: rivers, lakes, coastal waters, transitional waters, artificial and heavily modified water bodies (Borja, 2005). Within these categories, all water bodies are grouped in similar typologies according to their geo-morphological, physical and chemical features (Moog *et al.*, 2004). For each typology, reference conditions, *i.e.* "sites that show near-natural or un-impacted conditions", have to be identified so that water quality assessment is calculated as Ecological Quality Ratio (EQR) between the observed and the reference conditions (Von de Ohe *et al.*, 2007). The result is expressed in five quality classes (High, Good, Moderate, Poor and Bad) with High status meaning no differences between reference and observed conditions, while Poor and Bad classes are associated to strong differences (Birk *et al.*, 2012). Third, all member states were expected to achieve the "Good Ecological Status" for their water bodies by 2015 (Heiskanen *et al.*, 2004), encouraging the adoption of specific actions (Programmes of Measures)

in each country (Logan and Furse, 2002). In the WFD context, three different types of monitoring are employed for different purposes (Buffagni and Erba, 2007). The surveillance monitoring is performed when the assessment of the overall condition of a water body is needed. When results indicate that a risk of failing to achieve the Good Ecological Status occurs, then the operative monitoring is implemented. Finally, in those occasions where further insight on the effects of specific alterations or pollution phenomena is necessary, the investigative monitoring must be adopted. Based on the type of monitoring, specific BQEs must be employed. Due to this comprehensive approach, all biological quality elements are used in the surveillance monitoring, while just one or two BQEs must be considered in operational and investigative ones.

Although the WFD introduced many innovative inputs and challenges, it was undoubtedly clear that a strong effort was necessary to make the assessment procedure adoptable, coherent and comparable across Europe (Pollard and Huxham, 1998; Reyjol *et al.*, 2014). Major problems regarded the choice of sampling methods and the setting of boundaries among quality classes, because many Member States relied on their own sampling programs. Buffagni and Furse (2006) highlighted that the WFD did not give strict indications about the sampling system because the most important objective was the harmonization of findings rather than of methods. In this context, each single State was enabled by the WFD to choose whether to improve the national method or develop a new sampling procedure. However, with the aim of ensuring an acceptable level of standardization, two crucial European projects were developed: the AQEM (Development and testing of an integrated assessment system for the ecological quality of streams and rivers throughout Europe using benthic macroinvertebrates; 2000-2002; Buffagni *et al.*, 2001), and the STAR projects (Standardization of river classifications: Framework method for calibrating different biological survey results against ecological quality classifications to be developed for the Water Framework Directive; 2003-2005). The AQEM project focused exclusively on benthic macroinvertebrates, with the aim to define an operative and standard procedure for sampling and water quality assessment. The STAR project tried to solve some critical aspects related to the implementation of the Directive, especially those concerning the inter-calibration procedures (continuity with national methods, reliability of different taxa accounting for different stressors or stream types, setting procedures for the quality classes, *etc.*). Detailed information about these projects can be obtained by the reviews of Hering *et al.* (2004) and Furse *et al.* (2006), respectively; while in the next paragraph their main outcomes are briefly described with regard to biomonitoring in lotic environments.

The AQEM sampling method and the associated STAR_ICM index

The AQEM is a sampling procedure based on benthic invertebrates designed to assess the Ecological Status of running waters according to the WFD. More detailed information is available on the web site (www.aqem.de), in the associated manual (AQEM Consortium, 2002) and in an IRSA-CNR thematic publication (Buffagni and Erba, 2007). Briefly, the AQEM sampling method is a quantitative procedure that relies on a multi-habitat design, as already adopted in other contexts (*e.g.*, Rapid Bioassessment Protocol - USA; Barbour *et al.*, 1999). A defined number of samples are collected from different microhabitats according to their percentage of occurrence in the examined river section. The quantitative aspect is a fundamental requirement of the WFD, while the purpose of the proportional multi-habitat approach is to provide a reliable and standardized evaluation of the structure of the sampling site. According to the official protocol, two phases are generally needed: *a priori* assessment of the monitoring (river characteristics and other aspects) and the following field activity. The *a priori* phase is fundamental to define the preliminary conditions necessary to apply the type-specific biomonitoring required by the WFD. In particular, the type of monitoring and the Hydro-Ecoregion (HER) at which the watercourse belongs are identified in this phase. According to this information, the number of samples, the total sampling area (0.5 or 1 m²) and the mesohabitat to be sampled (riffle or pool) are defined. As previously stated, the AQEM method adopts a quantitative approach, with a set number of replicates that must be collected. This number varies according to the type of monitoring: 10 replicates are collected for the operative monitoring, while for the other types additional replicates may be added. By contrast, the selection of riffle rather than pool areas and the total sampling surface are type-specific and standardized. In the subsequent field activity, at first the percentage of occurrence of each microhabitat in the section must be recorded visually. With regard to this task, both biotic and abiotic microhabitats are considered as a continuous layer where only those with at least 10% of frequency are considered. It is important to note that the AQEM Consortium provided a list of coded microhabitats, representing a standard selection of substrata allowing an objective site description. The mineral substrata are classified according to the length of the median particle diameter (*i.e.*, gravel, megalithal, *etc.*), whereas biotic substrata are grouped according to the type of vegetation or organic matter (*i.e.*, algae, macrophytes, CPOM, *etc.*). Based on the visual estimates of microhabitat percentages, the 10 replicates are collected proportionally using the Surber net. Since most infrequent microhabitats may be ignored, additional samples may be

collected as optional replicates. This is especially suggested for surveillance monitoring and reference sites.

Benthic invertebrates are identified to family level for operative monitoring and genus or Operational Units (*i.e.*, sub-genus identification, only for some selected Ephemeroptera groups) for surveillance/investigative monitoring. In addition, the abundance of individuals of each taxon is reported (although the abundance can be estimated when beyond a threshold of 10 individuals). All benthic invertebrates sampled from each microhabitat are then pooled together in order to compose a unique list of taxa. On the basis of this taxonomic list, the Ecological Status is obtained applying the STAR_ICM (Intercalibration Common Metrics) index. This is a multimetric index developed after an intense process of inter-calibration (Verdonschot and Moog, 2006; Bennett *et al.*, 2011). Starting from over 50 different proposed metrics, only 6 were definitively selected to compose the index: ASPT (Average Score Per Taxon), $\text{Log}_{10}(\text{Sel_EPTD}+1)$ (where EPTD is the sum of selected Ephemeroptera, Plecoptera, Diptera and Trichoptera taxa), 1-GOLD (where GOLD is the sum of Gastropoda, Oligochaeta, and Diptera), total number of families, total number of EPT (Ephemeroptera, Plecoptera, and Trichoptera) families and the Shannon-Weiner diversity index (H'). The selected indices include different parameters of benthic communities: taxa sensitivity, abundance and diversity. Each metric is calculated separately and then they are combined into the overall index score, each metric with a specific weight. Finally, the values of each metrics and the final score also are normalized according to those of the reference conditions, giving the Ecological Status as an Ecological Quality Ratio (EQR) between the observed and the reference values.

Aim of the study

Aim of this study was to briefly review the history of benthic macroinvertebrates biomonitoring in Italy, with a special focus on the changes that occurred following the WFD implementation. In particular, we focused not on general, theoretical differences but on practical aspects, on the basis of our extensive experience and by applying comparatively the IBE and the STAR-ICMi to a large set of benthic macroinvertebrate samples. Since the post WFD method has been employed for a number of years (Hering *et al.*, 2010), we believe this is the right time to attempt some objective assessments, which can bring insights and ideas contributing to the future progress of biomonitoring with macroinvertebrates in Italy.

Direct comparison among indices?

Ideally, macroinvertebrate biomonitoring protocols, besides being sensitive to impacts, should be reliable, efficient, cost-effective, and easy to use: the search for a sat-

isfactory method has produced a number of comparisons, as shown by the abundant literature (Cao *et al.*, 1996; Buss *et al.*, 2015; Guareschi *et al.*, 2017). As reported above, the IBE has been used in Italy at the national scale since 1986, but afterwards it was considered inadequate because of its inconsistency with the WFD. In particular, the most common criticisms to the application of IBE were that this method did not consider 'reference conditions' and that was not type specific, because the same scoring system and quality class boundaries were applied to all types of rivers. Moreover, the IBE seemed not to satisfy some quantitative requirements because it did not take into account taxa abundances. For these reasons, this method (as happened for many others in Europe) was replaced by newer approaches. On the other hand, some European Countries tried to maintain a connection with the past, transforming or improving their pre-WFD method (Jáimez-Cuéllar *et al.*, 2002; Munné and Prat, 2009 for Spain). In Belgium, for example, sampling and identification procedures of the post-WFD Multimetric Macroinvertebrate Index Flanders were the same used in the "old" Belgian Biotic Index (Gabriels *et al.*, 2010). In Italy, this did not happen and the transition was a drastic clear-cut. Comparisons between IBE and STAR_ICMi results are scarce (Mancini *et al.*, 2010), but can be of some interest. Our unpublished data suggest that results of the two indices concur in general, with a modest tendency of IBE in overestimating the quality class. Nevertheless, after some years of application of the STAR_ICMi, it is possible to make some remarks, based on our research experience in different areas of the Italian peninsula, and from personal communications of ARPA (Regional Agencies for the Protection of the Environment), researchers and private consulting operators.

It is pointless to question which method performs better, because certainly the STAR_ICMi meets the WFD requirements while this cannot be said for the IBE. Moreover, the STAR_ICMi is the expression of a widely participated and complex process, based on the most modern findings and techniques.

Critical aspects in the current scenario

In our opinion, the main problem related to the hands-on, routine application of the post-WFD Italian method is that it is extremely more consuming in time and efforts than the one previously used. Considering the fact that Environmental Agencies (*e.g.*, ARPA agencies in Italy) and local governments are involved in extensive monitoring plans often carried out with scarce budgets and limited resources, the increased work required for a single sampling represents an important limiting factor. For example, as a consequence of the increased effort for each sampling point, the number of stations seasonally monitored by the ARPA in the Cuneo District (NW Italy) dropped and

reduced by a quarter from 2006 to 2016 after the WFD introduction. Moreover, the current classification process is based on 6-year cycles (subdivided in two 3-year campaigns covering all significant water bodies) while the previous classification of all watercourses occurred on a 1-year basis. This reduction may allow just a partial and scarcely updated and representative “picture” of the river ecosystem health.

Considering the fact that time-related issues are of the greatest importance in planning and realizing biomonitoring campaigns, we summarize below the elements that, in our opinion, are most responsible for the increase of workload required for each sampled station.

Quantitative approach

The quantitative approach and the use of Surber nets are the most relevant innovations in field work related to AQEM/STAR_ICMi. Regarding the quantitative approach, IBE takes into account only a numerical “threshold”, namely a limiting value below which the presence of a taxon is disregarded, and uses four semi-quantitative levels were considered (*i.e.*, * = below the fixed “entry” number; I = present; L = abundant; U = dominant). The new method introduces a quantitative approach, but in the first publication (Buffagni and Erba, 2007), it is stated that beyond a threshold of 10 individuals the abundance can be estimate (page 23; Buffagni and Erba, 2007). The introduction of subsamples involves an additional decrease of quantitative accuracy (ISPRA, 2014). A further reduction in quantitative precision results from the use of subsamples, adjusted in more recent publications (APAT, 2007; ISPRA, 2014). Laini *et al.* (2014) have already expressed concerns about the new quantitative approach. In our opinion, the underestimation of abundances (and taxa richness) is a critical point. In fact, estimating necessarily implies not using real quantitative data (needed when using metrics such as Shannon Index H'). In addition, the use of the same abundance threshold value (initially set to $n = 10$, then increased) for all groups seems largely inadequate. It can be very misleading to count 10 individuals of some groups, and to estimate the rest of their total population, which may amount to hundreds or even thousands of individuals in one sample in some cases. Moreover, it is very different to count 10 Chironomidae or 10 Perlidae, and then estimate the rest of their population: estimation errors are obviously greater for small and cryptic organisms that for large and clearly visible ones.

Taxonomic levels

Some reconsideration on the taxonomic aspects related to the calculation of the STAR_ICMi could be useful. In the previously cited manuals (Buffagni and Erba, 2007; ISPRA, 2014), the identification at the family level

is required for the operative monitoring, while the genus or Operational Units level are requested for the surveillance and investigative monitoring. However, the use of the Operational Units is limited to the Order Ephemeroptera, where this level coincides, in most cases, with the genus. *Rhithrogena*, *Caenis*, and *Baetis* represent interesting exceptions, whose identification needs a sub-genus level of detail (*e.g.*, species or groups of morphologically-similar species). Great taxonomic attention is mandatory also for other Ephemeroptera (such as *Procladius*, *Pseudocentropilum*) and not for other sensitive taxa such as Plecoptera and Trichoptera. In our opinion, this is an important point for future considerations. The sub-genus determination of these organisms is really time-consuming and not so easy for most of the technicians of the Environmental Agencies, so that hopefully a future update of the method could reconsider this particular aspect. Our main criticisms are the following:

- i) Is it really necessary to conduct the taxonomic identification at the sub-genus level? The methods proposed for France (I₂M₂; Mondy *et al.*, 2012), Spain (Jáimez-Cuellar *et al.*, 2002) and other European countries do not provide for such taxonomic detail, being family or genus the most detailed taxonomic resolution, and yet these methods allow achieving the monitoring goals of the WFD.
- ii) What is the scientific reason to focus only on Ephemeroptera? and precisely on some selected species-groups within Ephemeroptera? Although different sensitivity to environmental alterations have been reported within these families, Baetidae and Caenidae are considered, as a whole, examples of quite tolerant mayflies in most biomonitoring systems (such as BMWP, FBI-Hilsenhoff and many others). The same occurs for *Rhithrogena* (Heptageniidae), a relatively homogeneous group of generally rhithronic, reophilous and oligotrophic organisms, considered as wholly reliable indicators of good environmental quality. We suggest to exclude the introduction of a sub-genus level identification also for Plecoptera (or Trichoptera) for the same considerations reported above. However, if a more detailed taxonomic analysis is reputed to be useful or even mandatory, why not include or consider other groups? This choice could be made considering factors such as their large-scale (geographical areas) and small-scale (mesohabitat) distribution, and their relative abundance in benthic assemblages. Chironomidae, at present grouped all at the family level, are almost ubiquitous, and include taxa with different ecological requirements and could be a good choice (Adriaenssens *et al.*, 2004). A taxonomic deepening considering the tribes or sub-family units of Chironomidae could be more useful, as it would for example allow to separate Diamesinae, in-

habiting pristine waters, to the tolerant and even alpha-mesosaprobic Chironominae.

- iii) Finally, the ecological information gained by using the taxonomic resolution at specific or sub-genus level in selected Ephemeroptera is less important than the information lost by grouping all Plecoptera (in the IBE considered at genus level) into families and considering this important indicator group only in the metric “number of EPT” families. For a long time, biomonitoring reports have been considered an important resource also for studies on biodiversity and biogeography studies, and the record of “Perlodidae” instead of “*Besdolus*” or Capniidae instead of “*Capnopsis*” is scientifically very different.

CONCLUSIONS AND RECOMMENDATIONS

We are aware that, despite its long history, doubtless strengths, and innovative contribute, IBE is currently an outdated method. Anyway, we are also confident that the STAR_ICMi could be ameliorated.

A first, important issue that should be addressed for a better implementation of river biomonitoring is related to the direct and indirect training efforts. In fact, when the IBE was the official method in Italy, a training course was annually realized by the author of the method, Prof. Ghetti, and his collaborators for 19 years. In one week, in Trento, participants learned the use of the method with field samplings, laboratory analysis, data processing and discussions. In addition, the application of the method IBE was exhaustively explained in a single, comprehensive manual (Ghetti, 1997). The implementation of STAR_ICM perhaps lacks a similar teaching strategy, as many courses are organized independently by several institutions and associations, often without coordination and direct management of the authors. Furthermore, the information relating to STAR_ICMi application is distributed in a series of publications, while after some years of applications it would be better to concentrate them in a single, definitive manual.

Finally, we propose here two possible improvements to reduce the effort/time consumption for the data collection and processing and to better harmonize the history of biomonitoring and the comparability of the relative data in our country. Therefore, we would like to conclude this paper with some suggestions, hoping that they could be the subject of future discussion and applications.

- i) The first is related to the use of Surber nets and quantitative method. Since the STAR_ICM (unlike other methods such as the French, IBGN) is not strictly quantitative, because it involves numerical estimates of organisms, samples could be collected for example using fixed-time kick-nets. Also by using these devices, the multi-habitat approach could be maintained

as occurs in Denmark (Friberg *et al.*, 2005), Belgium (Gabriels *et al.*, 2010), and Spain (Munné and Prat, 2009). Interestingly, in a recent work Buss and collaborators (2015) reported that in the United States, kick-nets are used in more than 60% of the State/Federal biomonitoring protocols, whereas Surber, dredges, Hess, and other fixed-area samplers are used in ~9%.

- ii) Our second suggestion is related to the taxonomic detail. Since, according to us, it is not informative to use the sub-genus level for few mayfly taxa, here we propose to re-establish family-level determination for most groups, except for Ephemeroptera and Plecoptera, which should be considered at genus level. This would allow: a) to reduce and simplify taxonomic work; b) to obtain data that are comparable with diffuse and long-term data records. We are aware that this modification may have some consequences on the reference conditions already measured, but these changes may be introduced starting from the next verification of reference conditions.

- iii) Finally, in our opinion some changes could improve the reliability of the STAR_ICM. For example, there is some confusion about the source of ASPT (derived from BMWP) scores in the STAR_ICM index (Buffagni and Erba, 2007; Buffagni *et al.*, 2008), and some problems could arise from the adoption of scores originating from biomonitoring in the United Kingdom (Davy-Bowker *et al.*, 2008). The adoption of BMWP scores designed for Mediterranean countries (Jáimez-Cuéllar *et al.*, 2002) or the development of specific scores for Italian watercourses could be an improvement for the assessment of ecological status with the STAR_ICM index. The latter option would be possible by using information gathered with the IBE protocol during its 20 years of application. Moreover, the precision attained by the monitoring system is an essential requirement of the WFD (Clarke, 2013), and it is crucial to clearly discriminate between watercourses in good and less than good ecological status. To date little effort was made to assess the precision of the STAR_ICM (Laini *et al.*, 2014) and, more generally, to estimate the uncertainty of the biotic indexes adopted after the WFD (but see Clarke, 2013).

We are confident that the adoption of the suggested modifications, although not substantial in the architecture of the method, would have interesting repercussions. They would make it possible to shorten time, costs, and efforts required for each sampling (both in the field and in the lab), thereby increasing the number of stations that can be sampled seasonally by ARPA or local Agencies. In addition, they would re-establish a bridge with the past, allowing a better use of the long-time series of IBE data, that would share the same taxonomic detail.

ACKNOWLEDGMENTS

We are grateful to the eight colleagues who in various form provided suggestions and comments to this paper.

REFERENCES

- Adriaenssens V, Simons F, Nguyen LT, Goddeeris B, Goethals P, De Pauw N, 2004. Potential of bio-indication of chironomid communities for assessment of running water quality in Flanders (Belgium). *Belg. J. Zool.* 134:31-40.
- Allan JD, Castillo MM, 2007. *Stream ecology: structure and function of running waters*. Springer, Dordrecht: 436 pp.
- APAT, 2007. [Metodi biologici per le acque. Parte I]. [Book in Italian]. APAT, Rome. Available from: www.isprambiente.gov.it/it/pubblicazioni/manuali-e-linee-guida/metodi-biologici-per-le-acque-parte-i
- AQEM Consortium, 2002. Manual for the application of the AQEM system. A comprehensive method to assess European streams using benthic macroinvertebrates, developed for the purpose of the Water Framework Directive. Ver. 1.0: 202 pp. Available from: www.aqem.de
- Barbour MT, Gerritsen J, Snyder BD, Stribling JB, 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates and fish. EPA 841-B-99-002. U.S. Environmental Protection Agency.
- Bennett C, Owen R, Birk S, Buffagni A, Erba S, Mengin N, Murray-Bligh J, Ofenbock G, Pardo I, van de Bund W, Wagner F, Wasson JG, 2011. Bringing European river quality into line: an exercise to intercalibrate macro-invertebrate classification methods. *Hydrobiologia* 667:31-48.
- Birk S, Hering D, 2006. Direct comparison of assessment methods using benthic macroinvertebrates: a contribution to the EU Water Framework Directive intercalibration exercise. *Hydrobiologia* 566:401-415.
- Birk S, Bonne W, Borja A, Brucet S, Courrat A, Poikane S, Solimini A, van de Bund W, Zampoukas N, Hering D, 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.
- Bonada N, Prat N, Resh VH, Statzner B, 2006. Developments in aquatic insect biomonitoring: a comparative analysis of recent approaches. *Annu. Rev. Entomol.* 51:495-523.
- Borja A, 2005. The European Water Framework Directive: a challenge for nearshore, coastal and continental shelf research. *Cont. Shelf. Res.* 25:1768-1783.
- Buffagni A, Erba S, 2007. [Macroinvertebrati acquatici e Direttiva 2000/60/EC (W.F.D.) – Parte A. Metodo di campionamento per i fiumi guadabili]. [Article in Italian]. IRSA-CNR Notiziario dei metodi analitici 1:1-27.
- Buffagni A, Erba S, Belfiore C, Hering D, Moog O, 2001. A Europe-wide system for assessing the quality of rivers using macroinvertebrates: the AQEM project^o and its importance for southern Europe (with special emphasis on Italy). *J. Limnol.* 60:39-48.
- Buffagni A, Erba S, Pagnotta R, 2008. [Definizione dello stato ecologico dei fiumi sulla base dei macroinvertebrati bentonici per la 2000/60/EC (WFD): il sistema di classificazione MacroOper]. [Article in Italian]. IRSA-CNR Notiziario dei metodi analitici 1:25-41.
- Buffagni A, Furse M, 2006. Intercalibration and comparison-major results and conclusions from the STAR project. *Hydrobiologia* 566:357-364.
- Buss DF, Carlisle DM, Chon TS, Culp J, Harding JS, Keizer-Vlek HE, Robinson WA, Strachan S, Thirion C, Hughes RM, 2015. Stream biomonitoring using macroinvertebrates around the globe: a comparison of large-scale programs. *Environ. Monit. Assess.* 187:1-21.
- Cairns J Jr, Pratt JR, 1993. A history of biological monitoring using benthic macroinvertebrates, p. 10-27. In: M.D. Rosenberg and V.H. Resh (eds.), *Freshwater biomonitoring and benthic macroinvertebrates*. Chapman & Hall, London.
- Cao Y, Bark AW, Williams WP, 1996. Measuring the responses of macroinvertebrate communities to water pollution: a comparison of multivariate approaches, biotic and diversity indices. *Hydrobiologia* 341:1-19.
- Casellato S, Salin ML, Lorenzetti T, 1980. Applicability of Verneaux and Tuffery's Biotic Index to a study of the macrobenthos in the River Brenta (Northern Italy). *Boll. Zool.* 47:53-61.
- Chandler JR, 1970. A biological approach to water quality management. *Water Pollut. Control* 69:415-422.
- Clarke RT, 2013. Estimating confidence of European WFD ecological status class and WISER Bioassessment Uncertainty Guidance Software (WISERBUGS). *Hydrobiologia* 704:39-56.
- Collins AL, Anthony SG, 2008. Assessing the likelihood of catchments across England and Wales meeting 'good ecological status' due to sediment contributions from agricultural sources. *Environ. Sci. Policy* 11:163-170.
- Davy-Bowker J, Clarke R, Corbin T, Vincent H, Pretty J, Hawczak A, Blackburn J, Murphy J, Jones I, 2008. River Invertebrate Classification Tool. WFD72C SNIFFER.
- Friberg N, Baatrup-Pedersen A, Pedersen ML, Skriver J, 2005. The new Danish stream monitoring programme (NOVANA) - Preparing monitoring activities for the water framework directive era. *Environm. Monit. Assess.* 111:27-42.
- Friberg N, 2014. Impacts and indicators of change in lotic ecosystems. *WIREs Water* 1:513-531.
- Furse M, Hering D, Moog O, Verdonschot PFM, Johnson RK, Brabec K, Gritzalis K, Buffagni A, Pinto P, Friberg N, Murray-Bligh J, Kokes J, Alber R, Usseglio-Polatera P, Haase P, Sweeting R, Bis B, Szoszkiewicz K, Soszka H, Springe G, Sporka F, Krno IJ, 2006. The STAR project: context, objectives and approaches. *Hydrobiologia* 566:3-29.
- Gabriels W, Lock K, De Pauw N, Goethals PL, 2010. Multimetric Macroinvertebrate Index Flanders (MMIF) for biological assessment of rivers and lakes in Flanders (Belgium). *Limnologia* 40:199-207.
- Ghetti PF, Bonazzi G, 1977. A comparison between various criteria for the interpretation of biological data in the analysis of the quality of running waters. *Water Res.* 11:819-831.
- Ghetti PF, 1986. [I macroinvertebrati nell'analisi di qualità dei corsi d'acqua. Manuale di applicazione Indice Biotico: E.B.I. modificato]. [Book in Italian]. Provincia Autonoma di Trento, Italy.
- Ghetti PF, 1997. [Manuale di applicazione. Indice Biotico Esteso (I.B.E.). I macroinvertebrati nel controllo della qualità degli ambienti di acque correnti]. [Book in Italian]. Provincia Autonoma di Trento, Italy.

- Guareschi S, Laini A, Sánchez-Montoya MM, 2017. How do low-abundance taxa affect river biomonitoring? Exploring the response of different macroinvertebrate-based indices. *J. Limnol.* 76(Suppl.1):9-20. [Epub 7 Jun 2016].
- Heiskanen AS, van de Bund W, Cardoso AC, Nöges P, 2004. Towards good ecological status of surface waters in Europe - Interpretation and harmonisation of the concept. *Water Sci. Technol.* 49:169-177.
- Hellawell JM, 1986. *Biological indicators of freshwater pollution and environmental management.* Elsevier.
- Hering D, Buffagni A, Moog O, Sandin L, Sommerhäuser M, Stubauer I, Feld C, Johnson RK, Pinto P, Skoulikidis N, Verdonschot PFM, Zahrádková S, 2003. The development of a system to assess the ecological quality of streams based on macroinvertebrates - design of the sampling programme within the AQEM project. *Int. Rev. Hydrobiol.* 88:345-361.
- Hering D, Moog O, Sandin L, Verdonschot PFM, 2004. Overview and application of the AQEM assessment system. *Hydrobiologia* 516:1-20.
- Hering D, Feld CK, Moog O, Ofenböck T, 2006. Cook book for the development of a Multimetric Index for biological condition of aquatic ecosystems: experiences from the European AQEM and STAR projects and related initiatives. *Hydrobiologia* 566:311-324.
- Hering D, Borja A, Carstensen J, Carvalho L, Elliott M, Feld CK, Heiskanen AS, Johnson RK, Moe J, Pont D, Solheim AL, van de Bund W, 2010. The European Water Framework Directive at the age of 10: a critical review of the achievements with recommendations for the future. *Sci. Total Environ.* 408:4007-4019.
- ISPRA, 2014. [Linee Guida per la valutazione della componente macrobentonica fluviale ai sensi del DM 260/2010]. [Book in Italian]. ISPRA - Manuali e Linee Guida 107/2014, pp. 99. Available from: www.isprambiente.gov.it/it/pubblicazioni/manuali-e-linee-guida/linee-guida-per-la-valutazione-della-componente-macrobentonica-fluviale-ai-sensi-del-dm-260-2010
- Jáimez-Cuéllar P, Vivas S, Bonada N, Robles S, Mellado A, Álvarez M, Avilés J, Casas J, Ortega M, Pardo I, Prat N, Rieradevall M, Sáinz-Cantero CE, Sánchez-Ortega A, Suárez ML, Toro M, Vidal-Abarca MR, Zamora-Muñoz C, Alba-Tercedor J, 2002. [Protocolo GUADALMED (PRECE)]. [Article in Spanish].
- Kolkwitz R, Marsson M, 1908. [Ökologie der pflanzlichen Saprobien]. [Article in German]. *Ber. Dtsch. Bot. Ges.* 26:505-519.
- Laini A, Vorti A, Bolpagni R, Viaroli P, 2014. Small-scale variability of benthic macroinvertebrates distribution and its effects on biological monitoring. *Ann. Limnol. Int. J. Lim.* 50:211-216.
- Logan P, Furse M, 2002. Preparing for the European Water Framework Directive - making the links between habitat and aquatic biota. *Aquat. Conserv.* 12:425-437.
- Mancini D, Zanut E, Massarutto S, Piazza G, Tomasella M, Bertoli M, Pizzul E, 2010. [Valutazione biologica della qualità delle acque nel bacino del Fiume Stella (Friuli Venezia Giulia)]. [Article in Italian]. *Biol. Ambient.* 24:49-58.
- Metcalf JL, 1989. Biological water quality assessment of running waters based on macroinvertebrate communities: history and present status in Europe. *Environ. Pollut.* 60:101-139.
- Mondy CP, Villeneuve B, Archaimbault V, Usseglio-Polatera P, 2012. A new macroinvertebrate-based multimetric index (I2M2) to evaluate ecological quality of French wadeable streams fulfilling the W.F.D. demands: a taxonomical and trait approach. *Ecol. Indic.* 18:452-467.
- Moog O, Schmidt-Kloiber A, Ofenböck T, Gerritsen J, 2004. Does the ecoregion approach support the typological demands of the EU 'Water Framework Directive'? *Hydrobiologia* 516: 21-33.
- Munné A, Prat N, 2009. Use of macroinvertebrate-based multimetric indices for water quality evaluation in Spanish Mediterranean rivers: an intercalibration approach with the IBMWP index. *Hydrobiologia* 628:203-225.
- Nöges P, van de Bund W, Cardoso AC, Solimini AG, Heiskanen AS, 2009. Assessment of the ecological status of European surface waters: a work in progress. *Hydrobiologia* 633:197-211.
- Pollard P, Huxham M, 1998. The European Water Framework Directive: a new era in the management of aquatic ecosystem health? *Aquat. Conserv.* 8:773-792.
- Reyjol Y, Argillier C, Bonne W, Borja A, Buijse AD, Cardoso AC, Daufresne M, Kernan M, Ferreira MT, Poikane S, Prat N, Solheim AL, Stroffek S, Usseglio-Polatera P, Villeneuve B, van de Bund W, 2014. Assessing the ecological status in the context of the European Water Framework Directive: Where do we go now? *Sci. Total Environ.* 497:332-344.
- Rossaro B, Boggero A, Crozet BL, Free G, Lencioni V, Marziali L, 2011. A comparison of different biotic indices based on benthic macro-invertebrates in Italian lakes. *J. Limnol.* 70:109-122.
- Szivák I, Csabai Z, 2012. Are there any differences between taxa groups having distinct ecological traits based on their responses to environmental factors? *Aquat. Insects* 34:173-187.
- Verdonschot PFM, Nijboer RC, 2004. Testing the European stream typology of the Water Framework Directive for macroinvertebrates. *Hydrobiologia* 516:35-54.
- Verdonschot PFM, Moog O, 2006. Tools for assessing European streams with macroinvertebrates: major results and conclusions from the STAR project. *Hydrobiologia* 566:299-309.
- Von der Ohe PC, Prüß A, Schäfer RB, Liess M, de Deckere E, Brack W, 2007. Water quality indices across Europe - A comparison of the good ecological status of five river basins. *J. Environ. Monitor.* 9:970-978.
- Woodiwiss FS, 1964. The biological system of stream classification used by the Trent River Board. *Chem. Ind.* 11:443-447.