PRIMARY RESEARCH PAPER

A multimetric index based on benthic macroinvertebrates for assessing the ecological status of streams and rivers in central and southeast highlands of Ethiopia

Aschalew Lakew · Otto Moog

Abstract This study presents the development of a multimetric index using benthic macroinvertebrates (BMI) to assess the ecological health of highland rivers in Ethiopia. BMI were collected from 22 reference and 82 impaired sites determined based on hydro-morphological, land use, and physical and chemical criteria. Of 75 potential metrics tested to integrate the multimetric index, only nine core metrics were selected based on their abilities to distinguish reference and impaired sites, strength of correlation with pertinent environmental parameters, and their independence from other metrics. The metrics retained in the multimetric index were total number of taxa, EPT-BH > 1sp (Ephemeroptera, Plecoptera, and Trichoptera taxa where Baetidae and Hydropsychidae taxa are considered if they consist more than one taxon), % Oligochaeta and Red Chironomidae, % COPTE (Coleoptera, Odonata, Plecoptera, Trichoptera, and Ephemeroptera), % EPT-BCH (EPT without Baetidae, Caenidae, and Hydropsychidae), ASPT-SASS (Average South African Scoring System Per Taxa), FBI (Family Biotic Index), % shredders, and % collector gathering. The final index derived from these metrics was divided into five river quality class (high, good, moderate, poor, and bad). A validation procedure showed that the index is stable along different hydrological conditions and sensitive to the current range of anthropogenic disturbances in Ethiopian highland rivers.

Keywords Biomonitoring · Benthic macroinvertebrates · Multimetric index · Highland streams · Ethiopia

A. Lakew (✉)
National Fishery and Aquatic Life Research Centre,
Ethiopian Institute of Agricultural Research (EIAR),
P. O. Box 64, Sebeta, Ethiopia
e-mail: aschalewlh@yahoo.com

O. Moog
Department of Water, Atmosphere and Environment,
Institute of Hydrobiology and Aquatic Ecosystem
Management, University of Natural Resources and Life
Sciences, Max-Emanuel-Strasse 17, 1180 Vienna, Austria
e-mail: otto.moog@boku.ac.at

Introduction

Ethiopia is gifted with many streams and rivers that comprise diverse aquatic ecosystems of great scientific interest and economic importance. The rapid population growth, expansion of urban and suburban areas, industrialization, land use change, and removal of riparian vegetation have resulted serious ecological problems on surface water resources (Zinabu and Elias, 1989; Aschalew, 2014). To put a sustainable water resource management into practice, the first step is to understand the relationship between human-induced disturbance and their effect on aquatic

resources, followed by the development of appropriate assessment tools to undertake mitigation measures.

Biological monitoring methods have been developed largely to assess the health and level of stresses on an aquatic ecosystem (Birk et al., 2012). Unlike chemical monitoring, the advantage of biomonitoring lies in the ability of biological communities to reflect not only the quality of the water, but also the overall ecological status of the ecosystem (Rosenberg and Resh, 1993). Biological integrity in aquatic ecosystem represents the ability of the systems to support and maintain a balanced and integrated community of organisms comparable to the natural habitat (Karr and Chu, 1999).

Benthic macroinvertebrates are often the taxa group of choice for biomonitoring in streams and rivers (Moog, 1988; Rosenberg and Resh, 1993). They are ubiquitous and good indicators of several anthropogenic pressures such as water pollution (Armitage et al., 1983) and geomorphological alterations (Negishi et al., 2002). They consist of thousands of species that originate from different systematic categories with different environmental needs—thus they are differentially sensitive to pollutants of various types.

The conceptual basis for biomonitoring methods began in Europe after the saprobic system was developed by Kolkwitz & Marsson (1902). In the 1980s, most biotic indices were developed based on score systems (Armitage et al., 1983; Hilsenhoff, 1988). More recently, integrated assessment methods—multimetric approach (e.g., Karr, 1981; Barbour et al., 1995; Ofenboeck et al., 2004; Gabriels et al., 2010) and multivariate approach (e.g., Reynoldson et al., 1995; Kokes et al., 2006) were developed. Both multimetric and multivariate methods are based on the reference conditions approach and compare the deviation of an impaired sites from pristine sites. The use of multivariate approach is, however, still emergent in tropical areas but it constitutes the basis for the development of multimetric approach as it is used in temperate regions.

A multimetric index (MMI), first introduced by Karr (1981) using fish assemblages, has now gained increasing attention worldwide for its ability to include complementary information from a broad spectrum of stressors (Barbour et al., 1996; Hering et al., 2006). MMI integrates different biological measures (richness, composition, tolerance, and trophic status) into a single value that can potentially reflect impact of multiple anthropogenic pressures (Barbour et al., 1999; Hering et al. 2006). Moreover,

MMI are increasingly used for the purpose of conservation and restoration actions, as they provide relevant information for regulatory agencies and decision makers (Barbour et al., 1995; Karr and Chu, 1999). For example, multimetric approach has become a popular method for routine biomonitoring program in European member state countries (Hering et al., 2004; Ofenboeck et al., 2004). Recently, the multimetric approach is very occasionally applied in tropical countries (e.g. Ferreira et al., 2011; Bolivia: Moya et al., 2011; Couceiro et al., 2012; Vietnam: Nguyen et al., 2014; Brazil: Melo et al., 2015). An important advantage of MMI is that they are flexible and can easily be adjusted by adding or removing metrics or refining the index threshold values. However, MMI should be applied cautiously because of the differences in reference conditions, anthropogenic pressures, and regional species assemblages.

In Ethiopia, the use of biomonitoring in general and application of a multimetric approach in particular is still at its infancy. The evaluation of water quality is primarily focused on physicochemical data and optionally includes fecal coliforms as a biological component (Alemayehu, 2001). In recent years, however, research institutions and universities in collaboration with foreign scientists have tested biomonitoring systems using benthic macroinvertebrate communities to assess ecological condition of streams, rivers, and natural wetlands (Getachew et al., 2012; Mereta et al., 2013; Aschalew, 2014). The results showed that benthic macroinvertebrates are reliable indicators for biomonitoring purpose in the region. Thus, in this paper, we aimed to develop a MMI for assessing the ecological status of streams and rivers in a broader scale and under multiple stressors using benthic macroinvertebrates. The geographical focus comprises central and south east highlands of Ethiopia, one of the most diverse and threatened ecoregion in the country. We believe that this study provides a useful tool for the assessment of rivers in the highlands and also a methodological guideline for wider application of biomonitoring network in the country.

Materials and methods

Study area

The present study was conducted in the upper section of Awash, Wabe-shebelle, Genale, and Rift Valley

basins, lying between 6°57′N and 9°05′N latitude and 38°07′E and 40°06′E longitude (Fig. 1). Of 104 sampling sites, 95 sites were located in the Ethiopian montane grassland–woodlands ecoregion with altitude ranging from 1,900 to 2,500 m a.s.l, while the remaining 9 were distributed between 2,600 and 3,200 m a.s.l. in afro-alpine ecoregion of Bale highlands. The study area shows bimodal rainfall with a short rain from February to April and the main rainfall from June to September (National Meteorological Agency (NMA), 2012). The natural vegetation cover is restricted to state-protected forest areas in upper Awash (e.g., Chilimo, Suba forest) and central rift valley (e.g., Wondogent forest) basins. In Genale and Wabe-shebelle basins, the high mountains are dominated by small bushy shrubs and Erica forest, while the escarpments are covered by Junipers, *Hyginia* sp., mixed forest, and grass lands (e.g., Bale National Park and Adaba forest). In rural areas, the unprotected river catchments are highly populated and most of them are
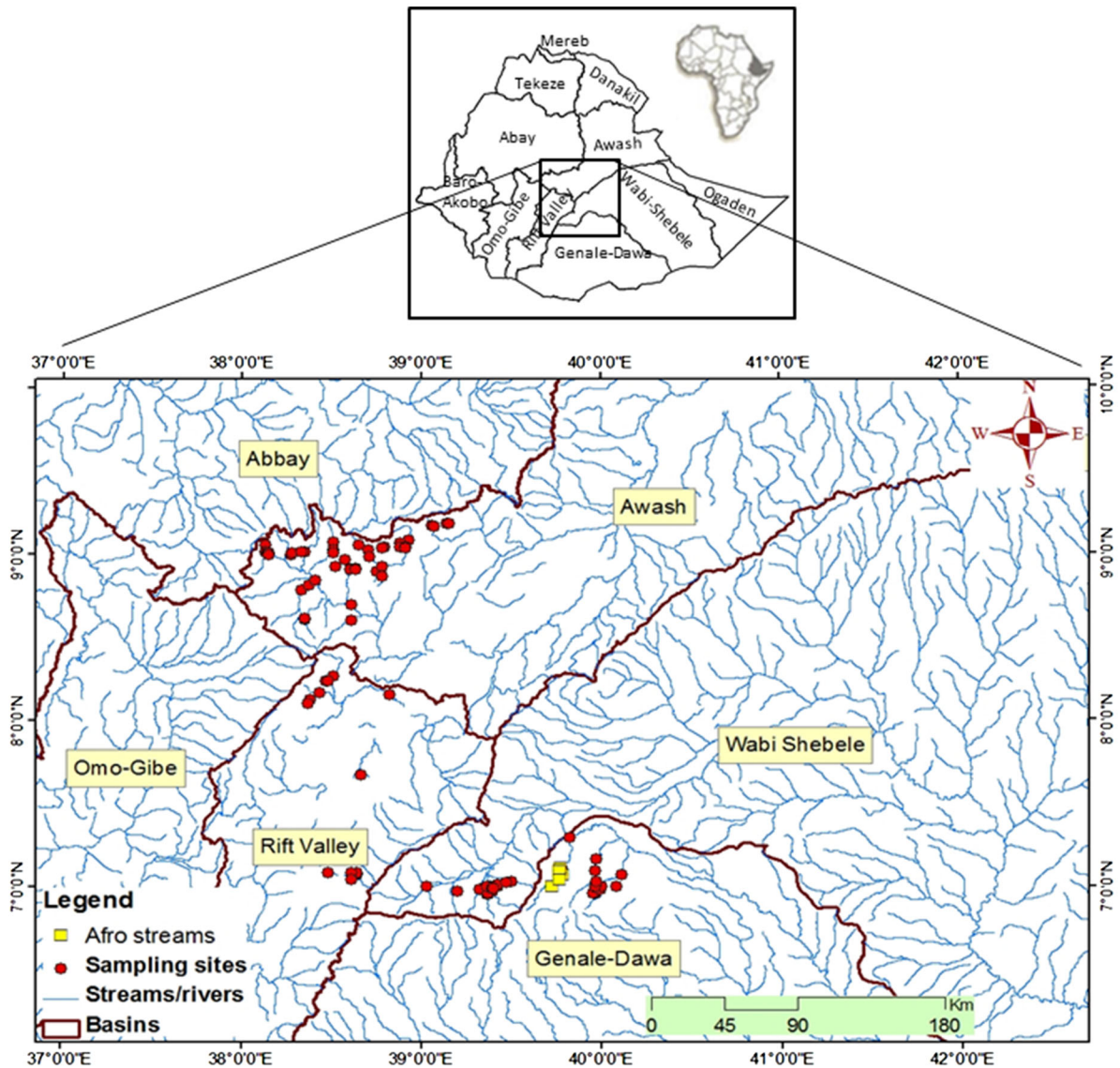

Fig. 1 Spatial distribution of study sites in four drainage basins of Ethiopia

subsistence farmers. Moreover, the demand on natural products and land for farming is huge, which could be the predominant reason for the widespread loss of natural vegetation in this ecoregion. In urban areas, especially in and around the capital Addis Ababa, streams and rivers serve as major domestic and industrial waste dumping sites.

Sampling sites were distributed strategically in the national park, protected forest areas, rural-agricultural areas and urban-industrial sites to include different stress gradients. The major threats in rural areas include removal of riparian vegetation, nutrient loading from farmlands and washing activities, siltation of stream bed due to sediment from catchment, sand excavation, and un-budgeted water withdrawal for irrigation. In urban areas, the major pollutants originate from diffuse and punctual loads of untreated domestic and industrial wastes.

Reference site selection

The development of a MMI requires establishing stress gradient into reference and impaired sites (Hering et al., 2006). Many authors defined reference condition as the condition that is representative of a group of 'least impaired' sites characterized by selected physical, chemical, and biological characteristics (Hughes, 1995; Barbour et al., 1996; Hering et al., 2003). We established reference sites based on 20 thoughtfully selected 'a priori' criteria during our extensive field sampling (Table 1). The 'a priori' criteria were representative for Ethiopian highlands and reflect some general aspects of naturalness and address wide range of human disturbances. To be considered as reference site, at least 16 criteria are required including the four compulsory 'a priori' criteria (* in Table 1) in the sense of the 'best attainable condition' (Stoddard et al., 2006).

Sample collection

Water quality parameters including temperature, pH, dissolved oxygen, and conductivity were measured using a portable WTW multi-parameter probe. From each site, two liters of water was collected and stored in ice box until return to the Laboratory of JIJE labo PLC and National Fishery and Aquatic Life Research center. Total phosphorus (TP) and biochemical

oxygen demand (BOD_5) were measured following the standard methodology described in APHA (1997). Macroinvertebrates were collected using standard square hand net with frame width of 25 cm and mesh size 500 μm following multi-habitat approach (Barbour et al., 1999; Moog & Sharma 2005). A composite sample was passed through a set of sieves to separate different size classes of macroinvertebrates and to wash off formalin using tap water. Benthic macroinvertebrates were identified to the lowest possible taxa using identification keys and the help of taxonomic experts in Vienna, Austria.

Selection of candidate metrics

In order to assess the ability of the potential metrics to discriminate between reference and impaired sites, two methodologies have been applied. (a) Mann–Whitney U test (P value at 0.05); (b) a graphical analysis using box-and-whisker plots combined with a sensitivity score. A five-class sensitivity score was assigned following Barbour et al. (1996) based on the degree of interquartile overlap in the Box-and-Whisker plots: a sensitivity score 3 was given if no overlap occurred in interquartile range; a sensitivity score 2 if the medians were outside the interquartile range overlap; a sensitivity score 1 if there was moderate overlap of interquartile range but one median was outside the interquartile range overlap; a sensitivity score 0a if one range was completely overlapping the other interquartile range but one median was outside the interquartile range overlap; and a sensitivity score 0b if both medians were inside interquartile range overlap.

Selection of core metrics and removal of redundant metrics

Selection of core metrics from the candidates was based on two applications: sensitivity score and discrimination efficiency (DE). Metrics with sensitivity score 3 ($P < 0.05$ in Mann–Whitney U test) were selected as core metrics. In the absence of sensitivity score 3 in a given metric category, a metric with sensitivity score 2 with higher DE (>50%) was selected to ensure the inclusion of primary metric categories (richness, composition, sensitivity, and feeding guild). DE was calculated as the percentage

Table 1 Criteria applied for selection of reference sites in Ethiopian highland streams

Category	Group	Criteria	Reference
Hydro-morphological	River morphology	1. Natural channel structure typical to the region (riffle, run, pool …)	Hughes (1995)
			Barbour et al. (1996)
		2. No migration barrier	Hering et al. (2003)
	Habitat composition	3. Representative diversity of substrate composition appropriate to the region	Hughes (1995)
			Barbour et al. (1996)
		4. Spawning habitats for the natural fish population	Moog & Sharma (2005)
	Hydrological condition	5. No alteration of the natural hydrograph and discharge regime	Barbour et al. (1996)
			Hering et al. (2003)
		6. No water extraction, damming, or diversion upstream	Nijboer et al. (2004)
Physicochemical	Point source pollution	7. No point source pollution and Eutrophication*	Hering et al. (2003)
		8. Natural color and odor*	Nijboer et al. (2004)
		9. Total phosphorus concentration < 0.035 mg/l	Moog & Sharma (2005)
		10. Oxygen concentration >8 mg/l and saturation range between 95–110%	Moog & Sharma (2005)
		11. $BOD_5 < 2$ mg/l	Moog (1988)
		12. No known cattle watering points upstream*	This study
		13. Minimal washing and bathing activities	This study
	Non-point source of pollution	14. Absence of known or expected diffusion input	Hering et al. (2003)
Land use		15. Crop farming <20% of the catchment area	Hering et al. (2003)
		16. Natural land use (forest, bushes, grassland) >60% coverage	This study
		17. Rural resident <0.5% in the catchment area	This study
		18. No sand or gravel excavation upstream*	Nijboer et al. (2004)
Biological	Riparian vegetation and wildlife	19. Having adjacent natural vegetation appropriate to the type and geographic location of the stream	Barbour et al. (1996)
		20. Presence of representative wild animals and drive some support from aquatic ecosystems	Barbour et al. (1996)

* Compulsory criteria

of impaired sites that score below 25th percentile of the reference sites for decreasing metrics with disturbance and above 75th percentile for increasing metrics with disturbance (Ofenboeck et al., 2004).

Metrics were considered as redundant if the Spearman rank correlation was higher than 0.8. The final core metrics were selected based on the following two major criteria: (a) a metric with an inter-correlation coefficient less than 0.8 (Hering et al., 2006) and (b) for highly inter-correlated metrics ($r > 0.8$); DE, ecological importance, wider applicability, and degree of correlation with environmental parameters were considered.

Metrics transformation, aggregation, and threshold values

The final core metrics were standardized between 0 and 1 (Hering et al., 2006). The 95th percentile of the data distribution of a metric was used as upper anchor to eliminate extreme outliers for metrics that decrease with increasing degradation (e.g., taxa richness, EPT taxa). Values close to the 95th percentile receive higher scores, while values having a greater deviation from this percentile receive lower scores. On the other hand, the 5th percentile of the data distribution of a metric was used as upper anchor for those metrics that

increase in response to degradation (e.g., % Red Chironomidae, Hilsenhoff Family Biotic Index). Transformation was performed using the following formula:

For metrics decrease with increasing disturbance:

$$\text{Value} = \frac{\text{Metric result} - \text{5th percentile of impaired sites}}{\text{95th percentile of reference sites} - \text{5th percentile of impaired sites}}.$$

For metrics increase with increasing impairment:

$$\text{Value} = \frac{\text{Metric result} - \text{95th percentile of impaired sites}}{\text{5th percentile of reference sites} - \text{95th percentile of impaired sites}}.$$

Whenever standardized value exceeds 1 or become negative, it was scaled to 1 and 0, respectively. A value close to 1 represents high ecological status, and value close to 0 represents bad ecological status. Such transformation was commonly applied to integrate different metrics into MMI and to simplify assessment of ecological status (Ofenboeck et al., 2004).

The numerical range of MMI (0–1) was divided into five ecological quality classes (high, good, moderate, poor, and bad) based on a literature survey and discussion of the data in a stepwise procedure.

Step 1 defines the lowest threshold value of the index for reference sites (high ecological quality class) through application of 25th percentile approach (Munne and Prat, 2009). This approach was applied (1) to be sure that only truly unimpaired sites serve as a reference and (2) to avoid that too many impaired sites with comparably high indexes will shift into reference sites.
Step 2 defines the upper threshold value for heavily impacted (bad ecological quality class) sites. The aim of this step was to clearly characterize the bad sites based on the observed index values following Moog & Sharma (2005) chemical water quality class approach (total phosphorus, dissolved oxygen, and biological oxygen demand).
Step 3 defines the application of equidistance approach (Hering et al., 2006) to define river quality

classes between good–moderate, moderate–poor, and poor–bad classes.

The sensitivity and robustness of a MMI was evaluated by assessing discrimination efficiency between reference and impaired sites. The percentage of correctly classified reference sites in high ecological quality class was calculated for performance evaluation. Moreover, the strength of relationship with relevant environmental variables and index stability across different hydrological conditions was evaluated.

Results

Metric selection

Of the 75 potential metrics initially considered as appropriate, only 39 metrics showed a significant difference between reference and impaired sites according to Mann–Whitney U test ($P < 0.05$). Most of these candidate metrics were widely recognized as being sensitive to various anthropogenic impacts. However, the sensitivity test from box-and-whisker plot followed by DE identified 18 core metrics (seven richness, five composition, three sensitivity, and three feeding guild-measures) (Table 2).

The redundancy test showed that four metrics namely % COPTE, % EPT-BCH, Family Biotic Index, and % shredders were not highly correlated with their respective metric category (Spearman rank correlation, $r < 0.8$). Since these metrics are representative for most sites and fulfilled all metric selection procedures, they were selected with the highest priority for the MMI (Fig. 2). When two metrics are

Table 2 Core metrics selected for the development of multimetric index, predicted response to degradation, sensitivity score, and discrimination efficiency

Metric type	Definition	Expected response to degradation	U test	P value	SS	DE
Number of family	Total number of family	Decrease	298	.001	2	72
Total taxa #	Total number of taxa	Decrease	233	.001	3	81
EPT	Total number of Ephemeroptera, Plecoptera, and Trichoptera taxa	Decrease	42	.001	3	98
EPT-BH	Total number EPT taxa without Baetidae and Hydropsychidae taxa	Decrease	99	.001	3	98
EPT-BH > 1sp	Total number of EPT taxa where BH taxa were considered if they consist more than one taxon	Decrease	63	.001	3	100
COPTE taxa	Total number of Coleoptera, Odonata, and EPT taxa	Decrease	123	.001	3	95
Margalef's	Margalef's diversity index	Decrease	312	.001	2	76
%Red Chironomidae	Percentage composition of Red Chironomidae	Increase	9	.001	3	96
% COPTE	Percentage composition of COPTE taxa	Decrease	211	.001	3	91
% non-insect & Chironomidae	Percentage composition of non-insect and Chironomidae individuals	Increase	78	.001	3	92
% Oligochaeta & Red Chironomidae	Percentage composition of non-insect and Red Chironomidae individuals	Increase	17	.001	3	98
% EPT-BHC	Percentage composition of EPT without BH and Caenidae	Decrease	125	.001	3	96
FBI	Hilsenhoff Family Biotic Index	Increase	159	.001	3	90
SASS	South African Scoring System	Decrease	122	.001	3	88
ASPT	Average score per taxa	Decrease	56	.001	3	99
% shredders	Percentage composition of shredder individuals	Decrease	421	.001	2	82
% Filtering collectors	Percentage composition of filtering collectors	Decrease	337	.001	2	68
% collector gathering	Percentage composition of gathering collector individuals	Increase	218	.001	2	71

SS sensitivity score (Barbour et al., 1996), *DE* discrimination efficiency (Ofenboeck et al., 2004)

strongly correlated, it is not justifiable to automatically disregard one metric if it is known that the two metrics represent two different aspects of the benthic invertebrate community. With this concept, total taxa and number of EPT-BH > 1sp from richness-diversity category, % Oligochaeta and Red Chironomidae from composition, ASPT from sensitivity, and % collector gathering from feeding group were retained based on their high DE, wide ecological representation, and application.

Environmental variables and final core metrics

The physical, chemical, and morphological characteristics were studied because in reference conditions they define the ambient quality in which aquatic biotas have developed, and in impaired conditions they correspond to the environmental pressures on aquatic species. Most of these parameters were significantly different between reference and impaired sites (Table 3). The final core metrics showed strong correlation with most environmental variables (Table 4). The correlation of physicochemical and land use variables were stronger than those variables related to substrate type, altitude, and catchment size. Dissolved oxygen showed strong positive correlation with EPT-BH > 1sp and ASPT ($r > 0.65$, $P < 0.05$), and strong negative correlation with % Oligochaeta and Red Chironomidae ($r > 0.62$, $P < 0.05$) which may indicate the importance of the metrics in response to oxygen depleting factors. Phosphorus concentration showed significantly high correlation with all core metrics in a predictable way ($r > 0.65$, $P < 0.05$). Furthermore, strong correlations of core metrics with % forest ($r > 0.5$, $P < 0.05$) and % urban ($r > 0.5$,
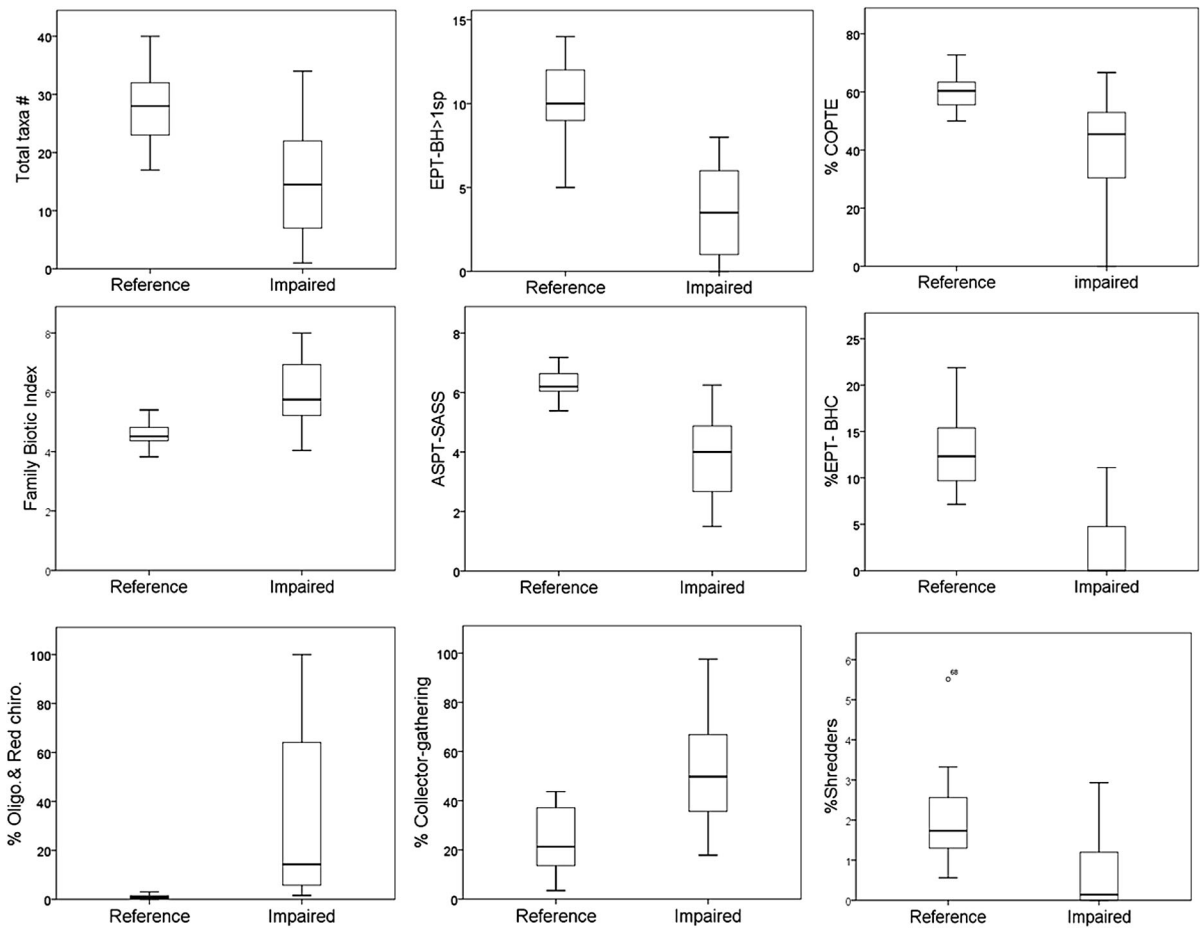
Fig. 2 Box-and-whisker plots of nine core metrics used to integrate multimetric index. *Bar line* within the *box* represents median number, *boxes* represent first- and third-quartile ranges (25th and 75th percentiles) and *range bars* show maximum and minimum of non-outlier numbers

$P < 0.05$) showed the efficiency of metrics in reflecting land use changes. With the exception of % fine sand, all substrate types showed either insignificant ($P > 0.05$) or weak correlation ($r < 0.5$) with core metrics which may show the sensitivity of the metrics in response to anthropogenic impacts than habitat variability.

Threshold values of ecological quality class

The range of MMI was divided into five ecological classes as described in the methodology. Following the 25th percentile approach of Munne and Prat (2009) to distinguish high–good ecological quality classes a value of 0.759 was obtained and we set at a value greater than 0.75 as a threshold. Since total phosphorus, dissolved oxygen, and BOD_5 were identified as important environmental variables, the 75th percentile of bad chemical water quality class was averaged to 0.149. Therefore, the upper threshold value for bad river quality class was set at a value less or equal to 0.15. The application of an equidistance approach resulted in values of 0.35 and 0.55 between poor–moderate and moderate–good river quality classes, respectively. To consider the effect of a different ecoregion in the dimension of a threshold value, 0.72 was set as high–good threshold for afro-alpine streams based on visual observation at the data distribution and physicochemical condition of the streams.

Discussion

The primary step in developing a MMI is the classification of sites into reference and impaired.

Table 3 Mean, standard deviation, and range (in bracket) values of environmental parameters in reference and impaired sites

Environmental parameters	'a priori' reference sites $n = 22$	'a priori' impaired sites $n = 82$	P value
Conductivity (µS/cm)	146.1 ± 66.5 (45.4–265)	464 ± 757 (60–6,500)	.053
DO (mg/l)	8.2 ± 0.7 (6.8–9.4)	6.5 ± 2.3 (0–11.7)	.001
DO %	107.8 ± 6.5 (94.6–120)	89 ± 32 (0–149)	.009
BOD$_5$ (mg/l)	1.9 ± 0.5 (1.4–2.8)	61 ± 190 (0.6–1,059)	.035
TP(mg/l)	0.033 ± 0.004 (0.03–0.05)	0.56 ± 1.1 (0.01–6.1)	.024
Temp.H2O (°C)	15.6 ± 3.4 (10–22)	19.4 ± 3.3 (11.9–25)	.000
Altitude (m)	2,629 ± 425 (1,904–3,449)	2,227 ± 264 (1,589–3,066)	.001
Catchment size (km^2)	112 ± 190 (8–794)	663 ± 1,094 (11–4,914)	.015
% Riffle	82.7 ± 9.7 (55–95)	73 ± 18 (10–100)	.024
Velocity (m/s)	0.5 ± 0.2 (0.2–0.85)	0.45 ± 0.2 (0.1–1.05)	.119
Substrate type			
% Megalithal	23.6 ± 15.9 (0–50)	8.4 ± 10.6 (0–40)	.001
% Macrolithal	37.8 ± 10.8 (20–55)	25 ± 16.8 (0–75)	.001
% Mesolithal	27 ± 9.5 (15–45)	26.5 ± 16.8 (0–70)	.897
% Microlithal	10.7 ± 11 (0–35)	11.5 ± 11.8 (0–50)	.782
% Akal	1 ± 2 (0–10)	10 ± 23 (0–100)	.061
% Psammal/fine sand	0	18.7 ± 27.3 (0–85)	.002
Land use			
% Forest/bush	82 ± 25 (55–100)	9 ± 18 (0–35)	.000
% Agriculture	17 ± 23 (5–35)	60 ± 35 (0–100)	.003
% Urban	0.4	24 ± 37 (0–100)	.003

This must be based on environmental features like physical, chemical, hydro-morphological, and eco-geographical data but in any case independently of the biological community. Although reference sites are defined as least impaired sites or minimal human disturbance (Hughes, 1995; Hering et al., 2003; Stoddard et al., 2006), there are difficulties in defining this condition in Ethiopia where most sites are easily accessible to humans. We adapted and used 20 'a priori' criteria to classify sites into two impairment levels for developing a MMI and this appeared to be an efficient method. For example, 82% of the 'a priori' reference sites were classified as 'high' and 18% as 'good' river quality classes, and all impaired sites were distributed from 'good' to 'bad' river quality classes. Specific but frequent pressures that occur in Ethiopian highland such as cattle watering point, washing, and rural residents are important criteria to characterize reference sites. The absence of point source of pollution from domestic and industries upstream of potential reference sites was a key criterion which should be fulfilled by all reference sites.

Taxonomic identification level is the other important issue in biomonitoring and has been widely discussed in previous studies (Schmidt-Kloiber and Nijboer, 2004). In Ethiopia, where the taxonomic knowledge of benthic macroinvertebrates is still incipient, defining the taxonomic resolution is particularly important for an index. Many assessment methods prefer family level identification for rapidity and cost benefits (Thorne and Williams, 1997). However, it is often claimed that different species or genera belonging to the same family exhibit different tolerance level (Stubauer and Moog, 2000). In the present study, we combined family level (e.g., Perlidae, Simuliidae) and genus level (e.g., genus of Hydropsychidae and Baetidae), and this appeared to be a good approach to improve the sensitivity of the MMI in Ethiopian highlands. In some cases, identification to the species level was also used.

After a detailed metric selection procedure, nine core metrics were designated for developing the MMI. Most of these metrics were widely used in developing MMI in different climatic regions (Ofenboeck et al., 2004; Stoddard et al., 2006; Baptista et al., 2007;

Table 4 Correlation between metrics selected for index integration and physical–chemical parameters using Spearman correlation ($n = 104$)

Environmental parameters	Metrics								
	Total taxa #	EPT-BH > 1sp	% COPTE	% Oligo. & Red Chiro	% EPT-BHC	FBI	ASPT	% Shredders	% Collector gathering
Conductivity	**−0.579****	**−0.650****	**−0.391****	**0.607****	**−0.597****	**0.595****	**−0.624****	**−0.495****	**0.491****
DO (mg/l)	**0.582****	**0.678****	**0.474****	**−0.623****	**0.518****	**−0.545****	**0.646****	**0.571****	**−0.569****
BOD5 (mg/l)	**−0.631****	**−0.707****	**−0.639****	**0.688****	**−0.512****	**0.653****	**−0.720****	**−0.516****	**0.591****
TP(mg/l)	**−0.787****	**−0.829****	**−0.702****	**0.800****	**−0.656****	**0.782****	**−0.801****	**−0.601****	**0.703****
Catchment	−0.443**	−0.498**	−0.286**	0.509**	−0.395**	0.452**	−0.466**	−0.335**	0.363**
Altitude	0.477**	0.439**	0.228*	−0.355**	0.459**	−0.351**	0.426**	0.279**	−0.314**
% Megalithal	0.379**	0.483**	0.312**	−0.440**	0.384**	−0.402**	0.426**	0.297**	−0.253**
% Macrolithal	0.403**	0.459**	0.253**	−0.422**	0.456**	−0.381**	0.445**	0.271**	−0.263**
% Mesolithal	0.348**	0.299**	0.117	−0.246*	0.217*	−0.216*	0.296**	0.148	−0.221*
% Microlithal	0.304**	0.283**	0.161	−0.307**	0.094	−0.187	0.239*	0.245*	−0.262**
% Akal	−0.044	−0.097	0.139	0.047	−0.227*	0.050	−0.058	0.025	0.004
% Fine sand	**−0.624****	**−0.703****	**−0.623****	**0.675****	**−0.456****	**0.548****	**−0.685****	**−0.515****	**0.566****
% Forest/bush	**0.432****	**0.736****	**0.339***	**−0.687****	**0.635****	**−0.585****	**0.715****	**0.469****	**−0.487****
% Agriculture	−0.364**	−0.386**	−0.133	0.309**	−0.368**	0.234**	−0.222*	−0.263*	0.259*
% Urban	**−0.751****	**−0.839****	**−0.857****	**0.802****	**−0.572****	**0.725****	**−0.794****	**−0.547****	**0.627****

* $P < 0.05$, ** $P < 0.01$

Bold indicates strong and high correlation between environmental variables and core metrics

Ofenboeck et al., 2010). Taxa richness decreased substantially at polluted sites and those species lost were usually pollution-intolerant. This supports the well-established view that sensitive species get reduced when water quality deteriorates (e.g. Barbour et al., 1996; Karr and Chu, 1999). We found elevated taxa richness in moderately impaired sites (good river quality class), which may be due to moderate increase in nutrient availability from nearby agricultural activities at a level to enhance taxa diversity and allow sensitive taxa to survive. The inclusion of this metrics in developing MMI has been reported by previous authors (e.g. Ofenboeck et al., 2004).

EPT richness is a widely used metric type in many biomonitoring systems since the components (Ephemeroptera, Plecoptera, and Trichoptera) are well known for their sensitivity to anthropogenic impacts (Rosenberg and Resh, 1993). However, in the present study, some taxa of Baetidae and Hydropsychidae were frequently collected in large number from range of impaired sites. The hardiness of Baetidae and Hydropsychidae to pollution was also previously reported by different authors (e.g. Harrison and Hynes, 1988; Malicky and Graf, 2012). Thus, we adapted EPT

metrics where Baetidae and Hydropsychidae were considered in EPT metric only if they consist more than a single species (EPT-BH > 1sp) which is highly sensitive and discriminant than EPT total. The diversity of Plecoptera is low in Ethiopian highland streams, as it was also reported by Harrison and Hynes (1988). In the present study, Plecoptera is represented only by a *Neoperla* sp. belonging to Perlidae. Many studies reported that Perlidae are highly sensitive to environmental degradation (e.g. Fore et al., 1996; Graf et al., 2002) but Baptista et al. (2007) reported some genus are partly tolerant to some pressures. Since Perlidae were mostly collected from 'a priori' reference sites characterized by natural riparian vegetation, undisturbed habitat with good physicochemical water quality parameters, it is worthy to consider them as highly sensitive taxa for assessment of Ethiopian highland streams.

Both Chironomidae and Oligochaeta were not sufficiently sensitive to be used separately in the development of the MMI, but when combined, the DE increased significantly. This could be due to the importance of these two groups of organisms in characterizing impacted sites. It was clearly observed

Table 5 Proposed river quality class boundaries for Ethiopian highland streams in comparison with ecological quality ration (EQR) by Hering et al. (2006) for European Water Framework Directives (WFD)

No	Color	Quality class	Proposed threshold	EU-EQRs (Hering et al., 2006)
1	Blue	High	>0.75[a]	≥0.8
2	Green	Good	>0.55 ≤0.75	≥0.6 < 0.8
3	Yellow	Moderate	>0.35 ≤0.55	≥0.4 < 0.6
4	Orange	Poor	>0.15 ≤0.35	≥0.2 < 0.4
5	Red	Bad	≤0.15	≤0.2

[a] Threshold value for afro-alpine streams was set at 0.72
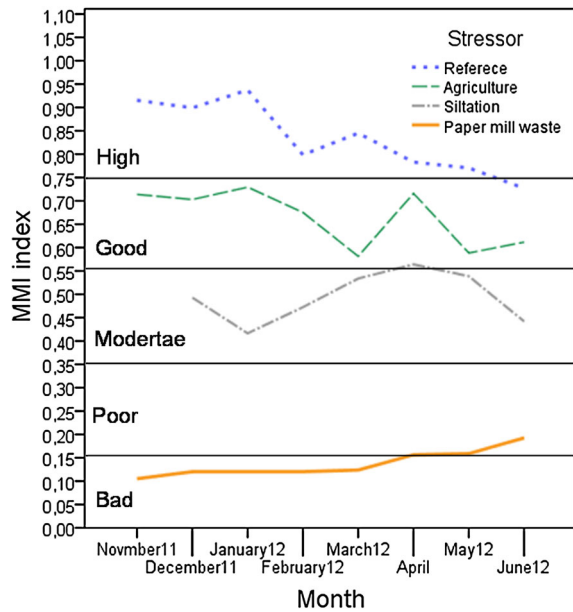
Fig. 3 Response of multimetric index to different hydrological conditions (dry and light rainy months) of year 2011/12 under reference and impaired sites (agricultural, siltation, and paper mill waste)

that the percentage abundance of Oligochaeta and Red Chironomidae changed very markedly, from being rare at reference sites to being common at slightly polluted sites and abundant at heavily polluted sites. Although numerous studies have also confirmed that Chironomidae abundance increases with increasing degradation (e.g. Fore et al. 1996; Botts, 1997), identification to genus or species level was suggested by many authors to use them as water quality indicators (e.g. Kerans and Karr, 1994; Wymer and Cook, 2003). We found that Red Chironomidae together with Oligochaeta is a useful metric, and above all the routine identification can be done by non-taxonomists.

Although some studies in tropical areas have shown that metrics based on feeding types yield variable responses to human perturbation (e.g. Thorne and Williams, 1997), many authors proposed the use of trophic measures for biomonitoring particularly in multimetric systems (e.g. Barbour et al., 1996; Karr and Chu, 1999; Hering et al., 2006). In the present study, % shredders and % collector gathering were incorporated into the final index. A consistent decrease of shredders was observed with decreasing vegetation cover and water quality deterioration, which is in agreement with previous studies (Anderson et al. 1978; Ofenboeck et al., 2004). We also observed low % shredder in *Eucalyptus* sp. and *Juniperus* sp. dominated streams, which may be because of low nutritional quality and high lignin contents of the leaves that made it hardy and unpalatable. In contrary, collector gathering increased toward increasing degradation as reported in Hilsenhoff (1988) and showed strong positive correlation with pollution indicator parameters such as TP, BOD_5, and % urban (Table 4).

The other key issue for the success of biological monitoring is to establish sound threshold values for river quality classes in the region. The methods applied in the present study are efficient for streams in Ethiopian grassland and woodland ecoregions (Table 5). Slightly lower threshold value needs to be established for 'high–good' river quality class in afro-alpine streams. This underlines the importance of ecoregions as useful spatial units for river typology (e.g. Moog et al., 2004) which can be explained in terms of natural factors that influence faunal structures and the metrics as a consequence. In afro-alpine ecoregion of Bale Highlands, we observed that the natural vegetation is dominated by short Erica forest and in between the soil is covered by various grass species. Thus, streams flowing in this region have

naturally limited autochthonous and allochthonous food sources to support high diversity of benthic macroinvertebrates. Furthermore, these streams are located at higher elevation (>2,600 m a.s.l.) characterized by low temperature throughout the year. For example, the mean maximum temperature at Sanetti Plateau was between 6 and 12°C, while mean minimum temperature was between 1 and 10°C (National Meteorological Agency (NMA), 2012) which may limit high diversity, reproduction, and growth of benthic macroinvertebrates.

The development of a MMI requires testing for stability with the natural seasonal variability (Barbour et al., 1996). Some studies on seasonal variability supported the argument on the importance of this step (Sporka et al., 2006; Baptista et al., 2007). The newly developed MMI is stable across different hydrological conditions of dry and light rainy months of the year (Fig. 3).

In conclusion, the newly developed MMI is robust and sensitive to a wide variety of anthropogenic impacts such as land use change, and physical and chemical river degradation in central and southeast highland rivers of Ethiopia. This index is adequate monitoring tool for the study area and could be applied for establishing a biomonitoring network in comparable highlands of the county. We also believe that this index provides scientifically sound and reliable information to develop and implement preservation and restoration measures by decision makers.

References

Alemayehu, T., 2001. The impact of uncontrolled waste disposal on surface water quality in Addis Ababa. SINET: Ethiopian Journal of Science 24: 93–104.

Anderson, N. H., J. R. Sedell, L. M. Roberts & F. J. Triska, 1978. The role of aquatic invertebrates in processing of wood debris in coniferous forest streams. American Naturalist 100: 64–82.

APHA, 1997. Standard Methods for the Examination of Water and Wastewater, 19th ed. American Public Health Association, Washington DC.

Armitage, P. D., D. Moss, J. F. Wright & M. T. Furse, 1983. The performance of a new biological water quality score system based on macroinvertebrates over a wide range of unpolluted running water sites. Water Research 17: 333–347.

Aschalew, L., 2014. Development of biological monitoring systems using benthic invertebrates to assess the ecological status of central and southeast highland rivers of Ethiopia. PhD Thesis, University of Natural Resources and Life Sciences, Vienna, Austria: 163.

Baptista, D. F., D. F. Buss, M. Egler, A. Giovanelli, M. P. Silveira & J. L. Nessimian, 2007. A multimetric index based on benthic macroinvertebrates for evaluation of Atlantic Forest streams at Rio de Janeiro State, Brazil. Hydrobiologia 575: 83–94.

Barbour, M. T., J. B. Stribling & J. R. Karr, 1995. Multimetric approach for establishing biocriteria and measuring biological condition. In Davis, W. S. & T. P. Simon (eds), Biological Assessment and Criteria. Tools for Water Resource Planning and Decision Making. Lewis Publishers, Ann Arbor: 63–76.

Barbour, M. T., J. B. Stribling & J. Gerritsen, 1996. Biological Criteria: Technical Guidance for Streams and Small Rivers. EPA/822/B-96/001. US. Environmental Protection Agency, Washington, DC.

Barbour, M. T., J. Gerritsen, B. D. Snyder & J. B. Stribling, 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers. EPA/841/B-99/002. US Enviromental Protection Agency, Washington, D.C.

Birk, S., W. Bonne, A. Borja, S. Brucet, A. Courrat, S. Poikane, A. Solimini, W. van de Bund, N. Zampoukas & D. Hering, 2012. Three hundred ways to assess Europe's surface waters: An almost complete overview of biological methods to implement the Water Framework Directive. Ecological Indicators 18: 31–41.

Botts, P. S., 1997. Spatial pattern, patch dynamics and successional change: Chironomid assemblages in a Lake Erie coastal wetland. Freshwater Biology 37: 277–286.

Couceiro, S. R. M., N. Hamada, B. R. Forsberg, T. P. Pimentel & S. L. B. Luz, 2012. Macroinvertebrate multimetric index to evaluate the biological condition of streams in the Central Amazon region of Brazil. Ecological Indicators 18: 118–125.

Ferreira, W. R., L. T. Paiva & M. Callisto, 2011. Development of a benthic multimetric index for biomonitoring of a neotropical watershed. Brazilian Journal of Biology 71: 15–25.

Fore, L. S., J. R. Karr & R. W. Wisseman, 1996. Assessing invertebrate responses to human activities: evaluating alternative approaches. Journal of the North American Benthological Society 15: 212–231.

Gabriels, W., K. Lock, N. De Pauw & P. L. M. Goethals, 2010. Multimetric macroinvertebrate index Flanders (MMIF) for biological assessment of rivers and lakes in Flanders (Belgium). Limnologica 40: 199–207.

Getachew, M., A. Ambelu, S. Tiku, W. Legesse, A. Adugna & H. Kloos, 2012. Ecological assessment of Cheffa Wetland

in the Borkena Valley, northeast Ethiopia: macroinvertebrate and bird communities. Ecological Indicators 15: 63–71.

Graf, W., U. Grasser & A. Weinzierl, 2002. Plecoptera Part III. In Moog, O. (ed), Fauna Aquatica Austriaca, Edition 2002. Wasserwirtschaftskataster, Bundesministerium für Land- und Fortwirtschaft, Umwelt und Wasserwirtschaft, Wien: 17.

Harrison, A. D. & H. B. N. Hynes, 1988. Benthic fauna of Ethiopian mountain streams and rivers. Archiv fur Hydrobiologie Supplement 81: 1–36.

Hering, D., A. Buffagni, O. Moog, L. Sandin, M. Sommerhaeuser, I. Stubauer, C. Feld, R. Johnson, P. Pinto, N. Skoulikidis, P. Verdonschot & S. Zahrádková, 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. International Review of Hydrobiology 88: 345–361.

Hering, D., O. Moog, L. Sandin & P. F. M. Verdonschot, 2004. Overview and application of the AQEM assessment system. Hydrobiologia 516: 1–20.

Hering, D., C. K. Feld, O. Moog & T. Ofenboeck, 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.

Hilsenhoff, W. L., 1988. Rapid field assessment of organic pollution with a family-level biotic index. Journal of the North American Benthological Society 7: 65–68.

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

Karr, J. R., 1981. Assessment of biological integrity using fish communities. Fisheries 6: 21–27.

Karr, J. R. & E. W. Chu, 1999. Restoring Life in Running Waters: Better Biological Monitoring. Island Press, Washington, DC.

Kerans, B. L. & J. R. Karr, 1994. A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. Ecological Applications 4: 768–785.

Kokes, J., S. Zahradkova, D. Nemejcova, J. Hodovsky, J. Jarkovsky & T. Soldan, 2006. The PERLA system in the Czech Republic: a multivariate approach for assessing the ecological status of running waters. Hydrobiologia 566: 343–354.

Kolkwitz, R. & M. Marsson, 1902. Grundsätze für die biologische Beurteilung des Wassers nach seiner Flora und Fauna. Mitteilungen der königlichen Prüfanstalt für Wasserversorgung und Abwasserbeseitigung 1: 33–72.

Malicky, H. & W. Graf, 2012. Eine kleine Trichopterenansbeute aus Äthiopien. BRAUERlA Lunz am See, Austria 39: 32–38.

Melo, S., C. Stenert, M. Schmidt Dalzochio & L. Maltchik, 2015. Development of a multimetric index based on aquatic macroinvertebrate communities to assess water quality of rice fields in southern Brazil. Hydrobiologia 742: 1–14.

Mereta, S., P. Boetsa, L. De Meesterc & P. L. M. Goethalsa, 2013. Development of a multimetric index based on benthic macroinvertebrates for the assessment of natural wetlands in Southwest Ethiopia. Ecological Indicators 29: 510–521.

Moog, O., 1988. Überlegungen zur Gütebeurteilung von Flussstauen.- Schriftenreihe der oberösterreichischen Kraftwerke AG. Umweltforschung am Traunfluss 3: 110.

Moog, O. & S. Sharma, 2005. Guidance for pre-classifying the ecological status of HKH rivers. Deliverable 7b for ASSESS-HKH, European Commission: 27 [available on internet at http://www.assess-hkh.at].

Moog, O., A. Schmidt-kloiber, T. Ofenboeck & J. Gerritsen, 2004. Does the ecoregion approach support the typological demands of the EU 'water framework directive'? Hydrobiologia 516: 21–33.

Moya, N., R. M. Hughes, E. Domingue, F. M. Gibon, E. Goitia & T. Oberdorff, 2011. Macroinvertebrate-based multimetric predictive models for evaluating the human impact on biotic condition of Bolivian streams. Ecological Indicators 11: 840–847.

Munne, A. & N. Prat, 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.

National Meteorological Agency (NMA), 2012. [available on internet at www.ethiomet.gov.et]. Accessed 7 Dec 2012.

Nijboer, R. C., R. K. Jhonson, P. F. M. Verdonschot, M. Sommerhauser & A. Buffagni, 2004. Establishing reference conditions for European streams. Hydrobiologia 516: 91–105.

Negishi, J. N., M. Inoue & M. Nunokawa, 2002. Effects of channelisation on stream habitat in relation to a spate and flow refugia for macroinvertebrates in northern Japan. Freshwater Biology 47: 1515–1529.

Nguyen, H. H., G. Everaert, W. Gabriels, T. H. Hoang & P. L. M. Goethals, 2014. A multimetric macroinvertebrate index for assessing the water quality of the Cau river basin in Vietnam. Limnologica 45: 16–23.

Ofenboeck, T., O. Moog, J. Gerritsen & M. Barbour, 2004. A stressor specific multimetric approach for monitoring running waters in Austria using benthic macro-invertebrates. Hydrobiologia 516: 251–268.

Ofenboeck, T., O. Moog, S. Sharma & T. Korte, 2010. Development of the HKH bios: a new biotic score to assess the river quality in the Hindu Kush-Himalaya. Hydrobiologia 651: 39–58.

Reynoldson, T. B., R. C. Bailey, K. E. Day & R. H. Norris, 1995. Biological guidelines for freshwater sediment based on BEnthic Assessment of SedimenT (the BEAST) using a multivariate approach for predicting biological state. Australian Journal of Ecology 20: 198–219.

Rosenberg, D. M. & A. P. Resh (eds), 1993. Freshwater Biomonitoring and Benthic Macroinvertebrates. Chapman and Hall, London: 461.

Schmidt-Kloiber, A. & R. C. Nijboer, 2004. The effect of taxonomic resolution on the assessment of ecological water quality classes. Hydrobiologia 516: 269–283.

Sporka, F., H. E. Vlek, E. Bulankova & I. Krno, 2006. Influence of seasonal variation on bioassessment of streams using macroinvertebrates. Hydrobiologia 566: 543–555.

Stoddard, J. L., D. P. Larse, C. P. Hawkins, R. K. Jonson & R. H. Norris, 2006. Setting expectations for the ecological conditions of streams: the concept of reference condition. Ecological Applications 16: 1267–1276.

Stubauer, I. & O. Moog, 2000. Taxonomic sufficiency versus need for information – comments on Austrian experience in biological water quality monitoring. Verhandlungen Internationale Verein für Limnologie 27: 2562–2566.

Thorne, R. S. T. J. & W. P. Williams, 1997. The response of benthic macroinvertebrates to pollution in developing countries: a multimetric system of bioassessment. Freshwater Biology 37: 671–686.

Wymer, D. A. & S. B. Cook, 2003. Effects of Chironomidae (Diptera) taxonomic resolution on multivariate analyses of aquatic insect communities. Journal of Freshwater Biology 18: 179–186.

Zinabu, G. & D. Elias, 1989. Water resources and fisheries management in the Ethiopian rift valley lakes. SINET: Ethiopian Journal of Science 12: 95–109.