Hydrobiologia
https://doi.org/10.1007/s10750-019-04042-x

PRIMARY RESEARCH PAPER

A macroinvertebrate multi-metric index for Ethiopian highland streams

Temesgen Alemneh · Argaw Ambelu · Benjamin F. Zaitchik · Simon Bahrndorff · Seid Tiku Mereta · Cino Pertoldi

Received: 8 May 2018 / Revised: 27 June 2019 / Accepted: 2 August 2019

Abstract Headwater streams in the highland regions of Ethiopia are exposed to severe human disturbances because of rapid population growth. Protection and management of headwater streams require the development of informative indices that can capture disturbance and assess streams water quality in highland regions. Therefore, we developed a multi-metric index based on macroinvertebrate communities, to assess the ecological condition of the headwater streams in the Ethiopian highlands. We reviewed 38 potential metrics, representing various aspects of the macroinvertebrate assemblages, and selected four of them [total family richness, Simpson index, percentage of shredder, and family richness of EOT (Ephemeroptera, Odonata, and Trichoptera)] as final core metrics. A trisection inter-quartile range system was applied to derive scores for each core metric at impaired and reference sites. The macroinvertebrate multi-metric index (MMI) distinguished well between impaired and reference stream sites and was negatively correlated with the degree of disturbance. The results suggest that the MMI is robust measure of disturbance and can be used as a sensitive tool for evaluating the ecological condition of headwater streams in Ethiopian highlands.

Keywords Headwater streams · Multi-metric index · Macroinvertebrates · Biological monitoring · Tropical · Ethiopia

Handling editor: María del Mar Sánchez-Montoya

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10750-019-04042-x) contains supplementary material, which is available to authorized users.

T. Alemneh (✉) · B. F. Zaitchik
Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, MD 21210, USA
e-mail: tyimani1@jhu.edu;
temu.1221@gmail.com

T. Alemneh · A. Ambelu · S. T. Mereta
Department of Environmental Health Sciences and Technology, Jimma University, P.O. Box 378, Jimma, Ethiopia

S. Bahrndorff · C. Pertoldi
Section of Biology and Environmental Science, Department of Chemistry and Bioscience, Aalborg University, Fredrik BajersVej 7H, 9220 Ålborg East, Denmark

C. Pertoldi
Aalborg Zoo, Ålborg, Denmark

Published online: 12 August 2019

Introduction

Headwater streams are among the most intensively human-influenced ecosystems in Ethiopian highland regions, because human population density and associated activities are highest here (Strayer & Dudgeon, 2010; Moya et al., 2011; Alemneh et al., 2017). The degradation of the stream channel and stream bank, modification of flowing systems, and organic pollution create significant problems on headwater streams (Dudgeon et al., 2006; Aerts et al., 2007). Therefore, understanding the severe impacts of anthropogenic activities on aquatic ecosystems has become a priority for research and management (Macedo et al., 2016).

Biological indices are often used to monitor the status of aquatic ecosystems (Gabriels et al., 2010; Baptista et al., 2011; Jun et al., 2012). These indices are intended to represent the quality of a given environment with the intention of being applicable across eco-regions. Different biological indices have been established based on assemblage structures, and functions related to species richness are widely used for biological assessment of streams and lakes (Moya et al., 2011).

Macroinvertebrates have been used successfully as bioindicators, as they can have a broad range of environmental and ecological requirements (Gabriels et al., 2010) and are affected by anthropogenic impacts such as changes in stream channel conditions, physical and chemical conditions, and the quality and quantity of food resources in the streams (Barbour et al., 1999; Cross et al., 2006). Environmental change in aquatic ecosystems also varies across macroinvertebrate life cycles, where larval stages of many insects, for example, can be sensitive to human-induced stresses in waterways (Feio et al., 2009; Ambelu et al., 2010).

Measurements of the status of macroinvertebrate assemblages can be used to develop multi-metric indices of highland stream health conditions at large scale (Baptista et al., 2011; Schoolmaster et al., 2012). Multi-metric methods were first applied to streams in the study of fish communities by Karr (1981) and have been the most widely applied method to assess freshwater assemblages. Multi-metric indices are composed of a number of metrics associated with biological attributes that can change in a predictable fashion in response to anthropogenic disturbances (Barbour et al., 1999; Jun et al., 2012). The purpose is to integrate different biological measures into a single value that captures multiple effects of anthropogenic impacts on the aquatic ecosystems (Menetrey et al., 2011). The development of such a multi-metric index is therefore based on the comparison of biological metrics related to predetermined criteria from impaired to reference or less impaired sites (Stoddard et al., 2006; Ode & Schiff, 2009; Mereta et al., 2013). Multi-metric indices are useful in the context of resource management and conservation actions because they provide relevant information for decision makers to develop a conservation strategy (Karr & Chu, 1999; Schoolmaster et al., 2012).

Multi-metric indices have proven to be one of the most useful approaches to evaluate water quality in freshwater ecosystem monitoring programs and have gained increasing attention in recent decades (Ferreira et al., 2011). A multi-metric index developed for a certain eco-region can be very helpful because it provides integrated information from various features of a community, provides a classification of degradation, and enables long-term monitoring and assessment of ecosystem health (Jun et al., 2012; Macedo et al., 2016). Such a metric can also be useful when classifying sites across ecological impairments based on comparisons of the actual condition of sites within similar eco-regions. However, an index developed for a certain region might not be able to capture relevant details of other ecological regions.

Macroinvertebrate-based multi-metric indices are widely used in the assessment of aquatic resources in developed regions, including Europe (Hering et al., 2006), the United States (Stoddard et al., 2008), and Latin America (Moya et al., 2011; Macedo et al., 2016). The literature on ecological monitoring of streams, rivers, and wetlands in Africa is relatively limited, where only a few studies using biological assessment and monitoring tools have been performed (e.g., Beyene et al., 2009; Okoth & Onderi, 2009; Raburu et al., 2009; Ambelu et al., 2010; Getachew et al., 2012; Patrick, 2015; Alemneh et al., 2017) and even a smaller number of studies have applied multi-metric indices to evaluate human disturbances on wetland ecosystems (e.g., Bird, 2010; Mereta et al., 2013; Moges et al., 2016). In addition, very few studies have applied multi-metric indices for assessing the ecological status of streams and rivers in highland Ethiopia. Lakew & Moog (2015) performed a study in the central and southeast highlands of Ethiopia, and Aschalew & Moog (2015) introduced the new biotic score "ETHbios" for assessing ecological conditions of highland streams and rivers in that region. But there is no research conducted so far in headwater streams of other Ethiopian highland regions, where ecologically relevant factors such as altitude, topography, temperature, humidity, rainfall, status of degradation and other climatic conditions are markedly different. Therefore, adequate multi-metric index development is needed for highland headwater streams in the western regions of Ethiopia. The main objective of the present study was thus to develop an effective macroinvertebrate multi-metric index (MMI) for headwater streams of the Blue Nile and Gilgle Gibe rivers in the western part of Ethiopia. This was achieved by: (1) developing and testing a multi-metric index based on macroinvertebrate assemblages in the headwater streams of Ethiopian highland regions, and (2) testing the reliability and applicability of the MMI in discriminating a less impaired (reference) site(s) from impaired headwater stream sites.

Materials and methods

Study area

The study was performed in two watersheds of the western Ethiopian Highlands: the Chemoga river, located on Choke Mountain, in the Blue Nile (Abay) River headwaters in the northwest of Ethiopia, and the Gilgel Gibe river at the upper Gilgel Gibe watershed, located in the southwest of Ethiopia (Fig. 1). The Altitude of the Choke Mountain watersheds varies from 4200 m.a.s.l. at the top of the mountain to 800 m.a.s.l. at the gorge of the main-stem Blue Nile (Zaitchik et al., 2012). Annual average precipitation ranges from 600 to 2800 mm, with a geographic average of 2000 mm and an average air temperature of 18°C. Similarly, the altitude of Gilgle Gibe watershed ranges from 3259 to 1096 m.a.s.l. (Ambelu et al., 2010), with different appearance of physiographic, ecological, socio-cultural, agricultural, and climatic conditions at different elevations. The area receives variable annual rainfall ranging from 1000 mm to 2600 mm with an average of 1550 mm rainfall and an average temperature of 19°C.

Chemoga river and the Gilgle Gibe river watersheds are located in the tropical afro-alpine and sub afro-alpine ecological regions and cover a large topographic area, with a total catchment of 10,423 km^2. Both watersheds are exposed to anthropogenic disturbance as a result of rapid human population growth and high livestock loads, resulting in severe land degradation and reduced water quality and quantity (Ambelu et al., 2010; Alemneh et al., 2017). Therefore, the study of the Chemoga River at the Blue Nile River headwaters and streams of Gilgle Gibe River in the Gilgle Gibe watershed gives a methodological guideline for a broader sense of application of biomonitoring in tropical African countries in general and Ethiopia in particular.

Site selection

In order to create an MMI that can be applicable for monitoring ecological degradation, it is first necessary to define each site as a reference condition (no or minimal exposure to human disturbance) or impaired condition (Hering et al., 2006; Ruaro & Gubiani, 2013; Silva et al., 2017). The reference condition is defined by many authors as a representative of minimally disturbed sites (Barbour et al., 1996; Hering et al., 2003; Sánchez-Montoya et al., 2009; Johnson et al., 2013; Silva et al., 2017) within the respective ecoregions. However, in tropical East African countries, there are no reliable cut-off levels that indicate references. Designation of reference sites can be done based on the prior measurement of environmental variables of each site and a field survey of the overall ecological conditions of the study area before sampling (Wiseman, 2003), or through a post measurement approach based on biological and environmental conditions (Mereta et al., 2013). In this study, the prior approach was used. Accordingly, stream sites were defined as less impaired or impaired sites based on the land use categories and level of anthropogenic disturbances (US EPA, 2002; Klemm et al., 2003; Moya et al., 2007), including land use intensity, local habitat alteration, and hydrological modification. These anthropogenic disturbances were carefully scored during each sampling episode based on the scoring criteria (Table 1). Each site was given a score from zero to five based on the degree of human activity. The higher the score, the greater was the human impact on the stream site. Hence, the study sites were categorized into three human impact classes on the basis of the level of disturbances (Wang et al., 1998; Sánchez-Montoya et al., 2009; Johnson et al., 2013; Chen et al.,

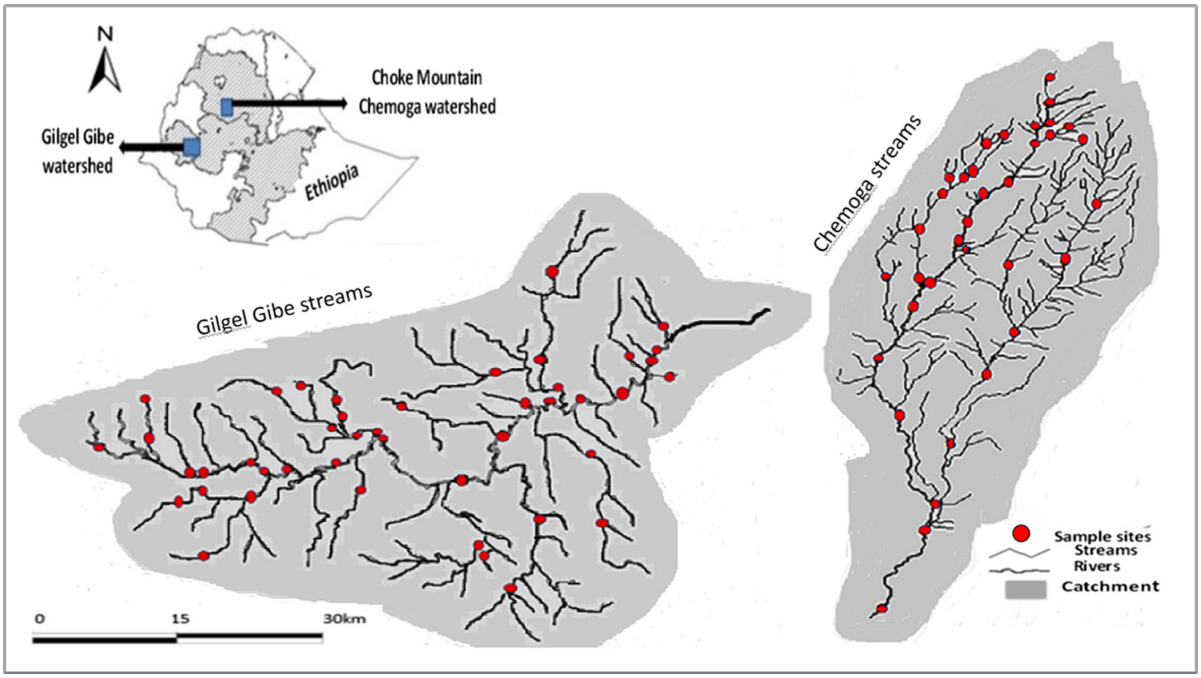

Fig. 1 Location of sampling sites of headwater streams in the upper Blue Nile and Gilgel Gibe watershed, Ethiopia

2017). Sites with an average impact score of ≤ 1 were less impacted (referenced), and sites with an average score of ≥ 3 were highly impacted. The remaining study sites had a moderate human impact with an average score between 1 and 3. The use of least disturbed and most disturbed sites comparison is essential in MMI development (Oliveira et al., 2011). Hence, 42 of the 96 sites were categorized as less and medium impaired and the remaining 54 were categorized as highly impaired. In addition, 14 from the pristine stream sites and 15 from the disturbed sites were used as a validation dataset. These validation sites were chosen based on their ease of access and representativeness to the reference and impaired sites. GIS and GPS (Magellan®, SporTrak Pro) were used to measure the distance between the sampling sites, altitude, and the total catchment of the study area.

Data collection

Biological data

Macroinvertebrate and environmental data were collected from the headwater streams of the Chemoga River and Gilgel Gibe River during dry and light wet seasons from late September 2014 through May 2015 (Fig. 1). Development of MMI requires testing for stability with natural seasonal variability (Baptista et al., 2007). Therefore, dry and light wet seasons were preferred to ensure constant discharges, stable assemblages of macroinvertebrate, and reduced signal-to-noise variability (Chen et al., 2017). A total of 182 samples were collected using a D-frame kick net with a mesh size of 300-µm diameter following Barbour et al. (1999). Samples were collected from each microhabitat reaches within a 10 m radius stretch along the streams (Ambelu et al., 2010). Macroinvertebrates were then sorted alive onsite (Mereta et al., 2013) and preserved in 75% ethanol for subsequent identification to family level. In total, 72 different macroinvertebrate families were identified at family level using Bouchard (2004) and all of them were assigned tolerance and sensitivity scores for the development of the MMI (see Barbour et al. (1996) and Gabriels et al. (2010)). The Functional Feeding Groups (FFG) were calculated based on the ratio of numerical abundance of different functional feeding groups of macroinvertebrates.

Table 1 Scoring criteria used in the selection of reference sites in the headwater streams of western Ethiopian highlands Modified from Barbour et al. (1996) Wang et al. (1998) and Alemneh et al. (2017)

Disturbance	No or less disturbed	Moderately disturbed	Highly disturbed
Land use intensity			
Farming	No farming, or farming > 100 m from the stream	Farming in 50–100 m from the stream	Farming, including buffer of the streams
Grazing	Minimal grazing < 10%	Medium grazing 10–50%	Intensive grazing, > 50%
Settlement	No settlement around streams or settlement at > 300 m	Settlement within 300 m	Settlement within < 300 m
Local habitat alteration			
Tree removal	Less tree removal, < 10%	Medium tree removal, 10–50%	High tree removal, > 50%
Landslide	No landslide or minimal landslide > 100 m	Landslide within < 100 m from the stream channel	Landslide next to streams channel
Irrigation	No irrigation or irrigation > 200 m	Irrigation within 100–200 m	Irrigation within < 100 m
Hydrological modification			
Weirs	No Weirs in the sampling site or weirs within > 100 m	Weirs within 100 m from the sampling site	Weirs within < 50 m from sampling site
Stream diversion	No stream diversions > 100 m	Stream diversion within 100 m from sampling site	Diversion within < 50 m from sampling site
Sand dredging	No or less sand dredging < 10%	Medium activities of sand dredging 10–50%	Intensive use of sand dredging > 50%
Physical and chemical			
Point source pollution	No or < 10% point source pollution	Medium pollution 10–50% pollution	High pollution > 50% pollution
Non-point source pollution	No or < 10% non-point source pollution	Medium or 10–50% non-point source pollution	Intensive or > 50% non-point source pollution

Physical and chemical parameters

Physical characteristics of streams were assessed at each sampling site following Barbour et al. (1999). Environmental features such as stream morphology (streams depth and width, channel depth and width, flow type, and substrate size), habitat alterations (tree removal, landslide), hydrological modification (irrigation, small dams), and land use intensity (farming, grazing) were measured in each sampling site. Similarly, ambient temperature, catchment area, and altitude of the study sites were measured. Each ecological attribute value was calculated for each site and compared against disturbance gradients. Similarly, physical and chemical variables such as conductivity of the water, dissolved oxygen, turbidity, pH, and water temperature were measured at each site using HACH multi-meter portable probe, model HQ40D. Environmental variables except pH were log $(x + 1)$ transformed prior to statistical analysis to normalize the distributions and homogenize the variance. Furthermore, two liter water samples were collected from each site and stored at 4°C in an insulated box containing ice packs. Subsequently, nitrate and total phosphate were measured in the laboratory according to standard methods (APHA, 1999).

Metric screening, scoring and statistical analysis

The multi-metric index must reflect the degree of stress caused by a specific impact, and all metrics composing the index should respond to relevant disturbances (Baptista et al., 2011). We screened candidate metrics following the same metric screening steps used in other MMI development studies (e.g., Barbour et al., 1996; Stoddard et al., 2008; Moya et al., 2011). More specifically, we screened our set of metrics by testing the performance of each candidate metrics with anthropogenic disturbances (Table 2). In addition, we screened metrics based on the evaluation of responsiveness of macroinvertebrates to natural environmental variables, and metric values were compared between reference and impaired streams with the aim of screening metrics that showed differences between categories. This was tested with a suite of descriptive statistical analyses. These included a non-parametric Mann–Whitney *U*-test for

Table 2 Metric selected in response to human impact and Mann–Whitney *U*-test between sites [considering number of taxa, composition, tolerance, and trophic measures (Ferreira et al. 2011)]

Metrics	Response to disturbance	*U*-test	*P*-value	Testing criteria
Relative abundance	Decrease	109	0.187	No
BMWP	Decrease	21	0.001*	Yes
Family richness	Decrease	22	0.001*	Yes
ASPT	Decrease	95	0.072	No
FBI	Increase	126	0.438	No
Dominance_D	Decrease	45	0.001*	Yes
Simpson index	Decrease	45	0.001*	Yes
Shannon_H index	Decrease	28	0.001*	Yes
Evenness_e^H/S index	Decrease	106	0.151	No
Menhinick index	Decrease	56	0.002*	Yes
Margalef index	Decrease	27	0.001*	Yes
Fisher alpha index	Decrease	31	0.001*	Yes
Non-insecta	Increase	125	0.285	No
%Odonata	Decrease	33	0.001*	Yes
%Ephemeroptera	Decrease	51	0.001*	Yes
%Trichoptera	Decrease	64	0.001*	Yes
%EOT	Decrease	10	0.001*	Yes
%Chironomidae	Increase	125	0.419	No
%Coleoptera	Decrease	136	0.668	No
%Diptera	Increase	108	0.176	No
%Oligochaeta	Variable	141	0.761	No
%Hirudinea	Increase	144	0.836	No
%Hemiptera	Increase	114	0.246	No
%Mollusca	Increase	147	0.863	No
%Predator	Variable	92	0.060	No
%Filterer-collector	Decrease	127	0.457	No
%Collectors-gatherers	Variable	140	0.754	No
%Scraper	Variable	144	0.866	No
%Shredder	Decrease	91	0.048*	Yes
Baetidae/Ephemeroptera	Increase	65	0.004*	Yes
Families of order Coleoptera	Decrease	98	0.092	No
Families of order Odonata	Decrease	38	0.001*	Yes
Families of order Trichoptera	Decrease	59	0.001*	Yes
Families of order Ephemeroptera	Decrease	26	0.001*	Yes
Families of order Hemiptera	Decrease	100	0.100	No
Families of order Diptera	Decrease	138	0.704	No
Families of order ET	Decrease	13	0.001*	Yes
EOT family richness	Decrease	7	0.001*	Yes

BMWP biological monitoring working party, ASPT average score per taxon, *FBI* family biotic index, *EOT* Ephemeroptera, Odonata, Trichoptera, *ET* Ephemeroptera, Trichoptera, % the percentage of the total number of macroinvertebrates in the sample

*Significant at $P < 0.05$

testing significant differences between medians. First, a Mann–Whitney *U*-test was used to test for significant difference in the abundance of macroinvertebrates between the reference and impaired sites for the selection of metrics (Ferreira et al., 2011). Second, a Spearman correlation analysis was used for identification of redundancy among suitable metrics. Metrics with Spearman correlation values of ($r \geq 0.70$) were considered as redundant and at least one of the redundant variables was eliminated. The one with the highest correlation coefficient with environmental variables and the most user-friendly to implement for monitoring purposes at the highland streams was selected (Barbour et al., 1996; Baptista et al., 2011). Third, we examined the degree of inter-quartile (IQ) overlaps in Box-and-Whisker plots, considering the overlaps among inter-quartile limits along with the direction and intensity of the response as the impact increased. The appropriate quartile was used as a threshold depending on the type of response to impairment (Ferreira et al., 2011). The range of metric values obtained from the Box-and-Whisker plots were divided into three possible scores (Barbour et al., 1996). Based on this, a score of 5 was given to metrics that represented a reference condition, a score of 3 represents an intermediate condition, and a score of 1 indicates the highest deviation from the expected values for the reference condition (Barbour et al., 1996; Baptista et al., 2011; Ferreira et al., 2011). Core metric values were determined based on the responses to habitat degradation; we used the threshold values of the core metrics box-and-whisker plots i.e., minimum, the 25th percentile, median, the 75th percentile and the maximum. Furthermore, to evaluate these disturbances, redundancy analysis (RDA) criteria were applied using CANOCO 4.5 (ter Braak & Smilauer, 2002) to see the appropriate response model between macroinvertebrate metrics and the natural environmental variables including altitude, catchment area, and physical and chemical data. The redundancy analysis (RDA) was used after discriminating the linear and unimodal ordination gradients using detrended correspondence analysis (DCA). Then, the result showed a short gradient (< 3) linear model that suggests using RDA. The relationship between core metrics, environmental variables, and the final MMI were analyzed using STASTICA software (version 7.1), where the P-value < 0.05 was considered statically significant.

Index development and validation

A multi-metric index that represents various aspects of macroinvertebrate community was developed based on the environmental and biological attributes (Klemm et al., 2003; Melo et al., 2015), such as responsiveness of metrics to several disturbance gradients (Land use, habitat alteration, hydrological modification), redundancy with other metrics, relationships with catchment area, altitude, and chemical attributes, to determine metrics that best discriminate least disturbed sites from most disturbed sites. All metrics that passed the screening were considered for MMI development. The selected metrics represent different structural and functional attributes of macroinvertebrate assemblages such as taxa richness metrics, diversity metrics, feeding group metrics and tolerance taxa metrics. These sets of metrics were chosen based on a set of non-redundant metrics that responded to a variety of environmental disturbance types and responses to anthropogenic disturbances. The candidate metrics were cross checked with both environmental and anthropogenic disturbances to generate the most promising MMI. This analysis was intended to confirm the appropriateness and potential utility of the MMI. The final macroinvertebrate multi-metric index (MMI) was developed by summing up the values of each core metric.

The performance of the developed MMI was then validated using the validation data set of 14 samples from reference streams and 15 from disturbed streams. The MMI was also determined in order to see if there was clear discrimination among the classified impaired and unimpaired sites using Box-and-Whisker plots and Kruskal–Wallis test (Baptista et al., 2011; Jun et al., 2012). In addition, the MMI scores were correlated with physical and chemical variables using Spearman correlation (r_s) to show its representativeness relative to other measures of stream impairment. Furthermore, redundancy analysis (RDA) was used to examine the responsiveness of the multi-metric index to environmental variables. Finally, MMI scores ranged from 4 to 20 and were grouped into the following five quality classes: very poor, poor, moderate, good, and very good, corresponding to the level of impaired highland streams. The low score (4) corresponds to highly impacted and 20 corresponds to less impacted sites. These quality classes were determined based on the extent of biotic integrity of the headwater streams.

Results

Metric selection

In the process of metric development, the candidate potential metrics should be selected from the total metrics considered in the evaluation. Based on this, from the total of 38 candidate potential metrics, only 19 (50%) showed a significant difference between the reference and impaired sties ($P < 0.05$). All selected metrics decreased with increasing human disturbance (Table 2). Of these 19 candidate variables, only four core metrics were retained for index development after excluding the redundant metrics using Spearman rank order correlation coefficients ($r_s \geq 0.70$ with $P < 0.05$; Table 3). The selected core metrics were total family richness, Simpson index, percentage of Shredder (%Shredder) and EOT family richness (the richness of sensitive taxa of Ephemeroptera, Odonata and Trichoptera). These metrics are representative for most sites and fulfilled all metric selection procedures; they were selected with the highest priority for the MMI development (Fig. 2). Each core metric was determined by using the threshold or inter-quartile range values from Box-and-Whisker plots of the reference streams and scoring criteria of trisection method of Barbour et al. (1996). The Box-and-Whisker plots of each selected core metric clearly showed the discrimination between referenced and impaired stream sites (Fig. 2; Table 4).

Relationships of environmental and physical and chemical parameters with core metrics

The environmental and physical and chemical parameters in reference conditions reflect the stream habitat quality, and in impaired conditions they can reflect the environmental disturbances. Most of these parameters are significantly different between reference and impaired sites (Table 5). The mean of physical and chemical parameters such as conductivity, turbidity, nitrate, phosphate, and environmental conditions such as land use, hydrological modification, habitat alteration, altitude, and catchment size in the impaired sites was higher than reference sites (Table 5). This was not the case for water temperature and dissolved oxygen. Physical and chemical parameters like primarily dissolved oxygen, turbidity, water temperature, and environmental conditions such as land use, hydrological modification, habitat alteration, altitude, and catchment size show significant differences between impaired and less impaired sites of the streams. In the higher altitude, there is a large number of settlements (see Alemneh et al., 2017) and the associated impacts of settlements such as land use, hydrological modifications, and habitat alterations are high in most of the study sites.

Multi-metric index development and validation

The multi-metric index was developed using core metrics derived from macroinvertebrate data collected from different land use categories of the study sites. The metric values used for MMI development were selected through repeated testing and validation procedures. Then, the final MMI scores ranged from 4 to 20, with the value of 4 corresponding to highly impacted and 20 corresponding to least impacted sites. The score range of MMI (4–20) was divided into 5 quality classes corresponding to highland stream sites: 4–6 = very poor, for highly impacted, severe disturbance and deterioration of habitat quality and significant anthropogenic disturbances on assemblages; 7–9 = poor, for slightly less than highly impacted sties; 10–13 = moderate, for moderately impacted sites; 14–17 = good, for low disturbed sites and 18–20 = very good, for desirable biological integrity and fewer problems of human disturbance with high habitat quality.

The Box- and Whisker plot and Kruskal–Wallis test were used to validate the MMI. Validation showed that sites initially classified as reference sites were categorized as less impacted habitat conditions and the majority of sites initially considered as impaired were categorized as in moderately and highly impacted habitat conditions (Fig. 3). Most validation datasets were correctly classified by the MMI ($R^2 = 0.93$) (Fig. 4). The Spearman correlation (Table 6) and RDA validation test (Fig. 5) also showed the efficiency of the MMI to discriminate between reference (less impaired) and impaired sites ($P < 0.05$).

Table 3 Spearman rank order correlation coefficients ($r_s \geq 0.70$) among the selected metrics to avoid redundant metrics (*BMWP* biological monitoring working party; *% EOT* % of Ephemeroptera, Odonata, Trichoptera; *ET fam.* Ephemeroptera, Trichoptera family, *Ephem. fam.* Ephemeroptera family, *EOT fam. rich.* Ephemeroptera, Odonata, Trichoptera family richness)

	BMWP	Family richn.	Dominance_D	Simpson index	Shannon_H	Menhinick index	Margalef index	Fisher alpha index	%Odonata	%Ephemeroptera
BMWP	1									
Family richnes	0.92*	1								
Dominance_D	− 0.69*	− 0.66*	1							
Simpson index	0.69*	0.66*	− 0.98*	1						
Shannon_H	0.83*	0.84*	− 0.93*	0.93*	1					
Menhinick index	0.73*	0.73*	− 0.88*	0.88*	0.90*	1				
Margalef index	0.92*	0.96*	− 0.78*	0.78*	0.92*	0.87*	1			
Fisher alpha index	0.88*	0.92*	− 0.83*	0.83*	0.94*	0.92*	0.98*	1		
%Odonata	0.54*	0.45*	− 0.49*	0.49*	0.52*	0.40*	0.46*	0.46*	1	
% Ephemeroptera	0.44*	0.31*	− 0.53*	0.53*	0.49*	0.45*	0.39*	0.42*	0.54*	1
%Trichoptera	0.71*	0.55*	− 0.39*	0.39*	0.46*	0.42*	0.56*	0.52*	0.39*	0.18
%EOT	0.66*	0.49*	− 0.61*	0.61*	0.61*	0.54*	0.55*	0.56*	0.77*	0.82*
%Shredder	0.64*	0.50*	− 0.49*	0.49*	0.52*	0.53*	0.56*	0.54*	0.29	0.12
Baetidae/Ephem.	0.55*	0.48*	− 0.33*	0.33*	0.39*	0.26	0.41*	0.40*	0.31*	0.65*
Odonata fam.	0.51*	0.43*	− 0.48*	0.48*	0.48*	0.39*	0.44*	0.49*	0.96*	0.47*
Trichoptera fam.	0.74*	0.60*	− 0.39*	0.39*	0.49*	0.42*	0.59*	0.54*	0.42*	0.16
Ephem. fam.	0.67*	0.57*	− 0.42*	0.42*	0.51*	0.35*	0.54*	0.50*	0.57*	0.81*
ET fam.	0.77*	0.63*	− 0.48*	0.48*	0.56*	0.40*	0.60*	0.56*	0.56*	0.65*
EOT fam. rich.	0.79*	0.64*	− 0.51*	0.51*	0.60*	0.45*	0.63*	0.59*	0.68*	0.68*

	%Trichoptera	%EOT	%Shredder	Baetidae/Ephem.	Odonata fam.	Trichopterafam.	Ephem. fam.	ET fam.	EOT fam. rich.
BMWP									
Family richnes									
Dominance_D									
Simpson index									
Shannon_H									
Menhinick index									
Margalef index									
Fisher alpha index									

Table 3 continued

	%Trichoptera	%EOT	%Shredder	Baetidae/Ephem.	Odonata fam.	Trichopterafam.	Ephem. fam.	ET fam.	EOT fam. rich.
%Odonata									
% Ephemeroptera									
%Trichoptera	1								
%EOT	0.56*	1							
%Shredder	0.75*	0.43*	1						
Baetidae/Ephem.	0.25	0.58*	0.12	1					
Odonata fam.	0.36*	0.72*	0.28	0.27	1				
Trichoptera fam.	0.99*	0.56*	0.73*	0.29	0.39*	1			
Ephem. fam.	0.36*	0.78*	0.15	0.84*	0.52*	0.39*	1		
ET fam.	0.69*	0.84*	0.44*	0.75*	0.52*	0.71*	0.88*	1	
EOT fam. rich.	0.66*	0.90*	0.44*	0.71*	0.65*	0.68*	0.88*	0.97*	1

*Significant at $P < 0.05$

Discussion

In this study, we developed a macroinvertebrate multimetric index (MMI) to assess the ecological condition of headwater streams in the western Ethiopian highland regions. The candidate monitoring metrics were sensitive to anthropogenic impacts, and applicable to biomonitoring of headwater streams in highland regions and useful in the context of management and conservation actions (Stoddard et al., 2008; Ferreira et al., 2011; Schoolmaster et al., 2012; Aschalew & Moog, 2015). Four core metrics were selected to reflect the overall ecological status of the headwater streams in Ethiopian highland regions and combined into a single monitoring index: (1) the total family richness represents broad phylogenetic richness; (2) the Simpson index metrics is a measure of biodiversity; (3) percentage of shredder encompasses functional feeding groups (FFG) and provides information on the balance of feeding strategies in the benthic assemblages and their trophic roles; and (4) EOT family richness targets the diversity of sensitive taxa (Table 4; Fig. 2). All selected core metrics showed a negative response to habitat degradation and water quality deterioration and thus have the potential to contribute to discriminating between reference and impaired sites. Studies have shown that Ephemeroptera, Plecoptera, and Trichoptera are useful in biomonitoring because of their sensitivity to disturbances (Ode et al., 2005). However, in this study, the frequency of occurrence of Stoneflies (order Plecoptera) was very low due to high habitat disturbances in most Ethiopian highland regions; instead order Odonata was included with Ephemeroptera and Trichoptera because of their sensitivity to anthropogenic disturbance (Hornung & Rice, 2003) and their frequent occurrence in most sampling sites.

Sensitive taxa are known to decrease with increased pollution, so they are a useful indicator of ecological condition of streams (e.g., Barbour et al., 1996; Karr & Chu, 1999). In this regard, the species belonging to the order of Ephemeroptera are important aquatic insects for bioassessment and biomonitoring of freshwater ecosystems (Edsall et al., 2005; Arimoro & Muller, 2010; Silva et al., 2017). Ephemeroptera are the most abundant insects in highland streams with relatively high dissolved oxygen concentration (Shelly et al., 2011; Lakew & Moog, 2015) and their abundance significantly decreases with increasing habitat

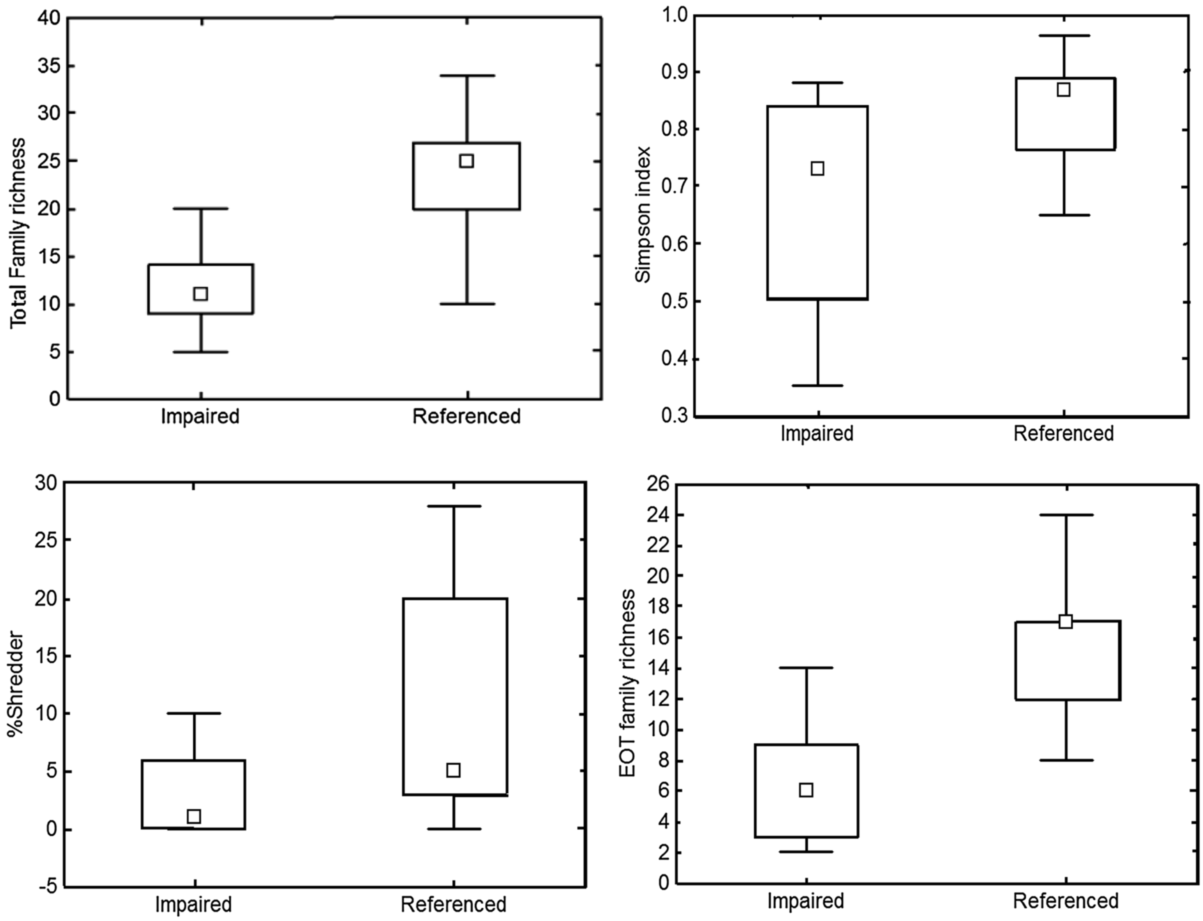

Fig. 2 Box-and-Whisker plots showed the core metrics of total family richness, Simpson index, %Shredder and EOT family richness across the reference and impaired sites. The middle small box is the median value, the large box represents the inter-quartile ranges, the lower and upper whisker represent the first and fourth quartile values, respectively; *EOT* Ephemeroptera, Odonata, Trichoptera

Table 4 Threshold values and scoring criteria of the standard core metrics using trisection scoring method

Core metrics	Threshold value based on box plot scores							
	Min.	25%	Median	75%	Max.	5	3	1
Total Family richness	10	20	25	27	35	≥ 20	20–10	< 10
Simpson index	0.65	0.77	0.87	0.89	0.98	≥ 0.77	< 0.77 < 0.65	< 0.65
%Shredder	0	4	6	20	27	≥ 4	< 4 > 0	0
EOT family richness	8	12	17	17	24	≥ 12	12–8	< 8

A score of 1 represents severely impaired site condition; 3, moderate site condition; and 5, represents the reference site condition (Barbour et al., 1996)

Min. minimum, *Max.* maximum

disturbances. They are thus used as a potential indicator of anthropogenic disturbances, except some groups (e.g., Baetidae, Caenidae) which are tolerant for disturbances (Lakew & Moog, 2015). Similarly, Odonata are important components of freshwater ecosystems and are abundantly found in the presence

Table 5 The means and standard deviation (SD) of environmental conditions and physical and chemical parameters of the referenced (less impaired) and impaired sites to develop MMI and *P*-value of the Mann–Whitney *U*-test less than 0.05 were considered as significant

Environmental and physicochemical parameter	Mean ± SD values in reference samples (*n* = 28)	Mean ± SD values in impaired samples (*n* = 39)	*P*-value
Water temperature (°C)	15.9 ± 3.3	12.6 ± 3.0	0.00*
Conductivity (μS/cm)	67.9 ± 19.6	77.9 ± 19.2	0.08
DO (mg/l)	7.1 ± 0.2	6.9 ± 0.3	0.03*
pH	7.2 ± 0.2	7.2 ± 0.3	0.92
Turbidity (NTU)	18.7 ± 12.6	222.6 ± 228.2	0.00*
Nitrate (mg/l)	1.7 ± 0.8	2.1 ± 0.8	0.09
Phosphate (mg/l)	1.1 ± 0.4	1.3 ± 0.7	0.49
Altitude (m.a.s.l.)	2677.5 ± 386.6	3030.3 ± 324.9	0.00*
Catchment (km^2)	2681.4 ± 2402	3607 ± 2560.7	0.02*
Land use intensity	2.8 ± 1.1	3.6 ± 0.7	0.02*
Habitat alteration	1.6 ± 1.1	3.0 ± 0.9	0.00*
Hydrological modification	2.7 ± 1.3	3.8 ± 0.8	0.00*

DO dissolved oxygen, *SD* standard deviation

*Significant at $P < 0.05$

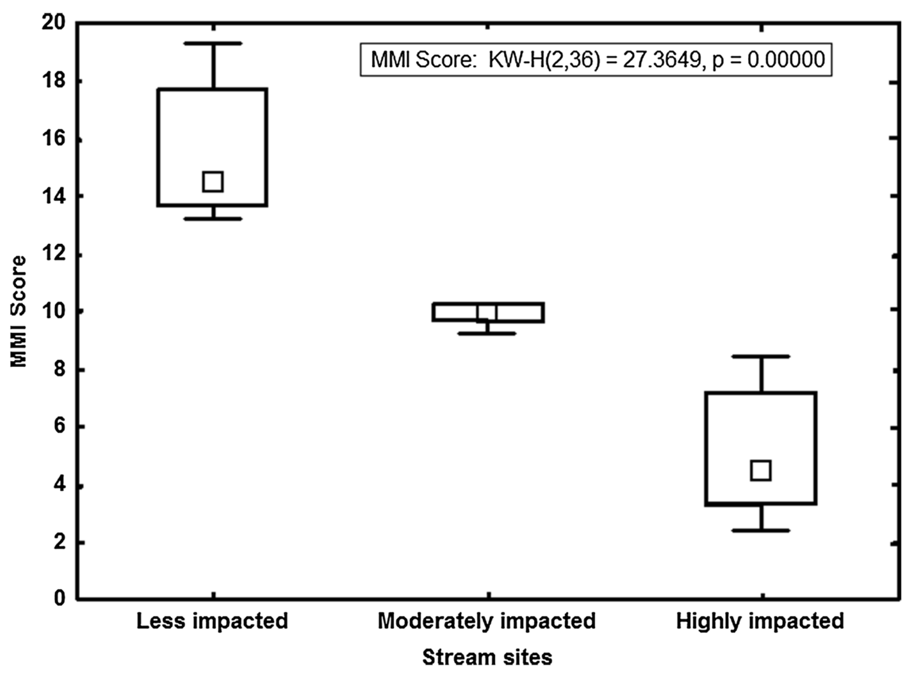

Fig. 3 Box-and-Whisker plots of MMI scores for headwater streams across different sites. Each plot shows the 25–75th percentiles of the boxes, medians, whiskers delaminating the maximum and minimum values of the MMI. Strong variation ($P < 0.05$) between headwaters stream sites

of riparian vegetation in waterways. This makes them useful indicators of, disturbance of riparian zone, which affects the biological composition and water quality of the streams. Some Odonata families are tolerant (e.g., Protoneuridae) and some others are sensitive to disturbances (e.g., Gomphiidae) can be used as an indicator of freshwater quality (Hornung & Rice, 2003; Mereta et al., 2013). Trichopterans are among the most sensitive to human disturbance and are important taxa for biomonitoring in many types of

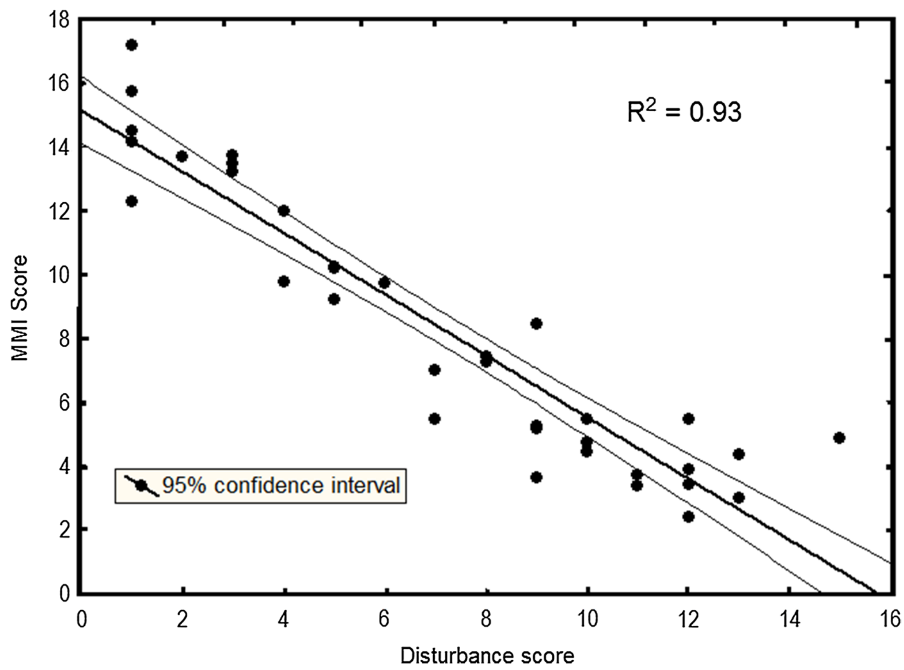

Fig. 4 Relationships of MMI score and disturbance gradient scores of the validation dataset ($P < 0.05$)

Table 6 Core metrics correlation with environmental variables using Spearman rank order correlation (*HA* habitat alteration; *HM* hydrological modification, *DO* dissolved oxygen)

Environmental parameters	Core metrics			
	Total family richness	Simpson index	%Shredder	EOT family richness
Altitude (m.a.s.l)	− 0.47*	− 0.37*	− 0.28*	− 0.42*
Catchment (km^2)	− 0.36*	− 0.32*	− 0.29*	− 0.35*
Water temperature (°C)	0.41*	0.18	0.20*	0.41*
Conductivity (μS/cm)	− 0.31*	− 0.14	− 0.31*	− 0.08
DO (mg/l)	0.12	0.31*	0.24*	0.44*
pH	0.03	− 0.08	− 0.09	0.08
Turbidity (NTU)	− 0.58*	− 0.45*	− 0.31*	− 0.63*
Nitrate (mg/l)	− 0.51*	− 0.22*	− 0.02	− 0.12
Phosphate (mg/l)	0.06	0.10	− 0.04	0.08
Land use	− 0.49*	− 0.46*	− 0.08	− 0.28*
HA	− 0.44*	− 0.28*	− 0.23*	− 0.43*
HM	− 0.52*	− 0.45*	− 0.36*	− 0.45*

*Significant at $P < 0.05$

freshwater habitats (Hornung & Rice, 2003; Houghton, 2015). In agreement with this, EOT assemblages have proven to be effective human disturbance indicators and are useful in established metrics that are able to discriminate between reference and impaired sites. Similarly, the Simpson index targets a measure of diversity often used to quantify the biodiversity of a habitat. It takes into account the number of species present as well as the relative abundance of each species in the highland streams. The index shows a similar trend across anthropogenic disturbance. Feeding strategies are typical traits reflecting the adaptation of the benthic communities into their trophic guilds (Tomanova et al., 2006), and

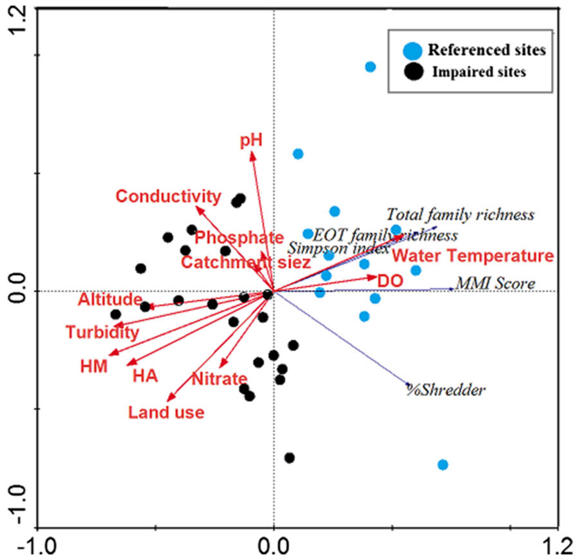

Fig. 5 Redundancy analysis (RDA) of macroinvertebrate metrics and environmental variables of highland streams of Ethiopia (*EOT* Ephemeroptera, Odonata and Trichoptera, *DO* dissolved oxygen, *HA* habitat alteration, *HM* hydrological modification)

the abundance of FFG has been related to gradients in environmental quality of streams. Similarly, the functional attributes of macroinvertebrate composition can be used to detect human disturbance independently of taxonomic composition (Tomanova et al., 2008), and hence can be considered for MMI development (Moya et al., 2011). Functional Feeding Groups, specifically %Shredders (Table S1), have been identified as important groups in previous work on stream assessment (Barbour et al., 1999; Baptista et al., 2011; Menetrey et al., 2011). These studies have found a significant relationship between %Shredder and habitat impairment such as tree removal and water quality deterioration.

Headwater streams in Ethiopia are extremely exposed to anthropogenic impacts such as agricultural activities (e.g., farming and grazing), hydrological modification (e.g., weirs, stream diversion for small scale irrigation), habitat alteration (e.g., riparian vegetation removal, landslides), and settlements (Beyene et al., 2009; Ambelu et al., 2010; Alemneh et al., 2017). Furthermore, increasing anthropogenic impacts can severely affect the natural morphology of stream channels and ecological functions of streams (Barbour et al., 1999), and have long been known to degrade water quality (Uriarte et al., 2011). Anthropogenic disturbances such as habitat alteration, land use, and hydrological modifications were found to be the most significant and negatively correlated to the selected core metrics. Similarly, physical and chemical parameters such as turbidity, conductivity, and nitrate were negatively correlated with the core metrics, while water temperature and dissolved oxygen were positively correlated with the core metrics (Table 6). In this study, our core metrics contain many taxa sensitive to changes in water quality and require a moderate to high concentration of dissolved oxygen and water temperature (Fierro et al., 2012). Altitude and catchment size were negatively correlated with core metrics (Table 6). At higher altitude and at large catchment size the disturbance of stream habitat were severe as a result of high anthropogenic impacts such as settlement, agricultural activities (see also Alemneh et al. (2017)).

An RDA was used to validate the core metrics and the final MMI. The correlation matrix of the RDA (Fig. 5) showed that environmental variables negatively correlated with core metrics and MMI except DO and water temperature. Similarly, the validation dataset clearly discriminates the MMI and the anthropogenic disturbance gradients (Fig. 4). This indicated that the selected core metrics were valid, as they were robust in detecting environmental variables resulting from ecological disturbances of the highland streams.

The macroinvertebrate multi-metric index we propose in this study is a summation of four metrics that reflect different aspects of the structure and function of macroinvertebrate assemblages in highland streams. Based on overall MMI scores, the Box-and-Whisker plot and the Kruskal–Wallis test ($H = 27.36$, $df = 2.3$, $P < 0.001$) showed significant variations among study sites and there was a strong negative correlation between the MMI score and disturbance gradient (Fig. 3). This indicates that the MMI responds appropriately to generalized measurement of headwater stream disturbance and represents a suitable tool to detect environmental degradation.

It is known that several biotic indices developed based on macroinvertebrates have been successfully used for biomonitoring of streams in different ecological regions. It is a relatively low cost and easy to apply in the highland regions of Ethiopia, once analysts are trained to identify macroinvertebrates at lowest classification level (Aschalew & Moog, 2015). However, the richness and diversity metrics of this MMI were not identified at the genus or species level due to limited taxonomic knowledge. In addition, our sampling was limited with two replicates per site. The natural landscape and the mountainous part of the study area presented logistical challenges, preventing us from obtaining more than two replicates at every study site. Another limitation of our MMI could relate to the site selection method. In this study, the importance of anthropogenic impacts was estimated based on prior observation of each site. Previous studies, for example, Klemm et al. (2003) and Silveira et al. (2006), have used physical and chemical parameters to distinguish between reference and impacted sites. However, we did not attempt these criteria as a preliminary analysis to discriminate reference and impacted sites. It is known that physical and chemical parameters may not be representative of stream quality because of occasional pollution (Moya et al., 2007).

Despite these limitations, the MMI developed here could be used for the biomonitoring of the highland stream quality. As shown here, the MMI includes a set of component metrics that are indicative of degradation status, and that reflects the status of invertebrates associated with diverse ecological functions. The MMI composites these diverse metrics, and it also correlates with degradation at sites across the study region. As such, the MMI offers a single, simple metric that can be used to assess disturbance status, monitor environmental quality, and inform conservation strategies in highland stream environments.

Conclusions

In summary, the MMI developed in this work provides a valuable tool for assessing different types of stream ecosystems, since it is responsive to the key degradation stressors affecting this highland stream region: anthropogenic disturbances and water quality degradation. This index is a potentially valuable monitoring tool for the western Ethiopian headwater streams and could be applicable for establishing biomonitoring techniques in similar headwater streams in the region. Furthermore, our study is essential for assessing the effects of environmental management practices on the headwater streams that feed into the Grand Ethiopian Renaissance Dam (GERD), currently under construction on the Blue Nile river catchment, and for the hydropower dam on the Gilgle Gibe River. This index could therefore provide helpful information for policy makers to develop headwaters stream resources protection strategies in the region.

Acknowledgements We are honored to acknowledge Jimma University and Debre Markos University for the Financial and Logistics support and John Hopkins University, Department of Earth and Planetary Science for visiting scholarship and providing other facilities for the PhD candidate Temesgen Alemneh (Grant No. ICER 1624335). We also forwarded our thanks to Dr. Admasu Moges for his technical support. Lastly, we acknowledge Aalborg Zoo Conservation Foundation (AZCF) for financial support.

References

Aerts, R., A. Negussie, W. Maes, E. November, M. Hermy & B. Muys, 2007. Restoration of dry afromontane forest using pioneer shrubs as nurse-plants for *Olea europaea* ssp. cuspidata. Restoration Ecology 15(1): 129–138.

Alemneh, T., A. Ambelu, S. Bahrndorff, S. T. Mereta, C. Pertoldi & B. F. Zaitchik, 2017. Modeling the impact of highland settlements on ecological disturbance of streams in Choke Mountain Catchment: macroinvertebrate assemblages and water quality. Ecological Indicators 73: 452–459.

Ambelu, A., K. Lock & P. Goethals, 2010. Comparison of modelling techniques to predict macroinvertebrate community composition in rivers of Ethiopia. Ecological Informatics 5(2): 147–152.

American Public Health Association, American Water Works Association, & Water Environment Federation, 1999. Standard Methods for the Examination of Water and Wastewater., Standard Methods American Public Health Association, Washington: 541.

Arimoro, F. O. & W. J. Muller, 2010. Mayfly (Insecta: Ephemeroptera) community structure as an indicator of the ecological status of a stream in the Niger Delta area of Nigeria. Environmental Monitoring and Assessment 166(1–4): 581–594.

Aschalew, L. & O. Moog, 2015. Limnologica Benthic macroinvertebrates based new biotic score "ETHbios" for assessing ecological conditions of highland streams and rivers in Ethiopia. Limnologica 52: 11–19.

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(1): 83–94.

Baptista, D. F., R. S. G. de Souza, C. A. Vieira, R. Mugnai, A. S. Souza & R. B. S. de Oliveira, 2011. Multimetric index for assessing ecological condition of running waters in the upper reaches of the Piabanha-Paquequer-Preto basin, Rio De Janeiro, Brazil. Zoologia 28(5): 619–628.

Barbour, M. T., J. Gerritsen, G. E. Griffith, R. Frydenborg, E. McCarron, J. S. White & M. L. Bastian, 1996. A framework for biological criteria for Florida streams using benthic macroinvertebrates. Journal of the North American Benthological Society 15: 185–211.

Barbour, M. T., J. Gerritsen, B. D. Snyder, J. B. Stribling & B. Macroinvertebrates, 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish, 2nd ed. US Environmental Protection Agency Office of Water, Washington, DC.

Beyene, A., W. Legesse, L. Triest & H. Kloos, 2009. Urban impact on ecological integrity of nearby rivers in developing countries: the Borkena River in highland Ethiopia. Environmental Monitoring and Assessment 153(1–4): 461–476.

Bird, M., 2010. Aquatic Invertebrates as Indicators of Human Impacts in South African Wetlands. University of Cape Town, Cape Town.

Bode, R. W., M. A. Novak & L. A. Abele, 1996. Quality assurance work plan for biological stream monitoring in New York State. NYS Department of Environmental Protection, Division of Water, Bureau of Monitoring and Assessment, Stream Biomonitoring Unit, Albany, NY.

Bouchard, R. W., 2004. Guide to Aquatic Macroinvertebrates of the Upper Midwest; Water Resources Center. University of Minnesota, St. Paul, MN.

Chen, K., R. M. Hughes, J. G. Brito, C. G. Leal, R. P. Leitão, J. M. B. de Oliveira-Júnior & J. Zuanon, 2017. A multi-assemblage, multi-metric biological condition index for eastern Amazonia streams. Ecological Indicators 78: 48–61.

Cross, W. F., J. B. Wallace, A. D. Rosemond & S. L. Eggert, 2006. Whole-system nutrient enrichment increases secondary production in a detrital-based ecosystem. Ecology 87(6): 1556–1565.

Dudgeon, D., A. H. Arthington, M. O. Gessner, Z.-I. Kawabata, D. J. Knowler, C. Lévêque & C. A. Sullivan, 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. Biological Reviews of the Cambridge Philosophical Society 81(2): 163–182.

Edsall, T., M. Bur, O. Gorman & J. Schaeffer, 2005. Burrowing mayflies as indicators of ecosystem health: status of populations in western Lake Erie, Saginaw Bay and Green Bay. Aquatic Ecosystem Health & Management 8(2): 107–116.

Feio, M. J., R. H. Norris, M. A. S. Graça & S. Nichols, 2009. Water quality assessment of Portuguese streams: regional or national predictive models? Ecological Indicators 9(4): 791–806. https://doi.org/10.1016/j.ecolind.2008.09.012.

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(1): 15–25.

Fierro, P., C. Bertrán, M. Mercado, F. Peña-Cortés, J. Tapia, E. Hauenstein & L. Vargas-Chacoff, 2012. Benthic macroinvertebrate assemblages as indicators of water quality applying a modified biotic index in a spatio-seasonal context in a coastal basin of Southern Chile. Revista de Biología Marina Y Oceanografía 47(3): 23–33.

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(3): 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(1): 63–71.

Hauer, F. & G. Lamberti, 1996. Methods in Stream Ecology. Academic Press, New York, NY.

Hering, D., A. Buffagni, O. Moog, L. Sandin, M. Sommerhäuser, I. Strubauer & S. Zahradkova, 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(3–4): 345–361.

Hering, D., C. K. Feld, O. Moog & T. Ofenbck, 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(1): 311–324.

Hornung, J. P. & C. L. Rice, 2003. Odonata and wetland quality in Southern Alberta, Canada: a preliminary study. Odonatologica 32(2): 119–129.

Houghton, D. C., 2015. Delineation and characterization of Michigan, USA, caddisfly (Insecta: Trichoptera) biodiversity and comparison with Minnesota. Journal of Freshwater Ecology 5060: 1–18.

Johnson, R.K., M. Lindegarth, J. Carstensen, 2013. Establishing reference conditions and setting class boundaries Deliverable 2.1-1, WATERS Report No.: 2. Havsmiljöinstitutet, Sweden

Jun, Y. C., D. H. Won, S. H. Lee, D. S. Kong & S. J. Hwang, 2012. A multimetric benthic macroinvertebrate index for the assessment of stream biotic integrity in Korea. International Journal of Environmental Research and Public Health 9(10): 3599–3628.

Karr, J. R., 1981. Assessment of biotic 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.

Klemm, D. J., K. A. Blocksom, F. A. Fulk, A. T. Herlihy, R. M. Hughes, P. R. Kaufmann & W. S. Davis, 2003. Development and evaluation of a Macroinvertebrate Biotic Integrity Index (MBII) for regionally assessing Mid-Atlantic Highlands streams. Environmental Management 31(5): 656–669.

Lakew, A. & O. Moog, 2015. A multimetric index based on benthic macroinvertebrates for assessing the ecological status of streams and rivers in central and southeast highlands of Ethiopia. Hydrobiologia 751(1): 229–242.

Lunde, K. B. & V. H. Resh, 2011. Development and validation of a macroinvertebrate index of biotic integrity (IBI) for assessing urban impacts to Northern California freshwater wetlands. Environmental Monitoring and Assessment 184: 3653–3674.

Macedo, D. R., R. M. Hughes, W. R. Ferreira, K. R. Firmiano, D. R. O. Silva, R. Ligeiro & M. Callisto, 2016. Development of a benthic macroinvertebrate multimetric index (MMI) for Neotropical Savanna headwater streams. Ecological Indicators 64: 132–141.

Melo, S., C. Stenert, M. S. 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): 1–14.
Menetrey, N., B. Oertli & J. B. Lachavanne, 2011. The CIEPT: a macroinvertebrate-based multimetric index for assessing the ecological quality of Swiss lowland ponds. Ecological Indicators 11(2): 590–600.
Mereta, S. T., P. Boets, L. De Meester & P. L. M. Goethals, 2013. Development of a multimetric index based on benthic macroinvertebrates for the assessment of natural wetlands in Southwest Ethiopia. Ecological Indicators 29: 510–521.
Moges, A., A. Beyene, E. Kelbessa, S. T. Mereta & A. Ambelu, 2016. Development of a multimetric plant-based index of biotic integrity for assessing the ecological state of forested, urban and agricultural natural wetlands of Jimma Highlands, Ethiopia. Ecological Indicators 71: 208–217.
Moya, N., S. Tomanova & T. Oberdorff, 2007. Initial development of a multi-metric index based on aquatic macroinvertebrates to assess streams condition in the Upper Isiboro-Scure Basin. Bolivian Amazon. Hydrobiologia 589(1): 107–116.
Moya, N., R. M. Hughes, E. Domnguez, 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(3): 840–847.
Ode, P. R., A. C. Rehn & J. T. May, 2005. A quantitative tool for assessing the integrity of southern coastal California streams. Environmental Management 35(4): 493–504.
Ode, P., & K. Schiff, 2009. Recommendations for the Development and Maintenance of a Reference Condition Management Program (RCMP_ to Support Biological Assessment of California's Wadeable Streams. *Swamp*, (March).
Okoth, P. & F. Onderi, 2009. Assessment of pollution impacts on the ecological integrity of the Kisian and Kisat rivers in Lake Victoria drainage basin, Kenya. African Journal of Environmental Science and Technology 3(4): 97–107.
Oliveira, R. B. S., D. F. Baptista, R. Mugnai, C. M. Castro & R. M. Hughes, 2011. Towards rapid bioassessment of wadeable streams in Brazil: develop- ment of the Guapiac¸ u-Macau Multimetric Index (GMMI) based on benthic macroinvertebrates. Ecol. Ind. 11: 1584–1593.
Patrick, M., 2015. Benthic macroinvertebrates as indicators of water quality: a case-study of urban Funa Stream (in Kinshasa, Democratic Republic of Congo). Open Journal of Water Pollution and Treatment 2015(1): 8–24.
Raburu, P. O., J. B. Okeyo-Owuor & F. O. Masese, 2009. Macroinvertebrate-based index of biotic integrity (M-IBI) for monitoring the Nyando River, Lake Victoria Basin, Kenya. Scientific Research and Essays 4(12): 1468–1477.
Ruaro, R. & É. A. Gubiani, 2013. A scientometric assessment of 30 years of the Index of Biotic Integrity in aquatic ecosystems: applications and main flaws. Ecological Indicators 29: 105–110.
Sánchez-Montoya, M. M., M. R. Vidal-Abarca, T. Puntí, J. M. Poquet, N. Prat, M. Rieradevall & M. L. Suárez, 2009. Defining criteria to select reference sites in Mediterranean streams. Hydrobiologia 619(1): 39–54.
Schoolmaster, D. R., J. B. Grace & E. W. Schweiger, 2012. A general theory of multimetric indices and their properties. Methods in Ecology and Evolution 3(4): 773–781.
Shelly, S. Y., Z. B. Mirza & S. Bashir, 2011. Comparative ecological study of aquatic macroinvertebrates of mangla dam and chashma barrage wetland areas. Journal of Animal and Plant Sciences 21(SUPPL. 2): 340–350.
Silva, D. R. O., A. T. Herlihy, R. M. Hughes & M. Callisto, 2017. An improved macroinvertebrate multimetric index for the assessment of wadeable streams in the neotropical savanna. Ecological Indicators 81(June): 514–525.
Silveira, M. P., D. F. Buss, J. L. Nessimian & D. F. Baptista, 2006. Spatial and temporal distribution of benthic macroinvertebrates in a southeastern Brazilian river. Brazilian Journal of Biology 66(2B): 623–632.
Stoddard, J. L., D. P. Larsen, C. P. Hawkins, R. K. Johnson & R. H. Norris, 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications 16(4): 1267–1276.
Stoddard, J. L., A. T. Herlihy, D. V. Peck, R. M. Hughes, T. R. Whittier & E. Tarquinio, 2008. A process for creating multimetric indices for large-scale aquatic surveys. Journal of the North American Benthological Society 27(4): 878–891.
Strayer, D. & D. Dudgeon, 2010. Freshwater biodiversity conservation: recent progress and future challenges. Journal of the North American Benthological Society 29(1): 344–358.
ter Braak, C.J.F., P. Smilauer, 2002. CANOCO Reference Manual and Canoco Draw for Windows User's Guide: Software for Canonical Community Ordination (version 4.5). Microcomputer Power, Ithaca NY, USA
Tomanova, S., E. Goitia & J. Helešic, 2006. Trophic levels and functional feeding groups of macroinvertebrates in neotropical streams. Hydrobiologia 556(1): 251–264.
Tomanova, S., N. Moya & T. Oberdorff, 2008. Using macroinvertebrate biological traits for assessing biotic integrity of neotropical streams. River Research and Applications 24: 1230–1239.
Uriarte, M., C. B. Yackulic, Y. Lim & J. A. Arce-Nazario, 2011. Influence of land use on water quality in a tropical landscape: a multi-scale analysis. Landsc. Ecol. 26: 1151–1164.
Us, E. P. A., 2002. Methods for Evaluating Wetland Condition: Developing an Invertebrate Index of Biological Integrity for Wetlands. Office of Water, U.S. Environmental Protection Agency, Washington, DC.
Wang, L., J. Lyons & P. Kanehl, 1998. Development and evaluation of a habitat rating system for low gradient Wisconsin streams. North American Journal of Fisheries Management 18: 775–785.
Wiseman, C. D., 2003. Multi-Metric Index Development for Biological Monitoring in Washington State Streams. Washington State Department of Ecology, Washington.
Zaitchik, B. F., B. Simane, S. Habib, M. C. Anderson, M. Ozdogan & J. D. Foltz, 2012. Building climate resilience in the Blue Nile/Abay Highlands: a role for Earth system sciences. International Journal of Environmental Research and Public Health 9(2): 435–461.