



A predictive index based on environmental filters for the bioassessment of river basins without reference areas in Atlantic Forest biome, Brazil

Natália Freitas de Souza^{1,2*}, Darcílio Fernandes Baptista¹ & Daniel Forsin Buss¹

¹Fundação Oswaldo Cruz, Instituto Oswaldo Cruz, Laboratório de Avaliação e Promoção da Saúde Ambiental, Rio de Janeiro, RJ, Brasil

²Universidade Federal do Rio de Janeiro, Departamento de Ecologia, Rio de Janeiro, RJ, Brasil

*Corresponding author: Natália Freitas de Souza, e-mail: souzanf00@gmail.com

SOUZA N.F., BAPTISTA, D.F., BUSS D.F. A predictive index based on environmental filters for the bioassessment of river basins without reference areas in Atlantic Forest biome, Brazil. *Biota Neotropica*. 19(1): e20180601. <http://dx.doi.org/10.1590/1676-0611-BN-2018-0601>

Abstract: Biological assessments that use the reference condition approach are based on the concept of comparing a site's observed biology to sites where disturbance is minimal or absent. However, in many regions of the world, such areas are scarce or nonexistent. In this study, an alternative approach proposed by Chessman and Royal for bioassessment without reference areas based on environmental filters was tested in Brazil. This approach assumes that key environmental features act in the selection of potential colonists, from a regional pool of taxa, based on the ecological traits (tolerances) possessed by each taxon. We developed the approach by: 1) determining the regional pool, based on a large Atlantic Forest biome database; 2) selecting environmental filters (elevation, original vegetation and soil type); and 3) including information on the tolerance and preferences of aquatic insects to these filters. With this information we were able to determine the expected taxon under natural conditions and compare with observed taxon, developing a predictive index (Observed/Expected). Although the model was intended to predict the fauna in regions without reference sites, we included reference areas to test the model responsiveness, precision and sensitivity. Our results indicated that the index was able to discriminate impairment classes ($F=56.9$; $p<0,001$), it has high precision due to low standard deviation across reference sites values ($SD=0.098$) and high sensitivity due to the correlation with environmental variables that are sensitive to human alteration ($r=0.74$, $p<0.01$). Also, it was strongly correlated with multimetric indices developed for multiple watersheds in the state, showing agreement between the methods in relation to ecological quality classification. Even though the predictive index had performed well in our study, we make some considerations that may help to improve its sensitivity of similar methods that are being tested using the environmental filters approach.

Keywords: biomonitoring, environmental management, aquatic insects, macroinvertebrate, neotropical region.

Índice preditivo baseado em filtros ambientais para o biomonitoramento de rios em bacias sem áreas de referência no bioma Mata Atlântica, Brasil

Resumo: Avaliações biológicas que usam a abordagem de condição de referência se baseiam na comparação entre a biota observada em locais teste e a presente em locais íntegros ou minimamente impactados. Entretanto, em muitas regiões do mundo, estes locais são escassos ou inexistentes. Neste estudo, a abordagem alternativa proposta por Chessman e Royal, que consiste no biomonitoramento baseado em filtros ambientais sem o uso de áreas de referência, foi testado para o Brasil. Esta abordagem assume que características ambientais chave agem na seleção de colonizadores potenciais, a partir do pool regional de táxons, baseado nos traços ecológicos (tolerância) de cada táxon. A abordagem foi desenvolvida através de: (1) determinação do pool regional, baseado em uma ampla base de dados do bioma Mata Atlântica; (2) seleção de filtros ambientais (elevação, vegetação original e tipo de solo); e (3) inclusão de informações sobre a tolerância e preferência ambiental dos insetos aquáticos a esses filtros. Essas informações possibilitam determinar os táxons esperados sob condições naturais (ausência de impacto) e comparar com os táxons observados em cada local, desenvolvendo um índice preditivo (Observado/Esperado). Embora esta abordagem proponha a predição da fauna na ausência de áreas de referência, estas áreas foram incluídas com o propósito de testar a resposta, precisão e sensibilidade desta abordagem. Os resultados indicam que o índice desenvolvido foi capaz de distinguir diferentes classes de impacto ($F=56.9$; $p<0,001$), obteve alta precisão, o que foi observado com o baixo desvio padrão obtido para os valores de locais de referência ($SD=0.098$) e alta sensibilidade, observado através da correlação significativa com variáveis ambientais sensíveis a alterações antrópicas ($r=0.74$, $p<0.01$). Além disso, o índice foi fortemente correlacionado com índices multimétricos desenvolvidos especificamente para diferentes regiões hidrográficas, mostrando concordância entre diferentes métodos no que se refere a avaliação da integridade ecológica. Apesar do bom desempenho do índice preditivo neste estudo, nós fizemos algumas considerações para melhorar a sensibilidade e o potencial da abordagem de filtros ambientais em aplicações futuras.

Palavras-chave: avaliação biológica, gestão ambiental, insetos aquáticos, macroinvertebrados, região neotropical.

Introduction

Biological assessments are generally based on measurements of attributes of biological assemblages, which are often characterized and expressed as indices. Different types of indices have been developed (Hawkins et al. 2010, Feio & Poquet 2011, Herman & Nejadhashemi 2015), but most of these methods share the need to use a benchmark (i.e. a reference condition), from which the measured biological condition will be compared. The impairment of a site is defined as how much their biological attributes differ from those found at the benchmark's condition.

Ideally, reference benchmark conditions represent historical, pristine conditions (Hawkins et al. 2010). However, in many regions of the world due to a long history of anthropogenic alteration of the aquatic ecosystems, those conditions are scarce or absent. In such cases, where "minimally disturbed" or "historical condition" sites (sensu Stoddard et al. 2006) are missing, some protocols relax the reference criteria and use "best attainable" or "least disturbed" sites as reference (Stoddard et al. 2006). The misuse of this practice can generate a series of problems (Chessman 2006, Labay et al. 2015, Elias et al. 2015). First, these definitions are often arbitrary and inconsistent, thus impractical to apply them in broad spatial scales (Cao & Hawkins 2011). Second, these sites have human-altered hydrological, physical and chemical conditions, that also affects their corresponding biological attributes. Thus, an index based on those sites may differ from one developed using natural benchmarks. This practice should be applied with great prudence, otherwise can influence indices by incorporating increasing degradation into the modelling.

The need to implement bioassessment programs in regions where reference sites are absent or scarce have fostered the development and testing of alternative methods (Chessman & Royal 2004, Carter & Fend 2005, Stranko et al. 2005, Chessman 2006, Blocksom & Johnson 2009, Hawkins et al. 2010, Birk et al. 2012, Schoolmaster et al. 2013, Labay et al. 2015, Milošević et al. 2016, Elias et al. 2016). Some researchers point that an approach that does not require the use of reference sites should be explored (Olden et al. 2006, Feio et al. 2009, Feio & Poquet 2011, Elias et al. 2015). Recent literature reviews on bioassessment (Dolédéc & Statzner 2010, Hawkins et al. 2010) classified Chessman and Royal's (2004) approach, the Observed Proportion of Potential – OPP, as promising, but lacking sufficient validation. This approach is based on the environmental filters concept (Poff 1997), which assumes that key environmental features act in the selection of potential colonists, from a regional pool of taxa, based on the ecological traits (tolerances) possessed by each taxon. The premise of the approach is ecologically intuitive: a taxon from the regional pool (i.e. known to possibly occur in a given area) that possess environmental tolerances that fit the 'natural' environmental conditions would occur in a site if those 'natural' conditions are found. If the environmental conditions are changed in a way they extrapolate a taxon's tolerance, they act as filters excluding the taxon from that site. The use of multiple key environmental filters would allow predicting the composition of the community potentially occurring in a site (Poff 1997, Chessman & Royal 2004, Stranko et al. 2005, Chessman 2006). Anthropogenic impacts can be viewed as either modifying the natural filters (allowing more, fewer, or different taxa to pass) or creating additional filters (Chessman & Royal 2004). Thus, the comparison of the observed taxa with the expected taxa provides a measure of the impairment level of a site. It results in an

Observed/ Expected index (O/E index), similar to the type used in a RIVPACS and AUSRIVAS (Feio & Poquet 2011), but Chessman and Royal's approach identify taxa expected to occur in each typology based on life history information, rather than through reference-based predictive modeling. This can be considered typological approach (sensu Hawkins et al. 2010), because the expected taxa are identical to all streams within each type listed.

In this study, we test the potential for environmental filters approach to provide a basis for bioassessment of regions without reference sites. The environmental filters predicted the natural potential distribution of aquatic insect families, which was used as benchmarks for comparisons with observed assemblages. The predictive index was tested for its ability to determine the impairment condition of streams sampled in southeast Brazil. Specifically, to assess index performance we tested for its (1) responsiveness, in order to verify the ability to discriminate impairment classes; (2) precision, to know the precision of filters approach worked; (3) sensitivity to stressor gradient through evaluation of the relationship between the OPP and environmental variables that are highly sensitive to human alteration. Then, we compared OPP with existing multimetric indices in order to verify the agreement between methods in relation to ecological quality classification.

Material and Methods

1. Study Area

This study was carried out with data collected from seven of the nine main river basins of Rio de Janeiro state, southeast Brazil (Figure 1). The geomorphology of the state is composed of coastal plains separated by hills and two mountain chains that run parallel to the ocean (Serra do Mar, ranging from elevations 0-2,000 m.a.s.l and Serra da Mantiqueira, ranging from 800-2,500 m.a.s.l). The state's main river, rio Paraíba do Sul, runs in the valley formed between the two mountain chains at an elevation of about 800 m.a.s.l. According to a recent review of Köppen's climate classification for Brazil most of Rio de Janeiro state's mid-to-lowland portions (44%) is classified as tropical with a summer rainy season (Aw type), and the mountainous regions and plateaus classified as humid subtropical zones with hot summer, without dry seasons (Cfa type) or with a dry winter (Cwa type) (Alvares et al. 2013). The temperature oscillates between 15oC and 28oC and the mean annual rainfall is around 1000-1500mm. The Atlantic Forest, which originally covered virtually the entire region, now represents less than 12% of its original extent, and is mostly spread in the higher parts of the mountains and in remnants interspersed with agriculture and pasture (Ribeiro et al. 2011).

2. Development of the predictive model

A model to predict the natural distribution of macroinvertebrate based on environmental filters was built using three main types of information: 1) the regional pool of macroinvertebrate taxa; 2) the environmental filters; 3) the preference and tolerance of each taxon to the three filters.

The 'regional pool' of aquatic insects was based on a database of both published and unpublished information. This database consisted of over 400,000 individuals representing 100 macroinvertebrates families from 370 stream sites distributed in five neighboring states (Rio Grande

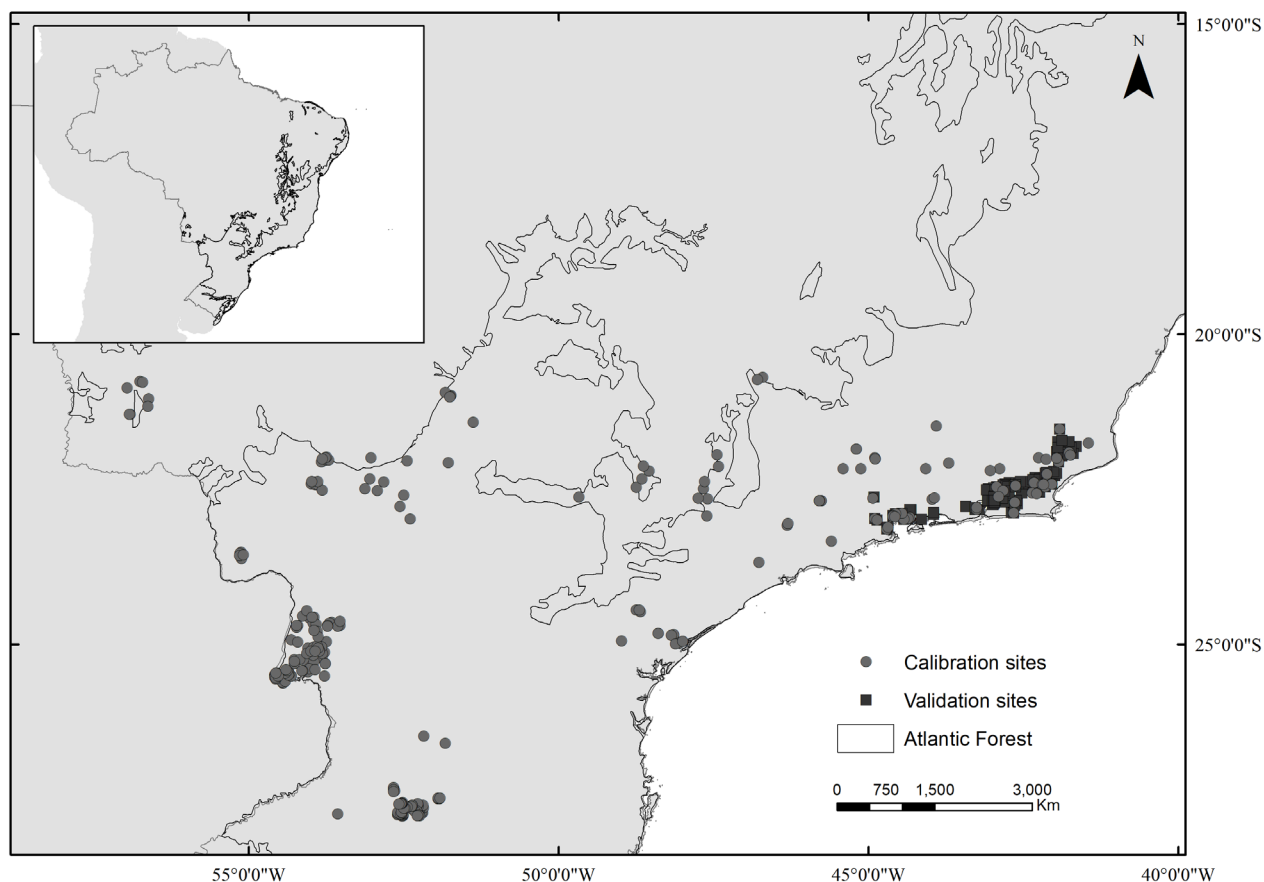


Figure 1. Locations of the macroinvertebrate sampling sites.

do Sul, Paraná, Mato Grosso do Sul, São Paulo and Rio de Janeiro) in south and southeast Atlantic Forest region. To define the regional pool, we identified the set of families that occurs in a region using the database. All families within the database were included as part of the regional pool and the list of families were confirmed with local experts. In next step, we excluded the rare families, which occurred in less than 10% of total sites in the database. These families that occurs with low frequencies could reduce model performance, because we couldn't define their preferences and environmental tolerances accurately or their absence in the following analyses are probably due to rarity rather anthropogenic degradation.

The environmental filters were used to generate a potential natural macroinvertebrate assemblage for each site by selecting from the regional pool the taxon which had the possibility (based on its environmental tolerance) that allowed it to occupy each test site. To obtain the local pool, we used a series of binary filters. Thus, a taxon was considered as a potential colonizer at a given site if it was listed in the regional pool and the site's characteristics were compatible with the environmental range described for the taxon. If a given taxon were filtered (excluded) by a single environmental attribute it was not considered to be potentially present in that location ("one-out-all-out" principle). We obtained a list of families potentially present in each combination of environmental filters (Appendix 1).

Ideally, the environmental filters should be relevant to the biological assemblages, must have enough available information regarding their biological tolerance ranges, and must not be strongly affected by stressors (Chessman & Royal 2004, Chessman 2006, Walsh 2006). In our study, the environmental variables that held these characteristics were elevation, original vegetation and soil type. These three filters are not directly influenced by human activities, they have been reported to influence aquatic insects assemblage distributions (Walsh 2006, Dudgeon 2012, Olson & Hawkins 2012), and there was enough information to map them and to calculate the tolerance range values for aquatic insects. The elevation range was divided into low (0-200m.a.s.l.), medium (>200-800m.a.s.l.) and high (>800m.a.s.l.), because these categories were reported as holding different aquatic insects assemblages in our studied region (Baptista et al. 2001, Henriques-Oliveira & Nessimian 2010). Original vegetation (Atlantic ombrophilous dense and semi-deciduous forest) and soil type (Cambisols, Ferralsols and Podzols) followed IBGE (<http://mapas.ibge.gov.br>).

The preferences and environmental tolerances of each taxon regarding the elevation range, vegetation and soil type were based on information obtained in the database and numerous publications. We used the database to define occurrence frequency of taxa in each category within the environmental filters. We used the database to define occurrence frequency of taxa in each category within the environmental

filters. Families which the occurrence was equal or more than 10% of the streams in that category was considered potentially present in the category of the environmental filter (Appendix 1). So, the probabilities were transformed to binary presence (1) and absent (0). We used the literature to avoid errors and the ones containing more information were Baptista et al. (2001); Hepp & Santos (2009); Henriques-Oliveira & Nessimian (2010); Suriano et al. (2011); Siqueira et al. (2012).

All taxa were identified to the lowest taxonomic level possible (mostly to genus level), but data were aggregated to family level for the modelling. The family level was chosen due to the lack of information about the preferences and environmental tolerances (ecological traits) on genus or lower levels. Also, this taxonomic level was shown to present similar responses to impairment as lower levels in Brazil (Buss & Vitorino 2010) and, due to practical reasons, family-level is recommended as a starting point for bioassessments in regions with taxonomic and resources constraints, such as Brazil (Buss et al. 2015). Other studies developing predictive modelling based on environmental filters (Chessman & Royal 2004, Stranko et al. 2005, Chessman 2006, Walsh et al. 2010, Davies et al. 2012) have found that those approaches may be applied successfully using family level.

3. Database used to apply and test the model

The dataset used to apply and test the model consisted of 106,088 individuals representing 67 families and 10 orders of insects, from 146 sites sampled in Rio de Janeiro state. The records from these 146 sites were excluded from model development so that they could be independently used to assess the model performance. Aquatic insects were sampled in streams of 1st to 4th orders (of which less than 15% were of 5nd to 6th orders), representing elevations from sea level to 1,700 m.a.s.l. and under different land uses (reference, pasture, agriculture or urban). The great majority of the samples were taken avoiding the wet season (thus, sampled from April to October) between the years of 2005 and 2010. All samples were taken and processed by the same research team. In the field, twenty samples (around 20m²) were collected proportionally to the microhabitats (substrates) available in each stream reach using a Kick sampler (30x30cm; 500 µm mesh size), following the multi-habitat method (Barbour et al. 1999). The percentage of available habitats was previously estimated by visual inspection and substrates with less than 5% of the site area were not sampled. Samples were obtained from a reach length of approximately 20 times the channel width. Samples were conserved in the field in 80% ethanol and taken to the laboratory for further inspection. In the laboratory, samples were washed to remove coarse organic matter, such as leaves and twigs and the remaining material was placed in a sub-sampler measuring 64 x 36 cm, divided into 24 quadrats, each measuring 10.5 x 8.5 cm, with area of 89.25 cm² (European patent number 2572576). Eight quadrats were chosen at random and processed entirely, following the procedures described in Oliveira et al. (2011a).

In the field, water was analyzed for pH (MPA 210p LabConte) and dissolved oxygen (mg/L; YSI 550A). Cooled samples were taken to the lab for further analysis. The parameters ammonia (mg/L NH₃), nitrate (mg/L NO₃), and total phosphorus (mg/L P-total) were analyzed using a spectrophotometer (HACH DR2500), following Standard Methods protocols (APHA 2000). Sampling sites were also classified in the field using the visual-based habitat assessment protocol (HAP;

Barbour et al. 1999). The HAP analyzes ten environmental parameters, such as substrate availability for colonization by benthic fauna, water velocity and embeddedness, channel condition, sediment deposition, margin stability and riparian vegetation extent and condition. For each parameter a score between 0 and 20 is assigned. Sites are classified according to the mean score obtained in the HAP, as follows: 0-5 "Poor", 5.1-9.9 "Marginal", 10-14.9 "Suboptimal" and 15-20 indicating an "Optimal" environmental condition (Barbour et al. 1999). Landscape variables (elevation, original vegetation and soil types) were obtained for each site using the ArcGIS 10.3 software and their corresponding digitalized 1:500,000 maps (IBGE <http://mapas.ibge.gov.br>).

Although our aim in this study was to develop a model to predict the fauna in regions lacking proper reference sites, we included reference areas to test the model. We hypothesize that if the model is robust enough it should yield higher scores to reference sites in comparison to other impairment conditions. Impairment classes were assigned based on physical, chemical and environmental parameters, the latter following Barbour et al. (1999). Sites were classified as reference if water had dissolved oxygen >6.0 mg/L, an "Optimal" or "Good" environmental condition according to the HAP, no sign of channelization locally or upstream and if <25% area upstream land-use were urban (based on recent satellite images). Most reference sites were within or in buffer zones of protected areas and were thus classified as "minimally disturbed" (sensu Stoddard et al. 2006). Impaired sites were classified if they had "Poor" condition according to HAP and if recent satellite images showed >40% of upstream area was affected by urban areas and/or agriculture. Intermediate sites had characteristics between these two classes. When in the field we noticed some intermediate sites had been reforested or were in process of recovery. Those sites were classified as "best attainable conditions" (sensu Stoddard et al. 2006), and we used this subset of intermediate sites to refine the model testing.

4. Comparison of observed and expected taxa

The suite of families observed at each site was compared with the suite that was attributed as potential colonizers of that site. This ratio (number of observed within the expected/total expected) is termed as "observed proportion of potential" (OPP; Chessman & Royal 2004). The OPP scores range from 0 to 1, where scores close to zero indicate few expected families were observed and scores close to one indicate the opposite. The lower the OPP score the higher the impairment level of a site.

5. Testing the model

The testing of the OPP was done threefold, aiming to verify: (1) responsiveness, if the OPP scores statistically discriminate sites of the four impairment classes (minimally disturbed reference, best attainable conditions, intermediate and impaired) using an ANOVA followed by a Tukey post-hoc; (2) precision, using standard deviation of scores across reference sites; (3) sensitivity, if OPP scores correlated with environmental variables that are highly sensitive to human alteration, which allow to define a stressor gradient; (4) agreement between the methods in relation to ecological quality classification, if OPP scores correlated with multimetric indices previously developed for river basins where those sites were located. To verify if the OPP scores were correlated to environmental variables, first a PCA were calculated

using environmental variables (ammonia, nitrate, P-total, and the HAP index) for 110 sites where all those informations were available. Prior to analysis, data were standardized by subtracting each value from its mean and dividing it by its standard deviation to reduce the effects of different scales used in the variables. Second, Pearson correlations were calculated using the OPP scores for each site and the obtained PCA first-axis values (environmental variables considered should be positively correlated in the PCA first-axis). To verify if the OPP scores were correlated with multimetric indices, Pearson correlations were calculated between the OPP score for each site and the indices: GMMI (Guapiaçu-Macacu Multimetric Index; Oliveira et al. 2011b), for 38 sites sampled in this basin; PPPMI (Paquequer-Piabanha-Preto Multimetric Index; Baptista et al. 2011), for 22 sites; MISB (Serra da Bocaina Mutimetric Index; Baptista et al. 2013) for 16 sites; IMMM (Macaé Multimetric Index; unpublished data) for 29 sites; and ECMI (East Coast Multimetric Index; Pereira et al. 2016), for 20 sites. Twenty-one sites were sampled in basins without prior developed indices and were not included in this analysis.

Results

Based on the criteria to determine impairment conditions, the 146 sampled sites were classified as follows: 35 reference sites, 20 with best attainable conditions, 55 intermediate and 36 sites classified as impaired. Physical and chemical parameters suggest a gradient of impairment related to organic origin (as the values of ammonia and P-total indicate) and to non-point source pollution and/or related to environmental degradation (based on the HAP index; Table 1). The HAP index classified reference sites as “optimal” or “good”, while impaired sites were classified as “regular” or “poor” environmental condition.

1. Environmental filters

From the 146 stream sites analyzed in this study, cambisols had the higher number of sites (104) among soil types, followed by podzols and ferralsols (26 and 16 sites, respectively). The great majority of sites (136) belonged to ombrophylous dense forest dominion, with the ten remaining belonging to the Atlantic semi-deciduous forest dominion. The elevation range separated sites in more even numbers: 51 sites from 0-200m.a.s.l., 60 sites in the 200-800m.a.s.l range and 35 sites >800m.a.s.l. It is important to notice that the distribution of streams do not follow the percentiles of each environmental condition in the state, being an artefact of the database. The three environmental filters were combined to determine the potential of occurrence of aquatic insects.

Sixty-seven families were identified as having the natural potential to occur at one or more sites. The combination with the higher number of families with potential to occur was cambisols covered by ombrophilous dense forest and in the 200-800m.a.s.l. elevational range (53 families; Table 2). The combinations with the lower numbers of families with potential to occur were those at low elevation (0-200m.a.s.l) covered by Atlantic semi-deciduous forest, both in ferralsols and podzols soil types (35 families; Table 2). Sampled sites represented most possible filters combinations. All combinations had at least one site sampled, with the exception of sites >800m.a.s.l. in Atlantic semi-deciduous forest on podzols and ferralsols, and in Atlantic ombrophilous dense forest on ferralsols (Table 2).

2. Testing the model

The OPP scores for the 146 sites ranged from 0.02 to 0.75. Reference sites had a high number of occasions where the expected families were also observed (E+O+, Table 3). This contributed for the highest OPP scores obtained by sites of this class (percentiles 25%-75% = 0.46 and 0.60, respectively; Figure 2). BAC sites scores were statistically similar to those of reference sites (ANOVA Tukey post-hoc test $p=0.11$), although 75% of BAC scores were lower than the median score of reference sites (Figure 2). On the other hand, in impaired sites, a high number of expected families were not observed (E+O-, Table 3). The lower number of E+O+ occasions resulted in lower OPP scores for impaired sites (percentiles 25%-75% = 0.12 and 0.32, respectively). Intermediate sites scored between the two extremes (ANOVA Tukey post-hoc test $p<0.01$ for all those pairs of data; Figure 2). The scores of reference sites had low standard deviation (0.54, SD= 0.098) which indicates high precision of filters approach. In general, sensitive taxa (Corydalidae, Grypopterygidae, Perlidae, Psephenidae and Pylalidae) and the shredders – group particularly vulnerable to riparian deforestation (e.g., Calamoceratidae, Leptoceridae and Tipulidae) – were observed with higher frequency in reference sites (Table 3). Those taxa, among others, had distinct presence patterns between reference and impaired sites. For example, Grypopterygidae was observed in 34 of the 35 reference sites it had a natural potential of occurring, but it was observed in only two of the 36 impaired sites it was expected. Some very abundant tolerant or moderately tolerant taxa were observed in most sites they were expected to occur (e.g., Elmidae, Chironomidae, Simuliidae, Baetidae, Leptohiphidae e Hydropsychidae), regardless the impairment condition (reference, BAC, intermediate or impaired; Table 3). In very few occasions, some families unassigned by the filters for occurring at a site were observed (E-O+, Table 3).

Table 1. Mean values (and standard deviation) of the physical, chemical and environmental parameters measured for stream sites classified as reference, best attainable condition (BAC), intermediate and impaired.

	Reference	BAC	Intermediate	Impaired
Habitat Assessment Protocol (HAP)	17.38 (± 2.39)	15.48 (± 3.62)	11.71 (± 4.77)	4.02 (± 2.04)
HAP classes	optimal-good	optimal-good	good-poor	regular-poor
NH ₃ -N (mg/L)	0.02 (± 0.02)	0.02 (± 0.02)	0.07 (± 0.25)	0.54 (± 0.93)
P-total (mg/L)	0.50 (± 0.22)	0.36 (± 0.07)	0.46 (± 0.22)	0.87 (± 0.69)
NO ₃ (mg/L)	1.99 (± 1.71)	2.60 (± 1.63)	1.90 (± 1.28)	1.59 (± 1.27)
Dissolved oxygen (mg/L)	6.83 (± 1.85)	7.80 (± 1.08)	7.15 (± 1.83)	5.38 (± 2.40)
pH	6.90 (± 0.45)	6.68 (± 0.29)	7.04 (± 0.51)	6.96 (± 0.45)

Table 2. Number of sampled sites and number of families with potential of occurring, according to environmental filters (soil types, original vegetation cover and elevation ranges).

Soil types	Original vegetation cover	Elevation (m)	Number of stream sites sampled	Number of families with potential of occurring
Cambisols	Atlantic ombrophilous dense forest	<200	31	40
		200-800	39	53
		>800	34	51
Ferralsols	Atlantic ombrophilous dense forest	0-200	4	36
		200-800	7	47
		>800	0	46
Ferralsols	Atlantic semi-deciduous forest	<200	1	35
		200-800	4	43
		>800	0	43
Podzols	Atlantic ombrophilous dense forest	<200	12	40
		200-800	8	51
		>800	1	50
Podzols	Atlantic semi-deciduous forest	<200	3	35
		200-800	2	45
		>800	0	44

Table 3. Number of occasions of aquatic insect families with/without a natural potential of occurrence in a site based on environmental filters (E+, E-), and observed/not-observed (O+, O-), at reference (**bold letters**) and impaired sites (*italic letters*). “E+O+” (number of occasions the family with potential was collected), “E+O-” (number of occasions the family with potential was not collected); “E-O+” (number of occasions the family with no potential was collected); “E-O-” (number of occasions the family with no potential was not collected).

	E+O+	<i>E+O+</i>	E+O-	<i>E+O-</i>	E-O+	<i>E-O+</i>	E-O-	<i>E-O-</i>
Aeshnidae	12	3	12	20	1	0	10	13
Baetidae	35	26	0	10	0	0	0	0
Belostomatidae	9	2	21	31	1	0	4	3
Blephariceridae	13	0	17	31	0	0	5	5
Caenidae	5	1	20	25	2	0	8	10
Calamoceratidae	34	4	1	32	0	0	0	0
Calopterygidae	21	18	14	18	0	0	0	0
Ceratopogonidae	18	13	17	23	0	0	0	0
Chironomidae	35	36	0	0	0	0	0	0
Coenagrionidae	9	5	26	31	0	0	0	0
Collembola	0	0	0	0	2	1	33	35
Cordullidae	3	1	18	22	1	3	13	10
Corixidae	0	0	4	5	0	0	31	31
Corydalidae	21	5	14	31	0	0	0	0
Culicidae	0	0	0	0	1	0	34	36
Dixidae	5	2	24	31	2	0	4	3
Dryopidae	4	0	20	23	2	1	9	12
Dytiscidae	1	0	23	23	1	0	10	13
Ecnomidae	0	0	2	3	0	0	33	33
Elmidae	35	27	0	9	0	0	0	0
Empididae	27	11	8	25	0	0	0	0
Ephrydidae	1	2	28	31	0	0	6	3
Euthyplociidae	1	0	17	20	1	0	16	16

A predictive index based on environmental filters

Continuation Table 3.

	E+O+	<i>E+O+</i>	E+O-	<i>E+O-</i>	E-O+	<i>E-O+</i>	E-O-	<i>E-O-</i>
Gerridae	0	0	0	0	3	0	32	36
Glossosomatidae	1	1	23	22	1	1	10	12
Gomphidae	17	12	18	24	0	0	0	0
Gryopterygidae	31	2	4	34	0	0	0	0
Gyrinidae	8	3	16	20	2	1	9	12
Hebridae	0	0	0	0	0	0	35	36
Helicopsychidae	24	3	11	33	0	0	0	0
Helotrephidae	9	2	1	8	9	1	16	25
Hydrobiosidae	16	1	15	35	2	0	2	0
Hydrophilidae	0	0	0	0	19	5	16	31
Hydropsychidae	35	24	0	12	0	0	0	0
Hydroptilidae	9	2	26	34	0	0	0	0
Leptoceridae	30	12	5	24	0	0	0	0
Leptohiphidae	35	23	0	13	0	0	0	0
Leptophlebiidae	31	15	4	21	0	0	0	0
Libellulidae	16	18	19	18	0	0	0	0
Lutrochidae	21	4	14	32	0	0	0	0
Megapodagrionidae	13	3	22	33	0	0	0	0
Mesoveliidae	1	0	9	11	7	1	18	24
Naucoridae	24	7	11	29	0	0	0	0
Noteridae	0	0	0	0	0	0	35	36
Notonectidae	0	0	0	0	0	0	35	36
Odontoceridae	13	3	11	20	3	0	8	13
Oligoneuriidae	0	0	9	10	0	0	26	26
Perilestidae	0	2	9	8	2	1	24	25
Perlidae	35	8	0	28	0	0	0	0
Philopotamidae	21	10	14	26	0	0	0	0
Pleidae	11	0	24	36	0	0	0	0
Polycentropodidae	4	2	31	34	0	0	0	0
Protoneuridae	3	1	16	21	2	1	14	13
Psephenidae	23	6	12	30	0	0	0	0
Psychodidae	4	11	31	25	0	0	0	0
Ptilodactylidae	0	0	20	21	2	0	13	15
Pyralidae	24	8	11	28	0	0	0	0
Scirtidae	0	0	0	0	2	0	33	36
Sericostomatidae	15	5	20	31	0	0	0	0
Simuliidae	34	31	1	5	0	0	0	0
Staphylinidae	16	9	19	27	0	0	0	0
Stratiomyidae	0	0	1	2	1	4	33	30
Syrphidae	0	0	0	0	0	0	35	36
Tabanidae	5	0	19	23	0	0	11	13
Tipulidae	25	9	10	27	0	0	0	0
Vellidae	30	13	5	23	0	0	0	0
Xyphocetronidae	0	0	0	0	1	0	34	36

Bold letters: Reference sites; *Italic letters:* Impaired sites.

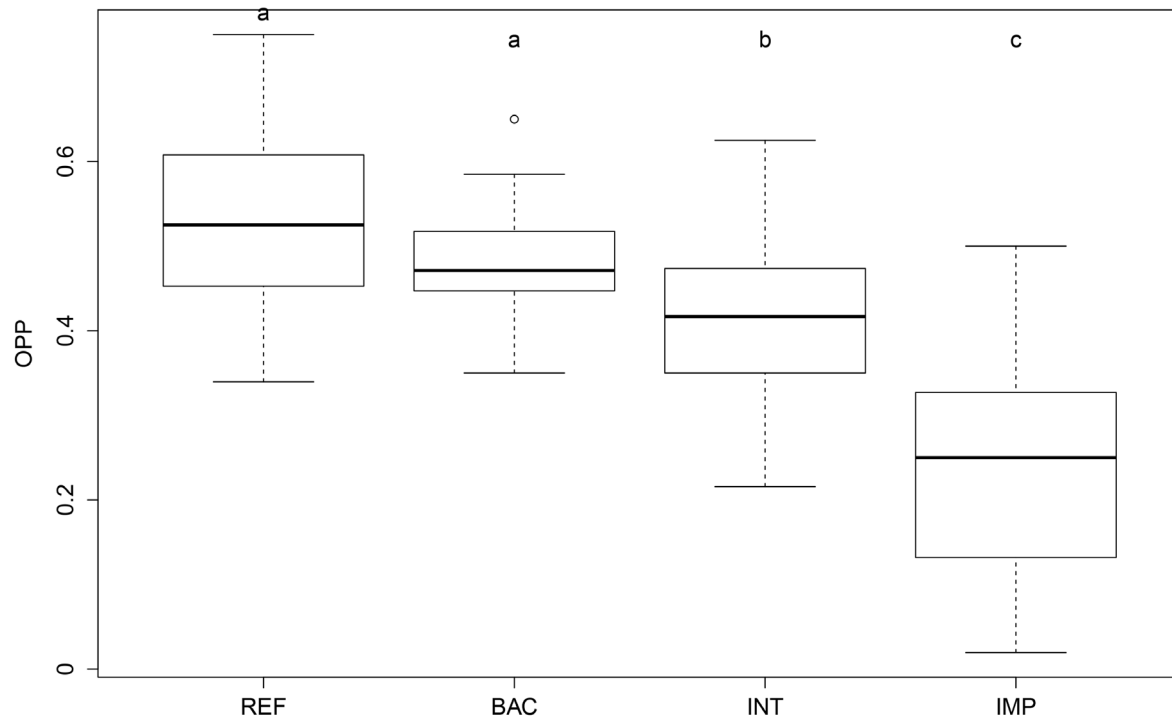


Figure 2. OPP scores calculated for the four impairment classes: Reference (Ref.), Best Attainable Condition sites (BAC), Intermediate (Inter.), and Impaired sites (Imp.). Different letters above the box-plots indicate significant difference, according to Tuckey post-hoc tests (ANOVA $F=56.9$; $p<0.001$).

The axis 1 of the environmental PCA represented 54.04% of the variance and it was the only significant one, according to the broken-stick model. PCA axis 1 discriminated Reference and BAC sites from intermediate and impaired conditions (Mann-Whitney test using PCA axis 1 scores, $p<0.01$ for all pairs except Ref x BAC, $p = 0.56$). The parameters ammonia and P-total had the highest eigenvalues (>0.50) for this axis. The OPP scores responded to this gradient of impairment conditions and were significantly correlated with the environmental PCA axis 1 values ($r=0.74$, $p<0.01$). The OPP scores were also highly correlated to all subsets of biological data, expressed by the multimetric indices calculated for each basin (Table 4).

Table 4. Pearson correlations (R and p-level) between the OPP scores and the multimetric indices for each stream basin.

Multimetric index	n	r	p-level
GMMI	38	0.82	<0.01
PPPMI	22	0.95	<0.01
MISB	16	0.72	<0.01
IMMM	29	0.70	<0.01
ECMI	20	0.92	<0.01

n = number of sites.

Discussion

The OPP index developed in this study was sensitive to discriminate a gradient of impairment conditions (reference, best attainable, intermediate and impaired sites; Figure 2). The OPP scores were

also correlated with a set of environmental variables, and it was strongly correlated with all multimetric indices developed for multiple watersheds in the state (Table 4). This latter result was unexpected, since those multimetric indices were developed separately for each watershed, incorporating their particular environmental and biological conditions. Chessman and Royal (2004) found that OPP was significantly correlated with different disturbance measures and with another biological index (SIGNAL2; Stream Invertebrate Grade Number Average Level). Chessman et al. (2006) reported the OPP had higher sensitivity to distinguish the impairment gradient than the AUSRIVAS O/E index (Australian River Assessment System Observed over Expected). Even though the OPP had performed very well in our study, we believe it could be improved if some of the following aspects is implemented: using a taxonomic level lower than family-level; using more refined environmental filters; and by incorporating more information about the auto-ecology of aquatic insects taxa.

1. Considerations on taxonomic level

The taxonomic level to be used for bioassessment and monitoring purposes is a topic of ongoing debate (e.g., Buss & Vitorino 2010, Mueller et al. 2013). Some researchers argue that the identification of aquatic insects to species-level have a higher sensitivity to detect small differences among sites (Heino 2014), but some others argue that in biomonitoring programs datasets are summarized in indices, which do not necessarily require species data and are often robust to taxonomic aggregation (Buss & Vitorino 2010, Whittier & Van Sickle 2010, Mueller et al. 2013). The original OPP index in Australia was also developed in family-level (Chessman & Royal 2004). Those authors

argued that this level was chosen because of the lack of taxonomic keys, due to practical reasons (lower cost and skill requirements), and because most bioassessments in the region has operated at family level. In Brazil and other parts of Latin America the situation is the same. The identification to the species-level is not always possible due to the limited taxonomic knowledge for many insect groups, and although taxonomic keys are being developed and becoming available (e.g., Domínguez & Fernandez 2009, Mugnai et al. 2010, Hamada et al. 2014) there is still a lack of taxonomic keys for most regions and an insufficient and decreasing number of experts in taxonomy. Still, if these limitations may be overcome, the use of lower taxonomic levels (e.g., genus) could hypothetically increase the precision of the model. Some neotropical families are composed by genera that differ considerably regarding their tolerance to stressors and ecological preferences (e.g., Chironomidae – Roque et al. 2010; Baetidae – Buss & Salles 2007). In such cases, a family may have been assigned as having the “potential of occurrence” based on the distribution and preferences of a given genus ‘A’, while the collected genus in one location was the ‘B’, incorrectly increasing the OPP score for that site. Higher taxonomic levels (e.g., family) could be used more consistently for families composed by few genera and/or where the genera share similar ecological traits. Since the tolerance and biological traits of many species or genera are still missing (see discussion below), an interesting starting point could be using mixed taxonomic levels – using the criteria above for the decision – instead of losing information by using data aggregated to family level.

2. Considerations on environmental filters

In our study, OPP scores for some reference areas were relatively low (OPP = 0.34 the lowest score for this condition), juxtaposing with some impaired condition scores (OPP = 0.46, the highest score for this condition). For the OPP, we used environmental filters at a “landscape-” or “regional-scale” (elevation, original vegetation and soil types). In part, we chose these filters based on Omernik’s (1987) ecoregional approach, which have already been shown to have a good response by the macroinvertebrate (Verdonschot & Nijboer 2004) and fish fauna (Pinto et al. 2009). Also, these features are not subject to human disturbance, which is an ideal condition for the use of the filters approach (Poff 1997, Walsh 2006). On the other hand, some researchers argue that macroinvertebrate faunal predictions based on filters should be based on multiple spatial scales, mixing both “regional” and more “local” features (Poff 1997, Olden et al. 2006). This is supported by research based on diversity partitioning analyses, which indicate that both scales help to explain macroinvertebrate assemblage distribution in Brazil (Ligeiro et al. 2010, Macedo et al. 2014). However, the use of local environmental filters has some limitations. There is an increasing difficulty for extracting patterns from large databases since local information are not widely available and may not be extracted by GIS procedures – rather, they must be gathered on site – and most importantly, that local features are strongly affected by anthropogenic impacts, even relatively mild ones. The problem of using predictor variables that are affected by human disturbance is that the modelling will incorporate part of the disturbance, and it may misclassify sites with some degradation as being undegraded (Walsh 2006). To correct this, simulation models of human influences may enable to predict natural features and they could be used instead of the observed features (Chessman et al. 2006). The use of GIS-based information coupled with

Self-organizing maps (SOM) techniques may aid further the prediction of local-level features based on larger-scale data (e.g., Davies et al. 2000, Snelder et al. 2011).

3. Considerations on taxa preferences and tolerances

The virtual lack of information on biological traits hinders the application of environmental filters approach in many regions (Helm et al. 2015). In recent years, there was a greater focus on this subject worldwide (Van den Brink et al. 2011, Culp et al. 2011, Mueller et al. 2013). Still, more data on the autecological characteristics, ecological preferences and biological traits are necessary, especially for neotropical benthic fauna. The way forward includes the compilation of bio/ecological traits in large databases, such as those developed for Europe (Bonada & Dolédec 2011, Schmidt-Kloiber and Hering 2015), North America (Vieira et al. 2006), South America (Tomanova & Usseglio-Polatera 2007), and New Zealand (Dolédec et al. 2006). Several researchers highlight the potential of using biological traits as metrics for bioassessments, since they can potentially reveal additional information concerning ecosystem properties beyond taxonomic composition (Poff et al. 2006, Dolédec & Statzner 2008, Culp et al. 2011).

In our study, we found few occasions where an unexpected family by the environmental filters analysis was, in fact, observed (E-O+, Table 3). Chessman & Royal (2004) argue that such cases may occur because of the model was based on inappropriate filters or incorrect information on biological traits, or according to Helm et al. (2015) they can indicate invasion of non-native species from different geographic regions or opportunistic species that historically do not occupy this particular habitat. Very little information is available on the historical distributions of the aquatic insect assemblages included in this study. However, this should be a concern when using this approach for the whole macroinvertebrate assemblages, especially given the many cases of mollusk and mosquito exotic species invasions in South America (e.g. Thomaz et al. 2014).

4. Considerations on reference sites

Based on our judgment, the “best attainable areas” (BAC) in this study would be classified as intermediate condition, or reference “on the fringe”. The OPP approach was sensitive to detect this subtle change: although BAC sites had OPP scores statistically similar to reference, their scores were lower than most reference sites’ scores. Also, both conditions were statistically different than intermediate and impaired conditions (Figure 2). These results indicate that best management practices (BMP) in these areas, such as maintaining or recovering the riparian zone and the natural stream habitat features, were efficient to partially sustain the aquatic insect fauna. We hope this can stimulate and guide managers toward BMPs, even though the recovery could be slow (Meals et al. 2010), and limited (Harding et al. 1998).

5. Further developments on environmental filters approach

Some alternative approaches were developed for the assessment of streams without reference areas. In general, those methods select from the population of sites the “least disturbed” ones – using physical, chemical and/or biological data – and use them as benchmarks. This can be achieved by using a pre-defined percentile of sites (Blocksom & Johnson 2009) or by multivariate analyses with abiotic data to

determine the “best” sites (Birk et al. 2012), or even through partitioned least squares regressions to identify the highest biotic scores along an impairment gradient (Carter & Fend 2005). Those approaches, however, include some disturbance in their reference database, thus reducing the sensitivity of indices to detect impairment in test sites. More recently, methods based on artificial intelligence (e.g., species distribution models) were employed to predict “null assemblages” from large historical databases and then compared with contemporary distributions to determine the ecological status of streams (Davies et al. 2012, Labay et al. 2015, Milošević et al. 2016). These tools rely on large, comparable and sufficiently representative databases, something that unfortunately are not available for many biological groups and regions around the globe. Still, similar approaches based on environmental filters, like the OPP, are promising alternatives for the bioassessment without true reference sites, and they should be tested further.

Supplementary material

The following online material is available for this article:
Appendix 1

Acknowledgments

We thank the research team from LAPSA/IOC/FIOCRUZ that helped building the database used in this study. We are also thankful to F.O. Roque and L. Hepp for sharing part of the data for the Atlantic forest database. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, CNPq/PROEP No 400107/2011-2, PAEF IOC-008-FIO-15-8 and PAEF IOC-023-FIO-18-2-4.

Author Contributions

Natália Freitas de Souza - Substantial contribution in the concept and design of the study, contribution to data collection, contribution to data analysis and interpretation, contribution to manuscript preparation, contribution to critical revision, adding intellectual content.

Darcílio Fernandes Baptista - Contribution to manuscript preparation and to critical revision, adding intellectual content.

Daniel Forsin Buss - Substantial contribution in the concept and design of the study, contribution to data collection, contribution to data analysis and interpretation, contribution to manuscript preparation, contribution to critical revision, adding intellectual content.

Conflicts of interest

The authors declare that they have no conflict of interest related to the publication of this manuscript.

References

ALVARES, C. A., J. L. STAPE, P. C. SENTELHAS, J. L. DE MORAES GONÇALVES, AND G. SPAROVEK. 2013. Kooppen's climate classification map for Brazil. *Meteorologische Zeitschrift* 22:711–728.

BAPTISTA, D. F., D. F. BUSS, L. F. M. DORVILLÉ, AND J. L. NESSIMIAN. 2001. Diversity and habitat preference of aquatic insects along the longitudinal gradient of the macaé river basin, Rio de Janeiro, Brazil. *Revista Brasileira de Biologia* 61:249–258.

BAPTISTA, D. F., R. S. G. DE SOUZA, C. A. VIEIRA, R. MUGNAI, A. S. SOUZA, AND 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 (Curitiba)* 28:619–628.

BAPTISTA, D. F., A. L. HENRIQUES-OLIVEIRA, R. OLIVEIRA, R. MUGNAI, J. L. NESSIMIAN, AND D. F. BUSS. 2013. Development of a benthic multimetric index for the Serra da Bocaina bioregion in Southeast Brazil. *Brazilian Journal of Biology* 73:573–583.

BARBOUR, M. T., J. GERRITSEN, B. D. SNYDER, AND J. B. STRIBLING. 1999. Rapid bioassessment protocols for use in streams and wadeable rivers. USEPA, Washington.

BIRK, S., L. VAN KOUWEN, AND N. WILLBY. 2012. Harmonising the bioassessment of large rivers in the absence of near-natural reference conditions - a case study of the Danube River. *Freshwater Biology* 57:1716–1732.

BLOCKSOM, K. A., AND B. R. JOHNSON. 2009. Development of a regional macroinvertebrate index for large river bioassessment. *Ecological Indicators* 9:313–328.

BONADA, N., AND S. DOLÉDEC. 2011. Do mediterranean genera not included in Tachet et al. 2002 have mediterranean trait characteristics? *Limnetica* 30:129–142.

BUSS, D. F., D. M. CARLISLE, T. S. CHON, J. CULP, J. S. HARDING, H. E. KEIZER-VLEK, W. A. ROBINSON, S. STRACHAN, C. THIRION, AND R. M. HUGHES. 2015. Stream biomonitoring using macroinvertebrates around the globe: a comparison of large-scale programs. *Environmental Monitoring and Assessment* 187:4132.

BUSS, D. F., AND F. F. SALLES. 2007. Using Baetidae species as biological indicators of environmental degradation in a Brazilian river basin. *Environmental Monitoring and Assessment* 130:365–372.

BUSS, D. F., AND A. S. VITORINO. 2010. Rapid Bioassessment Protocols using benthic macroinvertebrates in Brazil: evaluation of taxonomic sufficiency. *Journal of the North American Benthological Society* 29:562–571.

CAO, Y., AND C. P. HAWKINS. 2011. The comparability of bioassessments: a review of conceptual and methodological issues 1. *Journal of the North American Benthological Society* 30:680–701.

CARTER, J. L., AND S. V. FEND. 2005. Setting Limits: The Development and Use of Factor-Ceiling Distributions for an Urban Assessment Using Macroinvertebrates. *American Fisheries Society symposium* 47:179–191.

CHESSMAN, B. C. 2006. Prediction of riverine fish assemblages through the concept of environmental filters. *Marine and Freshwater Research* 57:601–609.

CHESSMAN, B. C., L. A. THURTELL, AND M. J. ROYAL. 2006. Bioassessment in a harsh environment: A comparison of macroinvertebrate assemblages at reference and assessment sites in an Australian Inland River System. *Environmental Monitoring and Assessment* 119:303–330.

CHESSMAN, B. C., AND M. J. ROYAL. 2004. Bioassessment without reference sites: use of environmental filters to predict natural assemblages of river macroinvertebrates. *Journal of the North American Benthological Society* 23:599–615.

CULP, J. M., D. G. ARMANINI, M. J. DUNBAR, J. M. ORLOFSKE, N. L. POFF, A. I. POLLARD, A. G. YATES, AND G. C. HOSE. 2011. Incorporating traits in aquatic biomonitoring to enhance causal diagnosis and prediction. *Integrated Environmental Assessment and Management* 7:187–97.

DAVIES, N. M., R. H. NORRIS, AND M. C. THOMS. 2000. Prediction and assessment of local stream habitat features using large-scale catchment characteristics. *Freshwater Biology* 45:343–369.

DAVIES, P. E., M. J. STEWARDSON, T. J. HILLMAN, J. R. ROBERTS, AND M. C. THOMS. 2012. Sustainable Rivers Audit 2: The ecological health of rivers in the Murray–Darling Basin at the end of the Millennium Drought (2008–2010). Murray–Darling Basin Authority, Australia.

DOLÉDEC, S., N. PHILLIPS, M. SCARSBROOK, R. H. RILEY, AND C. R. TOWNSEND. 2006. Comparison of structural and functional approaches to determining landuse effects on grassland stream invertebrate communities. *Journal of the North American Benthological Society* 25:44–60.

- DOLÉDEC, S., AND B. STATZNER. 2008. Invertebrate traits for the biomonitoring of large European rivers: An assessment of specific types of human impact. *Freshwater Biology* 53:617–634.
- DOLÉDEC, S., AND B. STATZNER. 2010. Responses of freshwater biota to human disturbances: contribution of J-NABS to developments in ecological integrity assessments. *Journal of the North American Benthological Society* 29:286–311.
- DOMÍNGUEZ, E. AND H.R. FERNÁNDEZ. 2009. Macroinvertebrados bentónicos sudamericanos: Sistemática y Biología. Fundación Miguel Lillo, Argentina.
- DUDGEON, D. 2012. Responses of benthic macroinvertebrate communities to altitude and geology in tributaries of the Sepik River (Papua New Guinea): the influence of taxonomic resolution on the detection of environmental gradients. *Freshwater Biology* 57:1794–1812.
- ELIAS, C. L., A. R. CALAPEZ, S. F. ALMEIDA, AND M. J. FEIO. 2015. Determining useful benchmarks for the bioassessment of highly disturbed areas based on diatoms. *Limnologia - Ecology and Management of Inland Waters* 51:83–93.
- ELIAS, C. L., CALAPEZ, A. R., ALMEIDA, S. F. P., CHESSMAN, B., SIMÕES, N., & FEIO, M. J. Predicting reference conditions for river bioassessment by incorporating boosted trees in the environmental filters method. 2016. *Ecological indicators* 69: 239-251.
- FEIO, M. J., R. H. NORRIS, M. A. S. GRAÇA, AND S. NICHOLS. 2009. Water quality assessment of Portuguese streams: Regional or national predictive models?. *Ecological Indicators* 9:791–806.
- FEIO, M. J., AND J. M. POQUET. 2011. Predictive Models for Freshwater Biological Assessment: Statistical Approaches, Biological Elements and the Iberian Peninsula Experience: A Review. *International Review of Hydrobiology* 96:321–346.
- HAMADA, N., NESSIMIAN, J.L. AND QUERINO, R.B. 2014. Insetos aquáticos na Amazônia brasileira: taxonomia, biologia e ecologia. 1ed. Manaus: Editora INPA. 724p.
- HARDING, J.S., E.F. BENFIELD, P.V. BOLSTAD, G.S. HELFMAN, AND E.B.D. JONES. 1998. Stream biodiversity: The ghost of land use past. *Proceedings of National Academy of Sciences* 95(25): 14843–14847.
- HAWKINS, C. P., J. R. OLSON, AND R. A. HILL. 2010. The reference condition: predicting benchmarks for ecological and water-quality assessments. *Journal of the North American Benthological Society* 29:312–343.
- HEINO, J. 2014. Taxonomic surrogacy, numerical resolution and responses of stream macroinvertebrate communities to ecological gradients: Are the inferences transferable among regions?. *Ecological Indicators* 36:186–194.
- HELM, A., M. ZOBEL, A. T. MOLES, R. SZAVA-KOVATS, AND M. PÄRTEL. 2015. Characteristic and derived diversity: implementing the species pool concept to quantify conservation condition of habitats. *Diversity and Distributions* 21:711–721.
- HENRIQUES-OLIVEIRA, A. L., AND J. L. NESSIMIAN. 2010. Aquatic macroinvertebrate diversity and composition in streams along an altitudinal gradient in Southeastern Brazil. *Biota Neotropica* 10:115–128.
- HEPP, L. U., AND S. SANTOS. 2009. Benthic communities of streams related to different land uses in a hydrographic basin in southern Brazil. *Environmental Monitoring and Assessment* 157:305–318.
- HERMAN, M. R., AND A. P. NEJADHASHEMI. 2015. A review of macroinvertebrate- and fish-based stream health indices. *Ecohydrology & Hydrobiology*:1–15.
- LABAY, B. J., HENDRICKSON, D. A., COHEN, A. E., BONNER, T. H., KING, R. S., KLEINSASSER, W. LINAM, AND K. O. WINEMILLER. 2015. Can Species Distribution Models Aid Bioassessment when Reference Sites are Lacking? Tests Based on Freshwater Fishes. *Environmental Management* 56:835–846.
- LIGEIRO, R., A. S. MELO, AND M. CALLISTO. 2010. Spatial scale and the diversity of macroinvertebrates in a Neotropical catchment. *Freshwater Biology* 55:424–435.
- MACEDO, D. R., R. M. HUGHES, R. LIGEIRO, W. R. FERREIRA, M. A. CASTRO, N. T. JUNQUEIRA, D. R. OLIVEIRA, K. R. FIRMIANO, P. R. KAUFMANN, P. S. POMPEU, AND M. CALLISTO. 2014. The relative influence of catchment and site variables on fish and macroinvertebrate richness in cerrado biome streams. *Landscape Ecology* 29:1001–1016.
- MEALS, D.W., S.A. DRESSING, T.E. DAVENPORT. 2010. Lag time in water quality response to best management practices: A review. *Journal of Environmental Quality* 39:85–96.
- MILOŠEVIĆ, D., D. ČERBA, J. SZEKERES, B. CSÁNYI, B. TUBIĆ, V. SIMIĆ, AND M. PAUNOVIĆ. 2016. Artificial neural networks as an indicator search engine: The visualization of natural and man-caused taxa variability. *Ecological Indicators* 61:777–789.
- MUELLER, M., J. PANDER, AND J. GEIST. 2013. Taxonomic sufficiency in freshwater ecosystems: effects of taxonomic resolution, functional traits, and data transformation. *Freshwater Science* 32:762–778.
- MUGNAI, R., J.L. NESSIMIAN, D.F. BAPTISTA. 2010. Manual de identificação de macroinvertebrados aquáticos do Rio de Janeiro. Ed. Technical Books. Rio de Janeiro.
- OLDEN, J. D., N. L. POFF, AND B. P. BLEDSOE. 2006. Incorporating ecological knowledge into ecoinformatics: An example of modeling hierarchically structured aquatic communities with neural networks. *Ecological Informatics* 1:33–42.
- OLIVEIRA, R. B., R. MUGNAI, C.M. CASTRO, AND D. F. BAPTISTA. 2011a. Determining subsampling effort for the development of a rapid bioassessment protocol using benthic macroinvertebrates in streams of Southeastern Brazil. *Environmental Monitoring and Assessment* 175:75–85.
- OLIVEIRA, R. B., D. F. BAPTISTA, R. MUGNAI, C. M. CASTRO, AND R. M. HUGHES. 2011b. Towards rapid bioassessment of wadeable streams in Brazil: Development of the Guapiaçu-Macau Multimetric Index (GMMI) based on benthic macroinvertebrates. *Ecological Indicators* 11:1584–1593.
- OLSON, J. R., AND C. P. HAWKINS. 2012. Predicting natural base-flow stream water chemistry in the western United States. *Water Resources Research* 48:1–19.
- OMERNIK, J. M. 1987. Ecoregions of the Conterminous United States. *Annals of the Association of American Geographers* 77:118–125.
- PINTO, B. C. T., F. G. ARAUJO, V. D. RODRIGUES, AND R. M. HUGHES. 2009. Local and ecoregion effects on fish assemblage structure in tributaries of the Rio Paraíba do Sul, Brazil. *Freshwater Biology* 54:2600–2615.
- POFF, N. L. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. *Journal of the North American Benthological Society* 16:391–409.
- POFF, N. L., J. D. OLDEN, N. K. VIEIRA, D. S. FINN, M. P. SIMMONS, AND B. C. KONDRATIEFF. 2006. Functional trait niches of North American lotic insects: traits-based ecological applications in light of phylogenetic relationships. *Journal of the North American Benthological Society* 25:730–755.
- PEREIRA, P. S., SOUZA, N. F., BAPTISTA, D. F., OLIVEIRA, J. L., BUSS, D. F. 2016. Incorporating natural variability in the bioassessment of stream condition in the Atlantic Forest biome, Brazil. *Ecological Indicators*, 69, 606-616.
- RIBEIRO, M. C., A. C. MARTENSEN, J. P. METZGER, M. TABARELLI, F. SACARANO AND M. J. FORTIN. 2011. The Brazilian Atlantic Forest: a shrinking biodiversity hotspot. *Biodiversity hotspot: Distribution and Protection of Conservation Priority Areas*. Springer Berlin Heidelberg.
- ROQUE, F. O., T. SIQUEIRA, L. M. BINI, M. C. RIBEIRO, L. R. TAMBOSI, G. CIOCHETI, AND S. TRIVINHO-STRIXINO. 2010. Untangling associations between chironomid taxa in Neotropical streams using local and landscape filters. *Freshwater Biology* 55:847–865.
- SCHMIDT-KLOIBER, A., AND D. HERING. 2015. www.freshwaterecology.info - An online tool that unifies, standardises and codifies more than 20,000 European freshwater organisms and their ecological preferences. *Ecological Indicators* 53:271–282.

- SIQUEIRA, T., L. M. BINI, F. O. ROQUE, S. R. MARQUES COUCEIRO, S. TRIVINHO-STRIXINO, AND K. COTTENIE. 2012. Common and rare species respond to similar niche processes in macroinvertebrate metacommunities. *Ecography* 35:183–192.
- SNELDER, T. H., N. LAMOUREUX, AND H. PELLA. 2011. Empirical modelling of large scale patterns in river bed surface grain size. *Geomorphology* 127:189–197.
- STODDARD, J. L., D. P. LARSEN, C. P. HAWKINS, R. K. JOHNSON, AND R. H. NORRIS. 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. *Ecological Applications* 16:1267–76.
- STRANKO, S. A., M. K. HURD, AND R. J. KLAUDA. 2005. Applying a large, statewide database to the assessment, stressor diagnosis, and restoration of stream fish communities. *Environmental Monitoring and Assessment* 108:99–121.
- SURIANO, M. T., A. A. FONSECA-GESSNER, F. O. ROQUE, AND C. G. FROEHLICH. 2011. Choice of macroinvertebrate metrics to evaluate stream conditions in Atlantic Forest, Brazil. *Environmental Monitoring and Assessment* 175:87–101.
- THOMAZ, S. M., K. E. KOVALENKO, J. E. HAVEL, AND L. B. KATS. 2014. Aquatic invasive species: general trends in the literature and introduction to the special issue. *Hydrobiologia* 746:1–12.
- TOMANOVA, S., AND P. USSEGLIO-POLATERA. 2007. Patterns of benthic community traits in neotropical streams: relationship to mesoscale spatial variability. *Fundamental and Applied Limnology/Archiv für Hydrobiologie* 170:243–255.
- VAN DEN BRINK, P. J., A. C. ALEXANDER, M. DESROSIERS, W. GOEDKOOP, P. L. GOETHALS, M. LIESS, AND S. D. DYER. 2011. Traits-based approaches in bioassessment and ecological risk assessment: Strengths, weaknesses, opportunities and threats. *Integrated Environmental Assessment and Management* 7:198–208.
- VERDONSCHOT, P. F.M., AND R. C. NIJBOER. 2004. Testing the European stream typology of the Water Framework Directive for macroinvertebrates. *Hydrobiologia* 516:35–54.
- VIEIRA, N. K. M., N. L. POFF, D. M. CARLISLE, S. R. MOUTON, M. L. KOSKI, AND B. C. KONDRATIEFF. 2006. A Database of Lotic Invertebrate Traits for North America. U.S. Geological Survey, Virginia. (Available from: <http://pubs.water.usgs.gov/ds187>)
- WALSH, C. J. 2006. Biological indicators of stream health using macroinvertebrate assemblage composition: A comparison of sensitivity to an urban gradient. *Marine and Freshwater Research* 57:37–47.
- WALSH, C. J., E. M. WALLIS, AND M. J. STEWARDSON. 2010. Macroinvertebrate Filters Project stage 3: revision and refinement of the Filters V2 macroinvertebrate analytical model. Murray-Darling Basin Authority, Australia.
- WHITTIER, T. R., AND J. VAN SICKLE. 2010. Macroinvertebrate tolerance values and an assemblage tolerance index (ATI) for western USA streams and rivers. *Journal of the North American Benthological Society* 29:852–866.

Received: 20/06/2018

Revised: 25/10/2018

Accepted: 17/11/2018

Published online: 11/04/2019