Ecological Indicators 78 (2017) 437–448

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Comparability of macroinvertebrate biomonitoring indices of river health derived from semi-quantitative and quantitative methodologies

Nicholas C. Everall[a,*], Matthew F. Johnson[b], Paul Wood[c], Andrew Farmer[a], Robert L. Wilby[c], Nick Measham[d]

[a] Aquascience Consultancy Limited, Chesterfield, Derbyshire, S42 7JS, UK
[b] School of Geography, University of Nottingham, NG7 2RD, UK
[c] Centre for Hydrological and Ecosystem Science, Department of Geography, Loughborough University, Leicestershire, LE11 3TU, UK
[d] Salmon and Trout Conservation, Burgate Manor, Fordingbridge, Hampshire, UK

ARTICLE INFO

Article history:
Received 4 August 2016
Received in revised form 19 February 2017
Accepted 19 March 2017
Available online 31 March 2017

Keywords:
Macroinvertebrate
Species richness
Biological monitoring
Biotic index
River

ABSTRACT

Aquatic macroinvertebrates have been the basis for one of the primary indicators and a cornerstone of lotic biomonitoring for over 40 years. Despite the widespread use of lotic invertebrates in statutory biomonitoring networks, scientific research and citizen science projects, the sampling methodologies employed frequently vary between studies. Routine statutory biomonitoring has historically relied on semi-quantitative sampling methods (timed kick sampling), while much academic research has favoured fully quantitative methods (e.g. Surber sampling). There is an untested assumption that data derived using quantitative and semi-quantitative samples are not comparable for biomonitoring purposes. As a result, data derived from the same site, but using different sampling techniques, have typically not been analysed together or directly compared. Here, we test this assumption by comparing a range of biomonitoring metrics derived from data collected using timed semi-quantitative kick samples and quantitative Surber samples from the same sites simultaneously. In total, 39 pairs of samples from 7 rivers in the UK were compared for two seasons (spring and autumn). We found a strong positive correlation (r_s = +0.84) between estimates of taxa richness based on ten Surber sub-samples and a single kick sample. The majority of biomonitoring metrics were comparable between techniques, although only fully quantitative sampling allows the density of the community (individual m^{-2}) to be determined. However, this advantage needs to be balanced alongside the greater total sampling time and effort associated with the fully quantitative methodology used here. Kick samples did not provide a good estimate of relative abundance of a number of species/taxa and, therefore, the quantitative method has the potential to provide important additional information which may support the interpretation of the biological metrics.

1. Introduction

Rivers and the ecological communities they support comprise some of the most biodiverse habitats on the globe but are also some of the most degraded as a result of anthropogenic activity (Dudgeon et al., 2006; Carpenter et al., 2011). River habitats and their ecosystems are threatened by ongoing human development (Vörösmarty et al., 2010), including the modification of channel morphology, dredging, changes to catchment land-use, pollution from diffuse and point sources, invasion by alien species, and alterations of the flow regime from abstraction, damming and flood risk management (Carpenter et al., 2011). The historic degradation of rivers has prompted the development of a range of biological monitoring tools to survey and quantify anthropogenic stressors over the past 40 years (e.g., Hering et al., 2004) and underpin calls to restore and improve the ecological health of lotic ecosystems (e.g., Geist, 2011).

In order to quantify trends in the health of riverine environments, the response of an organism or community is often characterised as a metric based on their known tolerances to 'stressors'. Biological monitoring, or biomonitoring, can be used to assess the effect of a known change to the state of a system by comparing the ecological community before and after the change or to routinely check compliance to nationally/internationally recog-

* Corresponding author.
E-mail address: rnaquaconsult@aol.com (N.C. Everall).

http://dx.doi.org/10.1016/j.ecolind.2017.03.040
1470-160X/

nised standards, such as the legal requirement for all waterbodies in the European Union to achieve 'Good Ecological Status' under the Water Framework Directive. The taxonomic resolution of such indices varies from family-level metrics that give broad indications of water quality (e.g., Walley and Hawkes 1997) to species/genus-level metrics that can provide information about specific stressors (Hubler et al., 2016); although some can be used at different taxonomic resolutions (Monk et al., 2012). Other metrics use higher resolutions; for example, the phenology of species or groups of species can be used to assess the impacts of climate change (Everall et al., 2015; Thackeray et al., 2016).

Aquatic macroinvertebrates are a fundamental component of freshwater ecosystems. Hence, maintaining macroinvertebrate communities, biodiversity and individual species populations contributes to the overall ecological integrity of the system (Spänhoff and Arle, 2007). Particular invertebrates (species, genus or families) have tolerance limits to specific environmental conditions, such as levels of salinity, pH, organic pollution, suspended sediment concentration, fine sediment deposition and flow velocity (e.g. Hellawell, 1986). Macroinvertebrate biomonitoring tools and assessment systems are widely used to assess water quality globally (e.g. North America – Barbour et al., 1999; Africa – Cummins et al., 2004; Asia – Morse et al., 2007; South America – Dickens and Graham, 2002), although there have been recent calls for methods of assessing ecological response to environmental changes and pressures to be more strongly rooted in ecological and biological theory (e.g. Friberg et al., 2011; Johnson and Rice, 2014). In Europe, macroinvertebrate biomonitoring forms an important part of compliance monitoring within the European Union Water Framework Directive (WFD). This Directive requires Member States to ensure that all freshwater bodies are of 'Good Ecological Status (GES) or Good Ecological Potential (GEP) for Heavily Modified Waterbodies (HMWB) and Artificial Waterbodies (AWB) by 2027 (EU Directive 2000/60/EC).

Biomonitoring techniques can be quantitative, semi-quantitative or qualitative, depending on the technique used. The most common method for sampling invertebrates in rivers is the semi-quantitative kick sample method, where invertebrates are sampled over a specified time period (typically three-minutes) supplemented by hand searches of larger substrate clasts; although the total area or proportion of the community sampled is typically unknown (Murray-Bligh, 1999; ISO 10870, 2012). Most macroinvertebrate biomonitoring indices have been developed to allow macroinvertebrate community composition to be analysed on a semi-quantitative basis where sampling effort (time) is standardised (Clements and Newman, 2002). Fully quantitative sampling is necessary for other forms of analysis that require information regarding the total abundance, density or diversity of organisms/communities within a specified area. This can be achieved with a Surber sampler (or other similar devices such as a cylinder sampler, or Hess sampler), where invertebrates are collected within a specified sampling area.

Whilst there is widespread agreement that the macroinvertebrate community provides a valuable tool to characterise the ecological health of rivers, there is less consensus about the most appropriate sampling methodologies to employ. Surprisingly, the degree to which biological metrics derived from semi-quantitative and quantitative samples differ has not been widely assessed in a systematic way. The largely untested assumption that biomonitoring scores are not comparable between these methods prevents both historic (e.g. Percival and Whitehead, 1929; Percival and Whitehead, 1930; Prigg, 2002) and contemporary fully quantitative data from being combined and used to characterise river health. Hence, the aim of this paper is to compare a semi-quantitative kick sampling methodology with a quantitative Surber sampling methodology at given sites by cross-matching: 1) derived biomonitoring scores/indices; 2) inferred water- and habitat-quality; and 3) the abundance and diversity of the taxa collected by each method.

2. Methodology

2.1. Sampling techniques

Kick sampling is a semi-quantitative method of surveying the invertebrate community, which is widely used internationally because it is cost effective and results are relatively consistent between operators (e.g. Carter and Resh, 2001; Metzeling et al., 2003). In this study, a 1 mm^2 mesh net with an opening 0.25 m wide and 0.22 m deep was held downstream of the operator who kicked the river bed and swept the net through, for example, submerged macrophytes. This action disturbs sediment and dislodges benthic invertebrates which are then carried by the river flow into the net. The duration of kick sampling here followed the Environment Agency of England (EA) best-practice standard, which requires three-minutes of kick sampling and one-minute hand search of larger substrates for macro-invertebrates (HMSO, 1985; Murray-Bligh, 1999; Environment Agency, 2009). The operator moved systematically across and upstream through the river reach being sampled, ensuring that all main habitat types were sampled (e.g. emergent and submerged macrophyte stands, woody debris, tree roots, different flow depth/velocities and bed substrate compositions). The amount of time spent in each designated habitat unit was proportionate to the surface area that each occupied.

To obtain a quantitative comparison, replicate Surber samples were collected. A Surber sampler is a rectangular quadrat, 0.33×0.30 m (area 0.1 m^2) that is placed on the river bed. The quadrat has a 1 mm^2 mesh net attached, with a 0.29×0.34 m opening. The operator disturbs by hand all surface material within the quadrat area. Total sampling times can vary but in the current study continued until all of the 0.1 m^2 quadrat area was fully sampled (Surber, 1937; Macan, 1958). Sediment was disturbed to a maximum depth of 0.1 m. Disturbance dislodges invertebrates that then drift into the downstream net and, with the aid of side curtains, captures dislodged animals that might otherwise avoid capture in the net. Traditional Surber net sampling tended to be micro-habitat specific but for some river types Surber net sampling can form part of a methodology that proportionally samples different microhabitats (Prigg, 2002; Everall, 2010). In this study, 10 Surber samples, distributed such that all habitat types within the site were represented, were undertaken at each survey site. As with kick sampling, the habitats sampled reflected the proportion of the area covered by each habitat type at the site. For ease of analysis, the 10 individual samples were aggregated into 5 sub-samples for identification. The data from these 5 sub-sample units were, in turn, aggregated prior to the calculation of the biomonitoring indices/scores used for comparison between methods.

All samples were collected following the EA best practice guides (Environment Agency, 2009) by an experienced operator (Everall). Kick and Surber sampling was undertaken on the same day, at the same site, one immediately after the other. The second sample was taken a few metres upstream of the first but spatially alternating between kick and Surber net sample reaches at survey sites to reduce any sampling sequence bias. Sample site reaches were selected for their similarity of instream habitat composition over the sampled reach and were divided into kick and Surber areas such that each had comparable proportions of the major habitat types.

2.2. Sampling times and locations

Sampling was undertaken on seven English rivers at a total of 20 sites (Fig. 1). These locations were chosen to provide a range