SITE LENGTH FOR BIOLOGICAL ASSESSMENT OF BOATABLE RIVERS

J. E. FLOTEMERSCH,a* J. B. STRIBLING,b R. M. HUGHES,c L. REYNOLDS,d M. J. PAUL, and C. WOLTERe

a U.S. Environmental Protection Agency, Office of Research and Development, 26 W. Martin Luther King Drive, Cincinnati, OH 45268, USA
b Tetra Tech, Inc., Center for Ecological Sciences, 400 Red Brook Blvd., Suite 200, Owings Mills, MD 21117, USA
c Department of Fisheries and Wildlife, Oregon State University, 200 SW 355th St., Corvallis, OR 97333, USA
d U.S. Environmental Protection Agency, Region 3, 1060 Chapline St. Ste. 303, Wheeling, WV 26003, USA
e Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Müggelseedamm 310, 12587 Berlin, Germany

ABSTRACT

There is increasing international interest by water resource management agencies worldwide in developing the capacity for quantitative bioassessments of boatable rivers. This interest stems from legal mandates requiring assessments, plus growing recognition of the threats to such systems from multiple and co-varying stressors (e.g., chemical pollutants, physical habitat alterations, altered flow regimes, channel modifications and alien species). The elevated cost and inefficiencies of jurisdictionally- and taxonomically-segregated assessments is widely recognized, as is the desire to obtain comparable data that can be easily shared among political jurisdictions and ecological regions. The objectives, sampling methods, indicators, site-scale sampling designs and geographic extent of the resources being sampled differ among programmes, thereby limiting such data exchanges. Our objective in this paper is to review major biological assessment design alternatives for boatable rivers, with special attention given to the sample site length from which data are collected. We suggest that sufficient site length determinations should be based on the survey objectives, the relative heterogeneity of the habitat template, and the quality of data necessary for meeting programmatic data quality objectives. Future sampling effort studies should be designed to allow separate samples of several short sub-sites at many diverse sites to generate multiple data points for each site. Data from those multiple sub-sites should be analysed using randomization-based data evaluation methods. We hope that our recommendations will be useful to the maximum number of institutions, including those with limited funds and a purely local focus, as well as those responsible for sampling at continental geographic extents. Copyright © 2010 John Wiley & Sons, Ltd.

KEY WORDS: reach length; non-wadeable; electrofishing distance; bioassessment; sampling effort; fish; benthic macroinvertebrates; algae

INTRODUCTION

Biological assemblages are the central focus of biomonitoring programmes, because they provide a direct measure of biological condition relative to biological integrity—a stated objective of the Clean Water Act of 1972 (USGPO, 1989) and the Water Framework Directive of the European Union (2000/60/EG, Abl. L 327 of 22.12.2000). In addition, biological assemblage assessments contribute to narrative water quality standards that are an important part of U.S. state water laws, and similarly, are essential for enforcement of the U.S. Endangered Species Act (16 U.S.C. 1531–1544), Canada’s Species at Risk Act (SARA; http://laws.justice.gc.ca/en/s-15.3/text.html) and the EU Habitats Directive (92/43/EEC, Abl. L 43 of 21.05.1992). Biota integrate the effects of multiple stressors in space and time, thus acting as environmental sentinels. They provide a way of detecting the effects of stressors that may be so variable that it is neither logistically nor economically feasible to monitor them directly. For example, episodic pollutants cause mortality or morbidity that is reflected by changes in assemblage structure long after the event (e.g., Dixit et al., 1999) and historical land uses continue having legacy effects on biota (Harding et al., 1998). Similarly, impassable barriers (Adams, 2000; Jackson and Marmulla, 2001; SEARIN, 2004) and sediment inputs associated with spatially variable erosion (Waters, 1995) may have biological impacts on multiple assemblages detected far from the sources of the stressors or pressures.

There are four primary survey constraints (objectives, funding, timeline and institutional) and three secondary constraints (survey design, indicators and logistics) that dictate sampling effort requirements (Hughes and Peck, 2008). For biological monitoring programmes focused on streams and rivers, where indicators of water body condition are based on assemblage-level data (e.g., fish, benthic macroinvertebrates and algae), the site-scale sampling design is typically defined by a combination of site length or extent and the sampling effort exerted therein, particularly for fish, but increasingly for macroinvertebrates and periphyton. In wadeable streams, extent essentially equals site area because all habitats are generally accessible to sampling, although
often only specific habitat types are sampled (e.g. pools, riffles, shorelines; Hughes and Peck, 2008). Site lengths for sampling fish in wadeable streams typically range from 100 m to 40 mean wetted channel widths (MWCW; Hughes and Peck, 2008). In studies of non-wadeable habitats, site lengths for fish sampling vary from 500 m fixed length (e.g. Gammon, 1976; Blocksom et al., 2009) to 100 MWCW (Hughes et al., 2002). Site lengths for sampling benthic macroinvertebrates and algae are commonly the same as those selected for fish at a given site. Likewise, within-site sampling effort may vary by the types of gears employed (Curry et al., 2009; guy et al., 2009), the crew size, and the habitat types sampled (e.g. shoreline only, all available, all that can be effectively sampled). While realizing the importance of all site-scale design elements, this paper will focus primarily on the question of site length in boatable rivers.

There is no single site length that addresses all research and applied questions; instead, the appropriate site length for stated objectives will be a compromise between the overall survey design, the intensity of data collection for a particular sampling event, and the sampling gear used. Additionally, with increasing channel size, the open-water column becomes important habitat, with some species only occurring as it becomes available (Wolter and Bischoff, 2001; De Leeuw et al., 2007) or as deep pools develop (Herzog et al., 2005). For example, effective sampling for fish species occurring in open-water areas frequently requires sampling approaches other than electrofishing. Such sampling may involve adjustments to the site length, as well as the inclusion of alternate gears, such as gill nets, fyke nets, push nets or trawls; however, the use of multiple gears must be balanced with data quality requirements and current and projected resource availability. Effective sampling of benthic macroinvertebrate species occurring in open-water areas may require the use of drift nets, bottom grab samplers or tow nets. Increased sampling effort and extent can be justified for increases in precision and accuracy (Lyons, 1992a; Cao et al., 2002; Reynolds et al., 2003); however, those increases may incur substantial increases in sampling costs as well. On the other hand, as Angermeier and Smoger (1995) point out, comparisons of species relative abundance or estimates of absolute numbers based on insufficient sampling effort can be misleading and a waste of time, because real differences in assemblage structure are likely to be indistinguishable from method error. Similarly, Dolph et al. (2010) found that insufficient and variable sampling effort increases the variability of multimetric indices. The same is true for estimates of species richness, largely because of the presence of a number of species that are rare and distributed in a discontinuous manner (Smith and Jones, 2005; Kanno et al., 2009; Haibo et al., in press). In an assessment context, this can translate to a decreased power to distinguish among sites of varying condition (Patton et al., 2000) and elevated potential for incorrect assessment decisions. Roset et al. (2007) and Bonar et al. (2009) recommended standardization of fish sampling protocols to increase the effectiveness of long-term, routine sampling and analysis for decision-making in water resource management. There is no reason these recommendations should not extend to other commonly samples assemblages as well. European electrofishing standards have been developed (CEN, 2003), but they were based on expert opinion rather than empirical studies. Lack of standardization confounds efforts at broad scale assessments that are based on multiple data sources (Hughes et al., 2000; Hughes and Peck, 2008), and weak standards lack precision, accuracy and statistical power.

In this paper, we review several approaches that were used for determining sample site length, as well as technical issues related to defining appropriate sample site extent for large river bioassessments. To a large extent, the review primarily relies on fish literature, as relatively few studies have been published that examine the question of appropriate site length for other biotic indicators. However, much of the rationale discussed is taken from basic principles of ecology, hydrology and geomorphology, and thus, should be applicable to other areas of stream and river research as well.

**SAMPLING SITE LENGTH**

In a hierarchical context, Frissell et al. (1986) defined the word ‘reach’ as a length of a small stream between breaks in channel slope, local side-slopes, valley floor width, riparian vegetation and bank material. They further added that the reach is sometimes the least physically discrete unit in the hierarchy, but an exceedingly useful scale for describing medium- and long-term effects of human activities on streams. In small wadeable streams, reaches are typically 10–100 m long, but in boatable rivers, reaches are typically hundreds to thousands of metres long. The next larger units, segments, are defined by major tributary junctions, waterfalls or major geomorphological changes (Frissell et al., 1986). Valley segments (Seelbach et al., 2006) also include catchment character, hydrologic regime and fish species associations. Both segments and valley segments in rivers are 10–100 km long (Frissell et al., 1986; Seelbach et al., 2006). To avoid confusion with these geomorphic terms in this paper, we use the term ‘site’ to describe the area, zone, stretch, station or locale of a river from which samples are collected or from which other measurements are taken (Flotemersch et al., 2006a). A site is not a geomorphic or cartographic reach, segment or valley segment, as defined above, although sample reach elsewhere is often used synonymously with sample site because of the long spatial extent of the sampling. Further, we define macrohabitats as the existence and spatial organization of the physical structure within a site, such as zones of banks with varying angles, large woody debris or accumulations of medium to small woody debris, large deep holes, riffles, macrophyte
bends, tributary inflows, riparian vegetation or inner and outer bends of channel meanders.

In linear systems, such as rivers and streams, a site is often quantified as some unit of channel length. For non-wadeable rivers, the site lengths sampled are generally larger than those for wadeable streams, as sampling efforts are often scaled up to accommodate the magnitude of the resource. The approach used can result in relatively long (e.g. 10 km) or short sampling sites (e.g. 500 m); as systems get larger (i.e. wider), the number of square metres of habitat increases exponentially and, therefore, may require more sampling. For example, a 500-m river section that is 10-m wide contains roughly 5000 m² of habitat, whereas an equal river length 50-m wide contains 25 000 m². The need for increased sampling is also a valid argument in places where the number of unique habitats increases as a function of increased area (e.g. deep water habitats or rapids), but not in systems where the number of unique habitats does not change in proportion to the width. As river size increases, the length of river required to encounter the same number of available habitats increases proportionately. Additionally, as systems increase in size, the occurrence of habitats difficult to sample may also increase (e.g. deep pools, rapids), thus potentially increasing the site length or number of gears required to meet data quality objectives.

In general, long sites (e.g. several kilometres) are considered advantageous for determining the overall condition of a larger section of river, as they are intended to minimize the influence of localized conditions. This advantage, however, can also be viewed as a disadvantage if long site lengths are sampled so that local habitat conditions are masked. Long sites that cannot be disaggregated diminish the sensitivity of the data to linkages between local river conditions and the drivers of those conditions. In addition, long sites may yield excessively large numbers of fish, substantially increasing sample processing. Conversely, single short sites (e.g. <1 km) may be too sensitive to localized conditions and tend to confound co-occurring natural and anthropogenic covariables, thus introducing bias to assessments of overall water body condition. This can occur, for example, when a 500-m site from a long bar or inside bend is compared with a 500-m site from an outside bend, especially when only one is receiving pollutants or has coarse versus fine substrate (Flotemersch et al., 2006a).

Clearly, the site length for a study should depend on both the ecological questions being addressed and the resource constraints present, but the sooner a standardized level of sampling effort can be established, the sooner data sharing and rigorous assessments can occur among and within states (Bonar and Hubert, 2002; Bonar et al., 2009).

Scientifically sound sampling designs include not only the distribution of sample locations around the watershed or region (Larsen, 1997; Smith and Jones, 2005), but also the intensity and distribution of site-scale sampling efforts (Lyons, 1992a; Angermeier and Smogor, 1995; Paller, 1995; Peterson and Rabeni, 1995; Patton et al., 2000; Cao et al., 2001; Cao et al., 2002; Hughes et al., 2002; Reynolds et al., 2003; Dauwalter and Pert, 2003a, 2003b; Maret and Ott, 2004; Wolter et al., 2004; Fayram et al., 2005; Flotemersch and Blocksom, 2005; Hughes and Herlihy, 2007; Kanno et al., 2009). Site-scale sampling design includes specifying the spatial scale over which the sample(s) will be collected (site length), the amount and types of habitats that will be sampled within that site, the field sampling methods to be used, the estimated number of person-hours (Reynolds et al., 2003; Flotemersch et al., 2006a; Hughes and Peck, 2008), and, sometimes, a minimum number of fish to be caught (Dußling et al., 2004; Hughes and Herlihy, 2007; Maret et al., 2007; USEPA, 2007).

Estimates and inferences regarding assemblage attributes (e.g. species richness, multimetric index scores) are sensitive to site-scale design and sampling effort, because riverine habitat is heterogeneous, with non-uniform distribution of organisms among habitat types (Angermeier and Smogor, 1995; Kanno et al., 2009). As a result, the number of taxa collected at a given site will increase with sampling effort, but will also vary with biogeography, behaviour and abundance of the species being sampled, and the patchiness of the macrohabitat types. Ideally, the sampling effort applied is the minimum that will allow the stated objectives to be precisely and accurately addressed (Angermeier and Smogor, 1995; Patton et al., 2000). As an example, estimates of species’ proportionate abundances have been shown to require less sampling effort for a given accuracy than estimates of the absolute number of species (Angermeier and Smogor, 1995; Reynolds et al., 2003; De Leeuw et al., 2007). For a biological assessment programme, potential cost savings realized through the use of efficient sampling protocols translate to opportunities to enhance other aspects of a study design or programme, such as increasing the number of indicators or sites sampled (Patton et al., 2000; Hughes and Peck, 2008).

Many factors that are relevant for establishing sampling site length in wadeable streams (e.g. Patton et al., 2000; Lyons et al., 2001) are also relevant for larger, non-wadeable rivers. If increased precision and accuracy in estimates of species richness are the focus of monitoring and assessment, adjustments in sampling effort relative to species richness are required (Paller, 1995; Cao et al., 2001; Hughes et al., 2002; Haibo et al., in press). Dolph et al. (in press) found that greater sampling effort is needed in larger water bodies than in smaller water bodies to capture rare species that influence richness measures. Paller (1995) found that the relative importance of depth for within-site sampling might depend on the behaviour of individual species (i.e. whether they are bottom, open-water, or littoral/sublittoral dwellers),
or upon width-to-depth ratios of the channel. Many large rivers have an abundance of macrohabitats that support fish species that are difficult to sample efficiently (e.g. those associated with deep, turbid or swift-moving waters, or off-channel habitats). For these kinds of rivers and species, which can be more frequent in some ecoregions than others, greater sampling effort with at least some coverage of all habitat types is necessary to adequately represent fish assemblage structure and species richness (Angermeier and Smogor, 1995; Hughes et al., 2002; Hughes and Peck, 2008). How one defines what is adequate is based on the objectives of the project, and the quantity and quality of data necessary to address them (Flotemersch et al., 2006a). Also, instead of composing the data from all the available habitat types in a single long site, an alternate approach could potentially focus upon recording data separately for each specific habitat type or sub-site (e.g. record data separately for every 4 MWCW sub-site vs. for the entire 40 MWCW; Hughes et al., 2002; Flotemersch and Blocksom, 2005). Such a data recording scheme is essential for understanding species-habitat relationships and estimations of sampling effort-species richness curves (e.g. Erös et al., 2008; Kanno et al., 2009).

**APPROACHES FOR DETERMINING SAMPLING SITE LENGTH**

In most applications, the channel length over which data are collected is the same for site-scale physical habitat measures and biota, particularly fish, benthic macroinvertebrates and algae. Advantages to using the same site length for multiple indicator measurements collected over the extent of the site include time and labour savings, simplified logistics and an enhancement of direct comparisons among indicators at the same spatial extent. However, different site lengths by indicator may be justified. For example, some biota migrate more than others or are affected differently by proximal versus distal environmental conditions. Certainly, the channel length over which land use/land cover data are collected should exceed that over which assemblage information is collected. It is useful to calculate land use and socioeconomic data for areas upstream of the site from which site-scale data have been collected (Hughes et al., 2006; Leprieur et al., 2008); however, Pringle (1997) and Hughes et al. (2005) have also recognized the importance of considering how downstream disturbances, particularly dams, affect upstream assemblages.

Different approaches have been used for establishing the channel length in biological assessments of large rivers. Most involve consideration of several factors, including the question(s) being addressed by the study, the level of resolution (precision and accuracy) required, the analytical approach, and the available resources. With exceptions (e.g. Lyons et al., 2001; Cao et al., 2001; Hughes et al., 2002; Reynolds et al., 2003 and references cited therein; Flotemersch and Blocksom, 2005; Hughes and Herlihy, 2007; Maret et al., 2007; LaVigne et al., 2008a, 2008b), site length designations have been based on judgment, tradition, untested assumptions or the need to match some other aspect of sampling or management activities. The purpose of this paper is not intended to be an exhaustive review of the topic; rather, it is to provide an overview of different approaches that have been used for site length determination, and to help provide a technical foundation for improving future studies of sampling effort.

**Biologically based site lengths**

Recent research has been conducted on the selection of sample site lengths by evaluating the response of biological parameters (e.g. species richness, assemblage metrics, multimetric index scores) as a function of channel characteristics (e.g. MWCW, meander wavelengths, riffle-pool sequences, macrohabitats). Most of these studies have used fish assemblages (Gammon, 1976; Lyons, 1992a, 1992b; Meador et al., 1993a; Penczak and Mann, 1993; Angermeier and Smogor, 1995; Paller, 1995; Yoder and Smith, 1999; Patton et al., 2000; Cao et al., 2001; Lyons et al., 2001; Hughes et al., 2002; Reynolds et al., 2003; Maret and Ott, 2004; Wolter et al., 2004; Flotemersch and Blocksom, 2005; Hughes and Herlihy, 2007), although a few have used benthic macroinvertebrates (Bartsch et al., 1998; Li et al., 2001; Poulton et al., 2003; Flotemersch et al., 2006b) or physical habitat structure (Kaufmann et al., 1999; Larsen et al., 2004).

The rationale for using biological measures to determine site length is that biota are used to assess the condition of the water body in biological assessment; therefore, it is logical to use them in determining the combination of site length and intensity of effort required to produce indicators that both reflect the biota a given length of river supports and are responsive to environmental factors of concern (both point and non-point stressors). In this approach, over-sampling is conducted at a series of sites that cover the gradient of conditions to be included in the study, and the site length at which the required data quality has been achieved is determined. This occurs when a specified indicator asymptote is reached (Lyons, 1992a; Angermeier and Smogor, 1995; Paller, 1995; Patton et al., 2000; Cao et al., 2001; Lyons et al., 2001; Hughes et al., 2002; Reynolds et al., 2003; Maret and Ott, 2004; Wolter et al., 2004), when some level of similarity has been attained (Cao et al., 2001; Cao et al., 2002), or variability of that indicator has been reduced to a specified level or a measurement quality objective (Flotemersch et al., 2006a, 2006b; Hughes and Herlihy, 2007).

Various researchers have used variability of biological data as the basis for determining site length sufficiency. Hughes et al. (2002) sampled 100 MWCW in 45 Oregon rivers...
(averaging 1 river/day), with the rivers ranging from 20- to 100-m wide. They determined through Monte Carlo and similarity analyses that 85 MWCW were needed to collect 95% of the fish species obtained in 75% of the sampled sites. Using the same data, Cao et al. (2001) used similarity analysis to suggest that an average of 286 MWCW are necessary to collect 100% of all fish species in a site. Those findings led to a field sampling design specification of 100 MWCW, sampled near alternating shorelines, being established for the U.S. Environmental Protection Agency’s Environmental Monitoring and Assessment Program (EMAP)-West study (Hughes and Peck, 2008). Hughes and Herlihy (2007) determined through Monte Carlo analysis that a site length of 50 MWCW was needed to obtain multimetric index scores that varied by <10% from those obtained using 100 MWCW. Flotemersch and Blocksom (2005) examined the effect of site length on the variability of fish assemblage metrics from samples covering up to 2 km and determined that at shallow river sites, sampling 1 km of shoreline was sufficient for limiting the variability in metric scores to 20%. Flotemersch and Blocksom (2005) further suggested that night electrofishing could be considered for sample sites with depths > 4 m, and if that was not feasible for logistical or safety reasons, increasing the daytime electrofishing distance to 2 km of shoreline would likely suffice. If the latter approach was used, the authors warned that metrics based on fish species prone to diel movements should be interpreted with additional caution.

The studies by Hughes et al. (2002), Hughes and Herlihy (2007), and Flotemersch and Blocksom (2005) utilized different maximum site lengths (100 MWCW vs. 2 km), different sized sub-sites (10 MWCW vs. 200 m [in most cases]), different values for acceptable variability (5, 10, and 20%), and different indicator endpoints (species richness, index of biotic integrity [IBI] metrics and composite IBI scores), and were conducted in rivers varying in condition from relatively natural to highly altered. The fact that the preceding studies produced different results should not be unexpected. For example, Kanno et al. (2009) reported that recommended sampling distances in eight different studies varied according to the length of the initial study site (40–100 MWCW), with recommended site lengths consistently 60–70% of the initial study site length, except where sampling efficiency was very high or very low.

**Physically based site lengths**

Some studies or manuals propose site lengths as a function of multiples of the MWCW (e.g. Lyons, 1992a; Lazorchak et al., 2000; Hughes et al., 2002; Reynolds et al., 2003; Maret and Ott, 2004; Hughes and Peck, 2008), while others support the use of a fixed distance (Yoder and Smith, 1999; Emery et al., 2003; Flotemersch and Blocksom, 2005). The MWCW approach follows the logic that as river size increases, the effort required to sample the system’s available habitat at an equivalent level should increase proportionally. Proponents of this approach assert that application of a fixed sample site length on larger systems (e.g. 500 m on a 100-m wide system) would potentially miss or under-represent habitat components that recur at longer intervals, such as bars, riffles, runs, glides, pools, inside bends and outside bends. An argument against the MWCW approach is that sampling extent and the intensity of effort differs among sites, by definition; however, this is not a simple argument. Some studies that use the MWCW approach for setting sample site extent distribute a prescribed level of effort (e.g. time or distance) over the span of the site. Nonetheless, difficulty may still be encountered with this approach in wide or impounded rivers, where long site lengths would result (e.g. 5 km for a 100-m wide river, if 50 MWCW is the protocol), unless a maximum site length has been set (e.g. Meador et al., 1993a, 1993b). In the case of impounded systems, unacceptably long site lengths may also be avoided by using the pre-impoundment wetted widths to set study site length; however, this information is often not readily available, and we are unaware of any examples of this approach.

An alternate physically based approach is to set site length based on the meander wavelength of the system. Leopold et al. (1964) proposed that 20 times the MWCW typically encompasses at least one complete meander cycle of a stream. Because fluvial characteristics are repetitive and cyclical (Dunne and Leopold, 1978), this distance should theoretically include all major physical habitat types within a given geomorphic reach and, by default, be available to all resident biota of those habitats. Adapting this guideline, the U.S. Geological Survey’s National Water Quality Assessment program (NAWQA; Meador et al., 1993a, 1993b; Meador, 2005) used 20 MWCW for setting sample site length, a minimum length of 500 m to help ensure sufficient biological data, and a maximum length of 1 km (Fitzpatrick et al., 1998). The designation of a maximum site length of 1 km is especially relevant when working in altered large rivers where the identity, extent and boundaries of habitat features in meander lengths are obscured by impoundments, obliterated by human alteration of the channel (straightening, bank-armouring and dredging), or modified by structures (locks, dams and flowages); these conditions render identification of a meander impractical for setting site length. In addition, not all natural rivers have a single-thread meandering morphology to begin with, not all major macrohabitat types are found in 20 MWCW, and bank full widths are inconsistently related to MWCW during base flow index periods (Kondolf, 2006; Walter and Merritts, 2008; Phil Kaufmann, USEPA, personal communication). Meador (2005) reported that site lengths greater than 20 MWCW or 1 km were needed to collect 95% of the fish species expected from many rivers in a large geographic region, likely because
of high species richness and discontinuous species distributions. Furthermore, setting a 500-m minimum and 1000-m maximum site length effectively amounts to a fixed site length in rivers with a <25-m MWCW and a >50-m MWCW.

Fixed site lengths

Proponents of fixed distance endorse the ease of application in the field and its usefulness in planning field activities (Patton et al., 2000; Lyons et al., 2001; Flotemersch and Blocksom, 2005). Opponents argue that using a fixed distance results in unequal sampling extents relative to river size, and typically results in greater variability regarding reference condition and relative level of effort (Hughes and Herlihy, 2007; Maret et al., 2007). If true, this would contribute to an increase in the variability attributable to the study design, thus, reducing the overall accuracy and precision of the data, and ultimately the usefulness of the data for detecting differences in condition.

Another argument against fixed lengths is that where the sites do not encompass all major habitat types, the biological differences detected may be an artefact of differences in the quantity and quality of the physical habitat sampled at the sites. This concern becomes greater as habitat features increase in proportion to increased river width. One strategy for dealing with this is to account for gross habitat differences in the design and in any subsequent analysis and reporting. For example, the Ohio River Valley Water Sanitation Commission (ORSANCO, 2008) conducts biological sampling on the 0.8- to 2-km wide Ohio River using 500-m sites. At this combination of river scale and site length, an individual macrohabitat or habitat feature (e.g. sandbars) may dominate, if not exceed, the length of the sampling site.

To account for this in an assessment context, ORSANCO categorizes sampling sites as soft-, hard- and mixed-bottom types. Using this approach, individual habitats may be allocated equal effort or sampled in proportion to their occurrence, each approach having its advantages and disadvantages in regards to statistical analysis and what can be gleaned from the data; however, a cautionary note is warranted. When partitioning habitats or analyzing data collected using a study design that partitions sampling effort among habitats, users are encouraged to fully consider the implications of the, often substantial, diel migrations of some fish species among habitats (Sanders, 1992; Copp and Jurajda, 1999; Hohausová et al., 2003; Wolter and Freyhof, 2004) in order to reduce the risk of misrepresenting the use of a particular habitat by the existent assemblage. In addition, focusing the sample site on a habitat type offers a poor design for assessing fish assemblages when the habitat type changes (Phil Kaufmann, USEPA, personal communication).

Another fixed length design option is to aggregate data from a series of 400- to 500-m long samples into a segment-scale assessment (Gammon, 1976; Yoder and Smith, 1999; Wolter et al., 2004; Blocksom et al., 2009). This approach has clear advantages for use in studies with the objective of sampling multiple or all available habitats in large rivers that are regulated and/or modified to enhance inland navigation. In such systems, the banks are often so monotonous that even sampling 100-MWCW may not cover all potential habitats in a given section of river. While sampling several kilometres of monotonous banks will substantially increase the fish yield and, accordingly, the processing effort (e.g. counting, measuring), the additional information collected may provide little added insight regarding the overall ecological condition of the site being studied. Field studies supporting this approach include Wolter et al. (2004), who found that in large, navigable, lowland rivers in Germany, electrofishing an average site of 400 m captures 95% of the fish species in the littoral zone, and that combining the data derived from individual 400-m sites provides an effective and efficient means of accomplishing assessments. In other studies, Dübling et al. (2004) and Diekmann et al. (2005) found that sampling and compositing data from multiple relatively homogeneous macrohabitats for 400 m each, with a cumulative sampling site of 40 MWCW (maximum 10 km), was more promising than sampling a single continuous river site. However, in a study that reviewed fish data collected using different site lengths, LaVigne et al. (2008b) reported that on the partially-navigable Willamette River (Oregon), 23–27 sites 500- to 1000-m in length yielded half as many fish species per site as 21 sites 50-MWCW long (4 to 11-km in length). Gammon, the early proponent of the 500 m fixed length in large U.S. rivers, reported that the number of fish species collected by increasing sampling from 350 to 1500 m in units of 350 m never approached an asymptote in the Wabash River, Indiana, although various diversity indices did plateau between 500 and 1500 m, with only 5–8% increases in scores with increased distance between 500 and 1000 m (Gammon, 1976). In the 300- to 600-m wide Danube River (Hungary), Erős et al. (2008) indicated that fish trait diversity was maximized with twenty 100-m long samples (2000 m; 3–7 MWCW) and assemblage autosimilarity reached 70% in 10–40 samples; however, 50–75 samples (5–7.5 km; 8–25 MWCW) were necessary to produce 90% of the species richness yielded by 125 samples (12.5 km). Erős et al. (2008) also found that during the day, rip-rip-supported more species in their study than natural gravel and sand habitat, but during the night, these differences diminished. Unless the explicit goal of data collection is to collect and evaluate data from all or individual habitats, care should be taken to sample habitats in proportion to their occurrence and record the data separately, in order to facilitate randomization-based data analyses and evaluate sampling effort.
ASSEMBLAGE-SPECIFIC SAMPLING ISSUES

Fish

Considerable research has been conducted on determining sufficient site lengths for sampling fish assemblages in large rivers (e.g. Gammon, 1976; Penczak and Mann, 1993; Yoder and Smith, 1999; Cao et al., 2001; Lyons et al., 2001; Hughes et al., 2002; Dubling et al., 2004; Wolter et al., 2004; Flotemersch and Blocksom, 2005; Hughes and Herlihy, 2007; Maret et al., 2007). Site lengths found suitable for sampling fish in rivers vary in length and approach (i.e. fixed distance vs. MWCW). Some of these differences can be attributed to the channel complexity (natural heterogeneous vs. channelized homogeneous), focus of study (site vs. segment), channel slope (high vs. low), productivity (mesotrophic vs. oligotrophic) and the indicator used to determine sufficient site length (e.g. species richness vs. IBI). However, such confusion can be reduced in future studies by sampling several short sub-sites at a diverse set of sites to generate multiple data points for each site. Data from these multiple short sub-sites facilitate randomization-based data evaluation methods, such as Monte Carlo analysis, and standard error estimates, which help one assess the effect of site size (length) and sample size (total number of sites) depending on the study objective (e.g. Cao et al., 2001; Hughes et al., 2002; Erös et al., 2008; Smith and Jones, 2008; Fischer and Paukert, 2009; Kanno et al., 2009).

While sample site lengths vary among studies, most sampling is generally designed to represent the diversity and complexity of habitats present in the site by sampling in relative proportion to their occurrence. In smaller systems, this is usually straightforward, because all habitats are accessible and, therefore, able to be sampled. This assumption is tested in river systems where similar macrohabitats span kilometres. If the focus of the sampling effort is truly proportional, the fact that the majority of the sample site is a single habitat should be of little concern. That is, a 1000-m site that is straightened with armoured banks for 900 m would likely produce fish samples with low diversity and an assessment rating of ‘impaired’, but still would be an accurate depiction of biological condition for that site. On the other hand, a 1000-m site with sampling targeted or focused on 100 m of large woody debris and 100 m of riffle, with little or no sampling of the remaining 800-m pool, would misrepresent the entire site.

Although most electrofishing designs specify fishing a continuous length of shoreline, other options exist. For example, the Tennessee Valley Authority (TVA) samples macrohabitats in rivers using a species and effort requirement. Riffles, runs and pools are all sampled using a fixed area for each sampling effort and can usually be found within a relatively short reach length in the Tennessee Valley. Each of the three habitat types sampled receives three passes, with each pass continuing until no new species are found (Charles Saylor, TVA, personal communication). Sampling specific habitats requires multiple gears (i.e. backpack electrofishing and boat electrofishing). Similarly, for reservoirs, TVA uses an electrofishing design that fishes fifteen 300-m long sites, each separated by 50 m (Jennings et al., 1995; Hickman and McDonough, 1996; McDonough and Hickman, 1999); the electrofishing catch is supplemented with 10 overnight experimental gill net sets within the larger sampling zone. Dominant habitat features, in each electrofishing pass and at each gill-net set, are recorded to determine habitat influences on metric results. In contrast to delineating different habitats, LaVigne et al. (2008a) electrofished river shorelines, alternating sides every 10 MWCW (2 sub-sites) of a 50-MWCW site, to account for potential bank-related habitat differences. The National Rivers and Stream Assessment (NRSA; USEPA, 2007) modified this approach, electrofishing alternating river shorelines for a minimum of 20 MWCW or a distance sufficient to collect 500 individuals (in ≥ 20 MWCW), or for the entirety of 40 MWCW (with a maximum site length of 4 km). Each sub reach is electrofished a maximum of 700 s for a total of 7000 s across 10 MWCW. Multiple electrofishing gears (e.g. boat-based, backpack, tote barge) are used when needed to sample specific habitats.

The European Fish Index (EFI) was calculated using data collected from a minimum sampling area of 1000 m² for boatable rivers (Schmutz et al., 2007; Pont et al., 2007), which approximates the European standard for electrofishing—10 MWCW and a minimum length of 100 m. However, the dataset from which that standard was developed was biased by wadeable streams and expert opinion; it was not based on data analyses. In contrast, the German fish-based assessment system (FiBS) requires a cumulative sampling site of 40 MWCW (maximum 10 km), with a minimum length of 200 m per sub-site, to assess the ecological quality of rivers (Diekmann et al., 2005). This length is increased if catches do not fulfill numerical preconditions for the assessment—30 times the number of species expected and an absolute minimum number of 101 fish (Dubling et al., 2004); lower numbers would overvalue a single fish, because most of the German assessment metrics are based on proportions.

Considering species richness in navigable large rivers essentially requires additional sampling gears and protocols (Galat et al., 2005; De Leeuw et al., 2007; Guy et al., 2009). The open-water column is an important habitat for larger specimens and specialized potamal fish species, as well as a migration pathway for migratory species, which cannot be efficiently sampled by electrofishing. Trawling the mid-channel section of the lower river Oder (Germany) yielded a very distinct fish assemblage dominated by potamal species (Wolter and Bischoff, 2001). Similar findings have been obtained by trawling Dutch sections of the Rhine and Meuse Rivers (De Leeuw et al., 2007) and drift-netting the river...
Elbe main channel in Germany (Fladung, 2002). Mixed gears are also substituted for long site lengths in U.S. rivers (e.g. Galat et al., 2005; Guy et al., 2009; Curry et al., 2009); however, such gears substantially increase the sampling effort and time spent per site.

Studies of diurnal fish migration patterns in the lower river Oder revealed significant inshore migrations of potamal species at night (Wolter and Freyhof, 2004). Inshore electrofishing at night yielded substantial proportions of large specimens and potamal species not, or infrequently, caught there during daytime. By assessing the ecological status of this river site using pooled electrofishing and trawl catches and day and night electrofishing catches, respectively, it was concluded that electrofishing at night may substitute for additional sampling gears in large rivers (Figure 1; Wolter and Freyhof, 2004). Simon and Sanders (1999) and Emery et al. (2003) arrived at a similar conclusion for the Ohio River. Erös et al. (2008) reported that night electrofishing in the summer more than doubled the number of species collected by day electrofishing in the Danube River, Hungary; species richness curves were strongly flattening by day electrofishing 17–33 MWCW, but not by night electrofishing the same distance.

Benthic macroinvertebrates

That there are multiple factors controlling both large- and small-scale distributional patterns of benthic macroinvertebrates is understood (Power et al., 1988; Bady et al., 2005), as is the nested hierarchical organization of their distribution within and among habitats in rivers (Parsons et al., 2003, 2004). Understandably, these patterns can influence the outcome of assessments. For example, Stepenuck et al. (2008) found differential responses of benthic macroinvertebrates to stressors from urban sources, depending on whether the sample was taken from woody snags or riffles. Related research suggests that analysis of samples from discrete targeted habitat types (e.g. riffles, macrophyte beds, large woody debris) could afford greater sensitivity to detection of site-specific responses to stressors (e.g. Chessman, 1995; Hewlett, 2000; Gerth and Herlihy, 2006; Chessman et al., 2007); however, it is acknowledged that routine monitoring on a large scale might be hampered by the greater resource requirements of processing habitat-specific samples. Furthermore, targeting species-rich habitats biases the sample of a site when such habitats are rare, just as does targeting only snags and riffles for fish, while ignoring the monotonous habitats (Hughes and Peck, 2008).

Li et al. (2001) found that new taxa continued to be added in stream sites that were 80-MWCW long, demonstrating that taxa richness is highly dependent on sampling effort. They also observed variability of metric response across spatial scales, which further emphasizes the importance of designing site-scale sampling strategies and analytical approaches that directly contribute to detection of differences at a specified geographic extent. Similar to Li et al. (2001), Bady et al. (2005) found that richness estimates were
strongly dependent on sampling effort. These findings suggest that determining an appropriate sampling site length for macroinvertebrates using species accumulation curves as a direct function of distance is impractical because of the large number of samples that would be required. Instead, it has become an accepted approach to sample with sufficient effort distributed over some standard site length to yield some pre-specified (fixed count) number of individuals (100–500), depending on the desired difference detection probability (Barbour et al., 1999; Cao et al., 2001; Cao et al., 2002; Flotemersch et al., 2006a; Hughes and Peck, 2008).

The approach that has evolved for benthic macroinvertebrate sampling is to have measured effort expended throughout the site length, either distributed in habitat types roughly proportional to their frequency of occurrence in the site (Barbour et al., 1999; Barbour et al., 2006), or at some level along discrete transects systematically distributed along the entire site (Flotemersch et al., 2006a, 2006b; De Pauw et al., 2006; USEPA, 2007; Hughes and Peck, 2008). In both approaches, the specific sub-samples are usually composited into a site-wide sample.

In Europe, large-scale assessment and calibration projects revealed that proportional multi-habitat sampling covering 20–50 m in small-sized rivers (i.e. 10- to 100-km² catchment area) and 50–100 m in medium-sized (100–1000 km²) rivers was representative of a 100-MWCW minimum area surveyed (Hering et al., 2004; Furse et al., 2006; Johnson et al., 2006; Johnson et al., 2007). Correspondingly, Lorenz et al. (2004) determined a misclassification ratio < 10% for the German Saprobic Index with a subsample size of 300 individuals.

Flotemersch et al. (2006b) found no consistent relationship between transect spacing (i.e. total site length) and indicator responsiveness, which they interpreted as robustness in the sampling protocol relative to site length. This suggested that the site length for the rivers sampled could (and should) be dictated by the spatial scale most relevant to meeting the objectives of a study. Further, it shows that the site length for benthic macroinvertebrate sampling could be set to correspond with that for other indicators or data collection efforts, for which a site length effect may have been demonstrated (such as for fish and physical habitat structure). Nonetheless, a few sites exhibited very strong differences among samples derived from different site lengths. This was likely the result of longer site lengths that encompassed more than one distinct macrohabitat type, which is common in heterogeneous western and northern U.S. rivers (Montgomery, 1999; Poole, 2002; Thorp et al., 2008). The few variable samples of Flotemersch et al. (2006b) also support the conclusions of Li et al. (2001) and Cao et al. (2002), who sampled streams with diverse macrohabitat types, and underscores the importance of carefully identifying and considering the geographic extent of interest (e.g. river segment to microhabitat) prior to pilot surveys and certainly before selecting and setting the sampling site length.

Algae

Stevenson (1997) emphasized that there is a hierarchical framework of environmental variables influencing the distribution of algal species within and among micro- and macrohabitats. Higher level constraints, such as climate, geology and biogeography, temper the capacity for understanding local or small geographic-extent effects, such as additions and deletions and changes in the relative dominance of different taxa for all assemblages (Tonn, 1990; Poff, 1997). Lavoie et al. (2005) found that variability of the diatom assemblage increased with geographic extent; thus, if one’s goal is to minimize variability of diatom data, the amount and distribution of sampling effort should be restricted. However, the likelihood that a highly restricted sampling effort would produce data or an indicator representative of a river site would be minimized; that is, unless multiple samples (subsamples) are collected and evaluated as a composite sample. Site length used for algal sampling most often matches that selected for other assemblages or for physical habitat data collection (e.g. Stevenson and Bahls, 1999; Fore and Grafe, 2002; Moulton et al., 2002; Flotemersch et al., 2006a; Hughes and Peck, 2008). Although algal assemblages might be considered ubiquitous, the distribution of individual species is often patchy. Therefore, appropriate distribution of sampling effort within the designated study site is warranted (see Hughes and Peck (2008) for examples).

DATA QUALITY OBJECTIVES AND SITE LENGTH OPTIMIZATION

If the principal measurement to be used for monitoring and assessment is a multimetric index, such as an IBI, then it is appropriate and necessary to base decision-making relative to site length on the variance in IBI scores, after appropriate classification of different river and stream types. Such assessments are typical objectives of agency bioassessment programmes. However, if the intent is to understand and evaluate patterns in species richness, variance in maximum species richness (MSR) would be most informative. Species presence or absence is often the focus of conservation groups, natural history surveys, and some natural resource agencies. A dataset optimized for documentation of biodiversity would have stronger potential for secondary uses, such as biogeographic investigations, documenting range extensions and contractions, life history studies, tracking the spread of invasive species, and monitoring diversity, and would also be applicable for IBI calculation.
To answer these questions, minimum detectable differences (MDD) can be calculated to allow comparison of two populations of IBI scores (either multiple regional river sites or a single site through time). Calculation of MDD requires an estimate of the variance in the condition measures. To illustrate this, we estimated regional variance based on the IBI scores across all sites, for each site length reported by Hughes and Herlihy (2007); the variance ranged from 543 to 607. For the site-specific question, we estimated variance based on five within-site sub-samples with replacement from the same dataset. This variance was much smaller, ranging from 2.6 to 8.4. Using these estimates of variance ($\sigma^2$), we then applied the standard Equation (1) below to estimate minimum detectable difference:

$$
\text{MDD} = \sqrt{\frac{(Z_{\alpha} + Z_{\beta})^2 \times 2\sigma^2}{N}}
$$

where $Z_{\alpha}$ and $Z_{\beta}$ are Z-scores for a specific two-tailed error rate (e.g. 1.645 for $\alpha$ or $\beta = 0.1$) and $N$ is the sample size (Ott, 1993). At a regional geographic extent, the large variability in conditions resulted in a large variance in scores, making detecting subtle changes problematic. In addition, since the variance in scores was relatively small among site lengths, there was little variation in detectable difference due to variance across site lengths. In fact, the range in scores was slightly less for site lengths of 10 MWCW than longer sites, resulting in marginally smaller detectable differences for 10-MWCW site lengths (Table I). Fischer and Paukert (2009) and Utrup and Fisher (2006) reached similar conclusions for prairie rivers. Smith and Jones (2008) estimated that an optimized design for detecting species in small Michigan catchments would be

<table>
<thead>
<tr>
<th>Site length</th>
<th>Regional scale</th>
<th>Variance ($\sigma^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 MWCW</td>
<td>20</td>
<td>543</td>
</tr>
<tr>
<td>10 MWCW</td>
<td>10</td>
<td>543</td>
</tr>
<tr>
<td>10 MWCW</td>
<td>5</td>
<td>543</td>
</tr>
<tr>
<td>20 MWCW</td>
<td>20</td>
<td>589</td>
</tr>
<tr>
<td>20 MWCW</td>
<td>10</td>
<td>589</td>
</tr>
<tr>
<td>20 MWCW</td>
<td>5</td>
<td>589</td>
</tr>
<tr>
<td>40 MWCW</td>
<td>20</td>
<td>589</td>
</tr>
<tr>
<td>40 MWCW</td>
<td>10</td>
<td>607</td>
</tr>
<tr>
<td>40 MWCW</td>
<td>5</td>
<td>607</td>
</tr>
<tr>
<td>80 MWCW</td>
<td>20</td>
<td>603</td>
</tr>
<tr>
<td>80 MWCW</td>
<td>10</td>
<td>603</td>
</tr>
<tr>
<td>80 MWCW</td>
<td>5</td>
<td>603</td>
</tr>
</tbody>
</table>

Table I. Number of samples ($N$) required at different site lengths to detect a 5-, 10- or 20-point change in mean regional IBI score

Data from Hughes and Herlihy (2007); $\alpha = \beta = 0.1$
4–7 sites (each 9–15 MWCW) per reach in first-, second- and third-order streams, with 70% of the effort in third-order segments. Based on our MDD analyses, approximately 30 samples would be needed to detect a significant \( (p < 0.1) \) 20-point change in IBI scores, 90% of the time (Table I), across a region with similar variability as the Oregon data (Hughes and Herlihy, 2007). This determination was not significantly affected by site length; the number of samples required to detect a 20-point change in IBI scores for 10–80 MWCW only varied from 29 to 33. Measuring smaller changes in IBI scores would require a substantially larger number of samples. Similarly, Smith and Jones (2005) reported that for wadeable Michigan streams draining catchments of 24–433 km² and supporting 23–58 species, 76–151 sites (each 30 MWCW) would need to be sampled to detect all species; 20–50 sites would be needed to detect 90% of the species. Clearly, sample number affects the ability to detect regional changes far more than does site length, leading EMAP statisticians to recommend sampling 50 sites per reporting region (Larsen, 1997).

With regard to site-specific detectable differences, the smaller variability in site replicates from the Oregon data (Hughes and Herlihy, 2007) makes detecting subtle change at a specific site through time easier than across the region (Figure 2). Because the variance in repeat-samples did improve for site lengths of 10–40 MWCW, there was improvement in the detectable difference across site lengths to 40 MWCW. However, site lengths of both 40 and 80 MWCW detected the same change in IBI score, again based on the sample variability observed in the Hughes and Herlihy (2007) dataset. This is why Hughes and Herlihy (2007) recommended sampling 40–50 MWCW in western U.S. rivers, when an IBI was the indicator of greatest interest. These results led to USEPA (2007) setting 40 MWCW as the default length for sampling sites in its National River and Stream Assessment (NRSA), although others argued that 20 MWCW sufficed. To assess a single segment (58- to 153-km long) of the Ohio River, Blocksom et al. (2009) estimated that 15 sites (each 500-m long) would be required to capture 90% of the species observed in the segment and to reduce IBI scores to a standard error <3.

Site length selection depends on the questions being asked, as well as the desired quality of data and resulting assessments. In EMAP and NRSA surveys, sites are often too distant to sample more than one per day, so the longer site lengths may be advisable for assessing site status and trends. However, in more intensive surveys of single rivers 40- to 100-m wide, 4 sites can be sampled per day when site length is 500 m (Hughes and Gammon, 1987) or 2 sites can be sampled per day, if site length is 50 MWCW (LaVigne et al., 2008b, 2008a).

**SUMMARY AND CONCLUSIONS**

When selecting a site length, or conducting research for setting site length, it is important to consider several factors, as suggested by Hughes and Peck (2008), including objectives, funding, timeline, institutional constraints, overall survey design, indicators, logistics, the level of resolution (precision and accuracy) required to address the questions, and the statistical approach that will be used to analyse the resulting data. It is especially important to consider what other indicators will be sampled (i.e. the

---

**Figure 2.** Change in minimum detectable difference (MDD) in IBI score for a single site at various sample sizes \( N = X\) as a function of site length. Data from Hughes and Herlihy (2007); \( \alpha = \beta = 0.1 \)

---

volume of site work) and to conduct pilot studies to determine how all the samples can be taken expeditiously (Plafkin et al., 1989; Hughes and Peck, 2008). Another trade-off requiring serious consideration is the potential desirability of maximizing the diversity of macrohabitats sampled, either through (a) adding different gear types, (b) targeting specific macrohabitats or (c) increasing site length; both with the limitation of sampling a site within a single day or less (Seegert, 2000; Bonar et al., 2009). Ideally, the sampling effort applied is the minimum that is required to allow stated study objectives to be addressed (Angermeier and Smogor, 1995; Patton et al., 2000).

We strongly recommend study designs that can address questions at multiple geographic extents. For example, in designs where a long sampling site is warranted for estimating spatially extensive characteristics, we strongly advocate sampling several shorter sub-sites, and keeping the data separate, to generate multiple data points that can be used to determine conditions at a smaller geographic extent (e.g. Cao et al., 2001; Hughes et al., 2002; Erös et al., 2008; Smith and Jones, 2008; Fischer and Paukert, 2009; Kanno et al., 2009). A split-geographic extent design, such as this, will provide data at two geographic extents, the smaller geographic extent also providing a means of estimating variability within the larger geographic extent, as long as the data are recorded by sub-site. Data from multiple shorter sub-sites facilitate randomization methods, such as Monte Carlo analysis; standard error estimates; and similarity analyses, such as Jaccard and Bray-Curtis; and are needed to help visualize the effect of site size and sample size (total number of sites). A split-geographic extent design also may be advantageous where crew fatigue is a concern; options here include having multiple crews or breaking the per-site sampling effort into multiple visits.

On large navigable rivers, further research seems warranted towards efforts to partition some of the habitat variability that longer site lengths (e.g. those based on geomorphic units) seek to encompass. Our recommended approach for these rivers is to sample multiple smaller sites based on the relative occurrence of given major macrohabitat types (Dußling et al., 2004; Wolter et al., 2004; Erös et al., 2008; Blocksom et al., 2009). Then, the optimal length that would allow all fish species present in a macrohabitat to be caught with 90–95% consistency (or some acceptable level of standard error) should be determined, adjusting the number of fishing sub-sites to the availability of different macrohabitats. This approach would also likely increase the ability to detect habitat-specific influences that may be masked by designs that include compositing across habitats. Accordingly, to specify an efficient sampling design for navigable rivers, one must first determine the amount of effort that is practical to apply in the time allotted for the survey. If successful, sampling efficiencies in these systems would likely increase, resulting in greater overall spatial coverage of the resource with the same level of effort expended.

Indeed, it is the spatial organization of catchments, from basin to macrohabitat geographic extents that establishes the geomorphic and physical template for biological processes and patterns (Frissell et al., 1986; Tonn, 1990; Gregory et al., 1991; Poff, 1997). Although much of this heterogeneity in terms of fish species richness can be captured by patterns in species-discharge relationships as long as the sample sites are of sufficient size and number (McGarvey and Hughes, 2008; McGarvey and Ward, 2008), biological interactions and movement patterns also contribute to heterogeneity at multiple levels (May, 1975; Crowl et al., 1997). To understand individual, population, metapopulation and assemblage responses to the environmental template, persons conducting surveys of river assemblages must take into account these multiple geographic extents (Norris, 1995; Fausch et al., 2002; LaVigne et al., 2008b). In addition, Scheiner et al. (2000) explained that the number of individuals, species and macrohabitats sampled all increase as the sample area or site size increases. Although, species richness tends to reach an asymptote in homogeneous rivers as site size increases, heterogeneous rivers and discontinuously distributed species would require much greater effort for species richness to reach an asymptote (Scheiner et al., 2000; Cao et al., 2001; Hughes et al., 2002; Erös et al., 2008; Kanno et al., 2009).

Ultimately, it is the quantity and quality of information required and the available funds that will dictate the level of effort expended at each sampling site. Thus, application of the data quality objectives process, including quantification of desired indicator performance (through multiple sub-site sampling and data recording, coupled with randomization-based analyses), and testing of the capacity of sampling design to meet those objectives (both site-specific and area-wide) should drive the appropriate site length. Increased standardization of such designs is needed to facilitate data exchange, indicate data quality, improve credibility and make assessments more comparable within and among political jurisdictions.

ACKNOWLEDGEMENTS

This paper received USEPA review and was approved for publication. USEPA cooperative agreement CR831682-01 to Oregon State University partially funded paper preparation. Jeroen Gerritsen, Alan Herlihy, Phil Kaufmann, Terry Maret, Blaine Snyder and Chris Yoder participated in many discussions on the topic that helped sharpen our thought process, and Jeroen Gerritsen assisted in the statistical re-analysis of the site length datasets. The paper was improved by critical comments from three anonymous reviewers and edited and formatted by Justicia Rhodus.
REFERENCES


Electrofishing, as a sampling technique, is widely used in biological monitoring programmes. However, the impact of sampling design on the estimation of metrics used to assess ecological integrity is often overlooked. To address this, the authors conducted a study to investigate how different sampling designs affect bioassessment metrics. They used simulations and real-world data to explore the effects of various sampling strategies on the estimation of biodiversity and other ecological metrics. The study highlights the importance of considering sampling design in bioassessment programmes to ensure accurate and meaningful results.


community in the channelized lower Missouri River. *Environmental Monitoring and Assessment* **85**: 23–53. DOI: 10.1023/A:10233016001


