Environ Monit Assess (2009) 155:157–167
DOI 10.1007/s10661-008-0425-0

Sampling effort needed to estimate condition and species richness in the Ohio river, USA

Karen Blocksom · Erich Emery · Jeff Thomas

Received: 17 January 2008 / Accepted: 3 June 2008 / Published online: 6 July 2008

Abstract The level of sampling effort required to characterize fish assemblage condition in a river for the purposes of bioassessment may be estimated via different approaches. However, the goal with any approach is to determine the minimum level of effort necessary to reach some specific level of confidence in the assessment. In the Ohio River, condition is estimated and reported primarily at the level of pools defined by lock and dam structures. The goal of this study was to determine the minimum level of sampling effort required to adequately characterize pools in the Ohio River for the purpose of bioassessment. We followed two approaches to estimating required sampling effort using fish assemblage data from a long-term intensive survey across a number of Ohio River pools. First, we estimated the number of samples beyond which variation in the multimetric Ohio River Fish Index (ORFIn) leveled off. Then, we determined the number of samples necessary to collect approximately 90% of the fish species observed across all samples collected within the pool. For both approaches, approximately 15 samples were adequate to reduce variation in IBI scores to acceptable levels and to capture 90% of observed species in a pool. The results of this evaluation provide a basis not only for the Ohio River Valley Water Sanitation Commission (ORSANCO) but also states and other basin commissions to develop sampling designs for bioassessment that ensure adequate sampling of all assessment units.

Keywords Biological assessment · Electrofishing · Field sampling effort · Fish assemblage · Large river · Nonwadeable river

K. Blocksom (✉)
U.S. Environmental Protection Agency,
National Exposure Research Laboratory,
Ecological Exposure Research Division,
26 West Martin Luther King Dr.,
Cincinnati, OH 45268, USA
e-mail: Blocksom.karen@epa.gov

E. Emery · J. Thomas
Ohio River Valley Water Sanitation Commission,
Cincinnati, OH, USA

Introduction

A commonly used approach to determining the adequacy of fish sampling in flowing waters is based on the relative proportion of species collected. Previous studies have examined the level of effort necessary to produce a representative

fish sample based on the proportion of the available species captured by that method in wadeable streams (Dauwalter and Pert 2003; Reynolds et al. 2003; Lyons 1992) and nonwadeable rivers (Meador 2005; Hughes et al. 2002; Lyons et al. 2001). Previous studies of nonwadeable rivers typically focused on estimating the sampling reach necessary to characterize fish assemblages in large sections of river, concluding that sampling reaches ranging from 500 m up to 100–200 times the wetted width were required (Meador 2005; Hughes et al. 2002; Lyons et al. 2001). However, since the Ohio River in the eastern USA has an average wetted width of approximately 700 m, conducting an assessment using a reach length of even 40 times the wetted width would result in only a few sampling reaches in most pools. Such a small sample would not allow for estimates of variability in condition within the pool and would not permit resource managers to estimate the extent of local impacts. As the navigational pools created by the 18 high-lift dams are the units of assessment in the Ohio River, this is not an appropriate scale at which to conduct assessments. Previous work has shown that a 500 m reach is adequate for bioassessment at a more local scale in the Ohio River (Simon and Sanders 1999). The multimetric Ohio River Fish Index (ORFIn) was developed based on samples collected at this scale (Emery et al. 2003). The ORFIn is combined with habitat-specific criteria for assessment of individual reaches (unpublished data). However, the number of reaches necessary to adequately represent an entire pool must be determined to fully develop a monitoring program for the river.

The approach of examining patterns of species richness with sampling distance is related to the idea of capturing spatial variability in a river. By looking at a series of 500 m reaches within a pool, we are essentially trying to do the same thing. Sampling more reaches allows us to quantify more of the variability within the pool while still allowing for a more local scale assessment for the purposes of detecting and diagnosing potential sources of impairment. In this scenario, we assume that a pool is a relatively closed unit (i.e., that fishes can move freely within the pool but have restricted access to other pools upstream or downstream). Thus, one measure of how well spatial variability within the pool has been captured can be represented by the cumulative proportion of all fish species in the pool observed in a set of samples. Another measure would be the number of samples beyond which the variability in the ORFIn score is reduced to an acceptable level.

Methods

Study area

The Ohio River begins in Pittsburgh, Pennsylvania, USA, at the confluence of the Monongahela and Allegheny rivers (rkm 0) and flows southwesterly for approximately 1,579 km through six states to the confluence with the Mississippi River (Fig. 1). Currently, there are 18 high-lift and two low-head dams on the Ohio River, each providing a minimum of 2.75 m depth for commercial navigation. These dams define major pools on the river and are suspected of limiting movement of fish populations among pools. For this reason, and because any watershed management actions would likely be carried out at this scale, the pool was viewed as the appropriate level of assessment in the Ohio River for this study.

For this study, we examined data collected previously in the Ohio River as part of a Long Term Intensive Survey (LTIS). We included data only from pools with intensive systematic sampling (approximately every 4 km (2.5 mi) for the length of the pool) (Table 1). Habitat data had only been collected at a subset of sites in the early years of the fish monitoring program, although habitat is required for assessing condition of a site. Thus, we excluded any sites that lacked habitat data, and we selected only seven pools having complete data from a large number of sites (i.e., at least 18 sites). Variations in condition were expected both across pools and among years, although the goal of this study was to identify patterns with sample size and not to specifically evaluate the condition of individual pools. For all analyses, if multiple rounds of sampling had been conducted at some sites, only the first round sampling event was included to avoid potentially biasing results for the whole pool.

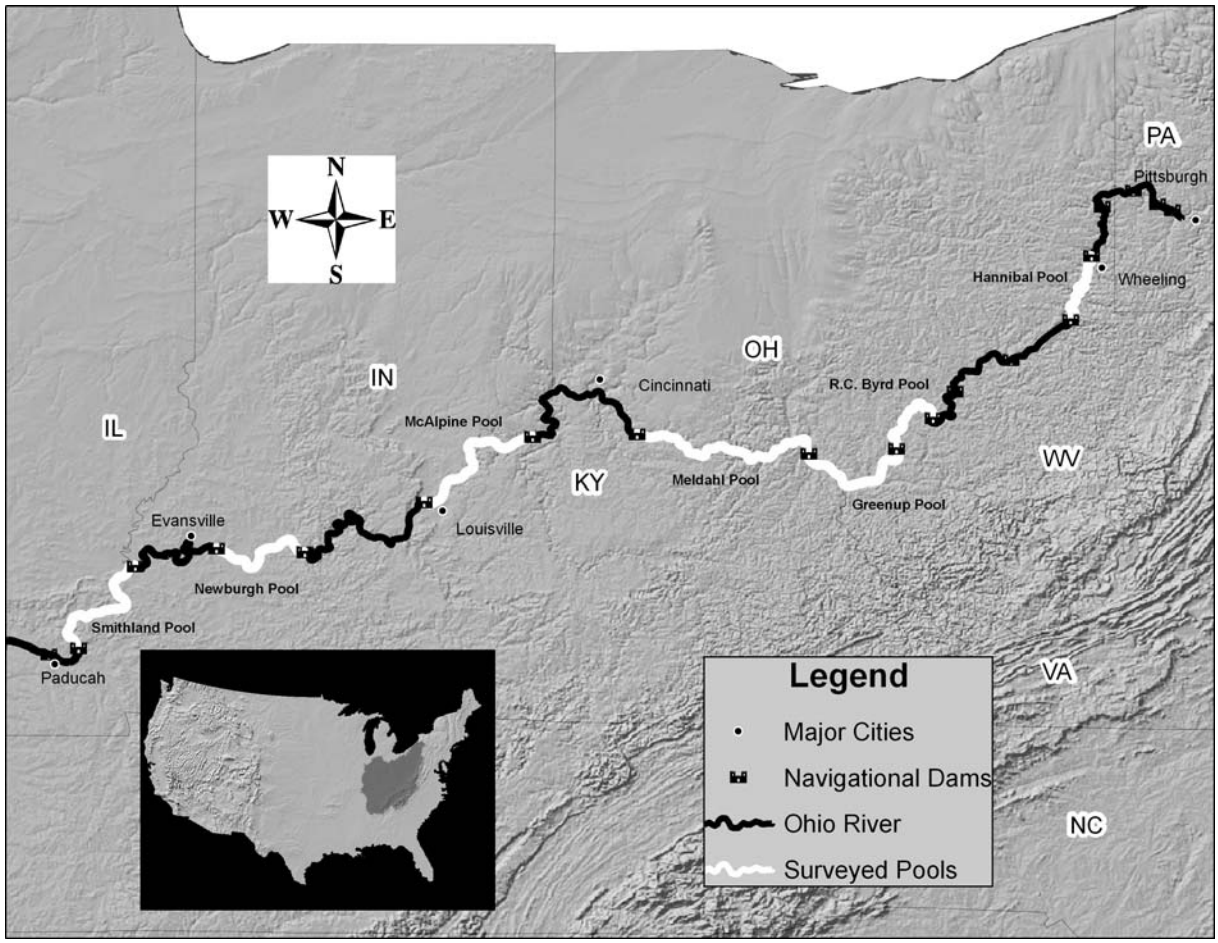

Fig. 1 Locations of intensive survey pools used in this study (highlighted in *white*) along the Ohio River

Sampling protocols

Electrofishing

All sampling events included in the analysis were conducted during the low-flow, more stable conditions of July through October and during water conditions meeting sampling criteria (i.e., minimum secchi depths of 38 cm; water levels within 61 cm of normal-flat-pool). Procedures for electrofishing followed that described by Emery et al. (2003). At each site, a 500 m reach was electrofished with a 5.5 m jon boat outfitted with an onboard generator. Electrofishing was conducted at night, as this is the established protocol used by the Ohio River Valley Water Sanitation Commission (ORSANCO) and has been documented to provide for a more representative sample of the resident fauna in deeper rivers when compared to day electrofishing (Sanders 1992). The onboard generator supplied AC power to 150-W floodlights on the bow of the boat, and to a Smith–Root Type VI-A alternator–pulsator used to convert the AC generator output to DC and then regulate the output for electrofishing. A single stainless steel ball suspended from a bow-mounted retractable aluminum boom served as the anode, with the aluminum boat hull serving as the cathode.

Each site was electrofished proceeding downstream along the shoreline at a speed equal to, or slightly greater than the prevailing current velocity. The electrofishing time at each site generally ranged from 1,800 to 5,000 s depending on the

Table 1 The seven pools of the Ohio River included in this study, along with sampling information associated with the LTIS, selected background information, and potential influences on water quality (ORSANCO 1994)

Pool	Year of sampling	Dam location (river km)	Length (km)	Average width (m)	Number of sites sampled	Tributaries (drainage > 200 mi^2 (518 km^2))	No. permitted discharges	General influences
R.C. Byrd	2002	449.3	67.1	352	24	2	30	Intersected by heavily industrialized Kanawha River
Greenup	1993	548.8	99.5	339	22	5	68	Heavily influenced by barge traffic, industry, moderate-sized towns, intersected by Big Sandy River
Hannibal	1996	203.4	58.3	345	20	2	52	Many small to moderate-sized towns, industrial sites and associated barge traffic
McAlpine	1997	976.5	118.3	622	32	2	44	Includes city of Louisville, Kentucky
Meldahl	2000	702.0	153.2	489	19	6	30	Intersected by heavily agricultural Scioto River and many smaller tributaries, several small towns
Newburgh	1994	1,249.0	89.2	755	25	2	53	Smaller industry, small towns, one moderate-sized city, minimal barge traffic or heavy industry
Smithland	1996	1,478.2	116.7	1,255	30	3	10	Intersected by agriculture-influenced Wabash River, few towns, no heavy industry

current velocity, available cover, and the number of fish encountered. Efforts were made to capture every fish sighted by the crew.

Upon capture, fish were placed in an aerated, recirculating on-board live well for processing. The majority of captured fish were identified to species, examined for external anomalies, weighed, measured for total length, and released in the field. Those requiring laboratory identification were preserved in buffered 10% formalin and later identified using regional ichthyological keys (e.g., Fishes of Ohio (Trautman 1981), Fishes of Missouri (Pflieger 1997), and Fishes of Tennessee (Etnier and Starnes 1993)). Fish measuring less than 20 mm in length (e.g., larval fish) were not recorded as they are difficult to identify accurately and offer data of questionable value to an assemblage assessment (Angermier and Karr 1986).

The occurrence of external DELT (deformities, eroded fins and body parts, lesions, and tumors) anomalies was recorded following procedures outlined by Ohio EPA (1989) and refined by Sanders et al. (1999). The frequency of DELT anomalies has been shown to be a good indication of stress caused by chronic agents, intermittent stresses,

and chemically contaminated sediments. As a result, it is a commonly used metric for assessment of rivers throughout the United States (Emery et al. 2003).

Habitat

Substrate information used in data analysis was collected in 2000 for sites sampled prior to that year, with the assumption that the basic characteristics of the habitat at a site have not changed significantly since the time of sampling. For sites sampled in 2000 and beyond, habitat was sampled within 4 months of fish sampling (i.e., during the fish sampling index period). Each 500-m sampling zone was divided into five 100-m segments, creating six points of reference for the zone along the shoreline (i.e., 0, 100, 200, 300, 400, and 500 m). At each interval, a 6 m copper pole was used to characterize the substrate at 11 points. The first measurement was taken at the shoreline with subsequent measurements at 3 m intervals towards mid-channel (total distance = 30 m). This resulted in a total of 66 point measurements within each 500-m fishing zone. Substrate was recorded as boulder, cobble, gravel, sand, fines, hardpan, or as a combination of these substrate types and used to estimate the percentage of each sediment-type within the 500-m sample area. Habitat data for some sites were unavailable.

Data analysis

There were two main objectives of the analysis. The first was to use a multimetric index of fish assemblage condition to determine the number of sites at which variation leveled off. As an additional way to gauge representativeness of sampling in the pool, the second objective was to determine the number of sites in each pool at which we captured 90% of fish species susceptible to the electrofishing configuration used.

Multimetric index

Analyses for this study involved the use of an existing multimetric index specifically developed for the Ohio River fish assemblage. The Ohio River Fish Index (ORFIn) was originally developed using LTIS data of the Ohio River and consists of 13 metrics describing various characteristics of the fish assemblage (Emery et al. 2003). Each individual sampling site was classified with respect to habitat characteristics as A (cobble substrate $\geq$ 14%), B (Cobble $<$ 14% and Sand $<$ 70%) or C (Sand $\geq$ 70%). The ORFIn score was then compared to a habitat-specific biological criterion value and the difference between the observed score and the criterion was calculated (DIFF). The three habitat-specific thresholds for ORFIn scores were quite different from one another. For habitat A, this threshold was 39, and for habitat B, it was 33 (of a possible score of 65). For habitat C, considered to be sand flats, the condition assessment was determined to be dependent on sampling date, with higher ORFIn scores obtained later in the sampling index period. Thus, the threshold was adjusted for Julian day of the sample collection based on a 25th percentile regression using a quantile regression method (Koenker and Bassett 1978). This adjustment resulted in the following formula for the criterion value: (0.12 $\times$ Julian Day -2.4) (unpublished data).

Bootstrap methods were used to simulate the random selection of sites within pools. A bootstrap approach randomly draws sites with replacement from the original dataset and assumes that the set of sites in the original dataset adequately reflects the distribution of conditions in the pool. Sampling with replacement means that during each random draw, each site has an equal probability of being selected. For example, when a subset of five sites is selected with replacement, it is possible for any particular site to be drawn all five times. This methodology is appropriate for these data because, although pools were sampled intensively for LTIS, they were not sampled in their entirety. Thus, although the systematic nature of the sampling likely resulted in a set of samples representative of the distribution of conditions in the pool, all of the possible sample locations in the pool were not specifically included in the dataset. By creating a large number of sample sets and subsets (of sites) using bootstrapping, almost any population parameter and its variance can be estimated robustly from the original dataset (Chernick 1999).

For each pool, the analysis followed the same general process. A number of sites equal to that in the full set of LTIS sites (e.g., 24 sites in R.C. Byrd pool) were selected with replacement from the set of LTIS sites (*original set*) to create a bootstrap set of sites (*full set*). The first sample of the full set was used to estimate the mean value of DIFF. Then the next sample in the full set was added to the first and the mean DIFF estimated for that set of two. This process was repeated by adding the next site each time until the full set of bootstrap sites was included in the calculation of DIFF. The entire process was repeated 1,000 times and the mean and standard deviation across all runs were used to represent the mean and standard error of the mean (Manly 1997), respectively, for each sample size. From these estimates, a 95% confidence interval can be estimated around each mean difference (as mean ± SE × 1.96) (Manly 1997), and the degree of overlap with a difference of zero determined. In this case, the confidence intervals are based on the assumption that the population standard deviation is known and the appropriate critical value from the standard normal distribution (1.96) is used. The minimum number of samples at which we see no overlap indicates the number of sites required to detect a significant difference from the criterion at the pool level, based on a mean difference and confidence interval calculated across data from all sites sampled in a given pool.

Species richness

To estimate the number of samples required to obtain approximately 90% of the total species collected within a pool, EstimateS (version 7.5, R. K. Colwell, http://purl.oclc.org/estimates) was used to model the species richness as a function of the number of samples (Colwell 2005). For each simulation run in EstimateS, a sample was selected from the full set of samples with replacement. The bias-adjusted bootstrap estimate of species richness (S_{boot}) was calculated based on Smith and van Belle (1984) as:

$$S_{\text{boot}} = S_{\text{obs}} + \sum_{k=1}^{S_{\text{obs}}} (1 - p_k)^m$$

where S_{obs} is the number of species observed in the pooled samples, p_k is the proportion of samples containing species k, and m is the total number of samples. From this calculation, it can be shown that the value of S_{boot} is maximized when each species occurs in only one sample, and S_{boot} equals S_{obs} when each species occurs in all samples from a pool (Smith and van Belle 1984). After calculating S_{boot}, another sample was selected with replacement and combined with the first sample, and the species richness was again estimated. This process continued until a number of samples equivalent to the full set of samples had been selected and combined, with species richness estimated each time using the bias-adjusted bootstrap approach. The entire procedure was repeated 1,000 times, and the average value and standard deviation across runs was determined. The standard error of the mean was then calculated by dividing the standard deviation by the square root of the sample size (i.e., number of sites) (Manly 1997). From this information, the minimum sample size required to collect an average of 90% of observed taxa within a pool was determined.

Results

Multimetric index

The average difference between the ORFIn score and the habitat-specific criterion (DIFF) leveled off to a nearly flat line almost immediately (Fig. 2a), although that is expected over such a large number of bootstrap runs. The more important feature of the bootstrapping is the estimate of the standard error (SE) of DIFF. The SE declined steeply within the first five samples, then more shallowly with additional samples, although it never leveled off completely (Fig. 2b). There is no pre-defined desirable level of confidence around the value of DIFF for a pool. However, beyond approximately 15 samples, all pools exhibited SE values consistently below three, which results in a 95% confidence interval of ±5.88 (SE × 1.96). The minimum sample size for non-overlap of CIs was relatively high for Hannibal and Meldahl pools (17 and 18 samples, respectively) and very low for

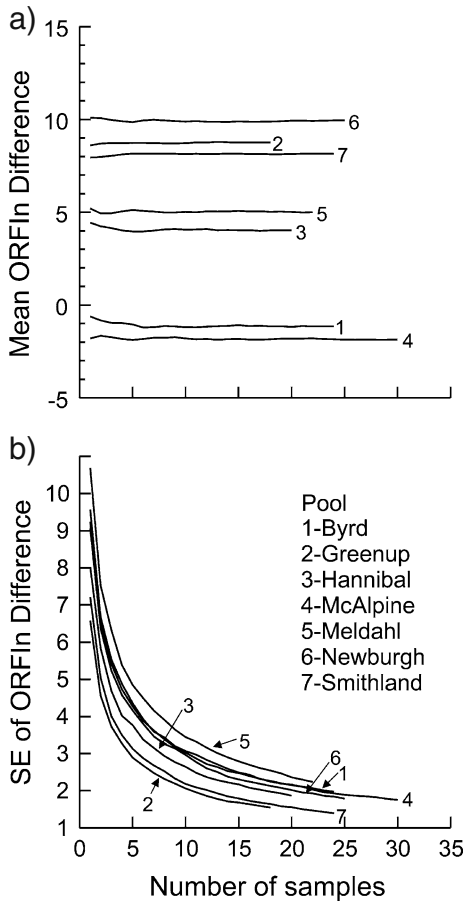

Fig. 2 (**a**) Mean and (**b**) standard error (*SE*) of difference between ORFIn score and the appropriate habitat-specific criterion as a function of sample size. Plots are based on 1,000 bootstrap randomizations for each pool

Greenup, Newburgh, and Smithland pools (three, four, and three samples, respectively). The confidence intervals for Byrd and McAlpine pools overlapped zero for all sample sizes.

Species richness

The plots of bootstrap species richness as a function of increasing sample size were shaped similarly across pools (Fig. 3a). The curves never level off, even at the maximum sample size for a given pool. However, the standard error around the estimate of mean species richness decreases sharply up to five samples and falls below one within 15 sites (Fig. 3b). The percent increase in species richness with increasing sample size leveled off below 3% within 10 samples and below 1% within 15 samples (Fig. 4a). From eight to 13 samples were required to collect 90% of the species richness found in the full set of bootstrap samples across all pools (Fig. 4b).

Discussion

Multimetric index

When conducting an assessment of water quality conditions as required by sections 305(b) and 303(d) of the Clean Water Act, it is very important to minimize Type I (pool is declared impaired when it is not) and Type II (pool is declared healthy when in fact it is impaired) assessment

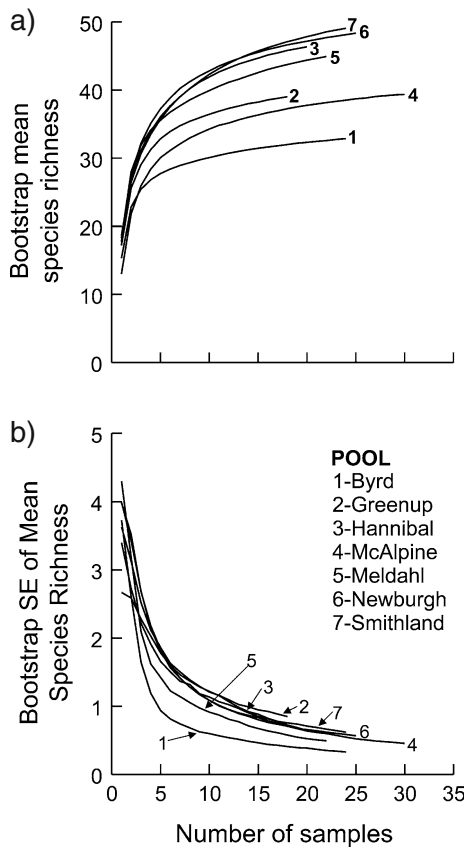

Fig. 3 (**a**) Mean and (**b**) standard error (*SE*) of bootstrap species richness as a function of sample size, plotted by pool. Plots are based on 1,000 bootstrap randomizations of species data

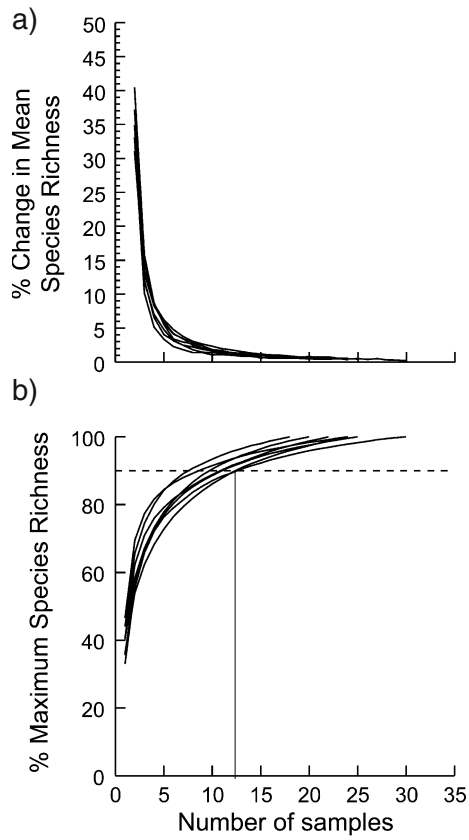

Fig. 4 Change in (**a**) bootstrap mean species richness and (**b**) percent of maximum bootstrap species richness as a function of sample size

errors when deciding if water quality standards are being met. Therefore, it is important to ensure that a sufficient number of samples have been collected so as to increase the level of confidence associated with the assessment, and decrease the likelihood of assessment error.

Mean ORFIn score differences (from criteria) showed little relationship with sample size, although variability around those mean differences did. The SE curve declined steeply up to approximately five samples, then shallowed to a negative slope close to zero within 13–15 samples. This was an important finding, providing sufficient justification of a 15-sample minimum for assessment of condition. There is no specific desired value of SE for the ORFIn or differences from criteria, so although examination of patterns is helpful, determination of the sample size at which SE is sufficiently small is difficult. However, we can use SE to calculate confidence intervals (CIs), and these CIs can be used directly in the biological assessment of a pool as a whole, assuming that mean difference is the measure of interest. For example, if the 95%CI for the mean difference from criteria does not include zero, we can conclude that on average a particular pool scores above or below the criteria. That is, we can make specific statements about the condition of that pool (i.e., meeting or not meeting criteria) with a known level of confidence. We are really only interested in cases where the difference is negative, or where the ORFIn falls below the criterion on average. However, if the CI for mean difference includes zero, we cannot make any definitive statements about the condition of that pool. For example, if the mean difference is less than zero, indicating that scores tended to fall below the criterion, but the CI still includes zero, for some random samples in that pool, the mean difference is still positive.

Note that the CI using a sample-based SE would be slightly larger than those calculated from bootstrapping. In the case of the bootstrap samples, we were estimating the true population standard error around the mean, whereas in typical practice the mean and standard error are simply estimated from a single set of samples. To calculate a 95%CI based on 15 samples (sites) in a pool, we would use the critical value for $n = 15$ and an upper tail probability of 0.025 from a t distribution. This results in a calculation of the 95%CI as: mean $\pm$ (SE $\times$ 2.131). The resulting CI is only about 8% larger than one based on critical values from the standard normal distribution. This is essentially the same as performing a one-sample t-test against a null hypothesis that the mean difference is zero.

In any case, as we add samples, we have more confidence in the true degree and direction of difference from the criteria for the pool, decreasing the potential for error when assessing water quality conditions. However, at some point, additional samples do not really provide much additional information for the purposes of assessment. In cases like the example where the CI overlaps zero, a relatively large number of samples might be required to achieve a small enough CI to provide a strong assessment. Resource managers will have to de-

cide when the sample size has reached a sufficient level as to balance effort expended and confidence gained. One could argue that once a sample size of 50 is reached, regardless of whether the CI for mean difference includes zero, an estimate of condition can be made. This limit would be justified as a maximum expenditure of effort to reduce the potential for error and describe the truest picture of condition possible, given resources available for sampling.

Species richness

Species richness curves never leveled off completely, but the slope became shallow enough that few new taxa were being added with each additional sample. This scenario is commonly seen with species richness curves (Hughes et al. 2002; Bady et al. 2005; Smith and Jones 2005; Lapointe et al. 2006; Melo et al. 2007), with new rare taxa being added as sample size increases. This is an inevitable situation when sampling is not intended to collect all species but rather a representative sample. In any case, the bootstrap SE dropped below one for all pools within 15 samples, and the estimate of species richness increased only marginally with each additional sample beyond about ten samples. Within about 13 samples, 90% of total species richness was achieved. From this information, we can conclude that approximately 15 samples are sufficient to estimate fish species richness at the pool level with minimal variability in our estimates.

Conclusions

The two analysis components of this study (i.e., ORFIn scores and species richness) both point to a sample size of 15 to initially estimate condition at the pool scale in the Ohio River. From the perspective of variability in ORFIn scores, however, subsequent sampling may be required to achieve adequate precision of condition estimates. The patterns that emerge from the data generally are consistent across pools, regardless of the level of variability within a pool. The seven pools included in this study varied in potential impacts to water quality and in habitat diversity, but patterns in species richness and in variability associated with the ORFIn were similar across these differences. For example, habitat classified as a mixture of sand and cobble (habitat B) was common in all seven pools, but the presence and relative abundance of cobble and sand habitats varied across pools. For example, Newburgh pool had no cobble (habitat A) and mostly sandy (habitat C) sites, whereas Hannibal pool tended to have more cobble than sand sites. Greenup, Meldahl, McAlpine, and Smithland pools all had more sandy than cobble sites, and Byrd pool had no sandy and few cobble sites. The influences in each pool that might affect water quality (e.g., contribution by tributaries, discharges) also differed across pools (Table 1), with varying levels of industry and sizes of towns along the banks of each pool. Still, these differences did not seem to directly affect the patterns seen.

Overall, the ability of ORSANCO to report on the biological condition of the Ohio River should improve with the use of a random sampling design, relative to the more intensive systematic one used in the past. This improvement is primarily due to the reduced effort required to employ a random design, allowing for more pools to be sampled with a limited set of resources. Although the typical recommended minimum sample size for this type of design is 50 sites per assessment unit (i.e., pool) (U.S. EPA Aquatic Resources Monitoring web site, http://www.epa.gov/nheerl/arm/surdesignfaqs.htm), a much smaller number of samples can provide enough information for assessment in some pools of the Ohio River. This study indicated that 15 sites may be adequate to draw conclusions about the overall condition of a navigational pool with the desired level of precision. However, the sufficiency of this smaller number of sites depends on the variation in condition of the fish assemblage within a given pool. The more consistent water and habitat quality are throughout a pool, the more consistent we would expect the quality of the fish assemblage to be, and the fewer samples that are needed to characterize the biological condition of that pool. In pools where greater variation exists, precision could be assessed at the end of 1 year, and additional sampling could be added the following year to reach the desired level of confidence

in the biological condition. Because there are a large number of pools in the Ohio River, an approach that allows ORSANCO to sample and assess more pools each year will result in a more robust assessment of the river for the purposes of reporting as required by U.S. EPA. The sampling procedures described in this paper are under consideration for formal adoption by ORSANCO as part of their biological assessment program.

Acknowledgements We thank Mark Barath, U.S. EPA Region III in Philadelphia, PA, and Tony Olsen and Dave Bolgrien, U.S. EPA Office of Research and Development in Corvallis, OR, and Duluth, MN, respectively, for the comments on a draft report on which this manuscript is based. We thank Joe Flotemersch, U.S. EPA Office of Research and Development in Cincinnati, OH, and an anonymous reviewer for the comments on an earlier draft of this paper. The data used in this research could not have been collected without the tireless efforts and skill of ORSANCO field crews. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the U.S. Environmental Protection Agency. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

References

Angermier, P. L., & Karr, J. R. (1986). Applying an index of biotic integrity based on stream-fish communities: Considerations in sampling and interpretation. *North American Journal of Fisheries Management, 6*, 418–427. doi:10.1577/1548-8659(1986)6<418:AAIOBI>2.0.CO;2.

Bady, P., Doledec, S., Fesl, C., Gayraud, S., Bacchi, M., & Scholl, F. (2005). Use of invertebrate traits for the biomonitoring of European large rivers: The effects of sampling effort on genus richness and functional diversity. *Freshwater Biology, 50*, 159–173. doi:10.1111/j.1365-2427.2004.01287.x.

Chernick, M. R. (1999). *Bootstrap methods: A practitioner's guide.* New York: Wiley.

Colwell, R. K. (2005). *EstimateS 7.5 user's guide*. Retrieved January 2, 2008, from University of Connecticut Web site: http://purl.oclc.org/estimates.

Dauwalter, D. C., & Pert, E. J. (2003). Electrofishing effort and fish species richness and relative abundance in Ozark highland streams of Arkansas. *North American Journal of Fisheries Management, 23*, 1152–1166. doi:10.1577/MO1-144.

Emery, E. B., Simon, T. P., McCormick, F. H., Angermeier, P. L., DeShon, J. E., Yoder, C. O., et al. (2003). Development of a multimetric index for assessing the biological condition of the Ohio River. *Transactions of the American Fisheries Society, 132*, 791–808. doi:10.1577/T01-076.

Etnier, D. A., & Starnes, W. C. (1993). *The fishes of Tennessee*. Knoxville: The University of Tennessee Press.

Hughes, R. M., Kaufmann, P. R., Herlihy, A. T., Intelmann, S. S., Corbett, S. C., Arbogast, M. C., et al. (2002). Electrofishing distance needed to estimate fish species richness in raftable Oregon rivers. *North American Journal of Fisheries Management, 22*, 1229–1240. doi:10.1577/1548-8675(2002)022<1229:EDNTEF>2.0.CO;2.

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica, 46*, 33–50. doi:10.2307/1913643.

Lapointe, N. W. R., Corkum, L. D., & Mandrak, N. E. (2006). A comparison of methods for sampling fish diversity in shallow offshore waters of large rivers. *North American Journal of Fisheries Management, 26*, 503–513. doi:10.1577/M05-091.1.

Lyons, J. (1992). The length of stream to sample with a towed electrofishing unit when fish species richness is estimated. *North American Journal of Fisheries Management, 12*, 198–203. doi:10.1577/1548-8675(1992)012<0198:TLOSTS>2.3.CO;2.

Lyons, J., Piette, R. R., & Niermeyer, K. W. (2001). Development, validation, and application of a fish-based index of biotic integrity for Wisconsin's large rivers. *Transactions of the American Fisheries Society, 130*, 1077–1094. doi:10.1577/1548-8659(2001)130<1077:DVAAOA>2.0.CO;2.

Manly, B. F. J. (1997). *Randomization, bootstrap, and Monte Carlo methods in biology.* London: Chapman and Hall.

Meador, M. R. (2005). Single-pass versus two-pass boat electrofishing for characterizing river fish assemblages: Species richness estimates and sampling distance. *Transactions of the American Fisheries Society, 134*, 59–67. doi:10.1577/ft03-094.1.

Melo, A. S., Bini, L. M., & Thomaz, S. M. (2007). Assessment of methods to estimate aquatic macrophyte species richness in extrapolated sample sizes. *Aquatic Botany, 86*, 377–384. doi:10.1016/j.aquabot.2007.01.005.

Ohio EPA (1989). *Addendum to: Biological criteria for the protection of aquatic life: Volume II: Users manual for biological field assessment of Ohio surface waters.* Columbus, Ohio: Ohio Environmental Protection Agency.

ORSANCO (1994). *Ohio river water quality fact book.* Cincinnati, Ohio: Ohio River Valley Water Sanitation Commission.

Pflieger, W. L. (1997). *The fishes of Missouri.* Jefferson City, Missouri: Missouri Department of Conservation.

Reynolds, L., Herlihy, A. T., Kaufmann, P. R., Gregory, S. V., & Hughes, R. M. (2003). Electrofishing effort requirements for assessing species richness and biotic integrity in western Oregon streams. *North American Journal of Fisheries Management, 23*, 450–461. doi:10.1577/1548-8675(2003)023<0450:EERFAS>2.0.CO;2.

Sanders, R. E. (1992). Day versus night electrofishing catches from near-shore waters of the Ohio and

Muskingum rivers. *The Ohio Journal of Science, 92*, 51–59.

Sanders, R. E., Miltner, R. J., Yoder, C. O., & Rankin, E. T. (1999). The use of external deformities, erosions, lesions, and tumors (DELT anomalies) in fish assemblages for characterizing aquatic resources: A case study of seven Ohio streams. In T. P. Simon (Ed.), *Assessing the sustainability and biological integrity of water resources using fish communities* (pp. 225–248). Boca Raton, FL: CRC.

Simon, T. P., & Sanders, R. E. (1999). Applying an index of biotic integrity based on great river fish communities: Considerations in sampling and interpretation. In T. P. Simon (Ed.), *Assessing the sustainability and biological integrity of water resources using fish communities* (pp. 475–506). Boca Raton, FL: CRC.

Smith, K. L., & Jones, M. L. (2005). Watershed-level sampling effort requirements for determining riverine fish species composition. *Canadian Journal of Fisheries and Aquatic Sciences, 62*, 1580–1588. doi:10.1139/f05-098.

Smith, E. P., & van Belle, G. (1984). Nonparametric estimation of species richness. *Biometrics, 40*, 119–129. doi:10.2307/2530750.

Trautman, M. B. (1981). *The fishes of Ohio*. Columbus, OH: Ohio State University Press.