Biological Assessment and Criteria

Tools for Water Resource Planning and Decision Making

Edited by
Wayne S. Davis
Thomas P. Simon

LEWIS PUBLISHERS
Boca Raton London Tokyo

Library of Congress Cataloging-in-Publication Data

Biological assessment and criteria : tools for water resource planning
and decision making / edited by Wayne S. Davis, Thomas P. Simon.
p. cm.
Includes bibliographical references and index.
ISBN 0-87371-894-1
1. Water quality biological assessment. 2. Water quality management.
3. Environmental monitoring. 4. Indicators (Biology). I. Davis, Wayne S. II. Simon, Thomas P.
QH96.8.B5B54 1994
363.73'94—dc20 94-27369
CIP

International Standard Book Number 0-87371-894-1
Library of Congress Card Number 94-27369
Printed in the United States of America 1 2 3 4 5 6 7 8 9 0
Printed on acid-free paper

CHAPTER 11

Biocriteria: A Regulated Industry Perspective

Robin J. Reash

1.0 BIOCRITERIA: NEW CHALLENGES FOR REGULATED INDUSTRY

Biocriteria (biological criteria) are being developed and implemented with increasing frequency as water quality management tools for state agencies responsible for administering the National Pollutant Discharge Elimination System (NPDES) program. Relying on U.S. Environmental Protection Agency's (USEPA) national guidance on biocriteria (USEPA 1990a) and national policy on biosurveys/biocriteria (USEPA 1991c), states are initiating programs that will result in the adoption of narrative and/or numeric biocriteria within water quality standards. The statutory mandate of required biocriteria development is somewhat controversial. Though USEPA's position is that the development and adoption of biocriteria is required by states within the water resource management program, it should be noted that some industry groups have questioned USEPA's authority to require biocriteria under the Clean Water Act. Nonetheless, USEPA has reaffirmed its position that states must initially adopt narrative biocriteria to comply with statutory requirements under Sections 303 and 304 of the Clean Water Act. Specifically, USEPA has advocated that states are to adopt narrative biological criteria into state water quality standards during the FY 1991–1993 triennium and then adopt numeric biocriteria by the end of FY 1996 (USEPA 1990). Regulated industry must become aware of this implementation "clock" and closely follow the development of biocriteria in individual states, whatever the implementation schedule happens to be. Because so few states have progressed to the point of proposing legally binding narrative or numeric biocriteria at this time, industry involvement has been relatively limited on a national scale.

The State of Ohio has led all other states in the derivation and regulatory adoption of numeric biocriteria, currently being the only state with approved biocriteria for tiered aquatic life use designations within ecoregions (see Southerland and Stribling, Chapter 7; Yoder and Rankin, Chapter 9). With the exception of facilities on specific waterbodies (e.g., acid mine drainage streams, the Ohio River, and Lake Erie) all regulated dischargers in Ohio must consider their compliance with numeric biocriteria during the NPDES permit renewal period.

American Electric Power Company (AEP) subsidiaries have several coal-fired electric generating facilities that are permitted by the Ohio EPA. Ohio Power Company's Muskingum River Plant and Columbus Southern Power Company's Conesville Generating Station are located on the Muskingum River (Figure 1) and are subject to Ohio EPA's inland biocriteria. Both Ohio EPA and AEP have conducted numerous biosurveys near these plants and each has assessed compliance with ecoregion-based biocriteria using these biosurvey results. A detailed discussion of these results is given in Section 2.3.

This chapter summarizes the technical challenges of biocriteria compliance from a regulated industry perspective. As previously mentioned, industry has not been forced to address issues related to numeric biocriteria on a national scale. At this time, only regulated industries in Ohio must comply with ecoregion-based numeric biocriteria. The facilities operated by AEP subsidiaries represent only a few of the many

0-87371-894-1/95/$0.00+$.50

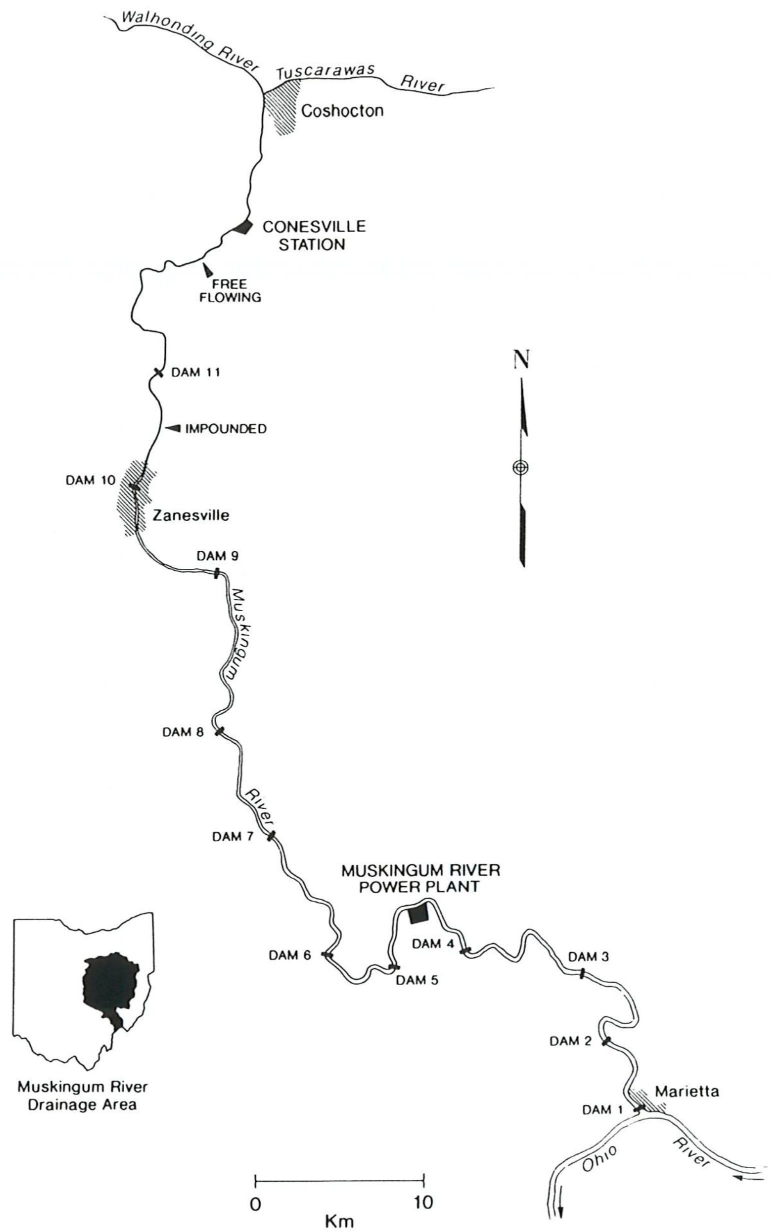

Figure 1. Location of two coal-fired power plants on the Muskingum River: Conesville Station (located on free-flowing portion) and Muskingum River Plant (located on impounded portion). Navigation dams are indicated relative to power plants and major cities.

facilities that are regulated under Ohio EPA's NPDES program. Therefore, permitting experiences for AEP facilities are unique and cannot be considered representative of Ohio's regulated industry as a whole. Biosurvey data are unique for individual facilities and Ohio EPA assesses compliance with all applicable criteria by site. The following discussion on technical advantages and deficiencies of biocriteria implementation in Ohio is intended to stimulate discussion among industry and regulatory agencies to address unresolved issues, and to encourage other industries to work with state agency staff at all stages in biocriteria development.

1.1 Biocriteria as a New Regulatory Tool

The development and adoption of biocriteria into state water quality standards poses challenges as well as opportunities. As a new regulatory tool, industry must determine the effects that biocriteria compliance will have on their individual facilities. Such assessments are familiar to industry. Projections of compliance have been, and are continually being conducted, for chemical-specific and whole effluent toxicity criteria in Ohio and all other states. The real challenge for industry and state agencies is that biocriteria, being an integrated biologically based regulatory tool, will vary tremendously both among states and within states. There can be no "national" biocriteria that all states or even several States may adopt as identical, in contrast to national water quality criteria. Even the adoption of "generic" narrative biocriteria within nearby states may not be valid because biological expectations may differ tremendously. It seems reasonable that states must assess at least some state-specific biosurvey data and existing land-use impacts before even a valid narrative biocriteria can be developed. During this early stage, it will be prudent for industry to closely follow agency developments. Some facilities may want to assume a proactive role by gathering data using state-approved methodologies, or methodologies that agencies are currently investigating. The risks and benefits of this proactive strategy are discussed in more detail below.

There are technical challenges for industry during the development of state-specific biocriteria (Van Hassel et al. 1992). Understanding the underlying concepts and theoretical assumptions of biocriteria may be a challenge to environmental managers that have little or no experience with biological assessments. Clearly, the regulation of complex wastewaters through biocriteria will be different compared with the traditional chemical-specific/end-of-pipe approaches. Ohio EPA (1987a, 1990a) has published comprehensive overviews of biocriteria rationale, whereas the overview by Karr et al. (1986) is generally regarded as the primary publication advocating using biological integrity within water quality management programs. Environmental managers within industry can become well acquainted with basic biocriteria concepts by reviewing the four documents cited above. Further discussion of specific challenges to industry are summarized below.

1.1.1 Differences With Chemical-Specific Criteria

Chemical-specific criteria are discrete values that should not be exceeded to protect a designated use. By using design flows and standard mass-balance equations, water quality-based effluent limits (WQBELs) can be reasonably predicted using chemical-specific criteria. Furthermore, the analysis of wastewaters for specific constituents is relatively rapid and a sizeable database can be generated within a short period. Biocriteria typically represent a composite of independent metrics or calculations that are summed to produce a value without units (e.g., the Index of Biotic Integrity and Modified Index of Well-Being). Predicting compliance with biocriteria will typically require more data, be more costly, and require staff biologists (or consultant biologists) who are familiar with data analysis techniques and interpretation. Instream biological data can vary considerably both temporally and spatially: thus, replicate samples are necessary to define the biological integrity near a particular facility. Though Ohio EPA has determined that instream biological assessments are cost-effective when compared with chemical-specific and toxicity testing methods (Ohio EPA 1990a), the agency has spent considerable years refining its field protocols and data analysis procedures (see Yoder and Rankin, Chapter 9).

Although there are fundamental differences in predicting compliance with chemical-specific criteria vs. biocriteria, industry should perform *both* chemical and biological analyses simultaneously. The collection of paired data allows testing the hypothesis that community-level measures respond predictably with varying concentrations of measured pollutants. Some regulatory agencies rely on a limited number of biosurveys often conducted in one year only and without measured water quality data to assess potential wastewater impacts. Long-term monitoring data, although costly to obtain, are invaluable for industries who need to refute or confirm perceived impacts based on small sample sizes.

Like chemical-specific criteria, assessment of whole effluent toxicity (WET) compliance for a particular facility may or may not be comparable with biocriteria compliance. Ohio EPA (1990b) studied the level of agreement between bioassay and instream biosurvey results using data collected near 43 separate facilities. The agency found good agreement between results in about 20% of cases when effluent (end-of-pipe) toxicity was assessed, and in about 30% of cases when mixing zone toxicity was measured. There is some empirical evidence that instream biological response and instream ambient toxicity are closely associated (Dickson et al. 1992). Marcus and McDonald (1992) caution against wide-scale extrapolation of these study results, citing biases in statistical design and a lowered association between instream measures and ambient toxicity when more subtle, or borderline, toxicity is occurring. Based upon a review of relevant studies, the author believes that WET results, especially instream toxicity assessments, will more closely parallel biocriteria compliance compared to using a chemical-specific approach. The reason is that WET and instream toxicity tests are a true biological-response assessment. Nonetheless, industry will likely be forced to utilize facility-specific strategies and data requirements when assessing biocriteria compliance. These strategies must rely on assessing various combinations of chemical-specific and WET criteria in the absence of instream biosurvey data.

1.1.2 Understanding How They Are Derived

In general, numeric biocriteria will be derived using one of two approaches: the site-specific reference condition, or regional reference condition (USEPA 1990a). Industry personnel should understand the difference between the two approaches and establish a dialogue with regulatory agencies in order to offer their insight on the most valid approach. Industry is already familiar with the site-specific reference approach, where a traditional upstream vs. downstream design is used to assess potential discharge impact. The regional reference design can result in biocriteria based on a paired watershed approach (one watershed impaired, one watershed unimpaired), or an ecoregion-based approach (USEPA 1990a). Evidently, some states are developing numeric biocriteria with inadequate sample sizes, e.g., one field season of sampling at ecoregion-specific reference sites. This "one year, one pass" sampling design has basic ecological flaws, and competent agency biologists should readily understand the pitfalls of such a meager database. Industry personnel should become aware of the process that states are using to derive biocriteria and communicate their concerns when technical flaws may result in unrealistic, or invalid, biological expectations outside of reference sites.

1.1.3 The Statutory Basis of Biocriteria

The statutory basis of biocriteria has been summarized by USEPA (1990a, 1991c). Industry should be aware of the principal statutory provisions pertaining to biocriteria. Sections 303, 304, and 308 of the Clean Water Act contain pertinent language for development of protective criteria and biological assessment methods. (Please see Adler, Chapter 22 for additional information on statutory bases for biological criteria.)

1.2 Compliance With Biocriteria

Once narrative or numeric biocriteria are developed, proposed, and adopted into state water quality standards (and subsequently approved by USEPA), industry must choose strategies that will ensure compliance with biocriteria, as well as any applicable chemical-specific or WET criteria. As previously discussed, projecting compliance with biocriteria may be difficult, especially if little or no instream data are available. Industries that choose a proactive strategy by performing biosurveys that are not required will obtain results that allow a direct assessment of biocriteria compliance. Moreover, if a regulatory agency performs biosurveys to assess attainment of applicable biocriteria, data collected by a permittee can be used to expand the site-specific database, and more importantly, validate or contest results obtained through agency studies.

1.2.1 Biocriteria as Part of "Independent Application"

USEPA has issued a policy of "independent application" for all applicable criteria to determine whether designated uses are being attained (USEPA 1991f). This policy affirms that an exceedance of any applicable chemical-specific, WET, or biocriteria will result in the nonattainment of a designated use despite evidence indicating that the other two criteria are, or will not be, exceeded. This policy not only integrates all three types of assessments but requires that the most stringent (i.e., limiting) criteria be applied in water quality-based toxics control programs. This position is summarized in USEPA's (1991c) guidance on biocriteria as follows:

> The failure of one method to confirm an impact identified by another method would not negate the results of the initial assessment. This policy, therefore, states that appropriate action should be taken when any one of the three types of assessments determine that the standard is not attained.

There are some significant technical flaws with this policy if implemented on a wide-scale basis, with no exceptions. Miner and Borton (1991) provide arguments against the independent application approach, emphasizing the underlying statistical requirements for such an approach. There may be instances when states are justified using this policy in the effluent characterization and hazard assessment process. When there is little or no information on potential receiving stream impacts and/or whole effluent toxicity, states should use a conservative approach in permitting the particular facility.

In contrast, there are some instances where the independent application approach is not justified because it may result in overprotective effluent limitations. In general, the independent application approach assumes that both the amount and quality (i.e., relevance) of data used to assess all three criteria types are equivalent (USEPA 1991f). In most instances, however, this will not be true. Many facilities will likely have considerable data for one assessment but relatively little for another. In such cases a risk assessment approach (not an independent application approach) should be taken and appropriate effluent limitations be required that (1) will protect, but not overprotect, the receiving stream, (2) are commensurate with the results of existing data, and (3) utilize a weight-of-evidence, best professional judgement process. Considerations of statistical power should be integrated into the risk assessment process. Thus, a weight-of-evidence approach can become increasingly obvious when hypothesis testing (statistical power) demonstrates the unnecessary usage of independent application.

Chemical-specific and whole effluent approaches are tools that only predict a level of protection for instream aquatic life whereas the biocriteria approach measures the actual level of protection. Logically, there will be instances when demonstrated compliance of biocriteria should take precedence over one or both of the other criteria. This is especially relevant when a chemical-specific criterion is being exceeded but the biological community indicates no impairment. Many water quality criteria were derived from toxicity tests that exposed test organisms to highly bioavailable fractions of toxicants. Unless site-specific adjustments of protective chemical criteria are made where toxicity mitigation is demonstrated, many facilities could have effluent limitations that are overly protective. Overly protective effluent limitations may require costly alternate treatment technologies that cannot reasonably be expected to provide real environmental benefits.

In summary, industries must understand the decision-making policies that govern the water quality-based toxics control program in each state. If biocriteria have not been adopted then the issue of independent application is absent, or less complicated. In states where numeric biocriteria have been adopted, the issue of independent application may be complicated. In general, regulatory agencies *and* industry are favored when states have some flexibility in administering the NPDES program. USEPA is correct in stating that required compliance with all criteria results in a powerful regulatory tool (USEPA 1991f). The independent application policy, however, seems to dismiss the professional judgment that states often utilize when a differing array of site-specific data are available.

1.2.2 Interaction With Regulatory Agencies

There are many opportunities where industry can interact with agencies during the biocriteria development process. Interaction during the early formulation stage is critical if regulated industry desires a voice in strategy and timetable. Besides attendance at agency-sponsored workshops and public meetings, industry biologists can inspect field-sampling and sample-processing techniques. Regulatory agencies should have well-established quality assurance procedures for field and laboratory protocols. For biocriteria development, expertise in fish and macroinvertebrate identification is crucial, and a standardized chain of custody procedure should be verified by industry. Ohio EPA (1987) has published a peer-reviewed standard operating procedures manual for fish and macroinvertebrate assessments. Because these standardized methods are used to gather data for biocriteria index calculations, regulated industry in Ohio has learned that close adherence to these methods is beneficial in resolving use attainment issues.

There are many advantages in establishing a constructive dialogue between regulated industry and agencies during the strategy planning phase. Resolving technical issues at the formulation stage is unmistakably far more profitable than trying to resolve disputes after strategy and methodologies are set.

2.0 THE OHIO EXPERIENCE

2.1 Overview of Ohio EPA Biocriteria Requirements

By adopting numeric biocriteria into Ohio's water quality standards, Ohio EPA explicitly requires one (and only one) approach in demonstrating attainment of the aquatic life use: compliance with biocriteria. In other words, Ohio EPA considers biocriteria indices a direct gauge of attainment or nonattainment of the aquatic life use. Unless it can be demonstrated that the designated aquatic life use of a waterbody cannot be attained due to habitat limitations or long-term irretrievable conditions, effluent limitations will be modified as appropriate when noncompliance of biocriteria has been demonstrated. Logically, there must be some empirical evidence indicating a relationship between a facility discharge and nonattainment of applicable biocriteria. This is a crucial question when *reasons* for biocriteria nonattainment are elucidated. Often, debates between regulated industry and Ohio EPA center on the interpretation of biosurvey data, especially regarding the dynamic relationship between instream pollutant levels and measured biological response. A robust statistical treatment of these data often can help determine whether or not the biological response is attributable to the measured pollutant.

Ohio's biocriteria vary among ecoregions and size of the waterbody. With the establishment of numeric biocriteria, however, industry should have no confusion as to what target biological community must be attained to comply with biocriteria. In Ohio, industry cannot change the regulatory implementation of biocriteria, but they can collect data which either confirms or challenges Ohio EPA's assessment of use nonattainment.

Ohio EPA has established a rule that attainment of biocriteria should be granted disproportionate weight in demonstrating overall use attainment, due to the fact that chemical-specific and WET criteria are only surrogate measures of biological integrity. This precedence for biological criteria is embedded in Ohio's statutory water quality standards:

> Demonstrated attainment of the applicable biological criteria in a waterbody will take precedence over the application of selected chemical-specific or whole-effluent criteria associated with these uses. (Ohio Revised Code 3745–1-07)

Industry should be aware that (1) demonstrated nonattainment of the biological criteria may result in more stringent chemical-specific limitations, and (2) USEPA has not yet approved this departure from its policy of independent application. Until the issue of independent application is resolved, Ohio's regulated industry must develop a strong technical case that biocriteria are met near a given facility or that nonattainment is due to other factors independent of a facility's operations.

2.2 Technical Validity of Ohio EPA Biocriteria

From an industry perspective, the most attractive feature of Ohio EPA's biocriteria is the empirical foundation. In general, Ohio EPA derived numeric biocriteria using a systematic standardized sampling of least impacted reference sites in each of the five ecoregions (see Yoder and Rankin, Chapter 9). From the standpoint of defining *regional* expectations of biological performance, the validity of actual Index of Biotic Integrity, Modified Index of Well-Being, and Invertebrate Community Index values has been demonstrated. Ohio EPA's Users Manual (Ohio EPA 1989) provides a good overview of how reference site data were translated into regional biocriteria. Ohio EPA has conducted biosurveys throughout the state for nearly 20 years. Replicate samples were taken at most sites, including ecoregion reference sites. The extensive database compiled by Ohio EPA has allowed the derivation of defensible biocriteria values.

Fore et al. (1994) analyzed the statistical properties of Index of Biotic Integrity values among several Ohio streams and concluded that, although IBI scores at a site are not distributed normally, the usage of a two-sample t-test and ANOVA model could be cautiously applied for hypothesis testing.

Ohio EPA's statewide database should serve as a model to other states who are developing numeric biocriteria. Though some states may be tempted to gather data quickly without an adequate number of replicates at several reference sites, this strategy is not sound and will only result in poorly validated metric cutoffs. Some states, in addition, tend to ignore data variability rather than let the variability determine the resultant criteria. The inherent variability of biological community data demands that adequate replicates be taken. More importantly, this variability must be defined empirically and incorporated into resulting biocriteria. Ohio EPA correctly analyzed the biological data variability and incorporated this variability into the numeric biocriteria.

There are technical aspects of Ohio EPA's biocriteria that had no precedent and thus required best professional judgement. Fausch et al. (1990) state that one disadvantage of using the IBI is the subjectiveness of defining metric criteria (e.g., maximum species richness lines). One notable item is the 25th and 75th percentile cutoff for ecoregion-specific biocriteria, which Ohio EPA uses. The specific percentile cutoff for biocriteria compliance or noncompliance varies with use designation. The 25th percentile of each biocriteria index, based on ecoregion-specific reference site data, is considered the minimum criterion for the Warmwater Habitat use designation. For waters that support highly diverse communities (the Exceptional Warmwater Habitat use), the 75th percentile value of the combined statewide reference site data is used as the minimum criterion. The designation of these percentile values to judge attainment of designated values seems reasonable. From an industry perspective, the percentile cutoffs are acceptable because they recognize the variation in index scores within a given ecoregion. For example, the designation of 25th percentile for the minimum attainment value acknowledges that 25% of the reference sites in a given ecoregion do not attain this minimum value, for a given index. Thus, even if a discharger has a facility located on a "least impacted" waterbody, there is a 1 in 4 chance that a given index score will fall below the ecoregion minimum due to factors independent of the facility's operations.

Even though sites that are not influenced by point-source discharges should not have an expectation of 100% biocriteria compliance, Ohio EPA still must interpret results carefully when upstream reference sites show a low compliance frequency. Ohio EPA uses an averaging of index scores to compare with the minimum ecoregion criterion. Thus, a discharger does not have to demonstrate 100% compliance with minimum criteria; rather, the average value is compared to the percentile cutoff value. When a relatively small database (e.g., three replicates or passes in a given field season) must be analyzed, comparisons of site-specific average scores to established biocriteria or upstream scores may be appropriate. When a larger database is available, however, the agency should use true hypothesis testing in accordance with USEPA policy (USEPA 1991c).

2.3 Technical Problems With Ohio EPA Biocriteria

Using standardized methodologies, Ohio EPA has developed three biocriteria indices that are implemented on an ecoregion-specific basis. With numeric biocriteria, the agency can assess numerous waterbodies for attainment of the aquatic life use based on regional expectations. Temporal trends in biological performance, especially useful for waterbodies that had varying degrees of historical impact, can be tracked with greater resolution when numerical community-based indices are used as a benchmark.

There are some technical problems with Ohio EPA's biocriteria. These flaws do not undermine the fundamental concepts of Ohio EPA's biocriteria, but they will require some rethinking of certain implementation steps. These problems are applicable to any state that develops ecoregion-based biocriteria, and are not unique to Ohio. Three problem areas discussed below are (1) site-specific modification of biological expectation, (2) biocriteria for regulated vs. unregulated waterbodies, and (3) derivation of large river biocriteria, with emphasis on the Ohio River.

2.3.1 Site-Specific Modification of Biological Expectation

Ohio EPA's biocriteria were developed to define *regional* expectations of biological performance. Several reference sites were sampled in each ecoregion to provide numerical estimates of "least impacted"

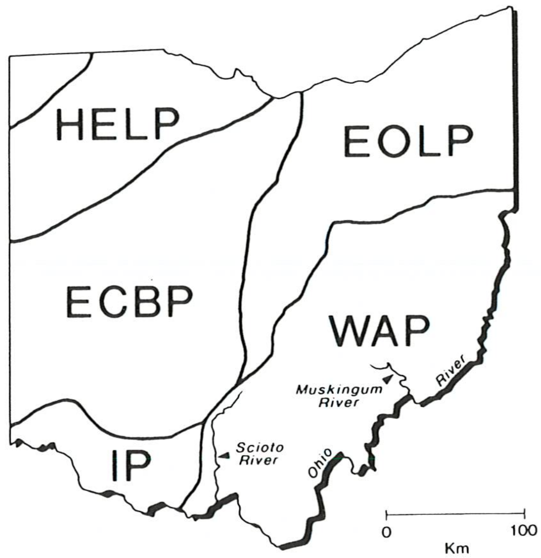

Figure 2. Location of lower Muskingum River and lower Scioto River relative to the Ohio River. Geographic locations of Ohio ecoregions are indicated. WAP = Western Allegheny Plateau ecoregion. (From Ohio EPA. 1992.)

conditions. Ohio EPA acknowledges, albeit briefly, that site-specific modifications may be appropriate in some circumstances:

> In situations where the biological criteria are not met because of the natural attributes of the surface water and/or watershed, a site-specific modification of the criteria may be performed. This procedure recognizes that there may be habitats that do not meet the ecoregional criteria due to unique, site and/or watershed specific characteristics. (Ohio EPA 1989, pp. 7–5)

Site-specific habitat constraints are a possible reason for nonattainment, but these can be addressed using an intensive habitat assessment or Ohio EPA's Qualitative Habitat Evaluation Index (see Rankin, Chapter 13). The extrapolation of minimum criteria to sites having drainage areas much larger than ecoregion-specific reference sites may also cause nonattainment. In such cases, the calibration of indices is exceeded due to a lack of comparable reference sites. An example of this technical problem is the experience at one of American Electric Power Company's coal-fired generating facilities, Conesville Station. Conesville Station is located on the Muskingum River at River Mile 102 (distance from Ohio River confluence), near Coshocton, Ohio (Figure 1). The drainage area of the Muskingum River at Conesville Station is 4882 mi^2. The mainstem Muskingum River (and most of the entire drainage area) lies within the Western Allegheny Plateau (WAP) ecoregion (Figure 2). Ohio EPA has adopted minimum biological criteria for the WAP ecoregion (pertaining to Warmwater Habitat aquatic life use) as 40 (IBI), 8.6 (Modified Index of Well-Being, or MIwb) and 36 (Invertebrate Community Index, or ICI). These criteria were derived from collections at least impacted reference sites, none of which were on the Muskingum River mainstem.

One relevant question is: how comparable are reference sites to the mainstem Muskingum River near Conesville Station? Ohio EPA does list the location and drainage area of all ecoregion reference sites (September, 1989 addendum to the Users Manual), as well as summary statistics for drainage area, species richness, and biocriteria index scores. Among least impacted reference sites within the WAP ecoregion, the sites with the highest drainage areas are those on the Tuscarawas River. These sites have drainage areas that are no more than 53% of the area at Conesville Station. Also, these sites were recently upgraded to the higher use of Exceptional Warmwater Habitat, further illustrating that they are less comparable to conditions at Conesville Station.

The only other reference sites within the WAP ecoregion that have similar drainage areas to that at Conesville Station are those on the lower Scioto River. The Scioto River is a tributary to the Ohio River, about 185 river miles downstream of the Muskingum River/Ohio River confluence (Figure 2). Conesville Station is an additional 102 river miles upstream from this point. Thus, a considerable distance (approximately 300 river miles) separates Conesville Station from the nearest ecoregion reference site having a similar drainage area. The comparability between the two areas is even more speculative considering that the lower Muskingum River is impounded, being regulated by navigation dams. In contrast, all sites on the lower Scioto River are free-flowing. Thus, even though both sites have similar drainage areas, some differences in biotic communities would be expected due to zoogeographic and habitat factors. For example, faunal similarity between midwestern river systems is dependent on river mile distance, along with other factors (Robison 1986).

A site-specific modification of biological expectation, or performance, would be reasonable if biocriteria index scores on the mainstem Muskingum River just upstream of Conesville Station have a different distribution than those for ecoregion reference sites. Biosurveys were conducted near Conesville Station from 1988 to 1991 using methods that conformed to Ohio EPA protocols or with minor deviations. Ohio EPA conducted fish sampling along the entire Muskingum River mainstem in 1988; AEP conducted subsequent studies from 1989 to 1991. Biocriteria scores (IBI, MIwb) were compiled for four sites just upstream of Conesville Station, using both Ohio EPA and AEP data. Samples collected in 1990 were not used due to high flow conditions. A total of 27 individual samples were compiled for sites just upstream of Conesville Station; 51 samples were taken within the entire WAP ecoregion. Thus, the site-specific sample size is slightly more than 1/2 of the entire ecoregion sample size.

Table 1 indicates the statistical parameters of IBI and MIwb data for WAP ecoregion reference sites and reference sites just upstream of Conesville Station. Table 1 provides evidence that the population of pooled ecoregion reference sites has distinct statistical parameters compared to sites in the mainstem Muskingum River near Coshocton, Ohio. The median values of the IBI and MIwb are consistently lower at site-specific reference sites. The lowered 75th and 25th percentile values is notable because Ohio EPA regards the 25th percentile in a given ecoregion as the minimum biological expectation. The above data suggests that the WAP ecoregion criteria of 40 IBI units and 8.6 MIwb units represent unrealistic biological performance expectations when compared to sampled reference sites near Conesville Station.

Though both point source and nonpoint source impacts have been documented upstream of Conesville Station, these impacts are intermittent (i.e., flow dependent) and thus there is no reason to believe that these have caused a continual, systematic decline in index values. Nonpoint sources probably have some effect on biological performance at many ecoregion reference sites (Whittier et al. 1987). Actually, from a biological perspective, the differences in index score parameters are not unexpected, because: (1) no sites on the mainstem Muskingum River were used for ecoregion criteria, (2) the relatively large sample size upstream of Conesville Station is much higher than at any individual ecoregion site, and (3) the only ecoregion reference site with a comparable drainage area is separated by about 300 river miles. From a regulatory perspective, these data provide good technical justification for a modified biological expectation. In this case, a discharger has met the burden of proof by providing an adequate database to evaluate the reasonableness of an ecoregion-based expectation.

2.3.2 Biocriteria for Regulated vs. Unregulated Waterbodies

Ohio EPA's biocriteria for the WAP ecoregion are identical for both free-flowing and impounded sections of the Muskingum River (Figure 2). As previously indicated, the agency did not use sampling results for the mainstem Muskingum River in the derivation of WAP ecoregion biocriteria. Thus, sampling results for least impacted reference sites on impounded river segments were never used to derive minimum biocriteria. This lack of reference stream data on impounded river segments presents problems for dischargers located on the impounded Muskingum River. The assumption of Ohio EPA's WAP biocriteria is that biological performance should be similar between sites on free-flowing reaches and sites on impounded reaches, assuming the drainage areas are similar. This assumption is not realistic, as discussed below.

The most significant problem with Ohio EPA's assumption is that the agency has adopted identical biocriteria for free-flowing and impounded reaches without testing the hypothesis of no differences in

Table 1. **Statistical Parameters of Fish Collection Sites and Biocriteria Index Scores for Ohio EPA's Western Allegheny Plateau Reference Sites and Site-Specific Locations on the Mainstem Muskingum River Near Conesville Station.**

Parameter	WAP ecoregion reference sites	Muskingum River reference sites (RM 103.5 to RM 105.8)
No. of samples	51	27
Drainage area of sites		
Mean (mi^2)	1860	4875
Minimum (mi^2)	90	4870
Maximum (mi^2)	6471	4880
Index of Biotic Integrity (IBI)		
Mean	44	38
(SE)	0.9	1.3
Median	44	36
Range	28–54	24–52
Quartile		
Lower	40	34
Upper	50	42
Modified Index of Well-Being (MIwb)		
Mean	9.3	8.1
(SE)	0.1	0.2
Median	9.4	8.1
Range	7.5–10.7	6.7–9.7
Quartile		
Lower	8.6	7.6
Upper	10.0	8.7

Table 2. **Summary of Fish Community Differences Found at Sites on the Free-Flowing and Impounded Portions of the Muskingum River During Electrofishing and Seining Studies in 1989**

Parameter	Study Location: Conesville Sta.[1] (free-flowing)	Study Location: Muskingum River[2] (impounded)
Total no. species	54	41
Total fishes collected	18,572	6,152
Relative abundances (%); number collected in parentheses		
Sand shiner	15.8 (2,904)	0.0
Bluntnose minnow	11.3 (2,078)	0.18 (11)
Quillback	0.04 (8)	1.7 (104)
Highfin carpsucker	0.17 (32)	0.02 (1)
Smallmouth buffalo	0.0	1.7 (104)
Silver redhorse	1.3 (231)	1.0 (59)
Northern hog sucker	0.35 (65)	0.0
Flathead catfish	0.10 (19)	1.7 (104)
Rock bass	0.71 (131)	0.1 (7)
Orangespotted sunfish	0.0	0.7 (40)
Longear sunfish	0.0	0.8 (51)
Smallmouth bass	4.1 (743)	0.3 (19)
Spotted bass	0.0	11.0 (679)
Largemouth bass	0.23 (43)	0.09 (6)
Greenside darter	0.92 (168)	0.0
Banded darter	0.88 (161)	0.0
Freshwater drum	0.02 (4)	2.2 (138)

[1] Results based on 32 electrofishing samples and 16 seine (riffle) samples.
[2] Results based on 48 electrofishing samples.

biological performance between the differing habitats. Recent fishery surveys conducted at Conesville Station (located in the free-flowing upper Muskingum River) and Muskingum River Plant (located at River Mile 28.0 in the impounded lower river) demonstrate the differences in both species composition and relative abundance among the sites. Results of the 1989 fishery surveys are given in Table 2.

As would be expected, the fish community found at each site reflects the predominant habitat features. A riverine/riffle community is present near Conesville Station whereas a large river/lentic community is present near Muskingum River Plant. Major compositional differences were observed in most of the dominant families. At Conesville Station, spotfin shiner dominated the cyprinid catch with sand shiner and bluntnose minnow being fairly common. At Muskingum River Plant, the family Cyprinidae was dominated by emerald shiner, but sand shiner and bluntnose minnow were absent or uncommon. Within the sucker family, the subfamily Ictiobinae was represented by highfin carpsucker at Conesville Station whereas quillback and smallmouth buffalo were fairly common at Muskingum River Plant. All redhorse species (subfamily Catostominae) and the northern hog sucker were more common near Conesville Station. Flathead catfish were considerably more common at Muskingum River Plant. The composition of sunfish species was different at each plant site. Rock bass, smallmouth bass, and largemouth bass were more abundant at free-flowing sites, whereas orangespotted sunfish, longear sunfish, and spotted bass were collected exclusively at Muskingum River Plant. Not surprisingly, darter species were either absent at Muskingum River Plant or were more common near Conesville Station. Freshwater drum were considerably more abundant in the impounded reaches of the river. The higher species richness and total fish abundance is largely due to the presence of riffle habitats near Conesville Station, but none near Muskingum River Plant.

These data clearly demonstrate the presence of two different fish communities on the same waterbody, yet Ohio EPA's biological criteria do not reflect these differences because all ecoregion reference sites were located on free-flowing streams and rivers. The fish community near Muskingum River Plant is actually more similar to assemblages in the nearby Ohio River compared to assemblages in inland rivers. Because Ohio EPA is currently conducting studies on the Ohio River for the purpose of deriving biological criteria (see Section 2.3.3), there appears to be no sound biological reason why the lower Muskingum River, having unique habitat characteristics, should have identical biological criteria with free-flowing portions, or with other free-flowing rivers in the same ecoregion. Ohio EPA, as a minimum, should have tested the hypothesis of no differences in biological performance for reference sites on impounded and free-flowing sections.

2.3.3 Derivation of Large River Biocriteria (Emphasis on Ohio River)

Ohio EPA has initiated sampling at nearshore zones on the Ohio River for selection of reference sites and eventual derivation of biocriteria for upper and middle river sections that border the State of Ohio (Sanders 1991). In addition, the Ohio River Valley Sanitation Commission (ORSANCO) has initiated sampling of nearshore zones along the entire Ohio River using an electrofishing procedure that generally follows Ohio EPA's protocols. The results of these surveys will apparently be pooled to derive proposed biocriteria for the mainstem Ohio River.

The derivation of numeric biological criteria for large rivers presents technical challenges for regulatory agencies. First, the definition and delineation of reference sites is problematic. For the Ohio River, the identification of "least impacted" reference sites (i.e., similar to those selected for ecoregion-specific inland watersheds) is probably not possible. Because of the enlarged physical dimensions of larger rivers and ease of faunal transfer within navigation pools, nonimpacted reference sites (which represent a target biological expectation) likely do not exist on large rivers. This factor changes the benchmark, or biological performance gauge, that will be used to assess attainment or nonattainment of the aquatic life use in the Ohio River. As a surrogate, a near-field reference site is probably the most valid target for regulatory assessments on large rivers.

A second problem is habitat comparability. Large rivers, such as the Ohio River, have extensive reaches of relatively unproductive habitat (e.g., shallow sloping banks with sand/muck substrate) with little or no attractive habitat feature (e.g., no log piles, overhanging vegetation, or gravel/cobble substrate) within these reaches. Productive and heterogeneous habitats are highly patchy and obviously these constraints will directly influence the diversity and abundance of fishes.

Another problem with large river biocriteria development is data variability. Temporal variations in fish community parameters must be expected in large rivers. Where significant seasonal variation is documented, this variability must be accounted for in the derivation of numeric biocriteria if the biocriteria are, in fact, based on multiseason sampling. This variability was demonstrated during a 1991 fisheries study at six power plant sites along the Ohio River (EA 1993). The plant sites were located on the upper, middle, and lower Ohio River between River Miles 54 and 946. Fishes were sampled in June or July, August, and September or October. For electrofishing and gill net samples there were 12 of 24 (50%) cases where a significant ($P < 0.05$) temporal variation was observed for either catch per unit effort or total biomass, at a particular plant site. The effects of temporal variation (both seasonal and year-to-year, if applicable) obviously must be accounted for when deriving numeric biocriteria for large rivers.

In addition to temporal effects on fishery parameters in the 1991 study cited above, statistical associations of calculated IBI and MIwb values with independent variables indicated that river flow, water temperature, and forage fish abundance significantly affected score values at combined locations (Reash, in press). For combined samples at six power plant sites ($N = 108$), statistically significant inverse correlations with calculated IBI values (using Ohio EPA inland methodology) were as follows: percent gizzard shad ($r = -0.43$; $P<0.001$) and river flow at time of sampling ($r = -0.21$; $P<0.03$). Significant inverse correlations with the Modified Index of Well-Being were: percent gizzard shad ($r = -0.31$; $P<0.002$) and water temperature ($r = -0.21$; $P<0.04$). These results indicate that (1) data interpretation of Ohio River biosurveys may be problematic due to stochastic factors, and (2) highly standardized methodologies will be required to minimize the influence of confounding factors regarding a site-specific assessment.

Ohio EPA has analyzed electrofishing data for Ohio River nearshore zones using the Modified Index of Well-Being (MIwb) and Index of Biotic Integrity (IBI). The MIwb appears to be suitable for the Ohio River as this index has been used in a wide range of waterbodies to assess changes in structural attributes of fish communities (Fausch et al. 1990). The application of Ohio EPA's inland version of the IBI to Ohio River mainstem sites is questionable, however. The most obvious reason that the inland metrics should not be applied to the Ohio River is that Ohio EPA's published reference sites do not encompass large, impounded rivers. Although the agency is currently in the process of recalibrating the IBI metrics for Ohio River sites, sampling data should not be interpreted using the inland IBI until this recalibration and validation is completed.

AEP, along with other electric utility companies, has sponsored annual ecological studies near coal-fired power plants since the early 1970s. Based on results of this long-term study, recommended modifications to Ohio EPA's inland IBI metrics (for application to the Ohio River) are listed in Table 3. New Ohio River-specific metrics may need to be developed if modifications to the inland IBI metric are not sufficiently sensitive to detect significant shifts in structural and functional community parameters. These new metrics may require a more generic approach. For example, a metric such as total number of trophic guilds (relative to a nearby reference site) may be sensitive enough to detect significant water quality degradation, but generic enough to prevent a "false positive" finding of use impairment that is actually caused by temporal effects or habitat effects. Whether established or new metrics are proposed for large river biocriteria, these metrics must be validated regarding their empirical foundation and ability to detect impairment beyond a reference condition.

Because the Ohio River is a large biological system with species composition constantly changing, interpretation of sampling data will require considerations of zoogeography, historical abundance and distribution, and historical ranges of variability (both population-specific and community-based parameters). Pearson and Pearson (1989) discussed historical trends of faunal composition and abundance. Van Hassel et al. (1988) reported significant distributional and temporal trends of Ohio River fishes in the upper and middle river, along with a segregation of species based on reproductive guilds, habitat preferences, and feeding habits. There appears to be sufficient evidence that IBI metrics will differ for the upper and middle Ohio River due to zoogeographic factors. The influence of zoogeographic and physicochemical factors on fish distribution in the upper and middle river was discussed by Reash and Van Hassel (1988). Navigation locks and dams restrict faunal transfer between navigation pools, thus creating a continuum of community similarity that is somewhat predictable. Reash (1992) developed a regression equation that predicts the similarity of Ohio River fish assemblages along the Ohio River. Using a two-variable model of river distance and drainage area difference, community similarity could

Table 3. Suggested Modifications to Ohio EPA's Inland Index of Biotic Integrity (Boat Method), for Potential Application to the Ohio River, an Impounded Large River

IBI metric	Current inland[a] cut-offs	Suggested modification
No. of species	>20, 10–20, <10	Cut-offs may need revision; sites with sparse habitat often yield <20 species
Percent round-bodied suckers	>38, 19–38, <19	Acceptable for upper river, possibly acceptable for middle river. Cutoffs will need modification
Sunfish species	>3, 2–3, <2	May require a lower expectation in middle river, where rock bass and pumpkinseed are rare
Sucker species	>5, 3–5, <3	Scoring cutoffs will require modification; hog sucker, white sucker, and black redhorse are rare in middle river
Intolerant species	>3, 2–3, <2	A new list of "Ohio River intolerant species" will be needed. Many intolerant species for inland metric are small stream forms
Percent tolerant	<15, 15–27, >27	Same comment as above. Scoring cutoffs will need modification
Percent omnivores	<16, 16–28, >28	Scoring cutoffs will need modification. A greater number of omnivore species would be expected in large, impounded rivers
Percent insectivores	>54, 27–54, <27	This metric is questionable for the Ohio River. Fewer insectivorous species present due to lentic-like hydrology. Benthic production of food organisms much less than in free-flowing systems
Percent top carnivores	>10, 5–10, <5	May require modification for upper and middle reaches
Percent simple lithophils	Varies w/drainage area	Questionable for use in Ohio River due to limited area with hard substrate
Percent DELT anomalies	<0.5, 0.5–3.0, >3.0	Scoring cutoffs will need modification based on Ohio River samples. A greater abundance of carp, catfish, and Ictiobinae suckers in impounded rivers may inflate the prevalence of DELT anomalies
Fish numbers	<200, 200–450, >450	Scoring cutoffs will need modification. Dense clusters of forage species will cause wide variation in total numbers

[a] Indicated cutoffs correspond to metric scores of 5, 3, and 1, respectively.

be predicted with a high degree of statistical confidence ($r^2 = 0.88$). Such factors should be considered when regulatory agencies assess the aquatic life use in the Ohio River using numeric biocriteria.

3.0 SUMMARY

Biological criteria will be developed and adopted in state water quality standards as generic narrative criteria or numeric criteria. According to USEPA, biocriteria will have no less legal weight as chemical-specific and whole effluent toxicity criteria, and thus regulated industry must understand the conceptual foundations of biocriteria and the technical aspects of data collection, data interpretation, and biocriteria derivation. Regulated industry should work with water resource agencies to ensure that standardized methodologies are used and especially that numeric biocriteria are valid and have a sound empirical foundation. Regulatory agencies must avoid the temptation of collecting a sparse amount of data to derive biocriteria indices.

The State of Ohio has utilized an extensive database of statewide biological surveys to derive numeric biocriteria based on the ecoregion approach. Ohio EPA's three biocriteria indices have a sound empirical foundation regarding incorporation of the broad historical database. The establishment of minimum biological performance in Ohio's five ecoregions allows a concise compliance target that industry can readily assess using Ohio EPA methodologies. Industry should take the initiative to confirm or challenge a regulatory agency's assessment of aquatic life use nonattainment, due to the fact that causes of nonattainment may be independent of instream pollutant levels. Technical flaws to Ohio EPA's biocriteria are discussed. Ohio EPA, and all other agencies, will need a concise mechanism for site-specific biocriteria modification. Large impounded rivers present technical challenges to biocriteria derivation. The assumptions (hypotheses) of biological expectation in large rivers must be tested before data are analyzed using methodologies applied to small and medium-sized rivers.

ACKNOWLEDGMENTS

I thank Wanda Vestermark and Toni Nijssen, who dutifully typed the manuscript. Greg Seegert provided insightful comments on Ohio River biocriteria. I thank the following individuals for contributing useful comments on a draft version: William Patberg, Alan Gaulke, Scott Matchett, Timothy O'Shea, Timothy Lohner, and John Van Hassel. Two anonymous reviewers provided improvements on a draft version. The cited Ohio River fishery data are from reports of the Ohio River Ecological Research Program, an interutility study sponsored by American Electric Power Company, Ohio Edison Company, Ohio Valley Electric Corporation, Cincinnati Gas and Electric Company, and Tennessee Valley Authority.