Introduction

The Pennsylvania Department of Environmental Protection (DEP) developed an index of biotic integrity (IBI) for benthic macroinvertebrate communities in Pennsylvania's wadeable, freestone, riffle-run type streams as a scientifically credible biological assessment tool. This indicator assists in guiding and evaluating legislation, policy and management strategies as well as setting goals for aquatic resources by enabling direct quantification of important ecological attributes along a gradient of biological conditions and ecosystem stressors (Davis and Simon 1995; Davies and Jackson 2006; Hawkins 2006). This indicator serves as a measure of the extent to which anthropogenic stressors impair the capability of a stream to support a healthy aquatic community (Davis and Simon 1995).

Biological Sampling Methods

This IBI applies to benthic macroinvertebrate samples collected any time of the year from wadeable, freestone, riffle-run streams in Pennsylvania using a D-frame net with 500-micron mesh. Field sampling and laboratory methods are more fully described in DEP's Standardized Biological Field Collection and Laboratory Methods, Section V (Pennsylvania Department of Environmental Protection 2003). Sampling biologists composite six kicks from riffle areas distributed throughout a 100-meter stream reach, working progressively upstream, with each kick disturbing approximately one square meter immediately upstream of the net for approximately one minute to an approximate depth of 10 cm, as substrate allows. Composited samples are preserved with 95% ethanol in the field and transported back to the laboratory for processing. In the lab, each composited sample is placed into a 3.5" deep rectangular pan (measuring 14" long x 8" wide on the bottom of the pan) marked off into 28 four-square inch (2" x 2") grids. Four of the grids are randomly selected, their contents are extracted using a four-square inch circular "cookie cutter," and placed into another identical empty pan. All the organisms are picked from this second pan. If less than 160 identifiable organisms are picked from the second pan, additional grids are randomly selected and extracted from the first pan, transferred to the second pan and picked until the target number of organisms (200 ± 40 organisms) is obtained. If more than 240 identifiable organisms are picked from the original four grids then the second pan is cleared of debris, the picked organisms are floated in the cleared pan and randomly-selected grids are picked until the target number of organisms is obtained. Any grids selected during this entire process are picked in their entirety and the total numbers of grids selected for each part of the sub-sampling process are recorded.

Organisms in the sub-sample are identified and counted. Midges are identified to the family level of Chironomidae. Snails, clams and mussels are all also identified to family levels. Roundworms and proboscis worms are identified to the phylum levels of Nematoda and Nemertea, respectively. Moss animacules are identified to the phylum level of Bryozoa. Flatworms and leeches are identified to the class levels of Turbellaria and Hirudenia, respectively. Segmented worms, aquatic earthworms, and tubificids are identified to the class level of Oligochaeta. All water mites are identified as

Hydracarina, an artificial taxonomic grouping of several mite superfamilies. All other macroinvertebrates are identified to genus level.

Most of the samples used to develop the IBI were taken from relatively small, mostly first through third order riffle-run type streams draining less than 25 square miles, so this IBI should be applied with discretion to other stream types (e.g., limestone type streams) and larger stream/river systems. Currently, DEP does not apply any regionally-based classification to wadeable, freestone, riffle-run streams in the Commonwealth for purposes of applying this IBI.

The Metrics

A number of different metric combinations were evaluated during index development and the following six metrics were selected for inclusion as core metrics in the IBI based on various performance characteristics. These six metrics all exhibited a strong ability to distinguish between reference and stressed conditions. In addition, these six metrics measure different aspects of the biological communities represented by the sub-samples, and when used together in a multimetric index, they provide a solid foundation for assessing the biological condition of benthic macroinvertebrate assemblages in Pennsylvania's wadeable freestone riffle-run stream ecosystems.

Total Taxa Richness

This taxonomic richness metric is a count of the total number of taxa in a sub-sample. Generally, this metric is expected to decrease with increasing anthropogenic stress to a stream ecosystem, reflecting loss of taxa and increasing dominance of a few pollution-tolerant taxa. Other benefits of including this metric include its common use in many biological monitoring and assessment programs in other parts of the world as well as its ease of explanation and calculation.

Ephemeroptera + Plecoptera + Trichoptera Taxa Richness
(Pollution Tolerance Values 0 – 4 only)

This taxonomic richness metric is a count of the number of taxa belonging to the orders Ephemeroptera, Plecoptera, and Trichoptera (EPT) in a sub-sample – common names for these orders are mayflies, stoneflies, and caddisflies, respectively. The aquatic life stages of these three insect orders are generally considered sensitive to, or intolerant of, pollution (Lenat and Penrose 1996); in fact, this metric only counts EPT taxa with pollution tolerance values (PTVs) of 0 to 4, excluding a few of the most tolerant mayfly and caddisfly taxa. This metric is expected to decrease in value with increasing anthropogenic stress to a stream ecosystem, reflecting the loss of taxa from these largely pollution-sensitive orders. This metric has a history of use across the world and is relatively easy to use, explain and calculate (Lenat and Penrose 1996).

Beck's Index, version 3

This taxonomic richness and tolerance metric is a weighted count of taxa with PTVs of 0, 1, or 2. The name and conceptual basis of this metric are derived from the water quality work of William H. Beck in Florida (Beck 1955). This metric is expected to decrease in value with increasing anthropogenic stress to a stream ecosystem, reflecting the loss of pollution-sensitive taxa. It should be noted that the version of the Beck's Index metric used for this project, although similar in name and concept, differs slightly in its calculation from the Beck's Index used in DEP's multihabitat protocol for assessing biological condition of low gradient pool-glide type streams.

Shannon Diversity

This community composition metric measures taxonomic richness and evenness of individuals across taxa of a sub-sample. This metric is expected to decrease in values with increasing anthropogenic stress to a stream ecosystem, reflecting loss of pollution-sensitive taxa and increasing dominance of a few pollution-tolerant taxa. The name and conceptual basis for this metric are derived from the information theory work of Claude Elwood Shannon (Shannon 1968).

Hilsenhoff Biotic Index

This community composition and tolerance metric is calculated as an average of the number of individuals in a sub-sample, weighted by PTVs. Developed by William Hilsenhoff, the Hilsenhoff Biotic Index (Hilsenhoff 1977, 1987, 1988; Klemm et al. 1990) generally increases with increasing ecosystem stress, reflecting increasing dominance of pollution-tolerant organisms.

Percent Sensitive Individuals (PTV 0 – 3)

This community composition and tolerance metric is the percentage of individuals with PTVs of 0 to 3 in a sub-sample and is expected to decrease in value with increasing anthropogenic stress to a stream ecosystem, reflecting loss of pollution-sensitive organisms.

Example calculations for each metric are provided below for a sample from Lycoming Creek.

Benthic macroinvertebrate sample from Lycoming Creek in Lycoming County taken on November 19, 2001		
Taxa Name	**Number of Individuals**	**Pollution Tolerance Value**
Acentrella	1	4
Isonychia	4	3
Epeorus	6	0
Leucrocuta	1	1
Rhithrogena	9	0
Stenonema	8	3
Ephemerella	32	1
Serratella	1	2
Paraleptophlebia	4	1
Pteronarcys	1	0
Taeniopteryx	1	2
Leuctra	2	0
Agnetina	1	2
Paragnetina	1	1
Chimarra	1	4
Dolophilodes	1	0
Cheumatopsyche	25	6
Hydropsyche	22	5
Rhyacophila	16	1
Glossosoma	2	0
Brachycentrus	3	1
Micrasema	1	2
Apatania	2	3
Psilotreta	1	0
Psephenus	3	4
Optioservus	7	4
Atherix	1	2
Antocha	2	3
Hexatoma	5	2
Prosimulium	1	2
Chironomidae	49	6
Ancylidae	2	7
Oligochaeta	1	10

Total Taxa Richness

There are ***33 taxa*** in this sub-sample, so

Total Taxa Richness = 33

EPT Taxa Richness (PTV 0 – 4 only)

There are ***9 Ephemeroptera taxa*** (Acentrella, Isonychia, Epeorus, Leucrocuta, Rhithrogena, Stenonema, Ephemerella, Serratella, Paraleptophlebia), ***5 Plecoptera taxa*** (Pteronarcys, Taeniopteryx, Leuctra, Agnetina, Paragnetina) and ***8 Trichoptera taxa*** (Chimarra, Dolophilodes, Rhyacophila, Glossosoma, Brachycentrus, Micrasema, Apatania, Psilotreta) in this sub-sample ***with PTVs ≤ 4***, so

EPT Taxa Richness (PTV 0 – 4) = 9 + 5 + 8

EPT Taxa Richness (PTV 0 – 4) = 22

Beck's Index, version 3

Beck's Index, version 3 =
(3 x (number of taxa with PTV = 0)) +
(2 x (number of taxa with PTV = 1)) +
(1 x (number of taxa with PTV = 2))

There are ***7 taxa in this sub-sample with PTV = 0***. There are ***6 taxa in this sub-sample with PTV = 1***. There are ***7 taxa in this sub-sample with PTV = 2***, so

Beck's Index, version 3 = 3(7) + 2(6) + 1(7)

Beck's Index, version 3 = 21 + 12 + 7

Beck's Index, version 3 = 40

Hilsenhoff Biotic Index

$$\text{Hilsenhoff Biotic Index} = \sum_{i=0}^{10} [(i * n_{indvPTVi})] / N$$

where $n_{indvPTVi}$ = the number of individuals in a sub-sample with PTV of i and N = the total number of individuals in a sub-sample

There are 22 individuals with PTV = 0
There are 57 individuals with PTV = 1
There are 11 individuals with PTV = 2
There are 16 individuals with PTV = 3
There are 12 individuals with PTV = 4
There are 22 individuals with PTV = 5
There are 74 individuals with PTV = 6
There are 2 individuals with PTV = 7
There are 0 individuals with PTV = 8 or 9
There is 1 individual with PTV = 10.

There are a total of 217 individuals in the sub-sample, so

$$\text{Hilsenhoff Biotic Index} = [(0 * 22) + (1 * 57) + (2 * 11) + (3 * 16) + (4 * 12) + (5 * 22) + (6 * 74) + (7 * 2) + (8 * 0) + (9 * 0) + (10 * 1)] / 217$$

Hilsenhoff Biotic Index = 3.47

Shannon Diversity Index

$$\text{Shannon Diversity Index} = [-\sum_{i=1}^{Rich} (n_i / N) \ln (n_i / N)]$$

where n_i = the number of individuals in each taxa (relative abundance); N = the total number of individuals in a sub-sample; and Rich = the total number of taxa in a sub-sample (total taxa richness)

There are 33 taxa in this sub-sample. The numbers of individuals in each taxa are shown in the table above. There are a total of 217 individuals in the sub-sample, so

$$\text{Shannon Diversity Index} = -(1 / 217) \ln (1 / 217) + (4 / 217) \ln (4 / 217) + (6 / 217) \ln (6 / 217) + (1 / 217) \ln (1 / 217) + (9 / 217) \ln (9 / 217) + (8 / 217) \ln (8 / 217) + (32 / 217) \ln (32 / 217) + (1 / 217) \ln (1 / 217) + \ldots$$

… *(do this for all 33 taxa)*

$$\ldots (1 / 217) \ln (1 / 217)$$

Shannon Diversity Index = 2.67

Percent Sensitive (PTV 0 – 3) Individuals

$$\text{Percent Sensitive (PTV 0 – 3) Individuals} = \left(\sum_{i=0}^{3} n_{indvPTVi}\right) / N * 100$$

where $n_{indvPTVi}$ = the number of individuals in a sub-sample with PTV of i and N = the total number of individuals in a sub-sample

There are 22 individuals with PTV = 0
There are 57 individuals with PTV = 1
There are 11 individuals with PTV = 2
There are 16 individuals with PTV = 3

There are a total of 217 individuals in the sub-sample, so

Percent Sensitive (PTV 0 – 3) Individuals = (22 + 57 + 11 + 16) / 217 *100

Percent Sensitive (PTV 0 – 3) Individuals = 106 / 217 * 100

Percent Sensitive (PTV 0 – 3) Individuals = 48.8%

The Index

An index is simply a means to integrate information from various measures of biological integrity, or various metrics (Barbour et al. 1999). In order to compare and combine sundry measures (e.g., percentage of individuals, counts of taxa, unitless numbers) of biological condition in a meaningful manner, it is necessary to standardize metrics with some mathematical transformation that results in a logical progression of values (Barbour et al. 1995).

The one selected core metric that increases in value with increasing anthropogenic stress (i.e., the Hilsenhoff Biotic Index) was standardized to the 5th percentile of metric scores for all samples in the IBI development dataset. Core metrics that decrease in value with increasing stress (i.e., total taxa richness, EPT taxa richness, % sensitive individuals, Shannon diversity, Beck's Index) were standardized to the 95th percentile of metrics scores for all samples in the IBI development dataset. The following table presents the standardization values used for each core metric.

Metric	Standardization value
Total Taxa Richness	33
EPT Taxa Richness (PTV 0 – 4)	19
Beck's Index, version 3	38
Hilsenhoff Biotic Index	1.89
Shannon Diversity	2.86
Percent Sensitive Individuals (PTV 0 – 3)	84.5

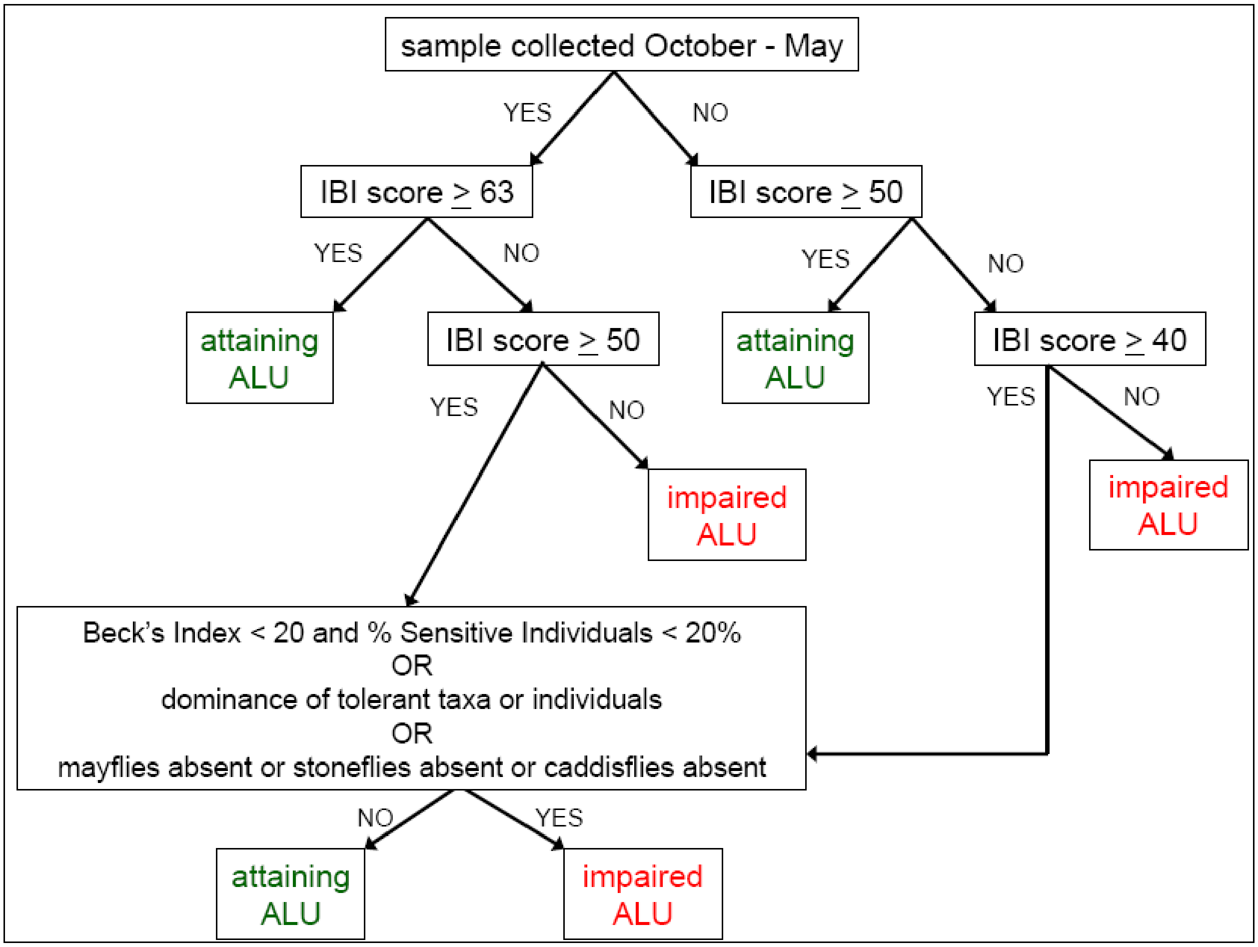
The values for standardized core metric values were set to a maximum value of 1.00, with values closer to zero corresponding to increasing deviation from the expected reference condition and progressively higher values corresponding more closely to the biological reference condition (Barbour et al. 1995). The adjusted standardized metric values for the six core metrics were averaged and multiplied by 100 to produce an index score ranging from 0 to 100. This number represents the multimetric index of biological integrity (IBI) score for a sample. The following table shows the standardized metric and index scoring calculations for the Lycoming Creek sample discussed above.

Metric	**Standardization Equation**	**Observed Metric Value**	**Standardized Metric Score**	**Adjusted Standardized Metric Score** Maximum = 1.000
Total Taxa Richness	observed value / 33	33	1.000	1.000
EPT Taxa Richness	observed value / 19	22	1.158	1.000
Modified Beck's Index	observed value / 38	40	1.053	1.000
Hilsenhoff Biotic Index	(10 – observed value) / (10 – 1.89)	3.47	0.805	0.805
Shannon Diversity	observed value / 2.86	2.67	0.934	0.934
Percent Sensitive Individuals	observed value / 84.5	48.8	0.578	0.578
Average of adjusted standardized core metric scores * 100 = **IBI Score =**				**88.6**

Aquatic Life Use Attainment Benchmarks

Based on the results of classification analyses (details available upon request), DEP decided not to establish separate reference conditions and thresholds for wadeable freestone, riffle-run type streams in separate regions of the Commonwealth. However, due to the influences of annual seasons and drainage area seen in the IBI development dataset, DEP recognizes different use attainment thresholds are appropriate for samples collected during different times of the year and from different size stream systems.

Based on the results of the analyses presented above, the results of workshops and feedback from DEP biologists and policy considerations, DEP implements a multi-tiered benchmark decision process for smaller wadeable freestone riffle-run streams in Pennsylvania that incorporates sampling season as a factor for determining aquatic life use (ALU) attainment and impairment for the cold water fishes (CWF), warm water fishes (WWF) and trout stocking (TSF) protected uses; this process is outlined in the diagram below.



The first step in the ALU assessment process for smaller wadeable freestone riffle-run streams in Pennsylvania considers sampling season (i.e. June through September versus October through May). These seasonal index periods are intended as general guidelines and may vary slightly year-to-year depending on climatological conditions; for example, a sample collected during the last week of May in a particularly hot, dry year may be more properly evaluated using procedures set forth for the summer months.

For samples collected from smaller streams between October and May, an IBI score $\geq$ 63 results in ALU attainment and an IBI score < 50 results in ALU impairment; an IBI score between 50 and 63 requires further evaluation to determine ALU impairment – three guidelines may be used: (1) if the Beck's Index score is < 20 and the % Sensitive Individuals in the sub-sample is < 20%, the ALU should be impaired without compelling reason otherwise; (2) if the sample is dominated by tolerant taxa or individuals, the ALU should be impaired without compelling reason otherwise; or (3) if mayflies, stoneflies or caddisflies are absent from the sub-sample the ALU should be impaired without compelling reason otherwise.

For samples collected between June and September from smaller streams, an IBI score $\geq$ 50 results in ALU attainment and an IBI score < 40 results in ALU impairment; an IBI score between 40 and 50 requires further evaluation to determine ALU impairment, guided by the same three guidelines outlined above for October to May samples scoring between 50 and 63 (although the absence of mayflies in samples collected immediately after spring hatches may be relaxed in some cases).

For larger wadeable freestone riffle-run type streams, DEP believes more samples are necessary to accurately establish ALU attainment and impairment benchmarks. Given the nature of flowing water bodies as gradually changing continuums, it is difficult to define a specific numeric cutoff to separate larger streams from smaller streams. However, the present dataset suggest that scores for some index metrics begin to decline for reference-quality streams drainage areas reach the 25 to 50 square mile range. Workshops conducted by DEP confirm that biological expectations or potential for most of the relatively pristine larger freestone streams in Pennsylvania are less than the biological expectations or potential for the relatively pristine smaller freestone streams.

The use assessment decision process and accompanying attainment/impairment benchmarks set forth above are intended as general guidelines, not as hard-and-fast rules. While the above guidelines will provide an accurate assessment of benthic macroinvertebrate community condition for the vast majority of samples collected from wadeable, freestone, riffle-run streams in Pennsylvania, there will be instances where a biologist's local knowledge of conditions may warrant a decision not arrived at using these guidelines. For instance, if a sample is heavily dominated by Simuliidae or Chironomidae larvae, often times this will make the metric and IBI scores difficult to interpret and the investigating biologist must rely on a more qualitative analysis of the metric scores and sample composition to arrive at an assessment decision. Similarly, samples from streams in areas receiving a substantial amount of flow from groundwater attributable to limestone geology are naturally expected to have less diversity than "true freestone" streams, so use attainment benchmarks may be justifiably relaxed for samples from these types of streams.

References

Barbour, M.T., J.B. Stribling, and J.R. Karr. 1995. Multimetric approach for establishing biocriteria and measuring biological condition. Chapter 6 in *Biological assessment and criteria: tools for water resource planning and decision making*, W.S. Davis and T.P. Simon, eds. (pp. 63 – 77). CRC Press, Boca Raton.

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling. 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates and fish, second edition. EPA 841-B-99-002. United States Environmental Protection Agency; Office of Water. Washington, D.C.

Beck, W.H., Jr. 1955. Suggested method for reporting biotic data. Sewage and Industrial Waste 27(10): 1193-1197.

Davis, W.S. and T.P. Simon. 1995. Introduction to *Biological assessment and criteria: tools for water resource planning and decision making*, W.S. Davis and T.P. Simon, eds. (pp. 3 – 6). CRC Press, Boca Raton.

Davies, S.P. and S.K. Jackson. 2006. The biological condition gradient: a descriptive model for interpreting change in aquatic ecosystems. Ecological Applications 16(4):1251-1266.

Hawkins, C.P. 2006. Quantifying biological integrity by taxonomic completeness: its utility in regional and global assessments. Ecological Applications 16(4): 1277-1294.

Hilsenhoff, W.L. 1977. Use of arthropods to evaluate water quality of streams. Technical Bulletin Number 100. Wisconsin Department of Natural Resources. 15 pp. Madison, Wisconsin.

Hilsenhoff, W.L. 1987. An improved biotic index of organic stream pollution. The Great Lakes Entomologist 20(1): 31-39.

Hilsenhoff, W.L. 1988. Rapid field assessment of organic pollution with a family-level biotic index. Journal of the North American Benthological Society 7(1): 65-68.

Hughes, R.M. 1995. Defining acceptable biological status by comparing with reference conditions. Chapter 4 in *Biological assessment and criteria: tools for water resource planning and decision making*, W.S. Davis and T.P. Simon, eds. (pp. 31 – 47). CRC Press, Boca Raton.

Klemm, D.J, P.A. Lewis, F. Fulk, and J.M. Lazorchak. 1990. *Macroinvertebrate field and laboratory methods for evaluating the biological integrity of surface waters*. Environmental Monitoring Systems Laboratory, United States Environmental Protection Agency. Cincinnati, Ohio. EPA-600-4-90-030.

Lenat, D.R. and D.L. Penrose. 1996. History of the EPT taxa richness metric. Bulletin of the North American Benthological Society 13(2).

Pennsylvania Department of Environmental Protection. 2003. Standardized Biological Field Collection and Laboratory Methods. Harrisburg, PA.

Shannon, C.E. 1948. A mathematical theory of communication. Bell System Technical Journal 27: 379-423 and 623-656.