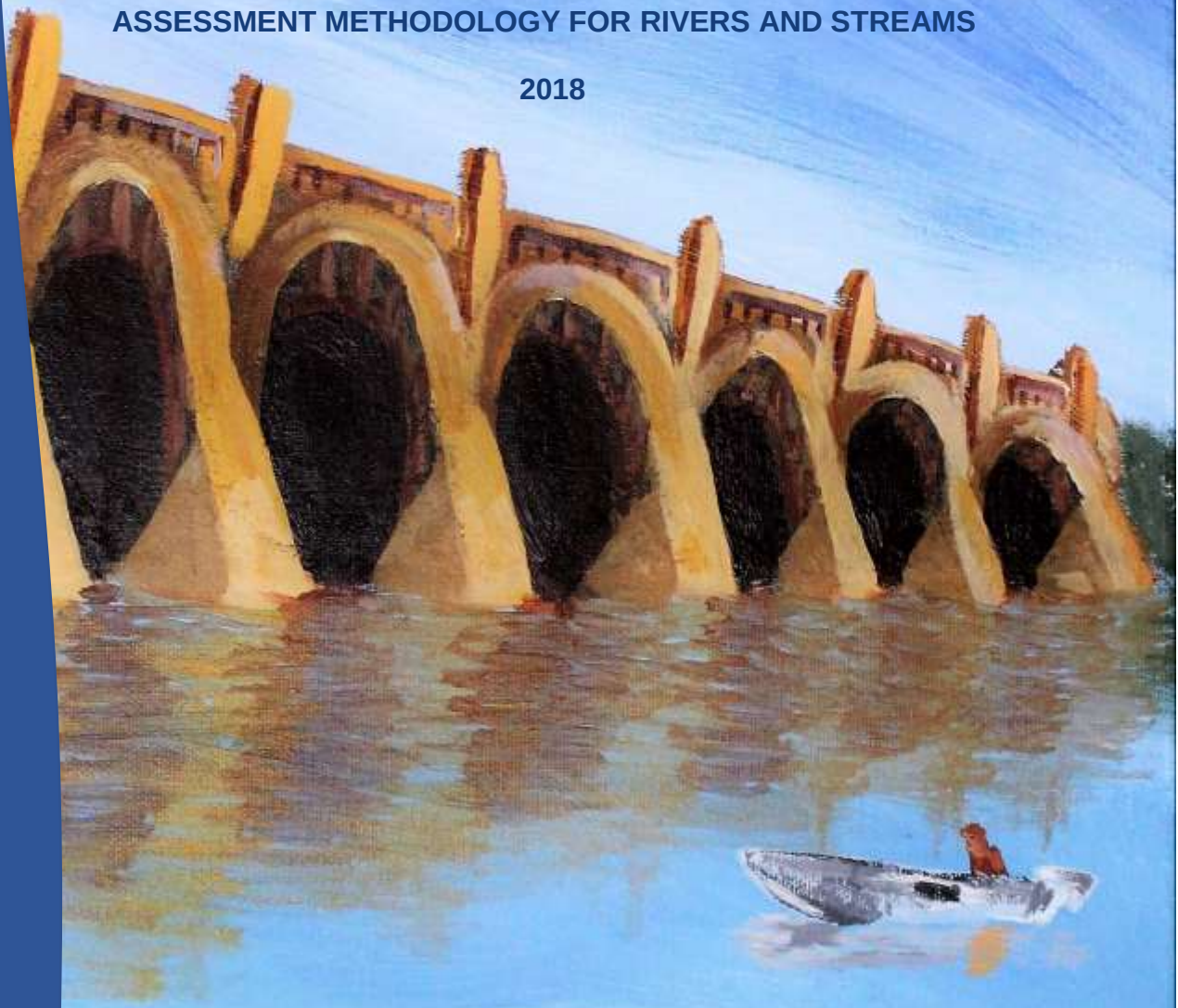
pennsylvania
DEPARTMENT OF ENVIRONMENTAL PROTECTION

OFFICE OF WATER PROGRAMS
BUREAU OF CLEAN WATER

ASSESSMENT METHODOLOGY FOR RIVERS AND STREAMS

2018



THIS PAGE INTENTIONALLY LEFT BLANK

OFFICE OF WATER PROGRAMS

BUREAU OF CLEAN WATER

ASSESSMENT METHODOLOGY FOR RIVERS AND STREAMS

2018

Edited by:

Dustin Shull and Molly Pulket
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

Cover image by Amy Williams

2018

TABLE OF CONTENTS

ABBREVIATIONS

ALU	Aquatic Life Use
AWS	Wildlife Water Supply
B	Boating
BOL	Bureau of Laboratories
CIM	Continuous Instream Monitoring
CWA	Clean Water Act
CWF	Cold Water Fishes
DEP	Pennsylvania Department of Environmental Protection
DRBC	Delaware River Basin Commission
DOH	Pennsylvania Department of Health
E	Esthetics
ECD	Eutrophication Cause Determination
EPT	Ephemeroptera, Plecoptera, and Tricoptera
EV	Exceptional Value
F	Fishing
FCATW	Fish Consumption Advisory Technical Workgroup
GIS	Geographic Information System
HQ	High Quality
HUC	Hydrologic Unit Code
IBI	Index of Biotic Integrity
IRS	Irrigation
IWS	Industrial Water Supply
LWS	Livestock Water Supply
MF	Migratory Fishes
MMI	Multimetric Index
MOU	Memorandum of Understanding
NELAP	National Environmental Laboratory Accreditation Program
NHD	National Hydrologic Dataset
NRC	National Research Council
ORSANCO	Ohio River Valley Water Sanitation Commission
PCB	Polychlorinated Biphenyls
PDA	Pennsylvania Department of Agriculture
PFBC	Pennsylvania Fish and Boat Commission
PWS	Potable Water Supply
RfD	Reference Dose
USEPA	United States Environmental Protection Agency
USFDA	United States Food and Drug Administration
USGS	United States Geological Survey
SAC	Standard Analysis Code
SIS	Sample Information System
SSWAP	Statewide Surface Water Assessment Program
SWMMI	Semi-wadeable Multimetric Index
TMDL	Total Maximum Daily Load
TSF	Trout Stocking

WC	Water Contact
WQN	Water Quality Network
WQS	Water Quality Standards
WWF	Warm Water Fishes

CHAPTER 1 INTRODUCTION

Prepared by:

Gary Walters and Dustin Shull
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2017

PURPOSE

Conceptual Framework

This document details the suite of methods currently used by the Pennsylvania Department of Environmental Protection (DEP) to assess surface waters for their attainment of protected designated and existing uses (further described below). This book is part of a larger conceptual framework (Figure 1) that DEP uses to collect quality data and make scientifically defensible decisions on various surface water matters across Pennsylvania. The data collections protocols (Monitoring Book, Shull and Lookenbill 2018) and evaluation methods that stem from them are based on peer reviewed technical reports and published literature. This foundation ensures that not only are DEP data collection protocols and evaluation methods founded on defensible scientific information, but that they also remain dynamic as new technology and innovation emerges. At this time, lake assessment methods are not included in this document.

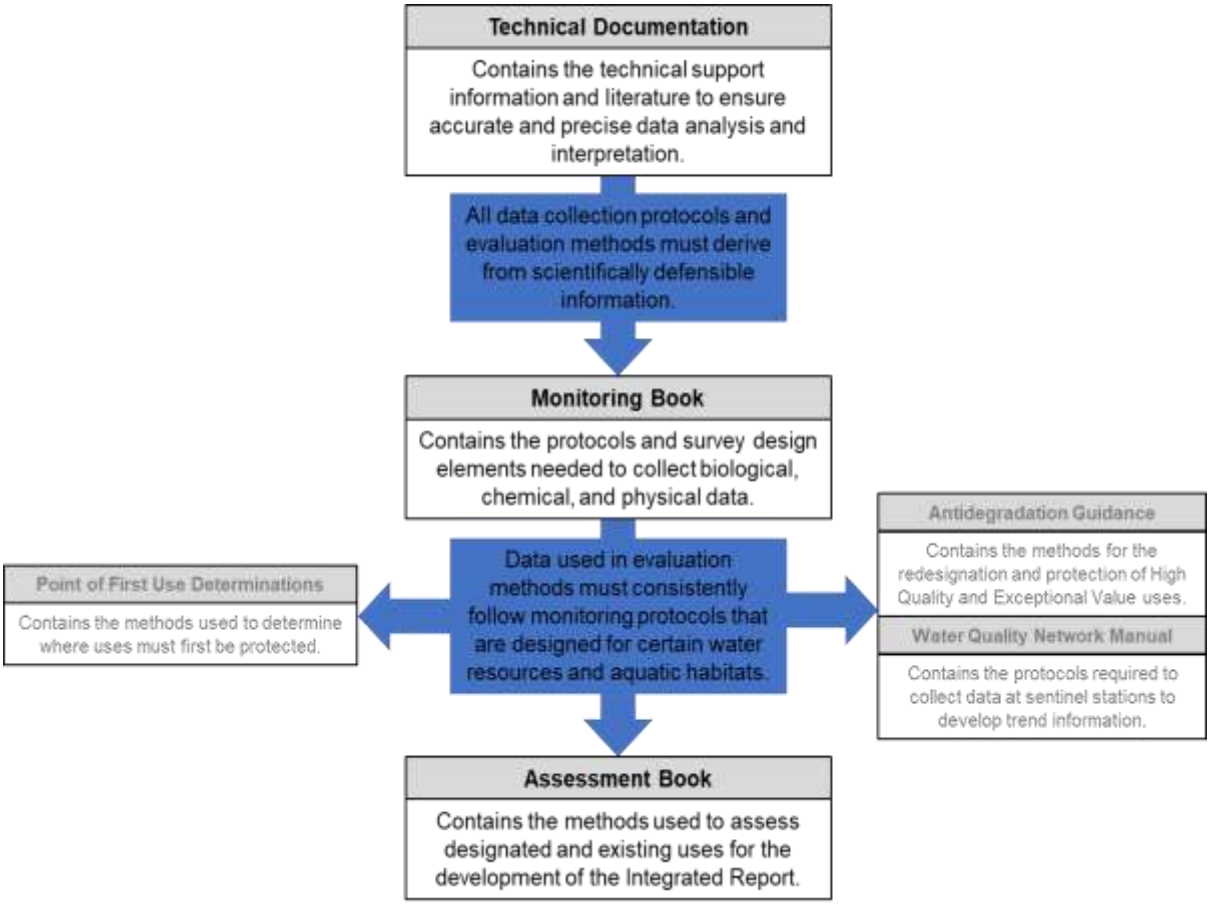


Figure 1. Conceptual framework that DEP uses to make scientifically defensible data collections and decisions. The conceptual progression for making use assessment determinations is highlighted for this book.

Federal Requirements

The primary purpose of this assessment methodology is to satisfy reporting requirements of the federal Clean Water Act (CWA) sections 303(d), 305(b), and 314. 33 U.S.C.A. §§ 1313(d), 1315(b) and 1324. Section 303(d) is composed of a list of waters that will not meet all water quality standards (WQS) after implementation of discharge controls. Consequently, 303(d) is a list of waters that require the development of Total Maximum Daily Loads (TMDL) or alternative restoration plans (Category 5 and 5alt in Table 1). Section 305(b) of the CWA is a description of the water quality of all waters within the state. As a part of section 305(b), section 314 also requires states to assess and report the status of significant publicly owned lakes. Pennsylvania's Integrated Water Quality Monitoring and Assessment Report (Integrated Report) is the vehicle by which the above information is submitted to the United States Environmental Protection Agency (USEPA). The CWA requires that the Integrated Report be submitted by April 1st of all even numbered years. The timely completion of the Integrated Report satisfies at least one requirement for states to receive Section 106 Water Pollution Control grant funding, which is a critical resource for DEP water quality protection programs. The Integrated Report is made up of a categorical list of assessment determinations. The current categories that waters fall under after they are assessed are provided in Table 1.

Table 1. Listing categories used to develop the Integrated Report.

Category	Description
1	Waters attaining all uses
2	Waters where some, but not all uses are attaining. Attainment status of the remaining uses may be unknown because data are insufficient to categorize the water or it may be impaired.
3	Waters for which there are insufficient or no data to determine if uses are met.
4a	Waters impaired for one or more uses, not needing a TMDL, because it has been completed.
4b	Waters impaired for one or more uses, not needing a TMDL, because uses are expected to be attained within a reasonable timeframe.
4c	Waters impaired for one or more uses, not needing a TMDL, because the impairment is not related to a pollutant.
5	Waters impaired for one or more uses by a pollutant that require the development of a TMDL.

5alt	Waters impaired for one or more uses by a pollutant that are selected for alternative restoration plan implementation.

Pennsylvania WQS and Implementation

DEP works under a state regulatory framework to make assessment determinations that ultimately create the Integrated Report. This framework includes 25 Pa. Code Chapter 93, (Pennsylvania's WQS) and Chapter 96 (WQS implementation). When making assessments, the most relevant parts of WQS are the protected uses and water quality criteria. Waters can have more than one protected use, and each use may have many applicable criteria. Each criterion may also have different implementation requirements through Chapter 96. Therefore, another purpose of this assessment methodology is to articulate how data are used within this regulatory framework to make assessment determinations.

Protected Uses in Pennsylvania

Protected uses fall into four main categories which include Aquatic Life, Water Supply, Recreation and Fish Consumption, and Special Protection. Aquatic Life Uses (ALU) include Cold Water Fishes (CWF), Warm Water Fishes (WWF), Trout Stocking (TSF), and Migratory Fishes (MF). Water Supply uses include Potable Water Supply (PWS), Industrial Water Supply (IWS), Livestock Water Supply (LWS), Wildlife Water Supply (AWS), and Irrigation (IRS). Recreation and Fish Consumption uses include Boating (B), Fishing (F), Water Contact (WC), and Esthetics (E). Special Protection uses include High Quality Waters (HQ) and Exceptional Value Waters (EV). In addition to the four main categories, an "Other" category lists the Navigation (N) use. See 25 Pa. Code § 93.3 for definitions of these uses.

For the purposes of this book, an "assessment determination" is the decision whether a waterbody is meeting the protected designated or existing use. DEP has an obligation to assess all surface waters of Pennsylvania for the uses listed in § 93.3. To date, the assessment methods in this book allow for assessment determinations in all four main use categories. DEP will continue to refine and develop new assessment methods to achieve the goal of accurate and complete water quality assessment determinations across Pennsylvania.

Criteria and Implementation

Assessment methods may also incorporate 25 Pa. Code Chapter 93 criteria including general and specific water quality criteria (§ 93.6 and § 93.7, respectively), as well as Chapter 96 implementation. This ensures that assessment methods are congruent with applicable WQS and implementation. Many methods in Chapters 2–6 of this book focus on the biological and physical aspects of water quality and have been shown to be

scientifically acceptable indicators for assessments. Assessment methods that do not address specific criteria correspond to general water quality criteria in § 93.6. Assessment methods corresponding to general water quality criteria can be independently applicable or used together (as weight of evidence) to make assessment determinations.

Other Purposes

The primary purpose of assessment methods is to create accurate and precise assessment determinations that make up Pennsylvania's Integrated Report. They are also meant to follow uses, criteria, and implementation found in state regulations. Yet, the production of useful tools (e.g., multimetric indices) that measure aspects of water quality may have additional and very relevant purposes. These purposes may include, but are not limited to, evaluation of permit compliance, source tracking, and measuring incremental progress. DEP encourages the use of assessment methods for additional purposes as long as the applicable data collection protocols (Shull and Lookenbill 2018) are followed. Using assessment methods for other purposes must also consider how each data collection protocol and assessment method was developed, because the divergence of purposes may not produce scientifically valid results. For instance, applying multimetric index calculations from a wadeable stream macroinvertebrate assessment method using data collected along the banks of a large non-wadeable river is not appropriate. This is because the macroinvertebrate data collection in the non-wadeable river is not consistent with wadeable stream data collection protocols, and because macroinvertebrate communities are naturally different between these waterbody types.

NAVIGATING THE BOOK

Chapters 2–4 constitute the specific methods DEP currently has developed and uses when making assessment determinations. These assessment methods fall into 3 categories of data collection: biological, chemical, and physical. The structure of this book is designed to reflect the structure of the Monitoring Book (Shull and Lookenbill 2018) so that users can easily transition between the data collection (monitoring) aspects and use assessment aspects. For accurate and consistent assessments to occur, data collection procedures must follow applicable protocols established in the Monitoring Book (Shull and Lookenbill 2018). As additional monitoring protocols are developed using the conceptual framework (Figure 1), they may subsequently become assessment methods through additional development. Once developed, new assessment methods are published for public and USEPA review and comment. Following any final revisions based on comments received, the method will be added as final to this book.

It is important to note, that all data submitted to DEP during the comment period for the Integrated Report is used in some form for making assessment decisions. More information on the types of data DEP accepts and how that information is used is found in Chapter 5 of this book.

Each assessment method must follow general guidelines for making assessment determinations so that decisions are transparent and consistently made throughout Pennsylvania. General guidelines for assessments are provided in Chapter 5. This Chapter also provides guidelines for conducting a "delisting," which is the determination that either the waterbody has been restored to meeting the use, or that a cause or causes of impairment no longer apply. A critical component in conducting good assessments is accurately identifying the source of the impairment (where the impairment originates from) and the cause of impairment (the pollutant or pollution that's degrading the waterbody). Without a source and cause defined, it is difficult for meaningful and effective restoration efforts to occur. Chapter 6 identifies and describes how DEP makes source and cause determinations. In some cases, sources and/or causes of impairment may require developing unique determination methods. For example, DEP has created the acid precipitation source and cause determination method to facilitate differentiation between anthropogenic impacts and natural conditions. DEP will continue to develop new source and cause determination methods when specific needs arise.

LITERATURE CITED

Shull, D. R., and M. J. Lookenbill (editors). 2018. Water Quality Monitoring Protocols for Streams and Rivers. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

CHAPTER 2 BIOLOGICAL ASSESSMENT METHODS

WADEABLE FREESTONE RIFFLE-RUN STREAM MACROINVERTEBRATE ASSESSMENT METHOD

Prepared by:

Brian Chalfant
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2015

Edited by:

Dustin Shull
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2017

INTRODUCTION

This assessment method is designed to make ALU assessment determinations using benthic macroinvertebrate communities in Pennsylvania's wadeable, freestone, riffle-run streams. Through direct quantification of biological attributes along a gradient of conditions, the index of biotic integrity (IBI) provided in this assessment method measures the extent to which anthropogenic activities compromise a stream's ability to support healthy aquatic communities (Davis and Simon 1995). This IBI may also help guide and evaluate aquatic resource legislation, policy, goals, and management strategies (Davis and Simon 1995; Davies and Jackson 2006; Hawkins 2006). Full technical documentation of this assessment method can be found in the technical report entitled 'A benthic index of biotic integrity for wadeable freestone streams in Pennsylvania' (Chalfant 2012). To use this method for assessment determination purposes data collection must follow applicable protocols established in the Monitoring Book (Shull and Lookenbill 2018).

THE METRICS

Several different metric combinations were evaluated during index development. The following six metrics were selected for inclusion in the IBI based on various performance characteristics. These six metrics all exhibited a strong ability to distinguish between relatively pristine and heavily impacted conditions. In addition, these six metrics measure different aspects of the benthic macroinvertebrate communities. When used together in a multimetric index, these six metrics 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, many types of pollution (Lenat and Penrose 1996), although sensitivity to different types of pollution varies among taxa in these insect orders. The version of this metric used here only counts EPT taxa with 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 pollution tolerance values 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 value 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 1948).

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 pollution tolerance values. 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 (Pollution Tolerance Values 0-3 only)
This community composition and tolerance metric is the percentage of individuals with pollution tolerance values 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 sub-sample from Lycoming Creek in Lycoming County collected November 19, 2001 (Table 1).

Table 1. Benthic macroinvertebrate sub-sample from Lycoming Creek in Lycoming County collected 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* $\leq$ *4*, so:
EPT Taxa Richness (PTV 0-4 only) = 9 + 5 + 8
EPT Taxa Richness (PTV 0-4 only) **= 22**

Beck's Index (version 3)

$$= 3 * (n_{taxaPTV0}) + 2 * (n_{taxaPTV1}) + 1 * (n_{taxaPTV2})$$

Where $n_{taxaPTV0}$ is the number of taxa with a PTV attribute of 0, $n_{taxaPTV1}$ is the number of taxa with a PTV attribute of 1, and $n_{taxaPTV2}$ is the number of taxa with a PTV attribute of 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

$$= \sum_{i=0}^{10} [(i * n_{indvPTVi})] \Big/ N$$

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

In this sub-sample,

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

There is a total of 217 individuals in this sub-sample, so

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

$$= -1 \left(\sum_{i=1}^{Rich} [(n_i/N) \ln(n_i/N)] \right)$$

where n_i = the number of individuals in each taxon (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 taxon are shown in the table above. There are a total of 217 individuals in the sub-sample, so

Shannon Diversity Index = – 1[(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) +
… *(do this for all 33 taxa)*
… (1 / 217) ln (1 / 217)]

Shannon Diversity Index = **2.67**

Percent Sensitive Individuals (PTV 0-3 only)

$$= \left(\sum_{i=0}^{3} n_{indvPTVi} \Big/ N \right) * 100$$

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

In this sub-sample,

there are 22 individuals with PTV = 0,
there are 57 individuals with PTV = 1,
there are 11 individuals with PTV = 2, and
there are 16 individuals with PTV = 3.

There are a total of 217 individuals in this sub-sample, so
Percent Sensitive Individuals (PTV 0-3 only) = (22 + 57 + 11 + 16) / 217 *100
Percent Sensitive Individuals (PTV 0-3 only) = 106 / 217 * 100
Percent Sensitive Individuals (PTV 0-3 only) = **48.8%**

THE INDEX

An index is simply a means to integrate information from various metrics of biological integrity (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).

To account for natural changes in benthic biota with stream size, different metric standardization values for samples from larger streams and smaller streams were developed for this IBI. Data suggest that the small stream approach is usually appropriate for first, second, and third order streams (using the Strahler stream ordering system) draining less than 25 to 50 mi^2, while the large stream approach is usually

appropriate for fifth order and larger streams draining more than 50 mi^2. More detailed guidelines for deciding whether to apply the large-stream or small-stream metric standardization values to a sample are discussed below.

The one selected core metric that increases in value with increasing anthropogenic stress – Hilsenhoff Biotic Index – was standardized to approximately the 5th percentile of metric scores for all samples from smaller streams and for all samples from larger streams in the IBI development dataset to arrive at the respective small-stream and large-stream standardization values. Core metrics that decrease in value with increasing stress – Total Taxa Richness, EPT Taxa Richness, Beck's Index, Shannon Diversity, and Percent Sensitive Individuals – were standardized to approximately the 95th percentile of metrics scores for all samples from smaller streams and for all samples from larger streams in the IBI development dataset to set the respective small-stream and large-stream standardization values (Table 2).

Table 2. The small-stream and large-stream standardization values used for each core metric.

Metric	Metric Standardization Values	
	smaller streams most 1st to 3rd order < 25 square miles	**larger streams** most 5th order and larger > 50 square miles
Total Taxa Richness	33	31
EPT Taxa Richness	19	16
Beck's Index	38	22
Hilsenhoff Biotic Index	1.89	3.05
Shannon Diversity	2.86	2.86
Percent Sensitive Individuals	84.5	66.7

To calculate the index of biological integrity, observed metric values are first standardized using the standardization values shown in the table immediately above and the following standardization equations.

The Hilsenhoff Biotic Index metric values are expected to increase in value with increasing anthropogenic stress and are standardized using the following equation:

Hilsenhoff Biotic Index standardized score =
(10 – observed value) / (10 – standardization value) * 100

The other five core metrics values are expected to decrease in value with increasing anthropogenic stress and are standardized using the following equation:

Standardized metric score = observed value / standardization value * 100

Once the observed metric values are standardized, the standardized metric scores are adjusted to maximum value of 100 if necessary. By standardizing metrics and setting a maximum value of 100 for the standardized metrics, the resulting adjusted standardized metric scores can range from maximum values of 100 to minimum values of zero, with scores 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). This approach establishes upper bounds on the expected condition and moderate effects of metrics that may respond in some manner other than a monotonic response to stress. The index of biological integrity is calculated by calculating the arithmetic mean of these adjusted standardized metric values for the six metrics, resulting in a multimetric index of biological integrity score that can range from 0 to 100. To get a score of zero, a sample would have to contain no organisms at all.

In order to incorporate the variability of metric scores with annual seasons in setting biological expectations, DEP chose to implement different use attainment benchmarks as discussed below rather than adjust metric standardization values.

The sample from Lycoming Creek presented above was collected from a fifth order site draining approximately 173 mi^2 of land, so we will apply the large-stream metric standardization values in the example metric standardization and index calculations presented below (Table 3). For a small-stream sample, we would simply substitute the small-stream metric standardization values in place of the large-stream metric standardization values – the rest of the index calculation process is the same regardless of stream size.

Table 3. Index calculation process for Lycoming Creek.

Metric	Standardization Equation (using large-stream standardization values)	Observed Metric Value	Standardized Metric Score	Adjusted Standardized Metric Score Maximum = 100
Total Taxa Richness	(observed value / 31) * 100	33	106.5	100
EPT Taxa Richness	(observed value / 16) * 100	22	137.5	100
Beck's Index	(observed value / 22) * 100	40	181.8	100
Hilsenhoff Biotic Index	[(10 – observed value) / (10 – 3.05)] * 100	3.47	94.0	94.0
Shannon Diversity	(observed value / 2.86) * 100	2.67	93.4	93.4
Percent Sensitive Individuals	(observed value / 66.7) * 100	48.8	73.2	73.2
Average of standardized core metric scores = **IBI Score =**				**93.4**

AQUATIC LIFE USE ATTAINMENT BENCHMARKS

Due to the influences of annual seasons and drainage area seen in the dataset, DEP recognizes different assessment tools and use attainment thresholds are appropriate for samples collected during different times of the year and from different size stream systems. It is noted that some site-specific exceptions to any thresholds may exist because of local scale natural limitations (e.g., habitat availability) on biological condition (Hughes 1995).

Based on the results of technical analyses, professional workshops, feedback from DEP biologists and other colleagues, as well as policy considerations, DEP implements a multi-tiered benchmark decision process for wadeable, freestone, riffle-run streams in Pennsylvania that incorporates stream size and sampling season as factors for determining aquatic life use attainment and impairment based on benthic macroinvertebrate sampling. A simplified flowchart of this decision process is outlined in the diagram below (Figure 1). Although this simplified decision matrix should guide most assessment decisions for benthic macroinvertebrate samples from Pennsylvania's wadeable, freestone, riffle-run streams using the collection and processing methods

discussed above, situations exist where this simplified assessment schematic will not apply exactly as outlined – some such situations are discussed in the following text.

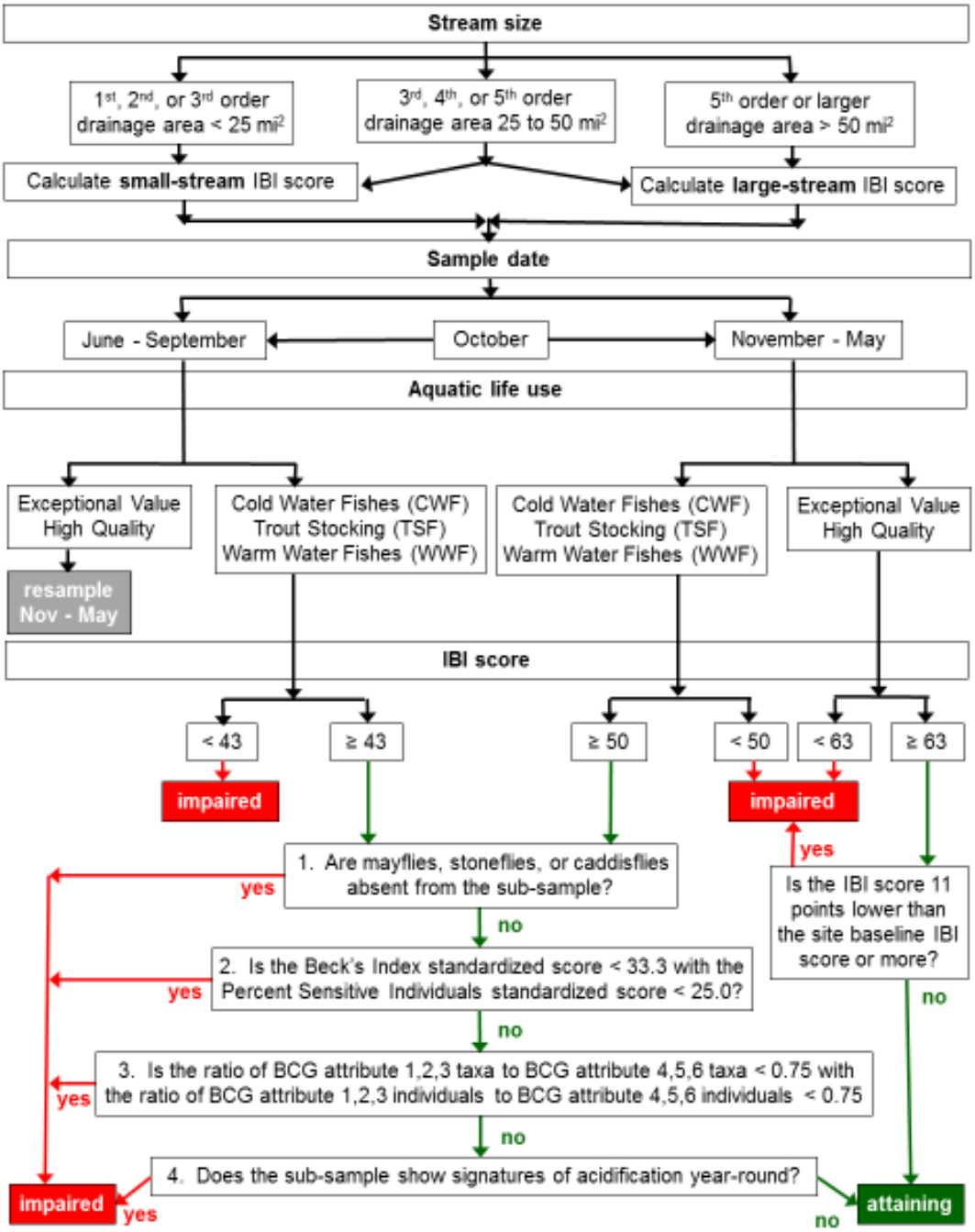


Figure 1. Assessment flowchart for wadeable, freestone, riffle-run streams.

The first step in the aquatic life use assessment process for wadeable, freestone, riffle-run streams in Pennsylvania based on benthic macroinvertebrate sampling considers stream size. DEP does not feel that it is appropriate to set a single cutoff drainage area or stream order threshold to define which set of metric standardization values and which resulting IBI (i.e., large-stream or small-stream) should be applied. However – as stated above – data suggest that the small-stream approach is usually appropriate for samples from first, second, and third order streams draining less than 25 mi^2 of land, while the large-stream approach is usually appropriate for samples from fifth order and larger streams draining more than 50 mi^2.

There are many important considerations when deciding whether to apply the small-stream or large-stream metric standardization values to a sample. Many stream systems experience a variety of changes as they flow from headwaters on downstream. These changes include, but are certainly not limited to changes in canopy shading, energy dynamics, algal growth, erosional and depositional patterns, habitat distributions, water temperature, and flow regimes. These shifts manifest themselves uniquely in each watershed. Streams in more northern, high elevation, high relief areas of the state may maintain cooler water, flashier flows, larger-particle substrates, and other characteristics typical of smaller streams at comparable drainage areas or stream orders when compared with streams in more southern, low elevation, low relief areas of the state. Local climatological and geological patterns also affect a stream's character.

When deciding which set of metric standardization values (i.e., small-stream or large-stream) to apply, care should be taken not to conflate human-induced changes to streams with natural landscape and climatological variations. For example, a stream draining 26 mi^2 of mostly corn and soybean fields with little forested riparian buffer may experience warmer water temperatures and more silted substrates than a stream of similar size draining a more forested watershed. The warmer water and more silted substrates of the agricultural stream may be characteristics typical of larger streams, but if those characteristics are primarily human-induced, then that argues against applying the large-stream metric standardization values based on the presence of those characteristics in the stream.

For streams of intermediate size (i.e., third, fourth, and some fifth order streams draining between 25 and 50 mi^2 of land), it will often be informative to consider both the small-stream and large-stream IBI scores and associated benchmarks. For example, if a sample from a fourth order site draining 30 mi^2 scores 77.0 on the small-stream IBI and 90.2 on the large-stream IBI and passes the additional screening questions, both

approaches indicate aquatic life use attainment, so the use assessment decision is the same regardless of which set of metric standardization values is applied. In another instance, a sample collected in mid-March from a site draining 36 mi^2 may score 44.1 on the small-stream IBI – indicating impairment – while scoring 51.2 on the large-stream IBI – indicating possible attainment. Here, the small-stream and large-stream IBI score assessment decisions diverge. In such situations, it may be especially useful to consider the additional screening questions – detailed below – when making an assessment decision.

The second step in the aquatic life use assessment process for wadeable, freestone, riffle-run streams in Pennsylvania based on benthic macroinvertebrate sampling considers sampling season. Samples collected during summer and early autumn months (i.e. June through September) are held to different IBI attainment thresholds than samples collected November through May since benthic macroinvertebrate communities in most wadeable, freestone, riffle-run streams in Pennsylvania exhibit consistent patterns of lower taxonomic diversity and organismal abundance during the summer and early autumn months compared with other times of the year. These seasonal index periods are intended as general guidelines and may vary slightly year-to-year depending on local climatological conditions. For example, a sample collected from a low elevation, low latitude stream 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 – especially if many mayflies have already emerged from the stream – while a sample collected from a high elevation, high latitude location during the first week of June in an uncharacteristically cool, wet year may be more properly evaluated using the November to May procedures – especially if many mayfly nymphs are still present in the benthos.

October often is a transitional time for benthic macroinvertebrate communities in Pennsylvania with samples from earlier in the month resembling late summer communities (e.g., relatively low diversity and abundance) and samples from later in the month resembling early winter communities (e.g., increasing abundance of winter stoneflies). Therefore, depending on local climate, basin geology, and other factors discussed above (e.g., latitude, elevation, basin relief) samples from October may be evaluated using the June to September benchmarks or the November to May benchmarks. DEP advises against sampling in mid-October to avoid these issues.

For samples collected between November and May, IBI scores < 50 result in aquatic life use impairment. Samples collected during these months scoring ≥ 50 on the appropriate IBI are subject to four screening questions before the aquatic life use can be considered attaining. **These additional screening questions are:**

1. **Are mayflies, stoneflies, or caddisflies absent from the sub-sample?** Organisms representing these three taxonomic orders are usually found in most healthy wadeable, freestone, riffle-run streams in Pennsylvania. If any or all of these orders are absent from a sample, this strongly suggests some sort of anthropogenic impact. Samples where one of these taxonomic orders is absent due to natural conditions (e.g., mayflies absent from a low-pH tannic stream) should be evaluated accordingly. *This question must be applied to small-stream samples collected between November and May, but does not have to be applied to samples from larger streams and samples collected between June and September.*

2. **Is the standardized metric score for the Beck's Index metric < 33.3 with the standardized metric score for the Percent Sensitive Individuals metric < 25.0?** Although these two metrics go into the IBI calculations, this screening question serves to double check that a sample has substantial richness and abundance of the most sensitive organisms. *This question must be applied to all samples.*

3. **Is the ratio of Biological Condition Gradient (BCG) attribute 1,2,3 taxa to BCG attribute 4,5,6 taxa < 0.75 with the ratio of BCG attribute 1,2,3 individuals to BCG attribute 4,5,6 individuals < 0.75?** This screening question evaluates the balance of pollution tolerant organisms with more sensitive organisms in terms of taxonomic richness and organismal abundance. By using the BCG attributes to measure pollution tolerance, this screening question serves as a check against the IBI metrics which account for pollution sensitivity based only on PTVs. *This question must be applied to small-stream samples collected between November and May, but does not have to be applied to samples from larger streams and samples collected between June and September.*

4. **Does the sub-sample show signatures of acidification year-round?** The primary acidification signatures in a sub-sample include low mayfly abundance and low mayfly diversity (i.e., scarce mayfly individuals and few mayfly taxa), especially when combined with high abundance of Amphinemura and/or Leuctra stoneflies, occasionally combined with high abundance of Simuliidae and/or Chironomidae individuals. A sub-sample with < 3 mayfly taxa, < 5% mayfly individuals, and > 25% Leuctra and/or Amphinemura stoneflies indicates likely acidification impacts. Acidification

effects on benthic macroinvertebrate communities are often most pronounced in small streams with low buffering capacity during the spring months when snowpacks melt and vernal rains are frequent. While it can be difficult to determine if low pH conditions in a stream are natural or more attributable to anthropogenic acidification, sampling of water chemistry and/or fish communities in addition to benthic macroinvertebrate communities can help inform assessment of acidic in-stream conditions. With this protocol, DEP will only impair sites that show persistent acidification signatures year-round. In other words, if a sample has no mayflies and is dominated by Leuctra and Amphinemura in the spring, but a November sample from the same site contains three or more mayfly taxa or over five percent mayfly individuals, the aquatic life use will not be considered impaired because the stream exhibits the ability to recover biological integrity in the fall and winter months. If a spring sample shows acidification signatures, a late fall or early winter sample must be collected before making an aquatic life use assessment decision. *This question must be applied to all samples.*

If the answer to these four screening questions (if applicable) is yes for a sample collected between November and May with an IBI score ≥ 50, then the sample is considered impaired without compelling reasons otherwise. If the answer to these questions (if applicable) is no for a sample collected between November and May with an IBI score ≥ 50, then the aquatic life use represented by the sample can be considered attaining unless other information (e.g., water chemistry) indicates the aquatic life use may not be fully supported at that location.

For samples collected between June and September, the same logic applies as for samples collected between November and May, but the attainment/impairment threshold is lowered to 43 instead of 50. For samples collected in the summer and early autumn time frame, the absence of mayflies – and in some instances stoneflies – in samples collected immediately after seasonal hatches may be relaxed in some cases. Because benthic diversity may be underrepresented in summer and early autumn samples DEP encourages monitoring in the November to May timeframe if possible. Benthic macroinvertebrate sampling for determining aquatic life use support should only be conducted from June to early October if sampling during other seasons is not possible due to hazardous conditions such as high, fast stream flow.

Limestone Influence

As discussed in the introduction, DEP deploys a different data collection protocol (Botts 2009b) and assessment methodology (Botts 2009a) for limestone spring streams whose flow is mostly or entirely derived from groundwater in areas with substantial primary calcareous geologies than for freestone streams. The sampling methodology and assessment protocol for these limestone spring streams incorporate the understanding that streams in areas receiving a substantial amount of flow from groundwater attributable to karst geologies often naturally have less diverse benthic macroinvertebrate communities than streams draining freestone geologies. This lower benthic macroinvertebrate community diversity in limestone spring streams is attributable in large part to less variable flow and thermal characteristics of such systems when compared with freestone streams that often exhibit flashier flows and a wider range of temperatures.

Some streams in Pennsylvania drain basins underlain partially by freestone geologies and partially by calcareous geologies. Such streams are often encountered in central regions of the state – especially in upper portions of the Juniata River basin – where they drain sandstone and/or quartzite upland ridges, steep shale slopes, and lower gradient calcareous valley floors. The calcareous valley geologies in these basins contributes to relatively high alkalinities and relatively high and consistent base flows in streams – characteristics of limestone spring streams – when compared with streams draining basins with no calcareous geologies. However, the upland sandstone, quartzite, and shale areas of these basins often contribute substantial surface runoff, which leads to surges in flow during rainfall and snowmelt events and dilution of alkalinity derived from the calcareous valleys. These streams – often referred to as “limestone-influenced” – exhibit some characteristics of limestone spring streams and some characteristics of freestone streams.

We often see substantial agriculture in the fertile valleys of these limestone-influenced streams, which makes it difficult to definitively establish reference conditions specific to these unique streams. However, there is evidence that the benthic macroinvertebrate communities in limestone-influenced streams are naturally less diverse than in freestone streams of similar size and with similar land uses. This lower diversity of benthic macroinvertebrate communities in limestone-influenced streams likely reflects the less variable flow and thermal patterns in these streams caused by the stabilizing influence of the substantial groundwater flowing into the streams through the calcareous valley geologies. Commonly, the benthic macroinvertebrate communities in limestone-influenced streams exhibit relatively low stonefly diversity and abundance when compared with streams of similar size and condition that drain freestone geologies.

In light of these considerations, use attainment benchmarks may be justifiably relaxed for samples from limestone-influenced streams. The June to September IBI benchmark of 43 for freestone streams can be applied to limestone-influenced streams year-round, but the four screening question should still be applied as outlined above to samples from limestone-influenced streams to make ALU assessment decisions.

Antidegradation, Special Protection Considerations

The assessment decision process is somewhat different for streams with special protection uses of high-quality (HQ) or exceptional value (EV) waters. DEP will protect special protection streams based on a baseline IBI score determined by previous surveys. Subsequent samples from HQ and EV streams will be compared to the baseline IBI score for a given site using the IBI temporal precision estimates (Table 4). For example, if Mill Creek is designated HQ and a previous sample from a given site on Mill Creek using the protocol described above results in a mid-April IBI score of 78.0, this IBI score of 78.0 would be the baseline IBI score for that site. Future samples from that site collected November to May that score more than 10.0 IBI points below 78.0, would be considered impaired. Since DEP's sampling season for special protection surveys is November to May, we need not be concerned about how June to October samples compare to the baseline IBI – DEP will only make assessment decisions for HQ and EV streams based on samples collected November to May. The temporal precision estimate of 10.0 points is used because it approximates the October to May temporal precision estimate calculated in the table below. DEP will apply the more restrictive March to May and October to February temporal precision estimates – about 9.0 and 8.0 IBI points, respectively – to special protection use assessments if the situation is appropriate (e.g., if the baseline IBI was established in April, future March to May samples that score more than 9.0 points lower than the baseline will be considered impaired). Furthermore, any sample from an HQ or EV stream that scores less than 63.0 on the IBI will be considered impaired without compelling reasons otherwise (e.g., a stream was designated HQ or EV for a reason other than assessment of the benthic macroinvertebrate community).

Table 4. Temporal precision estimates for IBI scores and core metrics based on ANOVA results. The ANOVA mean square error (MSE) estimates intrasite standard deviation. Coefficients of variation (CV) were calculated for each sample pair (or triplet or quadruplet…) and then averaged across all sample pairs. "*s*" indicates standardized metric values. "*r*" indicates raw metric values.

Metric		small-stream						large-stream					
		November to May 384 samples from 137 sites			June to September 26 samples from 12 sites			November to May 78 samples from 26 sites			June to September 26 samples from 7 sites		
		ANOVA MSE	90% CI (1 sample)	%CV	ANOVA MSE	90% CI (1 sample)	%CV	ANOVA MSE	90% CI (1 sample)	%CV	ANOVA MSE	90% CI (1 sample)	%CV
IBI score		48.9	**9.0**	**8.8**	95.7	**12.5**	**19.6**	69.0	**10.6**	**10.3**	18.5	**5.5**	**4.8**
Total Taxa	*s*	115.0	13.7	10.9	101.0	12.9	13.3	128.0	14.50	12.5	103.0	13.0	10.0
Richness	*r*	16.6	5.2	13.2	16.1	5.14	14.8	15.5	5.05	13.2	12.1	4.5	11.3
EPT Taxa	*s*	138.0	15.0	18.5	89.5	12.13	23.8	185.0	17.44	17.3	78.8	11.4	10.7
Richness (PTV 0-4 only)	*r*	6.3	3.2	19.7	4.8	2.81	24.7	7.9	3.59	20.8	2.0	1.8	10.7
Beck's	*s*	127.0	14.4	22.8	94.4	12.46	36.9	132.0	14.73	14.2	142.0	15.3	24.6
Index (version 3)	*r*	21.9	6.0	23.7	17.9	5.42	37.5	16.0	5.13	19.7	10.4	4.1	26.4
Hilsenhoff	*s*	53.1	9.3	7.3	222.0	19.10	22.6	71.3	10.83	8.3	18.5	5.5	4.5
Biotic Index	*r*	0.4	0.8	15.6	1.5	1.57	21.2	0.4	0.81	15.4	0.1	0.4	6.1
Shannon	*s*	96.1	12.6	10.1	131.0	14.67	14.1	120.0	14.04	10.5	33.5	7.4	5.3
Diversity	*r*	0.1	0.4	10.7	0.1	0.45	14.4	0.1	0.42	10.8	0.0	0.2	5.7
% Sensitive	*s*	215.0	18.8	23.6	361.0	24.36	65.7	337.0	23.53	27.7	133.0	14.8	16.5
Individuals (PTV 0-3 only)	*r*	157.0	16.1	23.8	258.0	20.59	65.7	197.0	23.53	30.2	59.1	9.9	16.5

Applications and Exceptions

If a sample results in fewer than 160 total organisms in the entire sample, the IBI and assessment procedures may not apply exactly as outlined above. The IBI and associated benchmarks are calibrated for use with sub-samples containing 160 to 240 organisms, so applications of the IBI to samples containing less – or more – than the target number of organisms, cannot necessarily be assessed using the procedures and benchmarks outlined above. Low abundance of benthic organisms often indicates toxic pollution or severe habitat alterations, which must be considered in making holistic stream assessments.

The use assessment decision processes set forth above are intended as general guidelines, not as hard-and-fast rules. The procedures and guidelines discussed above will provide tenable assessments – as required by federal and state law – of benthic macroinvertebrate community conditions for the vast majority of samples collected from wadeable, freestone, riffle-run streams in Pennsylvania. However, as noted by Hughes (1995), there will be exceptional circumstances – such as those outlined in the Pennsylvania Code (2011: Title 25, Section 93.4. (b) relating to less restrictive uses) – when the above assessment procedures do not apply (e.g., there are no obvious sources of impairment and natural factors such as habitat availability or water chemistry limit biotic potential). In some situations, a biologist's local knowledge of conditions may warrant a decision not arrived at using these guidelines. Although the large-steam IBI appears to work well when applied to samples from large rivers (i.e., sites draining over 1,000 square miles), discretion must be used when applying this IBI to samples from such large rivers. These methods do not apply if a stream/river is not wadeable in over 90% or more of its channel area under base flow conditions for the river segment to be sampled or other situations not consistent with riffle and run dominated habitat. The relatively small dataset of samples from such large rivers used in the IBI development limits analysis of variability (i.e., estimates of spatial and temporal precision) in metric and IBI performance with samples from such large rivers.

In other situations, like when samples are heavily dominated by Prosimulium larvae – as discussed above – often this will unduly lower metric and IBI scores, confounding the assessment decision procedures outlined above. In such situations, the investigating biologist may have to re-sample the site after the seasonal Prosimulium larval boom, or the biologist may have to rely on a more qualitative analysis of metric scores, sample composition, and site conditions to arrive at an assessment decision. In any instance, evaluating stream samples requires mindfulness of conditions, and is not always a definite, exact exercise. A certain section of stream may represent a transition between pool-glide, low-relief, marshy, glaciated uplands where the substrate is mostly fine-grained sand and higher-gradient lower reaches filled with cobble-strewn riffles and

runs. Some years see cooler, wetter springs than other years. Nevertheless, for the vast majority of cases involving benthic macroinvertebrate samples from wadeable, freestone (and limestone-influenced), riffle-run streams in Pennsylvania using the protocols described above, the assessment procedures described here will lead to tenable ALU assessment decisions.

LITERATURE CITED

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.

Botts, W. F. 2009a. Wadeable Limestone Macroinvertebrate Assessment Method. Chapter 2, pages 25–35. *In* Shull, D. R., and M. Pulket. (editors). Assessment Methodology for Streams and Rivers. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania.

Botts, W. F. 2009b. Wadeable Limestone Macroinvertebrate Data Collection Protocol. Chapter 3, pages 8–12. *In* Shull, D. R., and M. J. Lookenbill. (editors). Water Quality Monitoring Protocols for Streams and Rivers. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

Chalfant, B. 2012. A benthic index of biotic integrity for wadeable freestone streams in Pennsylvania. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania. http://files.dep.state.pa.us/Water/Drinking%20Water%20and%20Facility%20Regulation/WaterQualityPortalFiles/Technical%20Documentation/freestoneIBImarch2012.pdf

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):305–306.

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

Shull, D. R., and M. J. Lookenbill (editors). 2018. Water Quality Monitoring Protocols for Streams and Rivers. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

WADEABLE LIMESTONE STREAM MACROINVERTEBRATE ASSESSMENT METHOD

Prepared by:

William Botts
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2009

Edited by:

Amy Williams
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2017

INTRODUCTION

This assessment method is designed to make ALU assessment determinations using benthic macroinvertebrate communities in Pennsylvania's limestone streams. Limestone streams are streams formed or strongly influenced by limestone springs. All limestone streams are in limestone geology, but not all streams in limestone geology are limestone streams. To determine whether a stream is limestone, several parameters must be investigated. These parameters should be considered during the early stages of data collection and are consequently provided by Botts (2009). If these criteria are not met, the stream is likely not limestone, and another data collection protocol should be considered. To use this method for assessment determination purposes data collection must follow applicable protocols established in the Monitoring Book (Shull and Lookenbill 2018).

Limestone stream macroinvertebrate communities have low diversity, with only a few taxa showing high density. This is true to some degree in both reference and impaired streams. Table 1 lists the five most common taxa collected from limestone streams and shows how the composition of the macroinvertebrate communities changes from reference sites to impaired sites. This table clearly shows how different limestone macroinvertebrate communities are and how these differences could affect the metric selection process.

Table 1. Average percent of organisms per taxa collected per sample

Taxa	Common Name	TV	Reference Sites	Attaining Sites	Impaired Sites
Lirceus	Sowbugs	8	9.4 %	29.5 %	52.9 %
Gammarus	Scuds	6	25.0 %	10.7 %	12.7 %
Ephemerella	Mayflies	1	12.0 %	12.4 %	1.2 %
Optioservus	Riffle Beetles	4	11.6 %	5.8 %	1.8 %
Chironomidae	Midges	6	15.7 %	14.8 %	15.5 %
Total %	**Organisms**		**73.6 %**	**73.2 %**	**84.0 %**

These five taxa account for 45,967 out of 58,010 organisms, or 79.2%, collected from 188 samples. Tolerance Value = TV

This document outlines the procedures for interpretation of samples collected from true limestone streams. The protocol was modified from the IBI for limestone streams technical report (Botts 2009). Technical details of the metric selection process and scoring are presented here.

THE METRICS

The following describes the metrics used to evaluate the macroinvertebrate communities in a limestone stream sample (Table 2).

Table 2. Metrics used to evaluate limestone stream samples

Category	Metric	Definition	Response to Pollution
Richness Measure	Total Taxa	Number of taxa in the subsample.	Decreases
	EPT Taxa	Number of taxa in the orders Ephemeroptera, Plecoptera, and Trichoptera.	Decreases
Tolerance/ Intolerance Measures	HBI	The biotic index and abundance of each taxa are used to find a biotic index for the sample.	Increases
	% Tolerant	Percent of organisms considered to be tolerant of pollution (HBI > 6).	Increases
	Beck's Index, 4	Taxa with a Hilsenhoff Biotic Index (HBI) of 0 or 1 are given 2 points and HBI of 2, 3, or 4 are given 1 point.	Decreases
Composition Measures	Shannon Diversity	Uses both taxa richness and abundance to measure general diversity and composition.	Decreases

The following provides a detailed explanation on how to calculate the six metric scores for limestone streams (Table 3). After the field and lab procedures have been completed, a macroinvertebrate list of 300 +/- 10% organisms will be produced.

Table 3. Taxa List for Letort Spring Run (20160330-0900-ablascovic)

Taxonomic Level	Taxa Name	# of Individuals	Hilsenhoff Score
Coleoptera	Optioservus	1	**4**
Diptera	Antocha	3	**3**
Diptera	Chironomidae	35	6
Ephemeroptera	Baetis	19	6
Ephemeroptera	Ephemerella	34	**1**
Gammarus	Gammarus	34	**4**
Lirceus	Lirceus	55	8
Oligochaeta	Oligochaeta	10	10
Trichoptera	Cheumatopsyche	17	6
Trichoptera	Chimarra	2	**4**
Trichoptera	Hydropsyche	80	5
Turbellaria	Turbellaria	1	8

Total Taxa

This metric sums the total number of taxa identified in the sub-sample (count the number of rows in the above table). In the Letort Spring Run sample, there are **12** taxa.

EPT Taxa

To calculate this metric, sum the total number of Mayfly (Ephemeroptera), Stonefly (Plecoptera), and Caddisfy (Trichoptera) taxa found in the sub-sample. In the above sample, Ephemeroptera are colored red and Trichoptera are colored blue; there are no Plecoptera in the sample:

Letort Spring Run:
Ephemeroptera = 2
Plecoptera = 0
Trichoptera = 3
5

Hilsenhoff Biotic Index (HBI)

This community composition and tolerance metric is calculated as an average of the number of individuals in a sub-sample, weighted by pollution tolerance values. 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.

$$= \sum_{i=0}^{10} [(i * n_{indvPTVi})] \Big/ N$$

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

Letort Spring Run:
In this sub-sample:

There are 0 individuals with PTV = 0,
there are 34 individuals with PTV = 1,
there are 0 individuals with PTV = 2,
there are 3 individuals with PTV = 3,
there are 37 individuals with PTV = 4,
there are 80 individuals with PTV = 5,
there are 71 individuals with PTV = 6,
there are 0 individuals with PTV = 7,
there are 56 individuals with PTV = 8,
there are 0 individuals with PTV = 9,
there are 10 individuals with PTV = 10.

There is a total of 291 individuals in this sub-sample, so

Hilsenhoff Biotic Index = [(0 * 0) + (1 * 34) + (2 * 0) + (3 * 3) + (4 * 37) + (5 * 80) + (6 * 71) + (7 * 0) + (8 * 56) + (9 * 0) + (10 *10)] / 291 =

Hilsenhoff Biotic Index = **5.38**

% Tolerant
This metric is the percent of organisms in the sub-sample considered to be tolerant of pollution (HBI > 6).

Letort Spring Run:
In this sub-sample, there are 66 individuals with PTV > 6 (values 7 through 10).

(66 / 291) * 100 = **22.7%**

Beck's Index, Version 4
Beck's Index, Version 4 is a pollution weighted taxa richness measure, based on Hilsenhoff Biotic Index Scores (HBI). Hilsenhoff's index measures the pollution tolerance of an organism on a scale of 0 to 10, where the organisms' tolerance level decreases with the score. Therefore, it differs from the Beck's Index used in the DEP Riffle/Run Freestone protocol. For Beck's Index, 4, taxa with a HBI score of 0 or 1 are given 2

points and HBI scores of 2, 3, or 4 are given 1 point. In the table, taxa with a score of 0 or 1 are highlighted in blue and scores of 2, 3, and 4 are highlighted in purple.

Letort Spring Run:

Total # of taxa with HBI score of 0 or 1 = **1**
2 pts. x 1 = **2**

Total # of taxa with HBI score of 2, 3, or 4 = **4**
1 pt. x 4 = **4**

2 + **4** = **6**

Shannon Diversity
This index measures taxa abundance and evenness in the sub-sample by dividing the # of individuals in a taxon by the total # of individuals in the sub-sample and then multiplying by the natural logarithm of this proportion. This is done for all taxa in the sub-sample; the products are then summed and the answer multiplied by -1.

$$= -1\left(\sum_{i=1}^{Rich}[(n_i/N)\ \ln(n_i/N)]\right)$$

where n_i = the number of individuals in each taxon (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).

Letort Spring Run:

TaxaRich = 12
N = 291

(1/291) ln (1/291) + (3/291) ln (3/291) + (35/291) ln (35/291) + (19/291) ln (19/291) + (34/291) ln (34/291) + (34/291) ln (34/291) + (55/291) ln (55/291) + (10/291) ln (10/291) + (17/291) ln (17/291) + (2/291) ln (2/291) + (80/291) ln (80/291) + (1/291) ln (1/291)

See below for final answer:

-0.0194959562 + -0.0471619689 + -0.2547392859 + -0.1781745755 + -0.2508478806 + -0.2508478806 + -0.3148778505 + -0.1158329269 + -0.1659170745 + -0.0342280143 + -0.3549956378 + -0.0194959562 = -2.00615008 * -1 = **2.00615008 = 2.01**

INDEX OF BIOTIC INTEGRITY (IBI) SCORE

The individual metrics are scored using the standardization formulas as shown below (Tables 4 and 5). Table 4 scores the metrics that increase as conditions improve and Table 5 scores the metrics that increase as conditions degrade.

Table 4. Scoring metrics that increase with good stream conditions.

	Standard (Best Value)		
Metric	X_{95}	Xmin	Standardization Formula
Total Taxa	18.0	0	Score = (X/18.0) x 100
EPT Taxa	8.0	0	Score = (X/8.0) x 100
Beck's Index, 4	12.0	0	Score = (X/12.0) x 100
Shannon Diversity	2.13	0	Score = (X/2.13) x 100

Metrics such as % Tolerant and HBI increase with greater impairment. The lower the score for these metrics the better the ecological condition.

Table 5. Scoring metrics that increase with poor stream conditions.

	Standard (Best Value)		
Metric	X_5	Xmax	Standardization Formula
% Tolerant	1.5	100	Score = (100 - X/100 - 1.5) x 100
HBI	3.84	10	Score = (10 - X/10 - 3.84) x 100

Now that the six metric scores have been calculated, the scores are plugged into the normalized metric score equation: (Observed Value / 95th percentile) x 100. Some metrics may have a normalized score greater than 100 because normalization is based on the 95th percentile values of the statewide dataset. Normalized metric scores above 100 are adjusted to a score of 100. The adjusted metric scores for the six metrics are summed and then averaged to give the Total Biological Score. Table 6 below shows how to calculate the normalized metric scores and Total Biological Scores for the Letort Spring Run sample.

Table 6: Total Biological Score Calculation for Letort Spring Run

Metric	Equation	Observed Value	Normalized Metric Score	Adjusted Metric Score (Max = 100)

Total Taxa	(Observed / **18.0**) x 100	12	66.67	66.7
EPT Taxa	(Observed / **8.0**) x 100	5	62.5	62.5
Beck's Index, 4	(Observed / **12.0**) x 100	6	50	50
Shannon Diversity	(Observed / **2.13**) x 100	2.01	94.37	94.4
% Tolerant	[(100 - Observed) / (**100 - 1.5**)] x 100	22.7	78.4771574	78.5
HBI	[(10 - Observed) / (10 - 3.84)] x 100	5.38	75	75
Total Biological Score (IBI)				71.183333

AQUATIC LIFE USE BENCHMARK

The final score is compared to the values in Table 7, below, and assigned to one of four categories. Sites scoring less than 60 are considered impaired and should be placed on Integrated List Category 5 of impaired streams requiring TMDLs.

Table 7. Limestone Stream IBI Scoring Thresholds

	CWF	CWF	Impaired CWF	
Classification:	Reference	Attaining	Moderately	Severely
IBI Score	>73	73-60	<60-30	<30

Note: Less Than <60 is impaired

In the above example, Letort Spring Run has a final score of 71.2 and would be documented as attaining its ALU.

TEMPORAL PRECISION ESTIMATES

Temporal precision estimates were calculated to demonstrate the method's precision over time (Table 8). This used 193 temporally paired samples at 50 sites. Sites were sampled a minimum of two times and a maximum of nine times over a several-month to

twelve-year period. Only samples collected from January through May were used. Sites were determined to be from "true" limestone streams. Samples from reference, attaining, and impaired sites were included in this analysis. The 90% confidence interval was 13.5. This indicates that samples collected from the same site over time may differ by approximately 13.5 points, but differences greater than or less than that number of points may indicate an anthropogenic or other change in the site.

Table 8. Temporal precision estimates for IBI scores and core metrics based on ANOVA results. The ANOVA mean square error (MSE) estimates intrasite standard deviation. Coefficients of variation (CV) were calculated for each sample pair (or triplet or quadruplet…) and then averaged across all sample pairs. "*s*" indicates standardized metric values. "*r*" indicates raw metric values.

Metric		limestone January to May 193 samples from 50 sites ANOVA MSE	90% CI (1 sample)	%CV
IBI score		**111.5**	**13.5**	**13.9**
Total Taxa Richness	*s*	197.7	18	16.5
	r	6.4	3.2	16.5
EPT Taxa Richness	*s*	354.9	24.2	34
	r	2.3	1.9	34
Beck's Index (version 4)	*s*	271.4	21.1	29.8
	r	3.9	2.5	29.8
Hilsenhoff Biotic Index	*s*	115.3	13.8	14.3
	r	0.4	0.8	9.4
Shannon Diversity	*s*	159.7	16.2	15.7
	r	0.1	0.3	15.7
% Tolerant Individuals (PTV 7-10 only)	*s*	360.5	24.3	25.7
	r	349.7	24	51.9

CONCLUSION

In 2009, DEP finalized a macroinvertebrate bioassessment protocol for assessing Pennsylvania's limestone streams. Using the field and laboratory methods outlined in DEP's protocol, a macroinvertebrate taxonomic list is produced. The taxonomic data is then used to calculate metrics and produce IBI score that accurately reflects the ecological conditions of the waterway. This IBI score is compared to the ALU attainment benchmarks to determine if the sample reach is attaining or impaired.

LITERATURE CITED

Botts, W. F. 2009. Wadeable Limestone Macroinvertebrate Data Collection Protocol. Chapter 3, pages 9–12. *In* Shull, D. R., and M. J. Lookenbill. (editors). Water Quality Monitoring Protocols for Streams and Rivers. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

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.

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.

Shull, D. R., and M. J. Lookenbill (editors). 2018. Water Quality Monitoring Protocols for Streams and Rivers. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

WADEABLE MULTIHABITAT STREAM MACROINVERTEBRATE ASSESSMENT METHOD

Prepared by:

Charles McGarrell and Molly Pulket
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2007

Edited by:

Molly Pulket
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2017

INTRODUCTION

This assessment method is designed to make ALU assessment determinations using benthic macroinvertebrate communities in Pennsylvania's low-gradient streams. The USEPA's Rapid Bioassessment Protocols for use in Wadeable Streams and Rivers (Barbour et al.1999) describes two approaches to collecting macroinvertebrate community data. These approaches are the "riffle-run" approach and the "multihabitat" approach. Due to low-gradient streams typically lacking riffle-run habitat, the multihabitat approach is preferred. Multihabitat data collection involves sampling a variety of habitat types instead of sampling a single habitat. To use this method for assessment determination purposes data collection must follow applicable protocols established in the Monitoring Book (Shull and Lookenbill 2018). For detailed information on the development of this method, reference the full technical document available on the DEP website (McGarrell 2007).

METRICS

The six core metrics listed in Table 1 were chosen because they were the most powerful in differentiating between reference and impaired low-gradient sites. These metrics are used to calculate a station's IBI score.

Table 1. Six Core Low-Gradient Metrics

Category	Metric	Definition	Response to Pollution
Richness Measure	Taxa Richness	Total number of taxa	Decreases
	EPT Taxa	Number of taxa in the orders Ephemeroptera, Plecoptera, and Trichoptera.	Decreases
Tolerance/ Intolerance Measures	Beck4	Taxa with a Hilsenhoff Biotic Index (HBI) of 0 or 1 are given 2 points and HBI of 2, 3, or 4 are given 1 point.	Decreases
Abundance Measures	# Mayfly Taxa	Total number of Mayflies (Ephemeroptera)	Decreases
	# Caddisfly Taxa	Total number of Caddisflies (Trichoptera)	Decreases
Composition Measures	Shannon Diversity	Uses both taxa richness and abundance to measure general diversity and composition.	Decreases

The following provides a detailed explanation on how to calculate the six metric scores for two low-gradient streams, Saw Creek and Wiconisco Creek. After the field and lab procedures have been completed, a macroinvertebrate list of 200 +/- 10% organisms will be produced. The following taxa lists are color coded to help distinguish the taxa and information that will be used to calculate the metrics.

Table 2. Taxa List for Saw Creek (20040406-1705-CAM)

Taxonomic Level	Taxa Name	Number of Individuals	Hilsenhoff Score	Functional Feeding Group
Diptera	Chironomidae	109	6	CG
Isopoda	Caecidotea	8	6	CG
Trichoptera	Pycnopsyche	16	**4**	SH
Ephemeroptera	Eurylophella	4	**4**	SC
Trichoptera	Platycentropus	2	**4**	SH
Diptera	Ceratopogonidae	3	6	PR
Bivalvia	Sphaeriidae	3	8	FC
Oligochaeta	Oligochaeta	3	10	CG
Trichoptera	Oecetis	1	8	PR
Hirudinea	Hirudinea	1	8	PR
Ephemeroptera	Stenonema	3	**3**	SC
Plecoptera	Amphinemura	3	**3**	SH
Trichoptera	Lype	7	**2**	CG
Plecoptera	Isoperla	3	**2**	PR
Plecoptera	Leuctra	5	**0**	SH
Trichoptera	Diplectrona	3	**0**	FC
Trichoptera	Wormaldia	1	**0**	FC
Trichoptera	Rhyacophila	3	**1**	PR
Trichoptera	Lepidostoma	1	**1**	SH
Plecoptera	Prostoia	3	**2**	SH
Trichoptera	Molanna	7	6	SC
Diptera	Simulium	13	6	FC
Diptera	Prosimulium	2	5	FC
Diptera	Pseudolimnophila	1	**2**	PR
Diptera	Dicranota	11	**3**	PR
Diptera	Tipula	1	**4**	SH

Table 3. Taxa List for Wiconisco Creek (20050525-1030-CAM)

Taxonomic Level	Taxa Name	# of Individuals	Hilsenhoff Score	Functional Feeding Group
Diptera	Chironomidae	151	6	CG
Isopoda	Caecidotea	1	6	CG
Trichoptera	Platycentropus	1	**4**	SH
Diptera	Ceratopogonidae	2	6	PR
Bivalvia	Sphaeriidae	3	8	FC
Oligochaeta	Oligochaeta	35	10	CG
Amphipoda	Crangonyx	3	**4**	CG
Odonata	Calopteryx	1	6	PR
Plecoptera	Leuctra	1	**0**	SH
Megaloptera	Sialis	1	6	PR
Odonata	Lestes	1	9	PR
Odonata	Ischnura	1	9	PR

EPT

To calculate this metric, sum the total number of Mayfly (Ephemeroptera), Stonefly (Plecoptera), and Caddisfy (Trichoptera) taxa found in the sub-sample:

Saw Creek		Wiconisco Creek	
Ephemeroptera	= 2	Ephemeroptera	= 0
Plecoptera	= 4	Plecoptera	= 1
Trichoptera	= 9	Trichoptera	= 1
	15		**2**

Taxa Richness

This metric sums the total number of taxa identified in the sub-sample (count the number of rows in the above tables):

Saw Creek = **26** Wiconisco Creek = **12**

Beck4

Beck4 is a pollution weighted taxa richness measure, based on Hilsenhoff Biotic Index Scores (HBI). Hilsenhoff's index measures the pollution tolerance of an organism on a scale of 0 to 10, where the organisms' tolerance level decreases with the score. This metric is a modification of Beck's Index; it was chosen because this version works better for low-gradient streams. Therefore, it differs from the Beck's Index used in the DEP

Riffle/Run Freestone protocol. For Beck4, taxa with a HBI score of 0 or 1 are given 2 points and HBI scores of 2, 3, or 4 are given 1 point. In the tables, scores of 0 and 1 are highlighted in blue and scores of 2, 3, and 4 are highlighted in purple.

Saw Creek	Wiconisco Creek
Total # of taxa with HBI score of 0 or 1 = **5**	Total # of taxa with HBI score of 0 or = **1**
2 pts. x 5 = **10**	2 pts x 1 = **2**
Total # of taxa with HBI score of 2, 3, or 4 = 11	Total # of taxa with HBI score of 2, 3, or 4 = 2
1 pt. x 11 = **11**	1 pt. x 2 = **2**
10 + **11** = **21**	**2** + **2** = **4**

Shannon Diversity

This index measures taxa abundance and evenness in the sub-sample by dividing the # of individuals in a taxa by the total # of individuals in the sub-sample and then multiplying by the natural logarithm of this proportion. This is done for all taxa in the sub-sample; the products are then summed and the answer multiplied by -1.

$$= -1\left(\sum_{i=1}^{Rich}[(n_i/N)\ \ln(n_i/N)]\right)$$

where n_i = the number of individuals in each taxon (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).

Saw Creek	Wiconisco Creek
TaxaRich = 26	TaxaRich = 12
P = 217 (sum the 'Number of Individuals' column in Tables 2 and 3)	P = 201

p_i = this value is listed in the above tables in the Number of Individuals column.

Saw Creek

(109/217) ln (109/217) + (8/217) ln (8/217) + (16/217) ln (16/217)……(1/217) ln (1/217) = -2.12946 * -1 = **2.12946**

Wiconisco Creek

(151/201) ln (151/201) + (1/201) ln (1/201) + (1/201) ln (1/201)……(1/201) ln (1/201) = -0.875322793 * -1 = **0.87532**

Number of Caddisfly Taxa

To calculate this metric, sum the number of Caddisfly taxa present in the sub-sample.

Saw Creek	Wiconisco Creek
Trichoptera = 9	Trichoptera = 1

Number of Mayfly Taxa

Sum the total number of Mayfly taxa identified in the sub-sample.

Saw Creek	Wiconisco Creek
Ephemeroptera = 2	Ephemeroptera = 0

INDEX OF BIOTIC INTEGRITY (IBI) SCORE

Now that the six metric scores have been calculated, the scores are plugged into the normalized metric score equation: (Observed Value / 95th percentile) x 100. Some metrics may have a normalized score greater than 100 because normalization is based on the 95th percentile values of the statewide dataset. Normalized metric scores above 100 are adjusted to a score of 100. The adjusted metric scores for the six metrics are summed and then averaged to give the Total Biological Score. Tables 4 and 5 below show how to calculate the normalized metric scores and Total Biological Scores for Saw Creek and Wiconisco Creek.

Saw Creek's Raw Metric Scores	Wiconisco Creek's Raw Metric Score
EPT = 15	EPT = 2
Taxa Richness = 26	Taxa Richness = 12
Beck4 = 21	Beck4 = 4
Shannon Diversity = 2.12946	Shannon Diversity = 0.87532
# Of Caddisfly Taxa = 9	# Of Caddisfly Taxa = 1
# Of Mayfly Taxa = 2	# Of Mayfly Taxa = 0

Table 4. Total Biological Score Calculation for Saw Creek

Metric	Equation	Observed Value	Normalized Metric Score	Adjusted Metric Score (100 Max)
EPT	(Observed / **17**) x 100	15	88.2	88.2
Taxa Richness	(Observed / **31**) x 100	26	83.9	83.9
Beck4	(Observed / **22**) x 100	21	95.5	95.5
Shannon Diversity	(Observed / **2.43**) x 100	2.13	87.6	87.6
# Of Caddisfly Taxa	(Observed / **11**) x 100	9	81.8	81.8
# Of Mayfly Taxa	(Observed / **6**) x 100	2	33.3	33.3
Total Biological Score (IBI)				**78.4**

Table 5. Total Biological Score Calculation for Wiconisco Creek

Metric	Equation	Observed Value	Normalized Metric Score	Adjusted Metric Score (100 Max)
EPT	(Observed / **17**) x 100	2	11.8	11.8
Taxa Richness	(Observed / **31**) x 100	12	38.7	38.7
Beck4	(Observed / **22**) x 100	4	18.2	18.2
Shannon Diversity	(Observed / **2.43**) x 100	0.88	36.2	36.2
# Of Caddisfly Taxa	(Observed / **11**) x 100	1	9.1	9.1
# Of Mayfly Taxa	(Observed / **6**) x 100	0	0	0
Total Biological Score (IBI)				**19.0**

AQUATIC LIFE USE BENCHMARK

Aquatic life use attainment status of a given sample reach is determined by comparing its Total Biological Score to a use attainment benchmark. If the Total Biological Score

of the sample reach is less than the benchmark score, the sample reach is impaired. If the score is greater than or equal to the benchmark the sample reach is attaining.

Table 6. ALU Benchmark for Low-Gradient Streams

Multihabitat ALU Benchmark
55 (10th percentile)

Therefore, Saw Creek would be documented as attaining its aquatic life use and Wiconisco Creek would be impaired for aquatic life use.

TEMPORAL PRECISION ESTIMATE

The temporal precision is calculated using the 90% confidence interval and is typically used to show confidence around a change in the biological condition of a site. Available for this calculation were 25 temporally paired samples collected at 12 sites between 2003-2010. The 90% confidence interval was 13.2, indicating that measured changes in index score of 14 or greater are not likely due to natural variation.

CONCLUSION

In 2007, Pennsylvania Department of Environmental Protection finalized a macroinvertebrate bioassessment protocol for assessing Pennsylvania's low-gradient streams. Using the field and laboratory methods outlined in DEP's protocol, a macroinvertebrate taxonomic list is produced. The taxonomic data is then used to calculate metrics and produce a total biological (IBI) score that accurately reflects the ecological conditions of the waterway. This IBI score is compared to the ALU attainment benchmark of 55 to determine if the sample reach is attaining or impaired.

LITERATURE CITED

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. U.S. Environmental Protection Agency; Office of Water; Washington, D.C.

McGarrell, C. 2007. An Index of Biotic Integrity for Pennsylvania Multihabitat Streams. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

Shull, D. R., and M. J. Lookenbill (editors). 2018. Water Quality Monitoring Protocols for Streams and Rivers. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania

SEMI-WADEABLE LARGE RIVER MACROINVERTEBRATE ASSESSMENT METHOD

Prepared by:

Dustin Shull
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2018

INTRODUCTION

Assessment of ALU in large semi-wadeable rivers can be a complex process. To appropriately assess biological communities in large rivers and to increase the efficiency of ALU assessments, DEP separates large rivers into two categories: semi-wadeable and non-wadeable. This assessment method is designed for semi-wadeable rivers within the Commonwealth. Semi-wadeable rivers are defined as predominantly free-flowing systems with drainage areas >1,000 mi^2, and have physical characteristics that allow for riffle and run sections to occur with relative frequency. These river systems tend to lack a well-defined and navigable U-shaped channel for any significant distance and frequently present difficulties for both wadeable and non-wadeable macroinvertebrate data collection methodologies. Well over half of the large rivers within the Commonwealth are considered semi-wadeable (Figure 1). Several studies have shown that semi-wadeable rivers can express substantial and reliable differences in water quality across their width for great distances. These chemical and physical differences drive variations observed in the macroinvertebrate communities that inhabit these regions (Guild et al. 2014, DEP 2014, Shull 2017). The water quality differences across the width of large semi-wadeable rivers are usually the result of major tributary inputs that do not mix. Additionally, each major tributary input is driven by both the natural and anthropogenic influences within the respective basin.

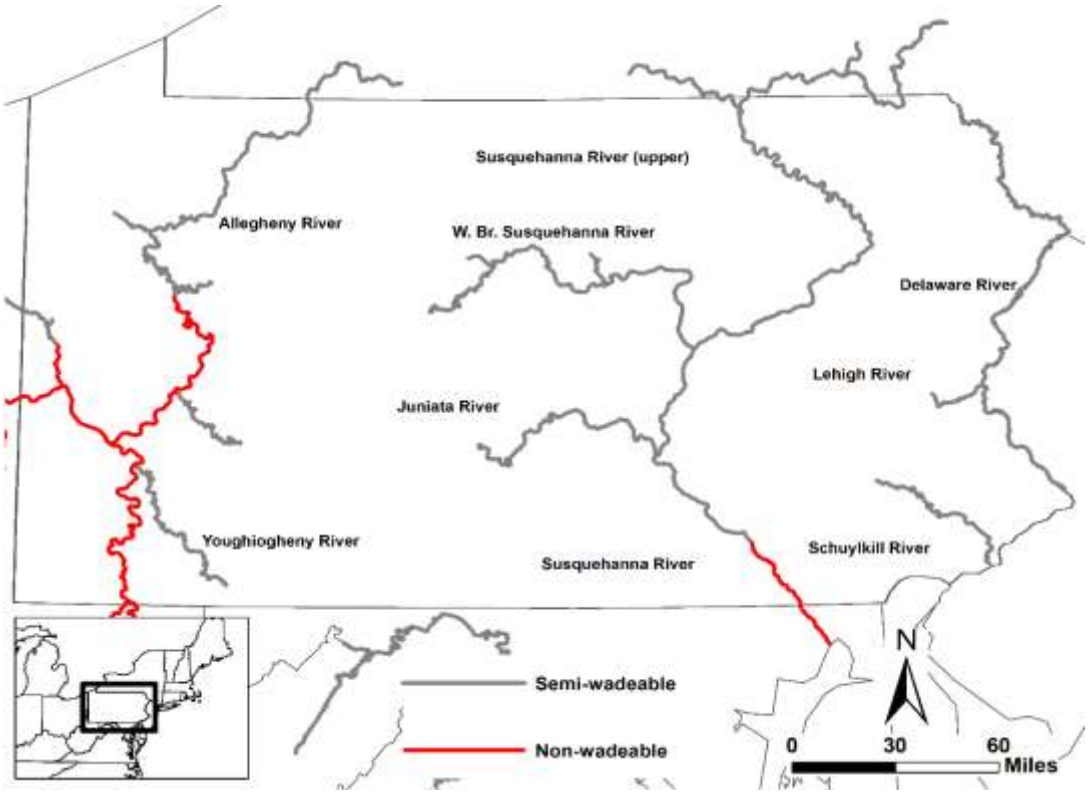


Figure 1. Large Rivers that are semi-wadeable and non-wadeable rivers throughout the Commonwealth. Assessment determinations will be made for semi-wadeable rivers using this assessment method. DEP continues to develop assessment methods for non-wadeable large rivers.

No other large river biological assessment tool has sought to understand and deal with these chemical, physical, and biological differences at one location on a river separately. Yet, many large river collection methods have been created to capture and composite these variables into one measure; thereby, accounting for, but not giving heed to these important differential aspects (Applegate et al. 2007, Wessell et al. 2008, Blocksom and Johnson 2009, Weigel and Dimick 2011). Final assessments using these tools average or generalize biological condition to provide valuable assessment information, but they do not consider potentially important details in the environment. This effectively obscures the ability to account for biological community degradation within large and important zones on each river. It also reduces the ability to track major sources of impacts driving degradation. Even more problematic are the large river biological collection methods that only collect data along the shoreline of a large semi-wadeable river (Merritt et al. 2005, Angradi 2006, USEPA 2013). These methods are particularly questionable when making large scale inferences about water quality conditions, because shoreline habitats are likely affected by minor tributary influences and point source discharges that follow the shoreline in semi-wadeable rivers (DEP 2014, Shull and Pulket 2015). Consequently, this assessment method does not use shoreline collection methods in large semi-wadeable rivers when making large scale assessment determinations. DEP spent several years developing and refining the transect collection method, and because of this method, each semi-wadeable multimetric index (SWMMI) can not only be used to make assessment determinations that are reflective of overall water quality, but also produce results that retain the unique aspects of water quality variations. This should greatly improve the validity of each assessment on large semi-wadeable rivers, as well as provide important source tracking information for future restoration efforts, if needed.

The goal of this document is to lay the framework for how DEP intends on making ALU assessment determinations in large semi-wadeable rivers. The semi-wadeable large river technical report (Shull 2017) goes into much detail about evaluating the complexity of large semi-wadeable rivers and how these assessment tools were developed to compensate. Making accurate and defensible assessment decisions requires both a sufficient number of data types (e.g., physical, chemical and biological) and a specificity of those data within a particular water influence (zone) – if needed – and season. Ultimately, ALU assessment determinations will be rather straightforward and similar to wadeable stream assessments if data can only be collected in one season and water influences are well mixed. However, ALU assessment determinations when water influences are not well mixed and when data are collected during both the summer and fall will require additional evaluation and discussion. To use this method for assessment purposes data collection must follow the protocols established in the Monitoring Book (Shull and Lookenbill 2017).

Each reach of river is assessed by the macroinvertebrate collection site immediately downstream. The length of each assessed reach is then determined by where the next potential impact to water quality exists upstream (i.e., major tributary or developed area). Therefore, the location of each upstream macroinvertebrate collection site should reflect this pattern. More explicitly, each macroinvertebrate collection site along the longitudinal gradient is determined by several factors including, where sufficient riffle-

run habitat exists, where changes in physiographic and demographic characteristics occur, and where additional major tributaries enter the system. Ideally, macroinvertebrate collection sites will occur at every viable riffle-run habitat, but at the very least, it is necessary to bracket major potential impacts to each system such as a major tributary or change in land use. In the example provided below (Figure 2), two semi-wadeable rivers converge to form another semi-wadeable river. Below this confluence, multiple water quality transects show that water influences do not mix so the non-mixing water influences were mapped. Hence, site 1 requires two unique 6D-200 samples composited completely within the delineated zone of each water influence. Additional macroinvertebrate collections sites (sites 2 – 5) are added upstream of the confluence to characterize each major water influence and to bracket the demographic characteristics across the drainage (e.g., communities, other land use transitions). It is important to note that water quality transect sites – used to delineate the area of specific water influences – can be collected at a higher frequency of locations than macroinvertebrate collections.

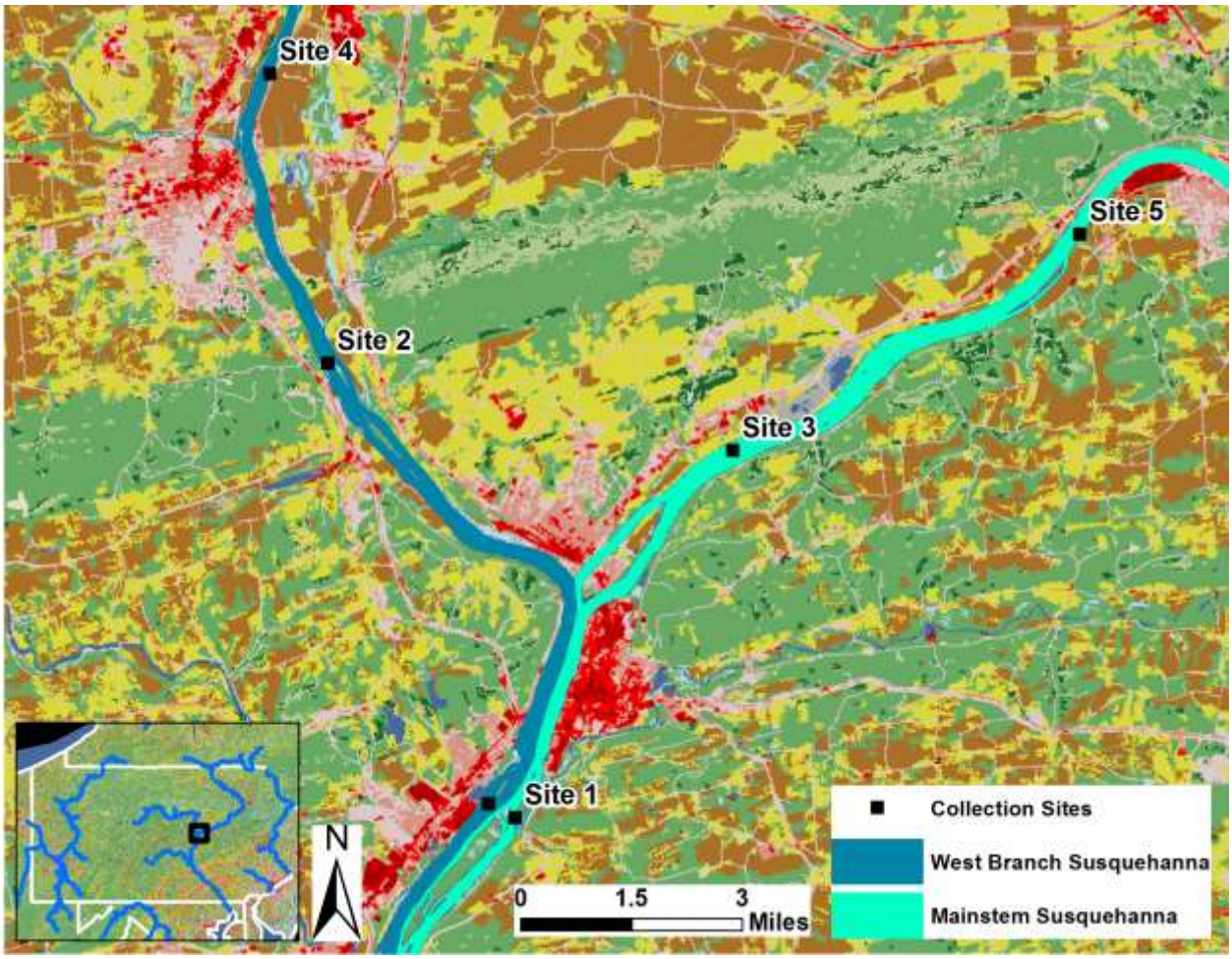


Figure 2. Macroinvertebrate collection sites on a large semi-wadeable river. Site locations were selected to bracket major land use changes and tributary inputs.

Habitat assessments are required with each semi-wadable macroinvertebrate sample. The DEP habitat data sheet for high gradient streams is used, which has undergone several iterations from Plafkin et al. (1989). This habitat evaluation uses a 12 parameter – 20-point scoring method. Currently, it is recommended that all 12 parameters are recorded when conducting habitat assessments in a semi-wadeable river. Although, instream parameters such as instream cover, epifaunal substrate, and embeddedness are the most reliable habitat indicators for large semi-wadeable rivers. Instream cover evaluates the percent makeup of the substrate (boulders, cobble, other rock material) and submerged objects (logs, undercut banks) that provide refuge for fish. Epifaunal substrate evaluates riffle quality, i.e. areal extent relative to stream width and dominant substrate materials that are present. Embeddedness estimates the percent (vertical depth) of the substrate interstitial spaces filled with fine sediments. These three instream habitat measurements can be summed to provide a possible range of 0 (indicating worst possible instream conditions) to 60 (indicating best possible instream conditions) points at each sampling site. Instream habitat totals that fall below 30 points may be an indication of poor physical habitat conditions. The other parameters in the habitat assessment are also useful for informational purposes, but tend to become difficult to measure as river size increases.

SWMMI CALCULATION AND PRECISION

The assessment method development process (Shull 2017) identified two different macroinvertebrate communities existing in large semi-wadeable rivers between the summer and fall seasons. The macroinvertebrate communities were shown to be different enough to justify creating two independent assessment tools for semi-wadeable rivers. In addition, October sampling is not recommended if the intent is to make ALU assessment determinations as this is a critical transition period for the macroinvertebrate communities. Examples for each SWMMI (Summer and Fall) are provided to show the metric and index calculation process step-by-step. The summer and fall SWMMI calculations are separated into their respective sections for clarity.

Many different metric combinations were evaluated during method development. Each SWWMI had six metrics selected for inclusion into the final index. All metrics, which are further defined and described in Shull (2017) exhibited a strong ability to distinguish between relatively unimpacted and heavily impacted conditions. In addition, these metrics measure different aspects of the benthic macroinvertebrate communities, but when used together in an index, they provide a solid foundation for assessing the biological condition of benthic macroinvertebrate communities in large semi-wadeable rivers. A complete list of taxa and their attributes is proved in Appendix B of Shull (2017).

Summer SWMMI

The following summer sample was collected in the Delaware River on September 9, 2016 and is used in the metric calculation and index standardization example below:

Taxa Name	Number of Individuals
Acroneuria	1
Agnetina	2
Baetisca	1
Brachycentrus	1
Cheumatopsyche	14
Chimarra	7
Chironomidae	10
Corbiculidae	8
Helicopsyche	14
Hydrobiidae	13
Hydropsyche	7
Isonychia	11
Lepidostoma	2
Leucrocuta	8
Maccaffertium	16
Micrasema	1
Oecetis	2
Oligochaeta	7
Optioservus	25
Physidae	1
Plauditus	7
Stenelmis	14
Teloganopsis	22
Tricorythodes	3

Percent Tolerant Individuals using Biological Condition Gradient (BCG) attribute 5 (BCGpct5)

$$= \left(\sum n_{indvBCG5}/N\right) * 100$$

Where $n_{indvBCG5}$ is the number of individuals in the subsample with a BCG value of 5, and N is the total number of individuals in the subsample.

Taxa Name	Number of Individuals	BCG
Acroneuria	1	3
Agnetina	2	3
Baetisca	1	2
Brachycentrus	1	3
Cheumatopsyche	14	5
Chimarra	7	4
Chironomidae	10	5
Corbiculidae	8	5
Helicopsyche	14	3
Hydrobiidae	13	4
Hydropsyche	7	5
Isonychia	11	3
Lepidostoma	2	2
Leucrocuta	8	3
Maccaffertium	16	3
Micrasema	1	3
Oecetis	2	3
Oligochaeta	7	5
Optioservus	25	4
Physidae	1	5
Plauditus	7	
Stenelmis	14	5
Teloganopsis	22	3
Tricorythodes	3	5

There are 64 individuals with a BCG of 5, and a total of 197 individuals in the subsample.

$(64/197)*100 = 32.5\%$

Percent Intolerant Individuals using Pollution Tolerance Value (PTV) attributes 0-3 (PTVpct03)

$$= \left(\sum_{i=0}^{3} n_{indvPTVi} / N \right) * 100$$

Where $n_{indvPTVi}$ is the number of individuals in a sub-sample with PTV of i, and N = the total number of individuals in the subsample.

Taxa Name	Number of Individuals	PTV
Acroneuria	1	0
Agnetina	2	2
Baetisca	1	4
Brachycentrus	1	1
Cheumatopsyche	14	6
Chimarra	7	4
Chironomidae	10	6
Corbiculidae	8	4
Helicopsyche	14	3
Hydrobiidae	13	8
Hydropsyche	7	5
Isonychia	11	3
Lepidostoma	2	1
Leucrocuta	8	1
Maccaffertium	16	3
Micrasema	1	2
Oecetis	2	8
Oligochaeta	7	10
Optioservus	25	4
Physidae	1	8
Plauditus	7	4
Stenelmis	14	5
Teloganopsis	22	2
Tricorythodes	3	4

There are 78 individuals with a PTV value of 0-3, and a total of 197 individuals in the subsample.

(78/197)*100 = 39.6%

Hilsenhoff Index using BCG attributes (BCGindex2)

$$= \sum_{i=1}^{6} [(i * n_{indvBCGi})] \Big/ N_{BCG}$$

Where $n_{indvBCGi}$ is the number of individuals in a sub-sample with a BCG of i, and N_{BCG} is the total number of individuals with BCG values in the subsample.

Taxa Name	Number of Individuals	BCG
Acroneuria	1	3
Agnetina	2	3
Baetisca	1	2
Brachycentrus	1	3
Cheumatopsyche	14	5
Chimarra	7	4
Chironomidae	10	5
Corbiculidae	8	5
Helicopsyche	14	3
Hydrobiidae	13	4
Hydropsyche	7	5
Isonychia	11	3
Lepidostoma	2	2
Leucrocuta	8	3
Maccaffertium	16	3
Micrasema	1	3
Oecetis	2	3
Oligochaeta	7	5
Optioservus	25	4
Physidae	1	5
Plauditus	7	
Stenelmis	14	5
Teloganopsis	22	3
Tricorythodes	3	5

There are 0 individuals with a BCG of 1, 3 with a BCG of 2, 78 with a BCG of 3, 45 with a BCG of 4, 64 with a BCG of 5, 0 with a BCG of 6, and a total of 190 BCG individuals in the subsample.

$[(1 * 0) + (2 * 3) + (3 * 78) + (4 * 45) + (5 * 64) + (6 * 0)]/190 = 3.89$

Percent Dominant Taxon (pctDOM)

$$= \left(\sum n_{indvDOM}/N\right) * 100$$

Where $n_{indvDOM}$ is the number of individuals of the dominant taxon in the subsample, and N is the total number of individuals in the subsample.

Taxa Name	Number of Individuals
Acroneuria	1
Agnetina	2
Baetisca	1
Brachycentrus	1
Cheumatopsyche	14
Chimarra	7
Chironomidae	10
Corbiculidae	8
Helicopsyche	14
Hydrobiidae	13
Hydropsyche	7
Isonychia	11
Lepidostoma	2
Leucrocuta	8
Maccaffertium	16
Micrasema	1
Oecetis	2
Oligochaeta	7
Optioservus	25
Physidae	1
Plauditus	7
Stenelmis	14
Teloganopsis	22
Tricorythodes	3

There are 25 individuals of the dominant taxon, *Optioservus* spp., and a total of 197 individuals in the subsample.

(25/197)*100 = 12.7%

Percent Ephemeroptera using BCG attributes 1-3 (pctEbcg13)

$$= \left(\sum_{i=1}^{3} n_{EphemBCGi} / N \right) * 100$$

Where $n_{EphemBCGi}$ is the number of Ephemeroptera individuals in a sub-sample with BCG of i, and N = the total number of individuals in the subsample.

Taxa Name	Number of Individuals	BCG
Acroneuria	1	3
Agnetina	2	3
Baetisca	1	2
Brachycentrus	1	3
Cheumatopsyche	14	5
Chimarra	7	4
Chironomidae	10	5
Corbiculidae	8	5
Helicopsyche	14	3
Hydrobiidae	13	4
Hydropsyche	7	5
Isonychia	11	3
Lepidostoma	2	2
Leucrocuta	8	3
Maccaffertium	16	3
Micrasema	1	3
Oecetis	2	3
Oligochaeta	7	5
Optioservus	25	4
Physidae	1	5
Plauditus	7	
Stenelmis	14	5
Teloganopsis	22	3
Tricorythodes	3	5

There are 58 Ephemeroptera individuals with BCG values of 1-3, and a total of 197 individuals in the subsample.

(58/197)*100 = 29.4%

Richness of Sensitive Ephemeroptera, Plecoptera, and Trichoptera taxa using BCG attributes 1-3 (richEPTbcg)

$$= n_{taxaEPTbcg}$$

Where $n_{taxaEPTbcg}$ is the number of taxa belonging to the orders Ephemeroptera, Plecoptera, and Trichoptera that have BCG attributes of 1-3.

Taxa Name	BCG
Acroneuria	3
Agnetina	3
Baetisca	2
Brachycentrus	3
Cheumatopsyche	5
Chimarra	4
Chironomidae	5
Corbiculidae	5
Helicopsyche	3
Hydrobiidae	4
Hydropsyche	5
Isonychia	3
Lepidostoma	2
Leucrocuta	3
Maccaffertium	3
Micrasema	3
Oecetis	3
Oligochaeta	5
Optioservus	4
Physidae	5
Plauditus	
Stenelmis	5
Teloganopsis	3
Tricorythodes	5

There are 5 Ephemeroptera taxa with BCG attributes of 1-3, 2 Plecoptera taxa with BCG attributes of 1-3, and 5 Trichoptera taxa with BCG attributes of 1-3.

5 + 2 + 5 = 12

Metric Standardization and Index Calculation

Final ceiling and floor standardization values are needed to standardize each metric. All standardized metrics are then multiplied by 100 to get the metric standardized score, and the score must range between 0 and 100. Final adjusted metrics scores are then averaged to get a final Summer SWMMI score on a 0 to100 scale.

Summer Metric Standardization Values

Metric	Floor Standardization (5^{th} percentile)	Ceiling Standardization (95^{th} percentile)
BCGpct5	28.5	80.6
PTVpct03	2.3	50.6
BCGindex2	3.76	4.76
pctDOM	14.4	46.8
pctEbcg13	0.4	49.7
richEPTbcg	1	10

For metrics like PTVpct03, pctEbcg13, and richEPTbcg (negative-response metrics), standardizations are calculated using the following equation:

(observed value - floor) / (ceiling - floor) * 100.

For metrics like BCGpct5, BCGindex2, and pctDOM (positive-response metrics) standardizations are calculated using the following equation:

(ceiling - observed value) / (ceiling - floor) * 100.

It is important to note that if a metric standardization score is < 0 then the score is set to 0, and if the metric standardization score is > 100 then the score is set to 100. This process creates the adjusted standardized metric score.

Metric / SWMMI	Observed Value	Standardized Metric Score	Adjusted Standardized Metric Score
BCGpct5	32.5	92.3	92.3
PTVpct03	39.6	77.7	77.7
BCGindex2	3.89	86.7	86.7
pctDOM	12.7	105.2	100
pctEbcg13	29.4	59.1	59.1
richEPTbcg	12	122.2	100
Summer SWMMI	--	--	**86.0**

Summer Precision Estimates

Summer SWMMI methodological precision is calculated using the coefficient of variation intrasite replicate samples (samples collected at the same site on the same day). The

summer SWMMI intrasite precision estimate was 8.8%, which was well below recommended limits (10 -15%, Stribling et al. 2008), indicated the summer SWWMI is a precise assessment tool. The summer SWMMI temporal precision is calculated using the 90% confidence interval and is typically used to show confidence around a change in biological condition at a site. The temporal precision estimate for the summer SWMMI using all available samples was 14.7, indicating that measured changes in index score of 15 or greater are not likely due to natural variation.

Fall SWMMI

The following fall sample was collected in the Delaware River on December 16, 2015 and is used in the metric calculation and index standardization example below:

Taxa Name	Number of Individuals
Acroneuria	5
Cheumatopsyche	3
Chimarra	2
Chironomidae	65
Cultus	1
Epeorus	9
Ephemerella	53
Helopicus	2
Hydropsyche	15
Isonychia	3
Lepidostoma	5
Leucrocuta	5
Maccaffertium	28
Nematoda	1
Neophylax	1
Oligochaeta	4
Ophiogomphus	2
Optioservus	15
Oulimnius	1
Paraleptophlebia	2
Psephenus	1
Rhyacophila	2
Stenacron	1
Stenelmis	4
Taeniopteryx	1
Teloganopsis	6

Beck's Index using PTV attributes 0-2 (PTVBeck3)

$$= 3 * (n_{taxaPTV0}) + 2 * (n_{taxaPTV1}) + 1 * (n_{taxaPTV2})$$

Where $n_{taxaPTV0}$ is the number of taxa with a PTV attribute of 0, $n_{taxaPTV1}$ is the number of taxa with a PTV attribute of 1, and $n_{taxaPTV2}$ is the number of taxa with a PTV attribute of 2.

Taxa Name	PTV
Acroneuria	0
Cheumatopsyche	6
Chimarra	4
Chironomidae	6
Cultus	2
Epeorus	0
Ephemerella	1
Helopicus	2
Hydropsyche	5
Isonychia	3
Lepidostoma	1
Leucrocuta	1
Maccaffertium	3
Nematoda	9
Neophylax	3
Oligochaeta	10
Ophiogomphus	1
Optioservus	4
Oulimnius	5
Paraleptophlebia	1
Psephenus	4
Rhyacophila	1
Stenacron	4
Stenelmis	5
Taeniopteryx	2
Teloganopsis	2

There are 2 taxa with PTV attributes of 0, 6 taxa with PTV attributes of 1, and 4 taxa with PTV attributes of 2.

3 * (2) + 2 * (6) + 1 * (4) = 22

Richness of Sensitive Ephemeroptera, Plecoptera, and Trichoptera taxa using PTV attributes 0-4 (richEPTptv)

$$= n_{taxaEPTptv}$$

Where $n_{taxaEPTptv}$ is the number of taxa belonging to the orders Ephemeroptera, Plecoptera, and Trichoptera that have PTV attributes of 0-4.

Taxa Name	PTV
Acroneuria	0
Cheumatopsyche	6
Chimarra	4
Chironomidae	6
Cultus	2
Epeorus	0
Ephemerella	1
Helopicus	2
Hydropsyche	5
Isonychia	3
Lepidostoma	1
Leucrocuta	1
Maccaffertium	3
Nematoda	9
Neophylax	3
Oligochaeta	10
Ophiogomphus	1
Optioservus	4
Oulimnius	5
Paraleptophlebia	1
Psephenus	4
Rhyacophila	1
Stenacron	4
Stenelmis	5
Taeniopteryx	2
Teloganopsis	2

There are 8 Ephemeroptera taxa with PTV attributes of 0-4, 4 Plecoptera taxa with PTV attributes of 0-4, and 4 Trichoptera taxa with PTV attributes of 0-4.

8 + 4 + 4 = 16

Percent Intolerant Individuals using PTV attributes 0-3 (PTVpct03)

$$= \left(\sum_{i=0}^{3} n_{indvPTVi} / N \right) * 100$$

Where $n_{indvPTVi}$ is the number of individuals in a sub-sample with PTV of i, and N = the total number of individuals in the subsample.

Taxa Name	Number of Individuals	PTV
Acroneuria	5	0
Cheumatopsyche	3	6
Chimarra	2	4
Chironomidae	65	6
Cultus	1	2
Epeorus	9	0
Ephemerella	53	1
Helopicus	2	2
Hydropsyche	15	5
Isonychia	3	3
Lepidostoma	5	1
Leucrocuta	5	1
Maccaffertium	28	3
Nematoda	1	9
Neophylax	1	3
Oligochaeta	4	10
Ophiogomphus	2	1
Optioservus	15	4
Oulimnius	1	5
Paraleptophlebia	2	1
Psephenus	1	4
Rhyacophila	2	1
Stenacron	1	4
Stenelmis	4	5
Taeniopteryx	1	2
Teloganopsis	6	2

There are 125 individuals with a PTV value of 0-3, and a total of 237 individuals in the subsample.

(125/237)*100 = 52.7%

Percent Ephemeroptera using BCG attributes 1-3 (pctEbcg13)

$$= \left(\sum_{i=1}^{3} n_{EphemBCGi} / N \right) * 100$$

Where $n_{EphemBCGi}$ is the number of Ephemeroptera individuals in a sub-sample with BCG of i, and N = the total number of individuals in the subsample.

Taxa Name	Number of Individuals	BCG
Acroneuria	5	3
Cheumatopsyche	3	5
Chimarra	2	4
Chironomidae	65	5
Cultus	1	1
Epeorus	9	2
Ephemerella	53	2
Helopicus	2	3
Hydropsyche	15	5
Isonychia	3	3
Lepidostoma	5	2
Leucrocuta	5	3
Maccaffertium	28	3
Nematoda	1	
Neophylax	1	3
Oligochaeta	4	5
Ophiogomphus	2	3
Optioservus	15	4
Oulimnius	1	2
Paraleptophlebia	2	2
Psephenus	1	4
Rhyacophila	2	2
Stenacron	1	4
Stenelmis	4	5
Taeniopteryx	1	3
Teloganopsis	6	3

There are 106 Ephemeroptera individuals with BCG values of 1-3, and a total of 237 individuals in the subsample.

(106/237)*100 = 44.7%

Total Taxa Richness (Richness)

$$= n_{taxa}$$

Where n_{taxa} is the total number of taxa in the subsample.

Taxa Name
Acroneuria
Cheumatopsyche
Chimarra
Chironomidae
Cultus
Epeorus
Ephemerella
Helopicus
Hydropsyche
Isonychia
Lepidostoma
Leucrocuta
Maccaffertium
Nematoda
Neophylax
Oligochaeta
Ophiogomphus
Optioservus
Oulimnius
Paraleptophlebia
Psephenus
Rhyacophila
Stenacron
Stenelmis
Taeniopteryx
Teloganopsis

There are 26 taxa in the subsample.

Richness of taxa in the Functional Feeding Group (FFG) Scrapers (FFGrichSC)

$$= n_{sctaxa}$$

Where n_{sctaxa} is the number of scraper taxa.

Taxa Name	Number of Individuals	FFG
Acroneuria	5	PR
Cheumatopsyche	3	FC
Chimarra	2	FC
Chironomidae	65	CG
Cultus	1	PR
Epeorus	9	SC
Ephemerella	53	CG
Helopicus	2	PR
Hydropsyche	15	FC
Isonychia	3	CG
Lepidostoma	5	SH
Leucrocuta	5	SC
Maccaffertium	28	SC
Nematoda	1	CG
Neophylax	1	SC
Oligochaeta	4	CG
Ophiogomphus	2	PR
Optioservus	15	SC
Oulimnius	1	SC
Paraleptophlebia	2	CG
Psephenus	1	SC
Rhyacophila	2	PR
Stenacron	1	SC
Stenelmis	4	SC
Taeniopteryx	1	SH
Teloganopsis	6	CG

There are 9 scraper taxa in the subsample.

Metric Standardization and Index Calculation

Final ceiling and floor standardization values are needed to standardize each metric. All standardized metrics are then multiplied by 100 to get the metric standardized score, and the score must range between 0 and 100. Final adjusted metrics scores are then averaged to get a final fall SWMMI score on a 0 to100 scale.

Fall Metric Standardization Values

Metric	Floor Standardization (5^{th} percentile)	Ceiling Standardization (95^{th} percentile)
PTVBeck3	2	15
richEPTptv	2	15
PTVpct03	3.3	65.3
pctEbcg13	0	62.3
Richness	11	27
FFGrichSC	2	10

For all fall metrics (negative-response metrics), standardizations are calculated using the following equation:

(observed value - floor) / (ceiling - floor) * 100.

It is important to note that if a metric standardization score is < 0 then the score is set to 0, and if the metric standardization score is > 100 then the score is set to 100. This process creates the adjusted standardized metric score.

Metric / SWMMI	Observed Value	Standardized Metric Score	Adjusted Standardized Metric Score
PTVBeck3	22	153.8	100
richEPTptv	16	107.7	100
PTVpct03	52.7	79.7	79.7
pctEbcg13	44.7	71.7	71.7
Richness	26	93.7	93.7
FFGrichSC	9	87.5	87.5
Fall SWMMI	--	--	**88.8**

Fall Precision Estimates

Fall SWMMI methodological precision is calculated using the coefficient of variation intrasite replicate samples (samples collected at the same site on the same day). The fall SWMMI intrasite precision estimate was 14.1%, which was within recommended limits (10 -15%, Stribling et al. 2008), indicating the fall SWWMI is a precise and repeatable assessment tool. The fall SWMMI temporal precision is calculated using the 90% confidence interval and is typically used to show confidence around a change in biological condition at a site. The temporal precision estimate for the fall SWMMI using

all available samples was 12.8, indicating that measured changes in index score of 13 or greater are not likely due to natural variation.

AQUATIC LIFE USE ASSESSMENTS

Both SWMMIs (summer and fall) are accurate and precise tools for making ALU assessment determinations in semi-wadeable rivers. Ideally, assessment in large rivers will understand and compensate for the complexity of the biological communities that exist in these rivers. This assessment tool is a substantial step toward that ideal situation. It is important to note that the transect method can produce multiple SWMMI results at any given location based on the number of major water influences discovered during transect data collection. To address this issue, DEP will use transect data to create zones within each river to be assessed independently, if needed. For example, if transect data shows that 3 unique water quality zones exist, then DEP will use the SWMMI to assess each zone independently. This determination will result in more accurate assessments on large semi-wadeable rivers without ignoring major impacts, or averaging major impacts with better conditions. This method also creates the ability to source track major impacts. Linking large river impacts to sources will inform more appropriate Total Maximum Daily Load (TMDL) and TMDL alternative solutions. In addition, the transect method specifically targets observed variations in water quality and measures biological conditions within those regions; therefore, SWMMI scores between defined zones across the width of a river should not be averaged.

The summer SWMMI impairment threshold is 49 and the fall SWMMI impairment threshold is 57. More information on the development of these impairment thresholds is found in the development report (Shull 2017). SWMMI scores below these thresholds will indicate impaired ALU. Each SWMMI (summer and fall) is independently applicable when making ALU determinations. This is based on USEPA guidance, which mandates that all biological communities DEP has assessment methods for must be evaluated on a stand-alone basis (USEPA 2002). Consequently, each SWMMI is functionally equivalent to having two completely different biological assessment tools (e.g., fish MMI and a macroinvertebrate MMI). Therefore, it is not appropriate to average both SWMMI scores to obtain an overall result. It is also not appropriate to favor the results of one SWMMI over the other. DEP will always strive to collect as much information as possible to make the most accurate assessment decisions. However, based on independent applicability, it is also understood that only one SWMMI (summer or fall) is required to make an ALU determination for a semi-wadeable river.

The following situation provides an example of this biological assessment rule. Multiple summer and fall samples were collected at the same site (Figure 3). Based on transect analysis the site had one homogeneous influence, so each macroinvertebrate sample was collected evenly across the entire width of the river during each visit. A total of five samples were collected; two samples during the summer and three samples during the fall. The summer samples consistently showed reduced, but attaining SWMMI scores, yet the fall samples resulted in impaired scores. The fall biological community was not

supporting the Aquatic Life Use; therefore, DEP would determine that this section of the river is impaired. It may be concluded from this example that one SWMMI is more sensitive than the other; however, that is not the case. Examination of the entire development dataset showed no preference for one SWMMI consistently selecting for one assessment decision when biological communities were close to thresholds.

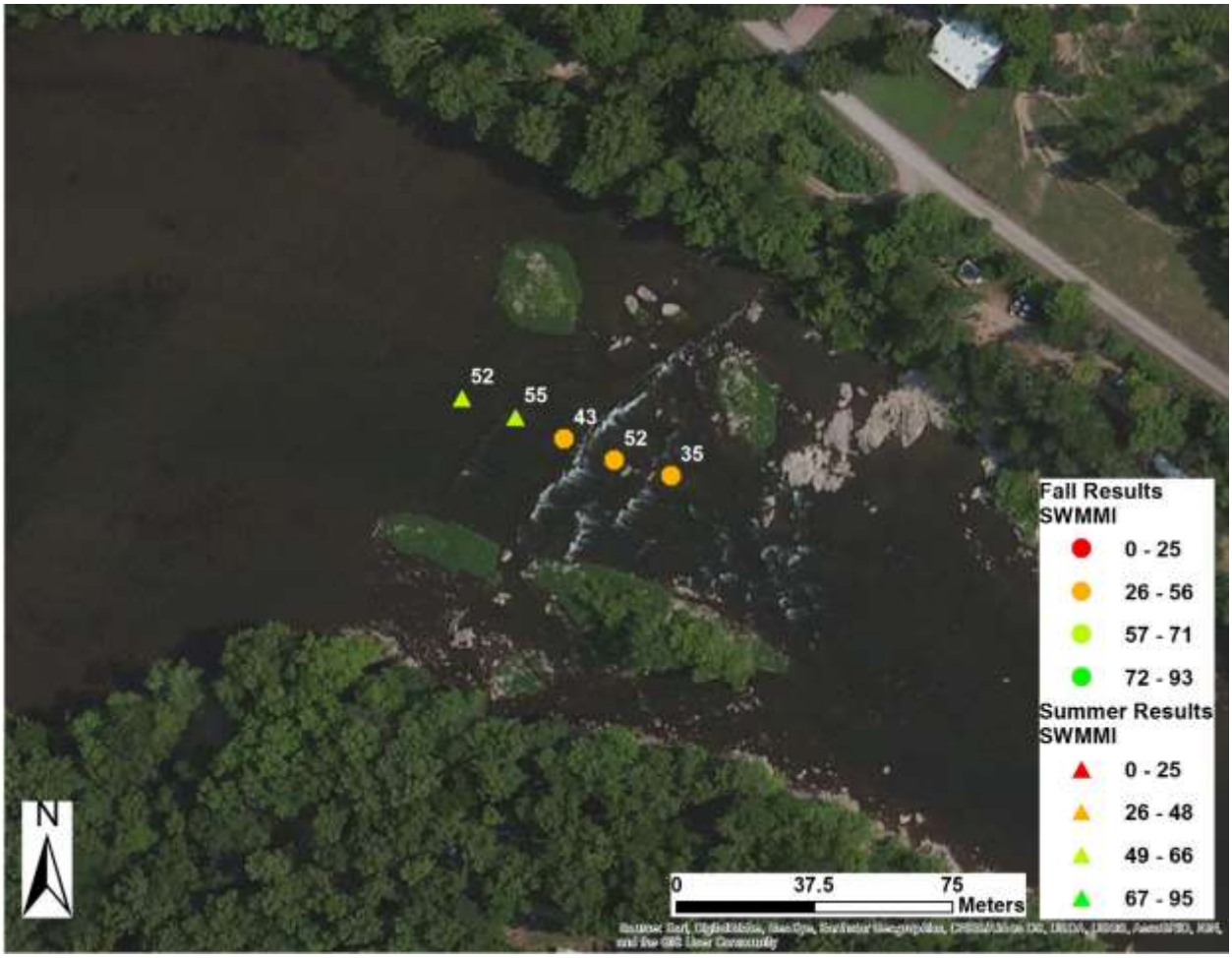


Figure 3. Multiple summer and fall SWMMI results over time at the same location on a semi-wadeable river. Location of points on the map do not indicate exact sample location; points were moved slightly to illustrate the results of sampling. Points are labeled with the respective SWMMI score.

ADDITIONAL APPLICATION CONSIDERATIONS

Data resulting from SWMMI scores may not be used in making ALU determinations in some situations. In fact, DEP uses the wadeable freestone riffle-run method developed by Chalfant (2012) for several other purposes, including, but not limited to cause and effect surveys and incremental improvement reports. These surveys can collect biological information in areas that are not appropriate for making ALU determinations. For example, two macroinvertebrate samples were collected on a semi-wadeable river

near a city in Pennsylvania, just downstream of a sewage treatment plant. In this example, the SWMMI results showed that a major portion of this semi-wadeable river (laterally) was being impacted by a facility, perhaps, not operating within permitted limits. Sampling locations specifically targeted one city's sewage treatment facility, but were not necessarily representative of river conditions in this area. Therefore, it would not be appropriate to use these results in making assessment decisions on this river. However, this example does illustrate the usefulness of the semi-wadeable biological collection method for other purposes. This example also illustrates the necessity to differentiate between ALU assessments and reports on local scale impacts. All ALU assessments on semi-wadeable rivers should examine the longitudinal scale that each macroinvertebrate sample represents. If a macroinvertebrate sample is determined to be more representative of a local scale impact, then consideration of appropriate compliance actions may be appropriate.

The SWMMIs may also be used to evaluate whether conditions are degrading or improving at a given site (e.g., trend analysis). It is important to note that this is a different type of analysis than making assessment determinations using an impairment threshold. Methodological error is already incorporated during the development of the impairment threshold, so using variability measurements as "gray areas" while making assessment determinations is not appropriate (Stribling et al. 2008). However, for analyses such as trend analysis, the temporal precision estimate can be used to decide whether a macroinvertebrate community changes over time. When SWMMI scores at the same site change over time beyond the temporal precision estimate, there is a high level of confidence that the biological community change was driven by human influences. The summer SWWMI temporal precision estimate for all sites (where repeat data were available) was 14.7 points, which suggests that observed score changes at a site over time of 15 points or more can be considered a change in condition. The fall SWWMI temporal precision estimate for all sites (where repeat data were available) was 12.8 points, which suggests that observed score changes at a site over time of 13 points or more can be considered a change in condition.

LITERATURE CITED

Angradi, T. R. (editor) 2006. Environmental monitoring and assessment program: Great River ecosystems field operations manual. EPA/620/R-06/002. Office of Research and Development, US Environmental Protection Agency, Washington, DC

Applegate, J. M., P. C. Baumann, E. B. Emery, and M. W. Wooten. 2007. First steps in developing a multimetric macroinvertebrate index for the Ohio River. River Research and Applications 23: 683–697.

Blocksom, K. A., and B. R. Johnson. 2009. Development of a regional macroinvertebrate index for large river bioassessment. Ecological Indicators 9: 313–328

Chalfant, B. 2012. A benthic index of biotic integrity for wadeable freestone streams in Pennsylvania. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania. (Available from:

http://files.dep.state.pa.us/Water/Drinking%20Water%20and%20Facility%20Regulation/WaterQualityPortalFiles/Technical%20Documentation/freestoneIBImarch2012.pdf)
DEP. 2014. 2012-13 Susquehanna River Sampling and Assessment Report. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania. (Available from: http://files.dep.state.pa.us/Water/Drinking%20Water%20and%20Facility%20Regulation/WaterQualityPortalFiles/Water%20Quality%20Standards/Susquehanna_Report_2013_FinalDraft.pdf)
Guild, K., A. Anthony, M. Bilger, and J. Holt. 2014. Assessment of Passive and Active Macroinvertebrate collection methods in adjacent reaches on the upper main stem of the Susquehanna River. Journal of the Pennsylvania Academy of Science 88: 47–56.
Merritt, R. W, J. D. Allan, K. W. Cummins, K. J. Wessell, and J. O. Wilhelm. 2005. Qualitative biological and habitat protocols for Michigan's non-wadeable rivers. Michigan Department of Environmental Quality, Lansing, Michigan.
Plafkin, J.L, M.T. Barbour, K.D. Porter, S.K. Gross, and R.M. Hughes. 1989. Rapid Bioassessment Protocols for Use in Streams and Rivers: Benthic macroinvertebrates and fish. EPA/440/4-89-001. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.
Shull, D. R., and M. J. Lookenbill. 2017. Water Quality Monitoring Protocols for Streams and Rivers. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.
Shull, D. R., and M. Pulket. 2015. Causal Analysis of the Smallmouth Bass decline in the Susquehanna and Juniata Rivers. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania. (Available from: http://files.dep.state.pa.us/Water/Drinking%20Water%20and%20Facility%20Regulation/WaterQualityPortalFiles/SusquehannaRiverStudyUpdates/SMB_CADDIS_Report.pdf)
Shull, D. R. 2017. A Benthic Macroinvertebrate Multimetric Index for Large Semi-wadeable Rivers Technical Report. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.
Stribling, J. B., B. K. Jessup, and D. L. Feldman. 2008. Precision of benthic macroinvertebrate indicators of stream condition in Montana. Journal of the North American Benthological Society 27: 58–67.
USEPA. 2002. Consolidated assessment and listing methodology: Toward a compendium of best practices. U.S. Environmental Protection Agency, Office of Wetlands, Oceans, and Watersheds. Washington, DC
USEPA. 2013. National Rivers and Streams Assessment 2013-2014: Field Operations Manual – Non-Wadeable. EPA-841-B-12-009a. U.S. Environmental Protection Agency, Office of Water, Washington, DC.
Weigel, B. M. and J. J. Dimick. 2011. Development, validation, and application of a macroinvertebrate based index of biotic integrity for nonwadeable rivers of Wisconsin. Journal of the North American Benthological Society 30(3):665–679.

Wessell, K. J., R. W. Merritt, J. G. O. Wilhelm, J. D. Allan, K. W. Cummins, and D. G. Uzarski. 2008. Biological evaluation of Michigan's non-wadeable rivers using macroinvertebrates. Aquatic Ecosystem Health and Management 11: 335–351.

BACTERIOLOGICAL ASSESSMENT METHOD FOR WATER CONTACT

Prepared by:

Megan Bradburn and Shawn Miller
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2015

Edited by:

Shawn Miller
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2017

INTRODUCTION

DEP implements bacteriological data collection protocols and assessment methods to protect the WC use. Except for Lake Erie Coastal Beaches and waters specified with exceptions to the water contact use in 25 Pa. Code § 93.9 a-z, all waters of Pennsylvania are subject to the criterion for fecal coliform bacteria in § 93.7. The criterion specifies that during the swimming season (May 1–September 30), the maximum fecal coliform level will be a geometric mean of 200 cfu/100 ml based on 5 samples collected in a 30-day period. In addition, no more than 10% of samples collected in a 30-day period will exceed 400 cfu/100 ml. Coastal Beach samples are evaluated for water contact recreational use attainment according to the *E. coli* standard referenced in 25 Pa. Code § 93.9x (the drainage list for Lake Erie). The standard specifies that a bathing beach will be considered contaminated for bathing purposes when either a 30-day geometric mean in all water samples collected exceeds 126 cfu/100 mL or a single sample exceeds 235 cfu/100 ml. Surface waters assessment priorities are given to those waters where water contact recreation is most likely to occur. To use this method for assessment determination purposes data collection must follow applicable protocols established in the Monitoring Book (Shull and Lookenbill 2018).

DATA PROCESSING

A geometric mean is calculated for each sampling group (5 samples collected on different days in a 30-day period, spanning at least 14 days) at a site. Geometric means are calculated by taking the natural logarithm (ln) of each sample result and then averaging the logarithm values. This average is then converted back to a normal value by computing the antilog. The following example illustrates this process (Table 1).

Table 1. Example of 5 samples collected at a single site for assessment determination purposes.

Sample	Result (cfu/100 ml)	ln(Result)
1	130	4.868
2	380	5.940
3	240	5.481
4	100	4.605
5	180	5.193
	Mean of ln(Results)	5.217
	Antilog of Mean	**184** cfu/100 ml

ASSESSMENT DECISIONS

The primary focus of WC assessments is to list waters that are impaired due to chronic long-term water quality problems, and not acute or transitory situations. Hence, for accurate assessments, there should be at least 5-day samples taken within any 30-day period. Additionally, these 5 samples must span more than 14 days.

DEP will assess waters as impaired for WC if there is one violation of the 30-day geometric mean fecal coliform criterion. If there is one violation of the 30-day geometric mean of the *E. coli* criterion in the case of Coastal beaches during the bathing season, DEP will assess the waters as impaired. The criterion also indicates that, no more than 10% of the grab samples collected from waters attaining recreational use exceed 400cfu/100 ml. This 10% exceedance will apply to assessment decisions on a case-by-case basis and only when there is enough data to support an impairment decision. Generally, the geometric mean will be used to assess waterbodies because it is more relevant to the long-term quality of a waterbody.

The primary waterbodies of concern are streams and rivers that are larger than Strahler Order 2. These streams are most frequently used by the public for swimming and full body immersion and thus the potential to contract illness due to waterborne pathogens is much more likely than small headwater streams. This will not preclude these smaller waters from assessment, however they will not be the primary focus.

In the example above, the geometric mean of all 5 samples was 184, which is below the criterion for fecal coliforms (200 cfu/100 ml), and no single sample was above 400cfu/100 ml. For this reason, this stream segment would be considered attaining the WC use.

Exceptions to the 5-sample limit

In the 2006 Integrated Report guidance (USEPA 2005) USEPA stressed the importance of not setting minimum sample size for data sets to assess attainment of WQS. DEP will evaluate incomplete data sets (i.e. at least 3 samples in a 30-day period). DEP will use its best professional judgment to evaluate the incomplete data and where samples document consistently low bacteria levels within mostly forested watersheds, DEP will consider attainment of the criterion if no likely sources of fecal coliform bacteria are present in the watershed. Conversely, for incomplete data sets that consistently document high bacteria counts (>400cfu/100 ml), DEP will consider whether the waterbody is impaired for the water contact recreational use.

LITERATURE CITED

Shull, D. R., and M. J. Lookenbill (editors). 2018. Water Quality Monitoring Protocols for Streams and Rivers. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

USEPA. 2005. Guidance for 2006 Assessment, Listing and Reporting Requirements Pursuant to Sections 303(d), 305(b) and 314 of the Clean Water Act. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

FISH TISSUE COMSUMPTION ASSESSMENT METHOD

Prepared by:

Michael Lookenbill and Timothy Wertz
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2015

Edited by:

Timothy Wertz
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2017

INTRODUCTION

To protect the Fishing (F) use, DEP implements fish data collection protocols and fish tissue assessment methods. Priority is given to surface waters that are targeted by anglers or subsistence populations. In surface waters that do not contain fishable populations of organisms, it may not be possible to assess fish consumption. To use this method for assessment determinations data collection must follow applicable protocols established in the Monitoring Book (Shull and Lookenbill 2018).

The importance of the fish tissue sampling and advisory issuance program was fully recognized in May 1986 with the signing of an interagency agreement between the Department of Environmental Resources (now DEP), DOH, and Pennsylvania Fish Commission (now PFBC). This agreement was developed because “the agencies desire to pursue a systematic approach for the detection and evaluation of fish tissue contamination and to develop coordinated procedures for informing the public that may consume such fish of possible adverse health impacts.” It listed the responsibilities of each agency and provided for the “timely joint issuance of a health advisory” when fish tissue contamination constituted a health risk. The first joint advisory was issued in June 1986 and included a number of waters throughout Pennsylvania. A new agreement, signed in 2002, added the Pennsylvania Department of Agriculture (PDA) to the fish consumption advisory program and established a two-tiered system for advisory decisions and issuance. A Fish Consumption Advisory Policy Workgroup was established to oversee the program and make management decisions. This workgroup includes deputy secretaries from the three cabinet agencies and the Executive Director of the PFBC. The existing staff-level workgroup was renamed the Fish Consumption Advisory Technical Workgroup (FCATW) and includes representatives of all four agencies. The technical workgroup coordinates routine program activities, such as sampling site identification and provides recommendations for advisory issuance or lifting to the policy workgroup.

DATA REVIEW

The annual data review process begins in late spring when the DEP Bureau of Labs (BOL) has finished analyzing the samples collected from the previous year. An initial review of the data is conducted to screen for anomalous results based on previous data and expected results for a species, sample size (average length and weight), lipid percentage or particular waterbody. If anomalous data are encountered, the BOL is requested to either verify the result or reanalyze the sample using a backup aliquot of the parent tissue. Once the results are final, the data is evaluated and compared to current advisory triggers. All recent tissue contaminant data is evaluated to determine the possible need for an advisory for a particular waterbody and fish species. Sample results that exceed the 1 meal per week statewide advisory, but do not exceed the “Do Not Eat” threshold, are subject to a second verification sample before an advisory can

be issued or lifted. A “Do Not Eat” advisory is issued if a single representative sample result exceeds the appropriate “Do Not Eat” trigger. The possibility of lifting or reducing a “Do Not Eat” consumption advisory also requires a verification sample.

ADVISORY TRIGGERS

PCBs and Chlordane

Currently, Pennsylvania’s program includes a mixture of risk assessment-based methods and United States Food and Drug Administration (USFDA) Action Levels that are used as the basis for issuing or lifting advisories. Risk assessment methods form the basis for meal-specific advisories due to PCBs, mercury, and chlordane. Advisories for other compounds use USFDA Action levels to issue “Do Not Eat” advice. Trigger levels for PCBs and chlordane are shown in Table 1.

Table 1. Trigger levels for PCBs and chlordane found in fish tissue and subsequent meal recommendations.

GROUP	MEAL ADVICE	PCB (ppm)	CHLORDANE (ppm)
1	UNRESTRICTED	0 - 0.05	0 – 0.15
2	1 MEAL /WEEK (52 MEALS /YEAR)	0.06 - 0.2	0.16 – 0.65
3	1 MEAL/MONTH (12 MEALS/YEAR)	0.21 - 1.0	0.66 – 2.82
4	6 MEALS/YEAR	1.1 - 1.9	2.83 – 5.62
5	NO CONSUMPTION	>1.9	>5.62

PCB Meal-specific advisories based on this method were issued for Lake Erie and Presque Isle Bay for 1997, and it was applied statewide in 1998. **Pennsylvania issued a general, statewide advisory recommending that anglers eat no more than one meal per week of recreationally caught sport fish in April 2001. As a result, only Groups 3, 4 and 5 from Table 1 are now applicable.**

Mercury

Consumption advisories due to mercury in fish tissue are based on a health risk assessment developed by USEPA. The USEPA risk assessment was originally released in 1997. As a result of a request from Congress, USEPA contracted with the National Research Council (NRC) to review the risk assessment and prepare recommendations on the appropriate reference dose for mercury exposure. In July 2000, the NRC reported that the Reference Dose (RfD) for mercury, developed by

USEPA, was a scientifically justifiable level for the protection of public health. As a result of this finding, USEPA recommended that sensitive individuals should eat no more than one meal per week of sport-caught fish. The USFDA and USEPA currently post these federal recommendations online. As noted above, Pennsylvania has issued a statewide one meal per week advisory that mirrors this federal advice. Pennsylvania also issues more protective mercury advisories on a site-specific basis, using the USEPA risk assessment and advisory triggers slightly modified from those in a September 1999 USEPA fact sheet. The trigger levels and meal recommendations are outlined in Table 2. Because a statewide one meal per week advisory has been issued, site-specific mercury advice begins at two meals per month. Meal-specific advisories for mercury were first issued at the same time as the general, statewide advisory, April 11, 2001.

Table 2. Trigger levels for mercury found in fish tissue and subsequent meal recommendations.

MEAL ADVICE	MERCURY (ppm)
UNRESTRICTED	0 – 0.12
1 MEAL/WEEK	0.13 – 0.25
2 MEALS/MONTH	0.26 – 0.50
1 MEAL/MONTH	0.51 – 1.0
6 MEALS/YEAR	1.1 – 1.90
DO NOT EAT	>1.9

USFDA Action Levels

USFDA Action Levels are regulatory standards applicable to commercial fish and other foodstuffs. These Action Levels are developed based on general consumption patterns and may include consideration of economic issues such as potential loss of food supply. The USFDA has acknowledged that Action Levels may not adequately protect sensitive individuals or those individuals who may consume larger quantities of recreationally caught sport fish. The work group has been unable to completely evaluate risk assessment-based methods for these contaminants due to resource constraints. In addition, evaluation of risk assessment-based methods for most of these contaminants has not been a priority because they are normally found in very low concentrations in Pennsylvania fish. The compounds for which USFDA Action Levels constitute advisory triggers are listed in Table 3.

Table 3. USFDA Action Level triggers for a recommendation of Do Not Eat.

Contaminant	FDA Action Level
Aldrin and Dieldren (sum)	0.3 ppm
Chlordecone (Kepone)	0.3 ppm
DDT, DDE, and TDE (sum)	5.0 ppm
Heptachlor and Heptachlor Epoxide (sum)	0.3 ppm
Mirex	0.1 ppm

ADVISORY DECISIONS

For the evaluation of advisories that are more restrictive than the statewide advisories (i.e., one meal per week), DEP evaluates all readily available tissue contaminant data to prepare for a meeting of the FCATW where final advisory decisions will be made. This meeting is held annually in early summer. These data are compared to the applicable advisory triggers to determine the possible need for an advisory for a particular waterbody and a specific species. The possibility of lifting or modifying an advisory is also considered during this evaluation. Once the advisories are agreed upon at the workgroup level, the FCATW considers the most appropriate spatial delineation of the advisory. The method for determining the advisory delineation area is based on the movement potential of fishes throughout a waterbody. The point or small reach where fish collection took place is located on a map, and major upstream and downstream landmarks (i.e., dams, roads, tributaries, other barriers) are located and evaluated as segment boundaries. Barriers, such as dams, are preferred because they block fish movement. Other boundaries are selected to be relatively easy for fishermen to recognize. Once the spatial delineation is determined, the official advisories are sent to the PFBC by August 1 for inclusion in the fishing regulations booklet for the next calendar year, and the advisory delineation is included on the 303(d) list of impaired waters. Additionally, DEP and the PFBC publish the advisories on their websites. Finally, a joint press release is usually issued in November to remind the public of the advisories.

LITERATURE CITED

Shull, D. R., and M. J. Lookenbill (editors). 2018. Water Quality Monitoring Protocols for Streams and Rivers. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

CHAPTER 3 CHEMICAL ASSESSESMENT METHODS

DISCRETE PHYSICOCHEMICAL ASSESSMENT METHOD

Prepared by:

Brian Chalfant
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2013

Edited by:

Heidi Biggs and Matthew Kundrat
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2017

INTRODUCTION

Water quality criteria, located in 25 Pa. Code Chapter 93, are implemented considering magnitude (concentration), frequency, and duration. This document discusses the technical, contextual, and conceptual aspects of procedures applied by DEP to inform water use assessment decisions based on discrete physical and chemical (physicochemical) data in relation to WQS. This document, however, contains relatively little discussion of the planning and execution phases of physicochemical water quality sampling, such as outlining study objectives, choosing sampling plan designs, and setting data quality objectives. These aspects are described in more detail within the data collection protocols of the Monitoring Book (Shull and Lookenbill 2018).

SAMPLING AND INFERENCE

Within the regulatory framework outlined above, DEP must determine if waterbodies meet WQS. Often, these determinations require evaluating if waterbodies meet WQS "at least 99% of the time," in accordance with 25 Pa. Code Chapter 96. A number of interrelated considerations – outlined in this document – must be addressed when assessing if water bodies meet WQS "at least 99% of the time" based on physicochemical samples.

Sampling Space-Time

Water quality conditions – like many perceptible phenomena – change in space and in time. Interpreting water quality sample results often requires considering how far in space and for how long in time observed conditions can reasonably be considered representative of unobserved conditions that were not sampled.

This inferential process of using discrete, spatiotemporally-limited observations (i.e., samples) to estimate a larger set of unobserved, continuously dynamic conditions can introduce uncertainty into the use assessment decision process. Such uncertainty is called sampling error. Uncertainty can also enter the use assessment decision process through variability attributable to analytical measurement techniques, or measurement error. Sampling error and measurement error subject decisions based on sampling to decision error by introducing the potential for inaccuracy and imprecision into the observational process via sampling and analytic quantification. The aim of much of the rest of this document is to describe how the inherent variation and uncertainty introduced by sampling (i.e., sampling error) is addressed by DEP in the use assessment decision process for discrete physicochemical water quality sampling data. Variability attributable to analytical and laboratory techniques and equipment (i.e., measurement error) is discussed elsewhere (DEP 2016). Both forms of error can be

minimized by applying quality assurance procedures during sample collection, processing, and analysis.

Here's a simple scenario to illustrate sampling error. Imagine someone dutifully collects a one-liter water quality sample from a set location in a creek each month for two years. That's 24 one-liter samples. For the sake of discussion, let's imagine that 600 billion liters of water flow past that sampling location each year. Many considerations arise when we evaluate how representative those 24 liters of sampled water are of all 1.2 trillion liters of water that flowed past that location in those two years. Imagine that none of the 24 samples show violation of a relevant criterion. When we look to that sampling data to inform our decision if the creek is meeting standards "at least 99% of the time," we must ask ourselves how representative we think the *observed* concentrations are of the *unobserved* concentrations during the monitoring period. We may not observe violations if they occur at times when – or in areas where – samples were not collected. How we deal with these sampling error considerations is the primary focus of much of the rest of this document.

Sampling Error Implications

Unless we continuously observe – or census – every quantum of water in a stream, and as long as we rely on limited observations derived from sampling, we have to acknowledge the possibility of sampling error.

Ideally, use assessment decisions for surface waters in Pennsylvania based on physicochemical data will be informed by sampling conducted frequently enough to definitively characterize the conditions for each parameter of concern over a long enough time frame to account for variations in concentration attributable to changes in all relevant factors. This may be possible for some parameters in some locations through deployment of automated, continuous instream monitoring devices or through extremely intensive monitoring efforts. However, many water quality sampling efforts require human beings to visit sites with chemical test kits or hand-held probes to measure water quality conditions or to collect grab samples of the water for later analysis at laboratory facilities. Such monitoring efforts require personnel, funding, and site accessibility among other considerations. As a result, chemical water quality sampling often provides limited windows into the dynamic continuum of water quality conditions at any given location at any given time.

Continuous instream monitoring devices can measure conditions on a relatively frequent basis (e.g., every 15 minutes, every hour). Monitoring water quality conditions at such high frequency minimizes how far sample results have to be extrapolated into unobserved time, thereby minimizing the potential for sampling error (Hoger 2018). With

less temporally-dense monitoring approaches (e.g., periodic grab samples collected by a person and analyzed in a laboratory) the amount of temporal extrapolation will likely need to be extended further in time, and the potential for sampling error may increase. Continuous instream monitoring devices can be deployed in remote locations and set up to report observations via telemetry or through occasional retrievals and downloads. While these devices can provide extremely detailed, temporally-dense observational records, many such devices can only measure a few water quality parameters (e.g., dissolved oxygen, temperature, conductivity, pH) for which WQS exist.

In the absence of temporally-dense observations, if we understand and have information on enough relevant variables (e.g., stream flow, precipitation, water temperature) as related to the parameter of interest, we may be able to confidently infer or extrapolate unobserved conditions from observed conditions based on empirical understanding of variability, and thereby reduce uncertainty attributable to sampling error. A wide variety of interrelated factors can contribute to spatial and temporal variation in the concentrations of water quality parameters of concern, including but certainly not limited to: precipitation rates, durations, and locations; thawing of ice and snow; stream flow; geologic and soil characteristics; annual and diel cycles of solar input; atmospheric conditions (e.g., cloud cover); discharges from permitted facilities; chemical spills; watershed drainage patterns; watershed land use; and hydrologic alterations (e.g., dams). Different water quality parameters often vary in unique ways relating to these and other factors. For example, dissolved oxygen concentrations in streams often exhibit strong annual and daily patterns attributable to interrelated patterns of solar flux, stream temperature, and photorespiratory activity. Meanwhile, total dissolved solids concentrations often vary much less with diel and annual patterns of solar flux, and more often vary primarily with stream flow and related patterns of surface runoff, geologies, and groundwater flow patterns. Knowledge and understanding of such patterns can strengthen inferences about unobserved conditions. USEPA (2005) recommends,

> *"… states should decide how far out in time to extrapolate from the time at which a particular single grab was collected. EPA recommends that such decisions be based on contextual information regarding conditions when and where the grab was taken. For example, such information might include: 1) precipitation, 2) streamflow, 3) location of point source discharges in relation to the monitoring site, 4) land use patterns in the vicinity, 5) expected patterns of pollutant loading from the different kinds of sources present in the watershed, 6) occurrence of a chemical spill or other unusual event, and 7) historic patterns of pollutant concentrations in the monitoring segment and/or waterbodies similar to it."*

In some situations, it may be possible to extrapolate patterns observed at nearby, physiographically similar, or hydrologically similar locations to unobserved locations for certain parameters. For example, data from past monitoring show that concentrations of total dissolved solids in many lotic systems in Pennsylvania often exhibit an inverse power type response to stream flow, so it may be possible to predict – or inferentially estimate – total dissolved solids concentrations with some confidence if we know stream flow and have reason to believe the often-observed relationship between stream flow and total dissolved solids holds for the particular situation at hand.

Even if we confidently observed or inferred the condition of every possible quantum of flow at a particular location and even if we quantified the concentration of the parameter of interest with immaculate accuracy and precision, the phrase “at least 99% of the time” allows for some temporary, rare exceedance of criteria, and introduces another set of considerations to the use assessment process aside from sampling error.

99% of Time

Along with considerations of uncertainty introduced by spatiotemporally limited sampling of continuously-dynamic conditions, the phrase “at least 99% of the time” introduces to the use assessment decision process the additional consideration that some temporally rare criteria exceedances are acceptable. Some criteria explicitly state they must be met “at all times,” and some criteria define averaging periods and minimal sampling requirements, but many criteria do not explicitly specify associated sampling requirements or temporal aspects.

The vast majority of physicochemical water quality criteria – as expressed in Pennsylvania regulation – are written as numeric values: concentrations of various parameters not to be transgressed. These specific numeric concentrations comprise the magnitude components of criteria. For standards that must be met “at least 99% of the time,” DEP must consider not only the magnitude components of standards expressed as concentrations, but also temporal components, namely the frequency with which violations of those concentrations occur and the duration of how long violations last.

The phrase “at least 99% of the time” addresses the temporal aspects (i.e., frequency and duration) of criteria for which these considerations are not otherwise specified in Chapter 93. However, the phrase “at least 99% of the time” needs further consideration and definition as it does not specify a time period to which it applies (e.g., 99% of one year; 99% of one month; 99% of one day, 99% of one minute). Contrast this with the comparatively exact specificity of the Water Contact criterion for fecal coliform. Therefore, interpreting the “at least 99% of the time” phrase requires context-specific

considerations that take into account the particular standard(s) being evaluated as well as site-specific evaluation of expected patterns of variability in the parameters of interest. The criteria for toxic substances are a bit different with regard to temporal aspects because toxic criteria are based on specified exposure durations and because USEPA recommends certain acceptable exceedance frequencies for some toxic criteria, as discussed below.

Considering these temporal aspects of water quality criteria, the underlying concept in the phrase “at least 99% of the time” is straightforward: there is some acceptable – albeit relatively low – frequency at which, and duration for which water quality criteria concentrations presented in Chapter 93 can be exceeded with the waterbody still being considered as “meeting standards.”

Note that paragraph 96.3(c), in reference to narrative criteria, stipulates that *general water quality criteria* contained in section 93.6 will be achieved in surface waters “***at all times at design conditions***.”

Sampling Design Considerations Specific to Discrete Chemical Samples

Sampling error is influenced by how, when, where, why, and by whom samples are collected. As such, these considerations play a critical role in the use assessment decision process. While sampling plan design is not the focus of this document, some considerations on this topic are discussed in this section because sampling plan design largely determines what analytical procedures can tenably be used to assess the sampling data. As stated above, DEP strongly suggests that anyone planning to collect physicochemical sampling data for consideration in the water use assessment decision process familiarize themselves with the Monitoring Book (Shull and Lookenbill 2018) for more thorough discussions of sampling plan design, data quality objectives, and other study planning considerations.

Thoughtful study design and execution are critical to assuring water quality sampling efforts provide the information necessary to address the study questions. USEPA (2006) details step-by-step considerations of study design and data assessment. USEPA (2002a) provides further specific details on designing a sampling plan. While all these sampling plan design considerations are not repeated here, DEP feels it is important to address a few interrelated implications of sampling plan design in light of Pennsylvania’s “at least 99% of the time” regulatory language. In a particularly germane excerpt, USEPA (2002a) recommends,

> *“…sampling design development should also take into account existing regulations and requirements (for example, state, municipal) if they apply.”*

In Pennsylvania, the phrase “at least 99% of the time” is a critical regulatory consideration in designing a sampling plan. Regardless of how the “at least 99% of the time” phrase is quantified in specific applications; the phrase implies that criteria exceedances are only acceptable for a very small proportion of time. As such, any study or investigation aiming to assess criteria must aim to observe an extreme end – or ends – of water quality parameter frequency distributions. The phrase “at least 99% of the time” implies we are concerned with assessing the 99th (or, conversely, the 1st) percentile of a given frequency distribution.

Due to interrelated considerations of decision error rates, sample sizes, and extreme percentiles of frequency distributions, it will very often be impractical to employ a probability-based sample design to assess against the “at least 99% of the time” stipulation without collecting large numbers of samples, at least at any reasonable decision error rates. It will often be the most resource-effective approach – especially when accounting for monitoring costs – to focus monitoring at times when violations are most likely to occur based on understanding of the factors affecting the parameter of interest in the particular monitoring situation being assessed. In the rest of this document, DEP refers to these times when criteria violations are most likely to occur as **critical sampling periods**. Sampling focused on these critical sampling periods will necessarily be based on human understandings of the variables at play. In the terminology used by USEPA (2002a), these critical sampling periods can be thought of as temporal “hot spots,” and sampling targeted to observe these “hot spots” based on understandings of context-specific variations is referred to as “judgment-based sampling” (as contrasted with probability-based sampling). Since this targeted, judgment-based sampling is not suited to rigorous quantitative statistical analyses, assessment processes based on such sampling will necessarily draw less on quantitative statistical tools than will assessment processes based on probability-based sampling designs.

Of the various sampling plan designs discussed by USEPA (2002a), DEP believes the so-called “judgment-based” sampling design is the most suited method to assess extreme, infrequent ends of distributions stipulated by the phrase “at least 99% of the time.” Other sampling plan designs (e.g., simple random sampling, systematic sampling) presented by USEPA (2002a) are unlikely to provide accurate, precise estimates of such extreme ends of distributions while maintaining reasonable decision error rates without requiring large numbers of samples. For example, systematic sampling (i.e., sampling at regular temporal intervals) may be useful for certain applications (e.g., determining temporal trends) and can be attractive in terms of scheduling personnel and logistics, but is unlikely to directly observe infrequent,

extreme events (i.e., heavy storm flows or conditions that occur 1% of the time or less) unless many samples are collected at relatively short intervals. Such systematic sampling will usually require a very large number of samples to accurately and precisely estimate extremely infrequent conditions.

Regarding systematic sampling, USEPA (2002a) notes,

> *"… if the scale of the pattern or feature of interest is smaller than the spacing between sampling locations [or times], then the systematic pattern of sampling is not an efficient design unless the spacing between sampling locations [or times] is reduced or some other procedure such as composite sampling is introduced into the design.*
>
> *Systematic sampling would be inappropriate if a known pattern of contamination coincides with the regularity of the grid design. Such a coincidence would result in an overestimation or underestimation of a particular trait in the target population of interest."*
>
> *"Systematic/grid sampling may not be as efficient as other designs if prior information is available about the population. Such prior information could be used as a basis for stratification or identifying areas of higher likelihood of finding population properties of interest."*
>
> *"… if nothing is known about the spatial characteristics of the target population, grid sampling is efficient in finding patterns or locating rare events unless the patterns or events occur on a much finer scale than the grid spacing. If there is a known pattern or spatial or temporal characteristic of interest, grid sampling may have advantages over other sampling designs depending on what is known of the target population and what questions are being addressed by sampling."*

For example, a systematic sampling plan for dissolved oxygen where samples are collected the 15th day of every month at midafternoon would be likely to sample the highest dissolved oxygen concentrations because photosynthetic activity usually peaks around midafternoon.

Regarding judgment-based sampling designs, USEPA (2002a) states that,

> *"In judgmental sampling, the selection of sampling units (i.e., the number and location and/or timing of collecting samples) is based on knowledge of the*

feature or condition under investigation and on professional judgment. Judgmental sampling is distinguished from probability-based sampling in that inferences are based on professional judgment, not statistical scientific theory. Therefore, conclusions about the target population are limited and depend entirely on the validity and accuracy of professional judgment; probabilistic statements about parameters are not possible. As described in subsequent chapters, expert judgment may also be used in conjunction with other sampling designs to produce effective sampling for defensible decisions."

As noted by USEPA (2002a), many commonly-used statistical analysis methods assume either implicitly or explicitly that data were obtained using a probability-based – often simple random – sampling design. Probability-based sampling designs allow for application of statistical tools, which allow for quantification and control of decision error rates. In short, probability-based sampling designs offer the benefit of being amenable to statistical results interpretation, but may require a lot of sampling. Judgment-based sampling designs may not be conducive to standard inferential statistical analyses – primarily due to sample selection bias – but offer the benefit of more resource-efficient sampling (i.e., needing fewer observations to achieve a given level of precision) by incorporating existing understandings of the site and systems being sampled. USEPA (2002a) notes that,

"Judgmental sampling is useful when there is reliable historical and physical knowledge about a relatively small feature or condition …."

"… whether to employ a judgmental or statistical (probability-based) sampling design is the main sampling design decision …."

"An important distinction between the two types of designs is that statistical sampling designs are usually needed when the level of confidence needs to be quantified, and judgmental sampling designs are often needed to meet schedule and budgetary constraints."

"Data obtained from convenience or judgment sampling cannot be used to make formal statistical inferences unless one is willing to assume that they have the same desirable properties as probability samples, an assumption that usually cannot be justified …."

"Although statistical methods for developing the data collection design … are strongly encouraged, not every problem can be resolved with probability-

based sampling designs. On such studies … the planning team is encouraged to seek expert advice on how to develop a non-statistical data collection design and how to evaluate the results of the data collection."

When designing a sampling plan, USEPA (2002a) recommends considering tradeoffs among considerations of desired data quality (e.g., characteristics of the parameters of interest, applicable analytical approaches, decision error estimates) and practical constraints (e.g., budgets, personnel, time, site accessibility) for a given parameter in a given situation.

USEPA (2002a) also suggests that some sampling plan designs – like stratified random sampling – draw on aspects of both probability-based designs and judgment-based designs. Stratified random sampling can be used to more efficiently focus sampling resources to critical sampling periods at a given location based on existing understanding about variability of the parameters of concern at the particular study location (i.e., where and when criteria violations are most likely to occur). For example, an understanding – or, to use a USEPA term, "conceptual model" – of dissolved oxygen concentrations could be used to define three temporal strata based on likely concentration ranges: likely low-level (pre-dawn, summer), likely mid-range (autumn and spring mornings and evenings), and likely high-level (mid-day, winter). Such a stratified approach may also incorporate spatial aspects with backwater, less-turbulent areas being more likely to have lower concentrations of dissolved oxygen than faster-flowing, more-turbulent areas mid-channel. In stratified random sampling designs, each member of the target population has a known – although perhaps unequal – probability of selection into the sample. Therefore, techniques of statistical inference can be applied to data resulting from stratified random sampling designs. Regarding stratified random sampling, USEPA (2002a) notes,

"When stratification is based on correlation with an auxiliary variable which is adequately correlated with the variable of interest, stratification can produce estimates with increased precision compared with simple random sampling or, equivalently, achieve the same precision with fewer observations. For increased precision, the auxiliary variable used to define the strata should be highly correlated with the outcomes being measured. The amount of increase in precision over simple random sampling depends on the strength of the correlation between the auxiliary variable and the outcome variable being measured."

"Stratified sampling needs reliable prior knowledge of the population in order to effectively define the strata and allocate the sample sizes. The gains in the

precision, or the reductions in cost, depend on the quality of the information used to set up the stratified sampling design. Any possible increases in precision are particularly dependent on strength of the correlation of the auxiliary, stratification variable with the variable being observed in the study."

USEPA (2002a) also acknowledges that no sampling plan design is completely objective, noting (emphasis original),

"Implementation of a judgmental sampling design should not be confused with the application of professional judgment (or the use of professional knowledge of the study site or process). Professional judgment should always be used to develop an efficient sampling design, whether that design is judgmental or probability-based. In particular, when stratifying a population or site, exercising good professional judgment is essential so that the sampling design established for each stratum is efficient and meaningful."

MAKING DECISIONS

In the past, DEP adopted an approach that stipulated minimum data requirements for chemical use assessment datasets that were applied across all criteria. These requirements stipulated a minimum number of samples (i.e., 8), sampling frequency (i.e., at least quarterly) and duration (i.e., at least one year) needed to assess sampling data against any criteria. While this approach attempted to direct sampling so that a variety of conditions would be observed (e.g., different flow conditions, different times of year), this approach did not address the idea of critical sampling periods, discussed above. Regarding data quantity, USEPA (2005) states,

"EPA encourages the collection of adequate data to make well-grounded attainment determinations. EPA has not established, required, nor encouraged the establishment of rigid minimum sample set size requirements in the WQS attainment status determination process. EPA is particularly concerned with application of such thresholds state-wide, without regard to key factors like the manner in which applicable WQC are expressed, variability in segment-specific conditions, and fluctuations in rates of pollutant loading. Rather if employed, target sample set sizes should not be applied in an assessment methodology as absolute exclusionary rules, and even the smallest data sets should be evaluated and, in appropriate circumstances, used. While it may be appropriate to identify target sample sizes as a methodology is developed, states should not exclude from further consideration data sets that do so solely because they do not meet a target

sample size. A methodology may provide for an initial sample size screen, but should also provide for a further assessment of sample sets that do not meet the target sample size. (EPA suggests that states avoid setting target sample set sizes higher than the amount of data available at most sites.)"

Presently

Presently, DEP recommends context- and site-specific approaches to evaluate various criteria, accounting for the fact that the water quality criteria in Chapter 93 are presented in different ways, and because some parameters vary in different ways with changing natural conditions (e.g., diel and annual cycles of solar radiation, changes in stream flow) and may exhibit variable responses to these factors at different locations. DEP believes it inappropriate to develop data requirement guidelines applicable to all criteria across the board since different monitoring efforts may utilize different means and may have different goals, and because different parameters, criteria, and situations call for different monitoring approaches. The present approach is consistent with recommendations from USEPA (2005) that,

"Any target sample set size thresholds must be consistent with the state's EPA-approved water quality standards. Hence, when making a determination based on comparison of ambient data and other information to a numeric WQC expressed as an "average" concentration over a specified period of time, a statement of a desired number of samples may be appropriate. Still, the methodology should provide decision rules for concluding nonattainment in cases where the target data quantity expectations are not met, but the available data and information indicate a reasonable likelihood of a WQC exceedance (e.g., available samples with major digressions from the criterion concentration, corroborating evidence from independent lines of evidence such as biosurveys or incidence of waterborne disease, indications that conditions in the waterbody and loadings of the pollutant into the waterbody have remained fairly stable over the period in question)."

All relevant data will be considered in DEP's use assessment decision process regardless of sample size, but – because waterbody assessments are made on a continual basis in an effort to document current conditions – more recent data take precedence over older data, especially in situations where conditions have recently changed (e.g., installation of pollution remediation projects, alteration of permit limits in the watershed, changing land use patterns, discontinuation of combined sewer overflows). In some instances, older and newer data may be considered in concert to document temporal trends.

DEP makes every effort to verify the accuracy of all data used in the use assessment decision process. DEP strongly encourages anyone submitting data to familiarize themselves with DEP Bureau of Laboratories quality assurance and quality control procedures (DEP 2016) regarding record keeping, methods documentation, sampling techniques, selection of analytic laboratories, chain of custody concerns and so forth. DEP will not drop extreme values (AKA outliers) from a dataset unless there is reason to believe the extreme value is invalid. For example, a dissolved oxygen concentration of 100 mg/L is physically impossible at tropospheric temperatures and pressures – it is likely that such a record is a typographical error meant to really be 1 mg/L or 10 mg/L. Similarly, in a water temperature dataset submitted in degrees Celsius where one value is recorded at 72, it is highly unlikely this is a valid reading and may be recorded in degrees Fahrenheit. DEP does not want to discount any data or information from consideration outright, so no strict guidelines are set forth with regard to what sampling designs are acceptable because different sampling approaches may be necessary to answer different questions in different situations depending on the particular parameter and waterbody in question.

DEP strongly recommends that any physicochemical water quality sampling datasets intended for consideration in the use assessment decision process be collected using a "judgment-based sampling" design – as discussed above – with sampling targeted to critical sampling periods when water quality violations are most likely to occur based on knowledge of the conditions affecting the parameter(s) of interest. Some considerations are common to sampling design decisions for many parameters.

Diel cycles of solar radiation – The rising and setting of the sun drives daily cycles of photosynthesis and respiration across much of the earth's surface, including in surface waters. During peak influx of solar radiation, photosynthetic activity tends to peak, resulting in higher instream dissolved oxygen concentrations and pH levels. Incident sunlight also increases instream temperatures, which affects levels of dissolved oxygen in the water and photosynthetic rates. Other parameters (e.g., TDS, alkalinity) may also exhibit some diel cycling.

Annual cycles of solar radiation – Northern hemisphere locations, such as Pennsylvania, receive their most intense and prolonged sunlight in the summer months of June and July, with less intense and shorter exposure to sunlight in winter months of December and January. Stream systems reflect these cycles in annual cycles of water temperature and dissolved oxygen. Fluctuations in pH are often less dramatic in winter months as well, likely due

to reduced photosynthetic activity with the colder temperatures. Some nutrient parameters may also exhibit variation with annual seasons.

Annual cycles of precipitation, evapotranspiration and stream flow – Across Pennsylvania, stream flow patterns vary annually with the lowest stream flows typically observed from July through September, gradually increasing through autumn and winter with peak flows often observed January through April and tailing off again May through June to summer base flows, although hurricane and tropical storm remnants occasionally dump heavy rain on Pennsylvania in early autumn. These stream flow patterns reflect annual cycles of precipitation, snow and ice melt, and evapotranspiration. Such patterns vary locally with a variety of factors (e.g., soil types, hill slopes) and fluctuate year to year. Some parameters of concern often exhibit predictable patterns of response with precipitation and stream flow. For example, TDS concentrations tend to decrease consistently with increasing stream flow; in many situations, this has to do with surface runoff containing lower dissolved ion concentrations than groundwater inputs to a stream, which tend to be higher in dissolved minerals. Likewise, some parameters, like various forms of phosphorus, usually enter stream systems attached to particulate matter washed in during periods of high surface runoff. Such patterns may also vary with surrounding land use, like fertilizer application, impervious surfaces and so on.

Conservative and non-conservative substances – Some water quality parameters – like sulfate – are considered conservative in that their concentrations are not directly affected by biological processes. These conservative substances do not decay, are not selectively incorporated by living organisms, and concentrations are affected mostly by sedimentation and dilution. Non-conservative substances – such as phosphorous – are removed from the water column by biological processes.

Clarifications

When compared to physicochemical sampling datasets used to make listing decisions, datasets used to inform **delisting decisions** (i.e., decisions to remove a waterbody from Category 5 of the 303(d) list) must: (1) have been collected more recently; (2) have been collected as frequent or more frequently; and (3) contain more samples.

For criteria written as time-average criteria, DEP encourages multiple sampling events within the specified time periods. For example, the total iron criterion is written as a 30-day average, so DEP encourages sampling multiple times in a given 30-day period

to compare existing conditions to the standard. Like any sampling, this sampling should represent the most likely violation times and spaces as discussed above. As a general guideline, DEP encourages at least three sampling events in the time period expressed in various time-averaged criteria.

Section 93.8 presents two types of water quality criteria for toxic substances designed to protect human health: threshold effect criteria and cancer risk criteria. Regarding threshold effect criteria, section 16.32 states:

> (a) A threshold effect is defined as an adverse impact that occurs in the exposed individual only after a physiological reserve is depleted. For these effects, there exists a dose below which no adverse response will occur. Threshold toxic effects include most systemic effects and developmental toxicity, including teratogenicity. Developmental toxicity includes all adverse effects in developing offspring resulting from prenatal exposure to a causative agent.

And regarding cancer risk, or non-threshold effect, criteria, section 16.33 states:

> (a) A nonthreshold effect is defined as an adverse impact, including cancer, for which no exposure greater than zero assures protection to the exposed individual. Thus, in contrast to the threshold concept discussed in § 16.32 (relating to threshold level toxic effects), the nonthreshold approach to toxics control is based upon the premise that there is no safe concentration of the toxic.

Currently, USEPA policy establishes that the duration for human health criteria for carcinogens should be derived assuming lifetime exposure, taken to be a 70-year time period (USEPA 2007).

All criteria for toxic substances – aquatic life acute criteria, aquatic life chronic criteria, human health threshold criteria, and human health non-threshold criteria alike – are expressed as maximum values.

For toxic pollutant assessment decisions, DEP follows the guidelines set forth by USEPA (i.e., CMC pollutants violating criteria more than once in a given 3-year period constitute an impairment; CCC pollutants with either a 30-day mean concentration violating a criterion more than once in a 3-year period or a 4-day mean concentration exceeding twice the CCC in any 4-week period constitute an impairment). The 3-year period stipulated for toxic pollutants starts at the time monitoring begins and not when

the first exceedance occurs, unless they coincide. For example, if a single violation is observed at the end of the year for a routinely monitored station, another violation must not be observed at that site for the next two years for the water to still be considered attaining its use for that parameter. Similarly, if the first sample at a previously unmonitored site shows a toxic criteria violation, another violation must not be observed at that site for the next 3 years for the water to still be considered attaining its use for that parameter. The 4-week period stipulation is applied in the same manner.

Guidance from USEPA defines the exposure period associated with CMC as one hour and with CCC as four days. Federal regulations specify that the acute criteria – or CMC – for a pollutant must not be exceeded more than once in a given 3-year period for any waterbody. For CCC, USEPA guidance states that: (1) The 30-day mean concentration of a pollutant in a body of water must not exceed the CCC for that pollutant more than once in a 3-year period; and (2) no 4 consecutive samples collected on different days during a 4-week period can exceed twice the CCC. For example, 10 samples are collected within a 4-week period, any 4 consecutive samples are averaged and compared against the CCC for Benzene (130 µg/L); if the average of those samples exceed 260 µg/L (twice the CCC of Benzene), then the criterion is violated and the water is impaired.

For conventional (i.e., not toxic) pollutants, DEP considers grab samples to be representative of a day unless convincing evidence exists to suggest otherwise (e.g., a documented spill, influence of a known biological process, supporting high-frequency monitoring data). Under this assumption, four days with conventional pollutant criterion violations in a year constitute use impairment for criteria expressed as maxima or minima because the standards are violated more than 1% of the time (i.e., 4 days / 365 days ≈ 1.1%, which means criteria are being met less than 99% of the time). For conventional pollutants expressed as 30-day or monthly averages, any one month or 30-day period showing a criterion violation will be considered an impairment based on the reasoning that the water is not meeting standards 99% of the time (i.e., 1 month / 12 months ≈ 8.3%).

LITERATURE CITED

Hoger, M.S. 2018. Continuous Physiochemical Assessment Method. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

DEP. 2016. Quality Assurance Manual for the Pennsylvania Department of Environmental Protection Bureau of Laboratories. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

Shull, D. R., and M. J. Lookenbill (editors). 2018. Water Quality Monitoring Protocols for Streams and Rivers. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

USEPA. 1985. Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses. PB85-227049. Office of Research and Development, Duluth, Minnesota.

USEPA. 2002a. Guidance on Choosing a Sampling Design for Environmental Data Collection for Use in Developing a Quality Assurance Project Plan. EPA QA/G-5S. U.S. Environmental Protection Agency, Office of Environmental Information, Washington, D.C.

USEPA. 2002b. Consolidated Assessment and Listing Methodology – Toward a Compendium of Best Practices. First edition. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

USEPA. 2005. Guidance for 2006 Assessment, Listing and Reporting Requirements Pursuant to Sections 303(d), 305(b) and 314 of the Clean Water Act. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

USEPA. 2006. Guidance on Systematic Planning Using the Data Quality Objectives Process. EPA QA/G-4. U.S. Environmental Protection Agency, Office of Environmental Information, Washington, D.C.

USEPA. 2007. Water Quality Standards Handbook. Second edition. EPA-823-B-94-005. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

CONTINUOUS PHYSICOCHEMICAL ASSESSMENT METHOD

Prepared by:

Mark Hoger
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2018

INTRODUCTION

Water quality data sondes record instream parameters such as temperature, specific conductance, pH, dissolved oxygen (DO), and turbidity at defined intervals. Because these intervals are sufficiently narrow (e.g., 30 minutes), these data are considered continuous, and are referred to as continuous instream monitoring (CIM) data. The number of samples in a CIM dataset approximates a census of the water quality conditions, and therefore should be evaluated differently than a dataset of only a handful of grab samples.

Some parameters directly recorded by CIM deployments (temperature, pH, and DO) have defined WQS in 25 Pa. Code Chapter 93 that are implemented through Chapter 96 (WQS Implementation). These criteria are expressed as either a minimum or maximum concentration, or an arithmetic mean concentration over a defined period (§ 93.1). In addition to the parameters directly recorded by CIMs, discrete grab samples of many other water quality measures can be expanded into continuous datasets by modeling the relationship between the discrete grab samples and recorded CIM parameters. These model-derived continuous datasets may also have defined WQS.

This document largely focuses on evaluating CIM data (both directly measured and model-derived) for the purposes of assessing protected water uses; however, these analyses are also important tools to be used for other objectives, including stream use evaluations or cause and effect surveys.

CRITERIA IMPLEMENTATION

99% Rule

The frequency, or acceptance threshold, associated with Chapter 93 water quality criteria is given in § 96.3(c), which states,

> *"To protect existing and designated surface water uses, the water quality criteria described in Chapter 93 (relating to water quality standards), including the criteria in §§ 93.7 and 93.8a(b) (relating to specific water quality criteria; and toxic substances) shall be achieved in all surface waters at least 99% of the time…"*

This WQS component introduces the allowance for temporary, rare exceedances of water quality criteria. Time, however, in "99% of the time" is not defined in § 96.3(c), leaving the frequency and duration of these allowed exceedances not fully described. Because CIMs record water quality parameters at such frequent intervals, the amount of interpolation between samples is very small, so the data are, in effect, a census of water quality, significantly reducing sampling error (see "Sampling Error Implications" section of Chalfant 2013). This more thorough dataset and the resulting reduction in sampling error means that the application of the 99% rule can be applied over a large period for temporally comprehensive protection under the WQS.

Therefore, for the purpose of the assessment of CIM data with respect to § 93.7 Table 3, time, as a part of § 96.3(c), is defined as a rolling year. The length of one year provides the inclusion of all seasonal variation found in a temperate ecosystem, including various life cycles and reproductive signals that are often strongly tied to season. The use of a *rolling* year affords the opportunity to monitor for stressful conditions that may span other, arbitrary yearly divides such as calendar year or water year (October 1).

Because water quality is to be achieved “99% of the time”, assessments will be made by determining whether recorded exceedances of criteria constitute greater than one percent of a year. To make this calculation the number of individual recordings exceeding the listed criteria will be summed and a percent of a year (%Y) that those readings represent will be calculated using the following equation:

$$\%Y = 100\left[\frac{n * i}{k}\right]$$

Where:

%Y = percent of a year
n = number of exceedances
i = recording interval in minutes
k = A constant (525,600) equal to the number of minutes in a year (365 days * 24 hrs/day * 60 min/hr)

If %Y > 1, then the criterion is not achieved 99% of the time as required by § 96.3(c), and the waterbody is not attaining WQS. A summary of common recording intervals and the number of readings that would constitute greater than one percent of a year is provided in Table 1.

Table 1. Common recording intervals and the number of readings necessary to represent greater than one percent of a year.

Recording Interval	Number of Readings
15 min	351
30 min	176
60 min	88

Because continuous monitoring equipment is left unattended for extended periods, CIM data are more susceptible to error from calibration drift or sensor fouling (see Quality Control Requirements below). This can result in the removal of portions of data from the dataset due to excess uncertainty in the accuracy of the data. The removal of data due to excess uncertainty makes it problematic to define time, with respect to the 99% rule, as the total number of readings in the dataset (rather than the rolling year as described above) because doing so would effectively lead to an increased likelihood of non-attainment for datasets with increased levels of uncertainty. For example, if a dataset had 10,000 readings, and 90 of the readings exceeded a criterion, 0.9% of the dataset

exceeded the criterion. But if there was a period during the deployment that the sensor became excessively fouled, some data may need to be removed due to an overabundance of uncertainty in the accuracy of the measurements. If 1,200 readings were removed from the 10,000 readings, there would only be 8,800 readings remaining, and the 90 exceedances would now represent 1.02% of the dataset. With increased uncertainty, data were necessarily removed; but, lowering the total number of readings increased the portion of the dataset that each exceedance represented, and in effect, increased uncertainty in the data made a non-attainment decision more likely. Therefore, the 99% rule is applied to a defined length of time (a rolling year) rather than the number of readings in a dataset.

The application of the 99% rule to a period of less than one year would also be problematic. Doing so would require a minimum period to be established because without a minimum, the period of record could be made so short as to necessitate only one reading to trigger non-attainment, and implementing a minimum period would, in effect, just establish that length of time as the *de facto* period to apply §96.3(c). In addition, application of §96.3(c) over a period shorter than a year (season, month, week, etc.) would drive the threshold for impairment exceptionally low, and cause a significant discrepancy between water quality criteria non-attainments—which were established for the protection of uses—and observed impacts to uses. The same discrepancy would result if §96.3(c) were applied as a percent of days for which the criterion was exceeded. This application would mean that a single reading outside a listed criterion on four days throughout the entire year would represent non-attainment [100 * (4 / 365) = 1.1%].

Data Collection Requirements and Critical Time Periods

Even though the assessment decision is based on a 365-day period, it is not necessary that a full year of CIM data is collected. In some circumstances, the number of exceedances, such that %Y > 1, may be observed in a rather short period. Focusing sampling effort during critical periods may give sufficient information to make an assessment decision while greatly reducing the amount of resources needed to conduct the survey.

If limited site-specific data are available, general knowledge of water quality processes can be used to determine critical periods and guide the period of record. For example, many water quality parameters are affected by seasonal change and their responses can, therefore, be predicted to a certain degree. DEP's CIM efforts have documented increases in pH values, increases in diel pH fluctuation, corresponding decreases in DO values, and increases in diel DO fluctuation from early spring through the fall. This correlates with increased photoperiod and increased air and surface water temperatures. The effect of increased temperature and photoperiod to increased instream production and respiration are well documented (Odum 1956, Strickland et al. 1970, Neori and Holm-Hansen 1982, Raven and Geider 1988). An increased photoperiod with adequate nutrition will increase the standing biomass of photosynthetic organisms (Valenti et al. 2011). Photosynthesis and respiration throughout the day and community respiration at night results in diel fluctuation of pH and DO (Odum 1956,

White et al. 1991, Wurts 2003). These processes indicate that during the growing season, pH is most likely to exceed the maximum criterion and DO to fall below the minimum criterion or 7-day average. If these criteria were the focus of the monitoring effort, the CIM deployment could be limited to this period to reduce resources while capturing the critical period.

DEP also recognizes that critical or limiting conditions may not be consistent year-to-year, and a single year of data may not accurately represent conditions that WQS were developed to protect. Typically, this is driven by the amount and timing of precipitation for a given period or year. Elevated precipitation will result in increased surface water discharge, which moderates limiting conditions characterized by temperature, pH and DO. DEP has documented in past surveys that elevated discharge can reduce daily DO, pH, and temperature fluctuations and increase daily minimum DO values and decrease maximum pH and temperature values. When multiple years of data are collected, assessment decisions will be based on years where the most critical or limiting conditions exist. For instance, if two years of data are collected, and in the first year there are exceedances of the maximum pH criterion greater than one percent of the time, and in the second-year exceedances are less than one percent of the time, it is likely that critical conditions existed in the first year that were not seen in the second, such as reduced amount or frequency of precipitation, or higher air temperatures. Therefore, the assessment decision will be based on the first year to be protective of critical periods. For this reason, it is also imperative to characterize conditions that drive critical or limiting conditions, and reference those conditions as part of the protected use assessment and subsequent reassessments.

CIM Parameters with Established Criteria

Table 3 of § 93.7(a) provides criteria for three parameters that can be directly measured by CIM deployments. The first two parameters are pH and DO. These parameters often have significant changes throughout the day, driven by photosynthesis, making CIMs particularly useful in assessments. The applicable criterion for pH is 6.0 to 9.0, inclusive. The minimum DO criterion is 5.0 mg/L, but other, more stringent criteria are applied for certain waters or times of the year, including minimum 7-day averages. A single 7-day average DO below the criterion indicates non-attainment of the criterion as seven days are more than one percent of a year [100 * (7 / 365) = 1.9%].

Maximum temperature criteria are provided for defined times of the year and water uses. Temperature criteria in § 93.7 are applied to heated waste sources regulated under 25 Pa. Code §§ 92a and 96. Temperature limits apply to other sources when they are needed to protect designated and existing uses. An appropriate thermal evaluation includes a biological assessment based on instream flora and fauna to determine whether the biological community is affected by the thermal regime. Typically, fish community evaluations have the best resolution in characterizing a waterbody’s thermal regime due to the effects to physiology and distribution patterns (Shuter et al. 1980, Ridgeway et al. 1991, Azevedo et al. 1998, Wehrly and Wiley 2003, Lyons et al. 2009). Continuous temperature data are not typically used to assess critical uses; however, to

qualify as a High Quality water under § 93.4b (a)(1)(i), a list of parameters, including temperature must be evaluated to meet the chemistry conditions. This regulation states,

> *"The water has long-term water quality, based on at least one year of data which exceeds levels necessary to support the propagation of fish, shellfish and wildlife and recreation in and on the water by being better than the water quality criteria in § 93.7, Table 3… at least 99% of the time…"*

In addition to the criteria listed in the Commonwealth's WQS, additional or more stringent criteria from the Delaware River Basin Commission (DRBC) water quality regulations, ORSANCO (Ohio River Valley Water Sanitation Commission) pollution control standards, and the Great Lakes Water Quality Agreement (GLWQA) are applicable as stated in §§ 93.2(b) and 93.9.

DETERMINATION OF SPATIAL EXTENT OF ASSESSMENT

While CIMs provide a thorough record of water quality conditions at a given point, additional data are necessary to understand the spatial extent to which the CIM data apply. To aid in this determination, discrete measurements should be collected throughout the area. The necessary spatial frequency of sampling will vary greatly depending on the stream, but these discrete measurements should target potential influences such as tributaries, discharges, or changes in land use that may significantly alter water quality (Walters and Pulket 2018). Transects at these additional points are helpful in determining any changes in mixing patterns (Hoger 2018). Most importantly, these measurements should be conducted during the critical periods of interest when water quality is suspected of exceeding criteria. Targeting these periods, at the periphery of expected values, provides the necessary information to characterize the spatial extent of an assessment.

Knowing when to sample is often informed by recent CIM readings and general knowledge of seasonal, daily, and weather-related trends in water quality. For example, if a CIM recently recorded exceedances of the maximum pH criterion in a stream, additional measurements should be taken upstream and downstream of that CIM, and in and around any tributary or discharge. In this example, exceedances are most likely to occur in the late afternoon as photosynthesis drives the pH higher. Reviewing the CIM record can give a more specific indication of the period of exceedance, perhaps it was between 17:00 and 19:00. The additional samples should then be targeted for that period. Samples taken in the morning, even if they align with the CIM readings, are likely to be below the criterion and would not provide sufficient information to extend the assessment.

QUALITY CONTROL REQUIREMENTS

All CIM data to be used for assessment must follow quality control methods as described in the Continuous Physicochemical Data Collection Protocol (Hoger et al. 2018) including regular fouling and calibration checks of the equipment, discrete

readings with a separate meter, appropriate corrections, and final independent approval of the data. Data that do not meet the usability threshold are removed from any assessment decisions. In addition, data are reviewed to determine if they are representative of the waterbody. Discrete water quality cross-section surveys (transects) are performed throughout the deployment, targeting various flows and water quality conditions to ensure that CIM data are representative of the targeted waterbody (Hoger 2018).

ASSESSMENT EXAMPLE

Continuous data were collected from late February to early November of 2016 on West Branch Octoraro Creek (WBOC), including temperature, specific conductance, pH, dissolved oxygen, and turbidity. After quality control checks were completed on the datasets, pH and dissolved oxygen data were compared to listed WQS criteria (Table 2). There were no exceedances of DO criteria, but the maximum pH criterion was exceeded numerous times (Figure 1). Because the number of exceedances represent greater than one percent of a year, WBOC was not attaining the pH criterion at this location.

Table 2. Comparison of continuous data from WBOC to WQS criteria.

Criterion	Number of Exceedances	Percent of Year
pH 6.0 to 9.0, inclusive	299	1.71%
DO minimum, 5.0 mg/L	0	0.00%

While performing a routine maintenance visit on April 5, 2016, it was noted that the sonde had recorded numerous exceedances of the maximum pH criterion. These exceedances had taken place between roughly 12:00 and 19:00, with peak pH occurring around 16:00. To aid in delineating a potential assessment of WQS criteria, a field visit was scheduled near the peak time of day to take discrete readings throughout the watershed. Because of a rain event on April 7, and the likely suppression of the maximum pH that would result, the visit was delayed. After several days without any rain, the visit was completed on April 21 (discrete at sonde shown in Figure 1-B). Numerous readings were taken on WBOC and its tributaries above the continuous station (Figure 2) from 14:57 to 16:10, the peak of pH readings in the continuous dataset. No readings were taken downstream of the continuous station due to the effects of Octoraro Lake just a short distance downstream. The discrete readings indicate that the impairment should extend along the entire mainstem of WBOC above Octoraro Lake. A reading was taken in each of the four largest tributaries. Three of the four discrete measurements were well below the maximum pH criterion, while the fourth was slightly above. The slight exceedance in Bowery Run may not provide enough evidence to extend the impairment into this stream, and more information may be necessary.

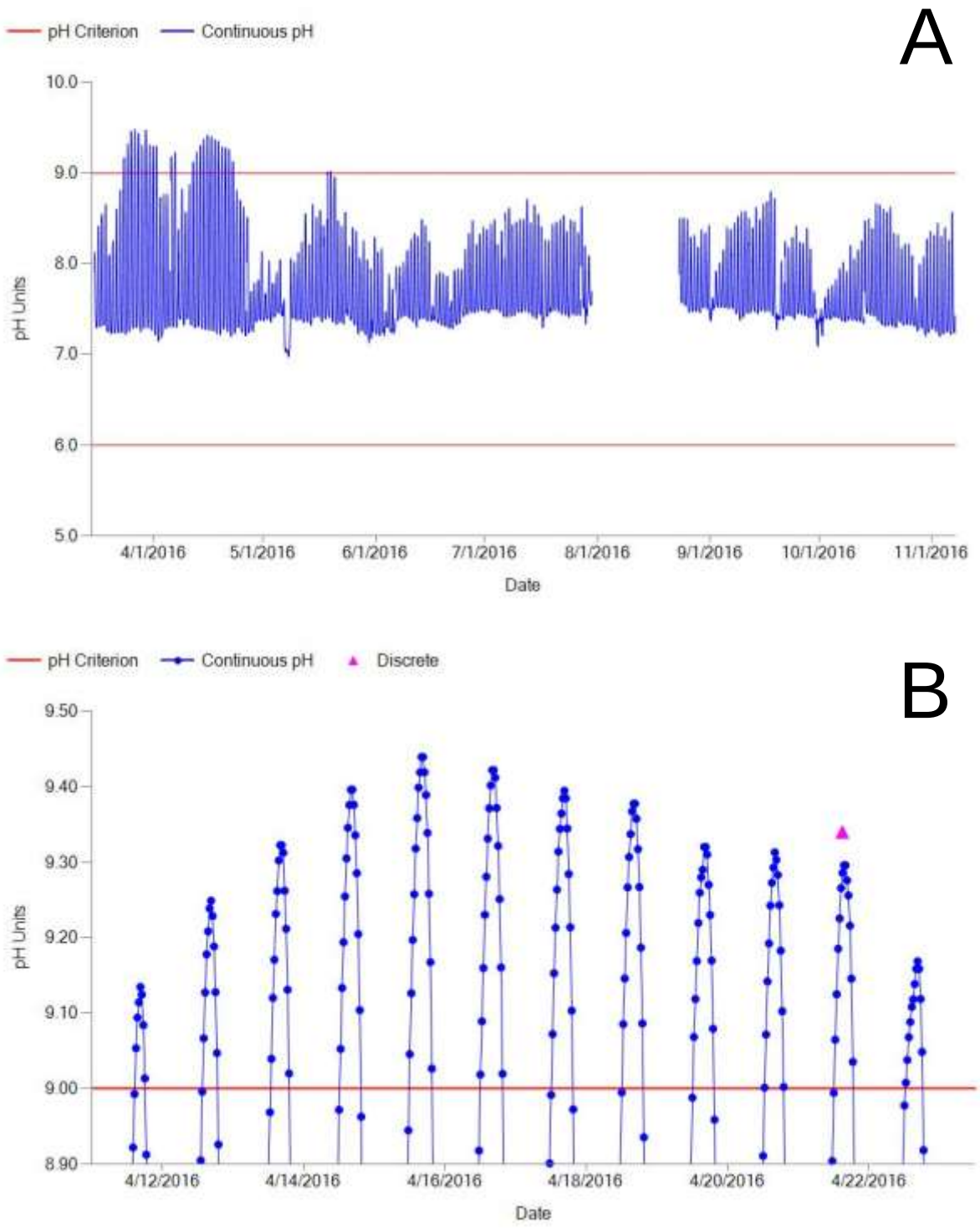


Figure 1. Continuous pH data (A) at WBOC shows exceedances of the WQS criterion. Individual points that exceed the criterion (B) are summed to calculate the percent of a year that these exceedances represent.

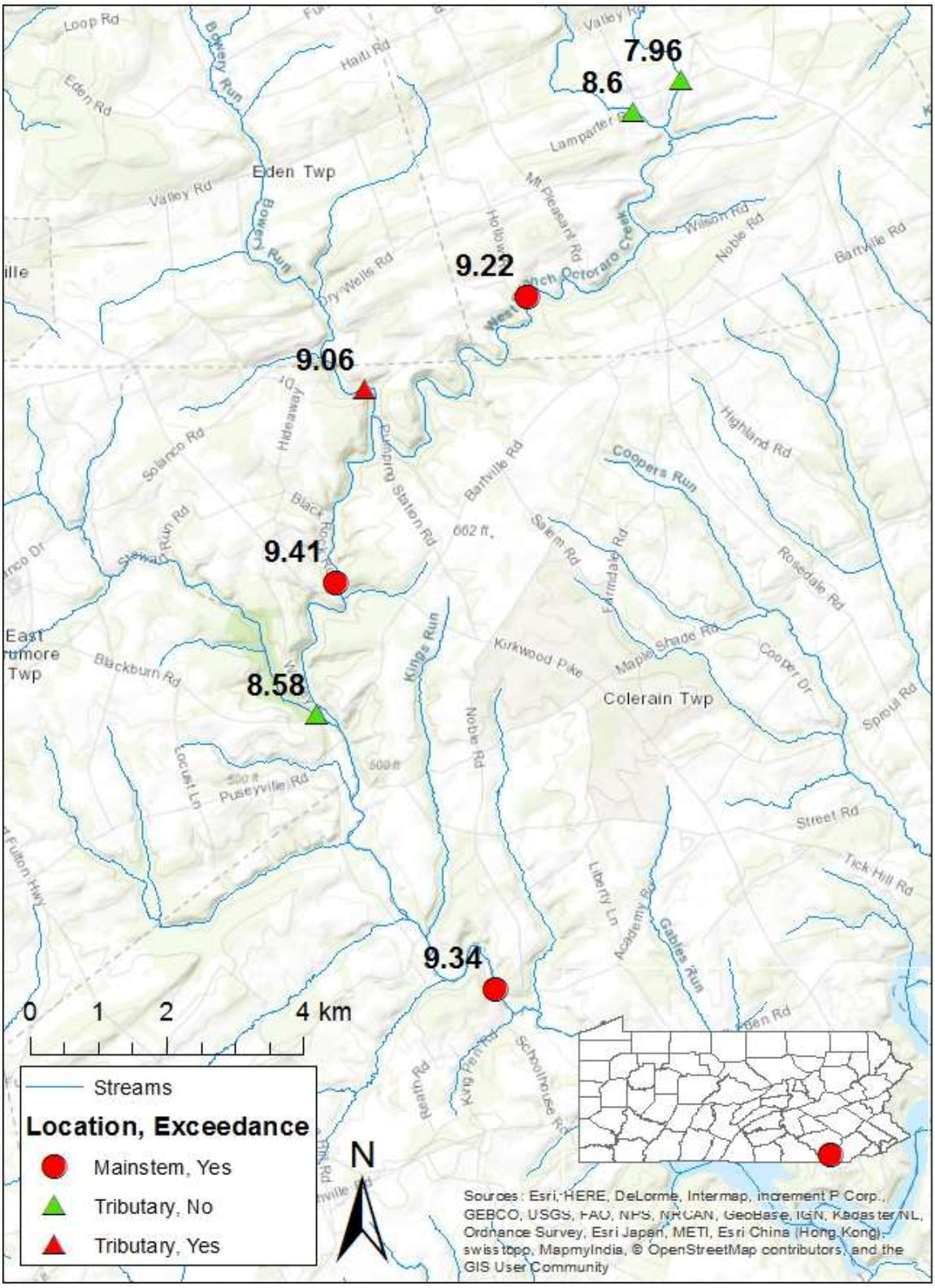


Figure 2. Discrete pH readings throughout the WBOC basin used to delineate an impairment. The furthest downstream location is the CIM station.

ASSESSMENTS USING DERIVED CIM DATA

Many water quality parameters of interest cannot be directly measured on a continuous basis. Auto-samplers could be used to collect discrete samples at regular intervals for analysis at the lab; however, this is a labor intensive and costly approach and is only realistic for relatively short periods of time. Though CIMs directly measure only a few parameters, some of the parameters have been shown to be highly correlated to other measures of water quality (e.g., Christensen et al. 2006, Foster and Graham 2016, Rasmussen et al. 2016). These relationships often have a strong physical basis, such as dissolved ions driving the specific conductance of water, or suspended sediment making water turbid. These relationships provide the opportunity to use easily-measured continuous parameters to accurately model numerous other parameters. The United States Environmental Protection Agency directs states to use all available water quality data in making assessment decisions (40 CFR 130.7(b)(5)) and, in guidance, specifies that models should be included in the data that are to be evaluated (US EPA 2005).

Models are developed by comparing discrete grab samples of the parameter(s) of interest to recorded CIM data. The discrete grab samples are collected directly over the CIM and should encompass the range of values observed in the CIM record. The number of samples necessary for the development of a strong model varies, but fewer samples are necessary if they are well distributed throughout the range of values (Rasmussen et al. 2009). Particular emphasis should be placed on collecting discrete samples when water quality is exceeding criteria. Inclusion of these samples adds critical support to models resulting in exceedances of criteria. Both discrete grab samples and CIM data should be collected following established DEP protocols and undergo all quality control procedures prior to final model development. Review of CIM data during the period of record can aid in the timing and collection of discrete samples that are distributed throughout the range of values. Models should be considered site-specific, recognizing the potential for differences in the relationship between water quality constituents at each site.

Most models are based on continuous specific conductance or turbidity data, though continuous water temperature, continuous streamflow, and Julian day (day of the year) have been used to strengthen models. Examples of constituents and the explanatory variables included in their model are shown in Table 3. While the table provides many examples, the list is not comprehensive and strong models are probable for many other parameters.

Table 3. Example response and explanatory variables for models of derived CIM. Citation listed in the explanatory variable column(s) that the study used in the model. All models listed in the table had R^2 values of at least 0.8.

Response Constituent	Specific Conductance	Streamflow	Turbidity	Temperature	Julian day
Actinomycetes			5		5
Alkalinity	2, 4, 5, 11	4, 11			
Atrazine		10	10		
Bicarbonate	4, 11	4, 11			
Boron	13				
Calcium	4, 5, 11, 12, 13, 14		14		
Chloride	1, 2, 4, 5, 10, 11, 12, 13, 14, 15	1			
Dissolved nitrate	1			1	
Dissolved nitrate + nitrite	10		10	10	
Dissolved orthophosphorus		10			
Dissolved phosphorus		14			14
Dissolved solids	1, 2, 4, 5, 10, 11, 12, 13, 14				
E. coli			12, 14		14
Enterococci bacteria			12, 14		14
Fecal coliform bacteria		14	12, 14		
Fluoride	1	1			
Hardness	4, 11				
Magnesium	5, 12, 13				
Particulate phosphorus			14		
Sodium	1, 2, 4, 5, 11, 12, 13, 14	1			
Sulfate	1, 2, 4, 5, 11, 12, 13, 14		14		
Suspended sediment		4, 6, 11	1, 2, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14		
Total nitrogen		3	3, 5	5	
Total organic carbon			11		
Total organic N + NH3	1		1, 2, 4	1	
Total phosphorus	3	2	2, 3, 10, 11, 12, 13	3	
Total suspended solids	4		1, 2, 4, 11, 12, 13, 14		

1. Christensen 2001
2. Christensen et al. 2006
3. Christensen et al. 2002
4. Christensen et al. 2003
5. Foster and Graham 2016
6. Juracek 2011
7. Lee 2009
8. Lee et al. 2008
9. Maloney and Shull 2015
10. Mau et al. 2004
11. Rasmussen et al. 2016
12. Rasmussen et al. 2008
13. Rasmussen et al. 2005
14. Stone and Graham 2014
15. Trowbridge et al. 2010

The development of models should follow strict guidelines to ensure a consistent, empirical approach at building and evaluating the strength of each model. Comprehensive guidance is provided in Rasmussen et al. (2009) which includes multiple tests of the uncertainty of the model such as root-mean-squared error (RMSE), model standard percentage error (MSPE), and prediction error sum of squares (PRESS). Decreases in RMSE, MSPE, and PRESS indicate reduced uncertainty in the model. Coefficient of determination (R^2) and adjusted coefficient of determination (R^2a) are measures of the strength of the relationship between variables. A higher R^2 or R^2a indicates that a higher portion of the response variable is described by the model. These values range from -1 to 1, where -1 indicates perfect negative correlation, 1 indicates perfect positive correlation, and 0 indicates no correlation.

Though these statistics can describe the relative strength or weakness of a model, it is problematic to define an appropriate threshold of strength that must be achieved by a model before it should be used for assessment of a derived parameter. Alternatively, the use of prediction intervals or probability of exceedance calculations incorporate the uncertainty of the model into the calculation.

To illustrate the importance of the difference in these approaches consider the following examples. First imagine a set of criteria (e.g., R^2, MSPE, PRESS thresholds) were established to determine whether a model was sufficiently strong to be used for assessment of a derived parameter. Then consider that a model just barely achieved that minimum standard for use, and that model generated a derived dataset that contained values beyond an established criterion. If many of the values were calculated at just slightly beyond the established criterion, it is possible that a significant portion of those values were below the criterion when the uncertainty of the model was incorporated. This would be analogous to a Type I error—a determination of non-attainment when the waterbody may be attaining.

The reverse could also happen. Imagine another model for a different set of data that just missed the standard for use, but, like the first example, calculated values that exceeded a criterion. If those exceedances were well beyond the criterion, instead of slightly past like the first example, the degree to which they exceeded could mean that they would likely be exceedances even if the higher uncertainty in the model was considered. This would be analogous to a Type II error—a determination of attainment when the water body is not attaining.

Because prediction intervals and probability of exceedance calculations incorporate the uncertainty of the model, these approaches reduce Type I and Type II error, and could lead to a determination of attainment of criteria in the first example and non-attainment in the second example. The first method suggested by Rasmussen et al. (2009, Appendix 3) to analyze derived data for criteria exceedance is to generate prediction intervals for cumulative frequency duration (CFD) curves. These curves show the proportion of values from the sample that fall below certain values. If 90 percent prediction intervals were then created around the CFD curve to assess based on a maximum criterion, the lower prediction curve could be used to determine the percent of

time that a criterion was exceeded with 90% confidence. This could then be compared to the 99% rule to determine if the waterbody is attaining; however, the percent of time in this calculation is based on the number of readings in the analysis and not one year. As discussed above, for the purposes of CIM assessment, the 99% rule should be applied to a 365-day period. Therefore, this method should not be used unless the calculation is adjusted.

An alternative method provided by Rasmussen et al. (2009, Appendix 3) is to generate a probability of exceedance for each data point of the derived series using the following equation:

$$P=1-D\left[\frac{x\text{-Criterion}}{RMSE}\right]$$

Where

P = probability the criterion was exceeded
D = cumulative distribution function for the standard normal curve (values found in tables provided in statistics textbooks)
x = model-computed value
RMSE = root-mean-square error, a measure of the variance between regression-computed and observed values

If the response variable was transformed in the model, both the model-computed value and the criterion must be transformed in the equation. For example, if the response variable was $\log_{10}$ transformed the equation would change to:

$$P=1-D\left[\frac{\log_{10}(x)-\log_{10}(\text{Criterion})}{RMSE}\right]$$

These probability of exceedance calculations can then be used to make assessment determinations on a derived dataset. All model-computed values with probability of exceedance greater than or equal to 0.9 (90%) are considered an exceedance of criteria. These exceedances are then summed and a percent of a year that they represent calculated. A number of exceedances such that the percent of a year is greater than one percent indicates non-attainment of WQS criteria.

The selection of a 90% probability threshold was chosen because it is a common break point for describing probable occurrence in statistical measures (confidence intervals, tests of significance, etc.). Ninety-percent prediction intervals are used by United States Geological Survey (USGS) in their presentation and analysis of model-derived continuous data (e.g., USGS National Real-Time Water Quality website: http://nrtwq.usgs.gov/, Juracek 2011, Mau et al. 2004, Rasmussen et al. 2009, Rasmussen et al. 2016). In addition, DEP has used 90% as a threshold of significance for assessment methods in the past. For example, the limestone stream protocol (Botts 2009) and both the wadeable (Chalfant 2012) and semi-wadeable (Shull 2018) macroinvertebrate protocols all use 90% confidence intervals for the determination of precision estimates.

DELISTING CIM ASSESSMENTS

As previously discussed, critical conditions can vary greatly year to year. Therefore, to properly delist impaired waters, reassessment data should encompass conditions similar to those that existed during the original assessment period. For example, if non-attainment of criteria was determined for a stream with frequent exceedances during a “dry” summer with infrequent rain, the reassessment data should also include a “dry” summer. It would be inappropriate to delist the stream using data from a “wet” summer with frequent rain and elevated flows as these are likely to moderate critical conditions.

A waterbody assessed with continuous data should not be delisted with discrete grab samples. Continuous data are temporally more comprehensive than discrete grab samples and should be used to delist a waterbody if a temporally-comprehensive, continuous dataset was the basis of the assessment. A waterbody assessed with discrete grab samples, however, can be delisted with continuous data, as continuous data are fundamentally discrete grab samples collected much more frequently.

LITERATURE CITED

Azevedo, P.A., C.Y. Cho, S. Leeson, and D.P. Bureau. 1998. Effects of Feeding Level and Water Temperature on Growth, Nutrient and Energy Utilization and Waste Outputs of Rainbow Trout (Onchorhynchus mykiss). Aquatic Living Resources 11:227-238.

Botts, W.F. 2009. Wadeable Limestone Macroinvertebrate Assessment Method. Chapter 2, pages 23-42. *In* Shull, D. R., and M. Pulket. (editors). Assessment Methodology for Streams and Rivers. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania.

Chalfant, B.A. 2012. An Index of Biotic Integrity for Benthic Macroinvertebrate Communities in Pennsylvania's Wadeable, Freestone, Riffle-Run Streams. Chapter 2, pages 2-22. *In* Shull, D. R., and M. Pulket. (editors). Assessment Methodology for Streams and Rivers. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania.

Chalfant, B.A. 2013. Discrete Physiochemical Assessment Method. Chapter 3, pages 2-17. *In* Shull, D. R., and M. Pulket. (editors). Assessment Methodology for Streams and Rivers. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania.

Christensen, V.G. 2001. Characterization of Surface-Water Quality Based on Real-Time Monitoring and Regression Analysis, Quivira National Wildlife Refuge, South-Central Kansas, December 1998 Through June 2001. U.S. Geological Survey Water-Resources Investigations Report 01-4248, 28 p.

Christensen, V.G., J.L. Graham, C.R. Milligan, L.M. Pope, and A.C. Ziegler. 2006. Water quality and relation to taste-and-odor compounds in the North Fork Ninnescah River and Cheney Reservoir, south-central Kansas, 1997-2003. U.S. Geological Survey Scientific Investigations Report 2006-5095, 43 p.

Christensen, V.G, P.P. Rasmussen, and A.C. Ziegler. 2002. Real-time water quality monitoring and regression analysis to estimate nutrient and bacteria concentrations in Kansas streams. Water Science Technology 45(9):205-211.

Christensen, V.G., A.C. Ziegler, P.P. Rasmussen, and Xiaodong Jian. 2003. Continuous real-time water-quality monitoring of Kansas streams, in Proceedings of 2003 Spring Specialty Conference on Agricultural Hydrology and Water Quality, May 12-14, 2003, Kansas City, Missouri. American Water Resources Association, Middleburg, Virginia, AWRA Technical Publication Series No. TPS-03-1, compact disc.

Foster, G.M., and J.L. Graham. 2016. Logistic and linear regression model documentation for statistical relations between continuous real-time and discrete water-quality constituents in the Kansas River, Kansas, July 2012 through June 2015. U.S. Geological Survey Open-File Report 2016–1040, 27 p.

Hoger, M.S. 2018. In-Situ Field Meter and Transect Data Collection Protocol. Chapter 4, pages 2-7. *In* Shull, D. R., and M. J. Lookenbill. (editors). Water quality monitoring protocols for streams and rivers. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania.

Hoger, M.S., D.R. Shull, and M.J. Lookenbill. 2018. Continuous Physicochemical Data Collection Protocol. Chapter 4, pages 22-87. *In* Shull, D. R., and M. J. Lookenbill.

(editors). Water quality monitoring protocols for streams and rivers. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania.
Juracek, K.E. 2011. Suspended-sediment loads, reservoir sediment trap efficiency, and upstream and downstream channel stability for Kanopolis and Tuttle Creek Lakes, Kansas, 2008–10. U.S. Geological Survey Scientific Investigations Report 2011–5187, 35 p.
Lee, C.J. 2009. Transport and sources of suspended sediment in the Mill Creek watershed, Johnson County, northeast Kansas, 2006-07. U.S. Geological Survey Scientific Investigations Report 2009-5001, 52 p.
Lee, C.J., P.P. Rasmussen, and A.C. Ziegler. 2008. Characterization of suspended-sediment loading to and from John Redmond Reservoir, east-central Kansas, 2007-2008. U.S. Geological Survey Scientific Investigations Report 2008-5123, 25 p.
Lyons, J., T. Zorn, J. Stewart, P. Seelbach, K. Wehrly, and L. Wang. 2009. Defining and Characterizing Coolwater Streams and Their Fish Assemblages in Michigan and Wisconsin, USA. North American Journal of Fisheries Management 29:1130-1151.
Maloney, K.O., and D.R. Shull. 2015. Sediment and discharge yields within a minimally disturbed, headwater watershed in North Central Pennsylvania, USA, with an emphasis on Superstorm Sandy. Water and Environment Journal 29(3):402-411.
Mau, D.P., A.C. Ziegler, S.D. Porter, and L.M. Pope. 2004. Surface-water-quality conditions and relation to taste-and-odor occurrences in the Lake Olathe watershed, northeast Kansas, 2000-02. U.S. Geological Survey Scientific Investigations Report 2004-5047, 95 p.
Neori, A., and O. Holm-Hansen. 1982. Effect of Temperature on Rate of Photosynthesis in Antarctic Phytoplankton. Polar Biology 1:33-38.
Odum, H.T. 1956. Primary Production in Flowing Waters. Limnology and Oceanography 1:102-117.
Rasmussen, P.P., P.J. Eslick, and A.C. Ziegler. 2016. Relations between continuous real-time physical properties and discrete water-quality constituents in the Little Arkansas River, south-central Kansas, 1998-2014. U.S. Geological Survey Open-File Report 2016–1057, 20 p.
Rasmussen, P.P., J.R. Gray, G.D. Glysson, and A.C. Ziegler. 2009. Guidelines and Procedures for Computing Time-Series Suspended-Sediment Concentrations and Loads from In-Stream Turbidity-Sensor and Streamflow Data. U.S. Geological Survey Techniques and Methods book 3, chap. C4, 53 p.
Rasmussen, T.J., C.J. Lee, and A.C. Ziegler. 2008. Estimation of constituent concentrations, loads, and yields in streams of Johnson County, northeast Kansas, using continuous water-quality monitoring and regression models, October 2002 through December 2006. U.S. Geological Survey Scientific Investigations Report 2008-5014, 103 p.
Rasmussen, T.J., A.C. Ziegler, and P.P. Rasmussen. 2005. Estimation of constituent concentrations, densities, loads, and yields in lower Kansas River, northeast Kansas, using regression models and continuous water-quality monitoring, January 2000 through December 2003. U.S. Geological Survey Scientific Investigations Report 2005-5165, 117 p.

Raven, J.A. and R.J. Geider. 1988. Temperature and Algal Growth. New Phytologist 110:441-461.
Ridgeway, M.S., B.J. Shuter, and E.E. Post. 1991. The Relative Influence of Body Size and Territorial Behaviour on Nesting Asynchrony in Male Smallmouth Bass, Micropterus dolomieui (Pisces: Centrarchidae). The Journal of Animal Ecology 60:665-681.
Shull, D.R. 2018. Semi-wadeable Large River Macroinvertebrate Assessment Method. Pages 2-43 – 2-69. *In* Shull, D. R., and M. Pulket. (editors). Assessment Methodology for Streams and Rivers. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania.
Shuter, B.J., J.A. Maclean, F.E.J. Fry, and H.A. Regier. 1980. Stochastic Simulation of Temperature Effects on First-year Survival of Smallmouth Bass. Transactions of the American Fisheries Society 109:1-34.
Stone, M.L. and J.L. Graham. 2014. Model documentation for relations between continuous real-time and discrete water-quality constituents in Indian Creek, Johnson County, Kansas, June 2004 through May 2013. U.S. Geological Survey Open-File Report 2014–1170, 70 p.
Strickland, J.D.H. 1970. The Ecology of the Plankton off La Jolla, California, In the Period April Through September, 1967. Bulletin of the Scripps Institution of Oceanography, Volume 17.
Trowbridge, P.R., J.S. Kahl, D.A. Sassan, D.L. Heath, and E.M. Walsh. 2010. Relating Road Salt to Exceedances of the Water Quality Standard for Chloride in New Hampshire Streams. Environmental Science and Technology 44:4903-4909.
US EPA. 2005. Guidance for 2006 Assessment, Listing and Reporting Requirements Pursuant to Sections 303(d), 305(b) and 314 of the Clean Water Act.
Valenti, T.W., J.M. Taylor, J.A. Black, R.S. King, B.W. Brooks. 2011. Influence of Drought and Total Phosphorus on Diel pH in Wadeable Streams: Implications for Ecological Risk Assessment of Ionizable Contaminants. Integrated Environmental Assessment and Management 7(4):636-647.
Walters, G., and M. Pulket. 2018. Assessment and Delisting Methods. Chapter 5, pages 1-10. *In* Shull, D.R., and M. Pulket. (editors). Assessment Methodology for Streams and Rivers. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania.
Wehrly, K.E., M.J. Wiley. 2003. Classifying Regional Variation in Thermal Regime Based on Stream Fish Community Patterns. Transactions of the American Fisheries Society 132:18-38.
White, P.A., J. Kalff, J.B. Rasmussen, J.M. Gasol. 1991. The Effects of Temperature and Algal Biomass on Bacterial Production and Specific Growth Rate in Freshwater and Marine Habitats. Microbial Ecology 21:99-118.
Wurts, W.A. 2003. Daily pH Cycle and Ammonia Toxicity. World Aquaculture 34(2):20-21.

CHAPTER 4 PHYSICAL ASSESSMENT METHODS

PHYSICAL HABITAT ASSESSMENT METHOD

Prepared by:

Gary Walters
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2013

Edited by:

Gary Walters
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2017

INTRODUCTION

Physical habitat of the aquatic environment is a critical component of the overall ecological integrity of the aquatic community (Barbour, et al. 1999). Therefore, assessment of the physical habitat is performed in conjunction with biological monitoring of all flowing waters. The instream habitat availability and condition is a major factor in determining the abundance and diversity of benthic macroinvertebrates and fish. Healthy, diverse aquatic communities in streams require a diversity of cover such as boulders, cobble and coarse woody debris (such as logs), interstitial space between cobble and boulder substrate largely free of fine sediment and sand, diverse flow regimes that provide slow and fast moving water as well as shallow and deep pools (Barbour, et al. 1999). Accumulation of fine sediment and sand or deposition of other pollutants such as iron precipitate (yellow boy) can reduce or eliminate cover, interstitial space and deep pools degrading the habitat and impairing the ability of the stream or river to support healthy aquatic communities.

To assess ALU of flowing surface waters, DEP includes assessments of the physical habitat, whenever possible. To use this method for assessment determination purposes data collection must follow applicable protocols established in the Monitoring Book (Shull and Lookenbill 2018). There are two habitat assessment protocols based on flow regimes related to gradient or slope. The DEP habitat assessment for high gradient streams and semi-wadeable rivers (waters dominated by riffle-run habitat) is based on the habitat assessment published in Plafkin et al. (1989). The DEP method is a revision of the Plakin et al. (1989) method, which had undergone several iterations during the 1990s. This habitat evaluation uses a twelve parameter – 20-point scoring method. The method for low-gradient streams and rivers (waters that lack riffles) is based on the Habitat Assessment and Physicochemical Parameters described in Barbour, et al. (1999). The DEP assessment uses nine of the ten parameters of Barbour, et al. (1999). More information on the data collection aspects of these parameters are found in Chapter 4 of DEP's Monitoring Methodology for Streams and Rivers (Shull and Lookenbill 2018).

AQUATIC LIFE USE ASSESSMENT

Qualitative Method for High Gradient Streams and Rivers

Wadeable Streams

The threshold for ALU assessment impairment for high gradient riffle/run dominated wadeable (<1000 mi^2) streams is a total habitat score of 140 or less. Certain instream and riparian area habitat parameters are strong predictors of habitat degradation leading to ALU impairment, and as a result, these parameters alone may warrant independent assessment decisions. These parameters are embeddedness, sediment

deposition, condition of banks, and bank vegetative protection. The impairment threshold for the parameters of embeddedness + sediment deposition, or condition of banks + bank vegetative protection is a total score of 24 or less for either combination.

Semi-wadeable Rivers

Habitat assessments are required with each semi-wadable survey, but certain habitat parameters (e.g., riparian vegetation zone width) are difficult to measure as river size increases. All 12 parameters are recorded when conducting habitat assessments in semi-wadeable rivers, but instream parameters such as instream cover, epifaunal substrate, and embeddedness are the most reliable habitat indicators. These three instream habitat measurements can be summed to provide a possible range of 0 (indicating worst possible instream conditions) to 60 (indicating best possible instream conditions) points. Instream habitat totals that score 30 or less are an indication of instream habitat impairment.

Qualitative Method for Multihabitat/Low Gradient Streams and Rivers

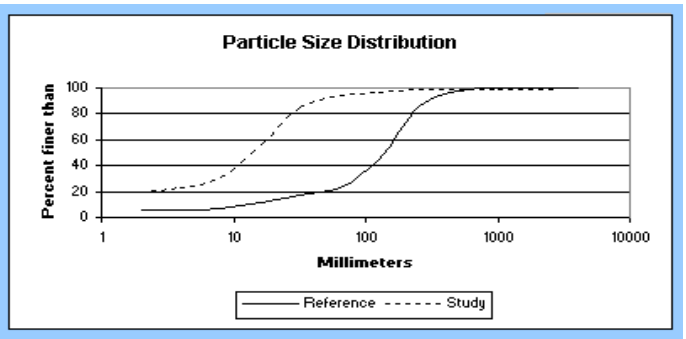
The threshold for ALU assessment impairment threshold for qualitative physical habitat of multihabitat/low gradient wadeable streams and rivers is 105 or less. Certain instream and riparian area habitat parameters are strong predictors of habitat degradation leading to ALU impairment and as a result these parameters alone warrant independent assessment decisions. These parameters are pool substrate characterization, sediment deposition, bank stability and bank vegetative protection. The impairment threshold for the parameters of pool substrate characterization + sediment deposition or bank stability + bank vegetative protection is a total score of 20 or less for either combination.

Quantitative Method for Stormwater Impacted Habitat

For stormwater-impacted sites where a pebble count analysis was conducted, collected data are plotted on graph paper or entered into Microsoft Excel spreadsheets and plotted electronically (Figure 1), as cumulative percentages for both reference and study streams. Particles 8 mm or smaller are of primary concern since they should have the most biological significance and are most likely to smother macroinvertebrate and fish spawning habitat. Reference streams should have no more than 15 percent of particles smaller than 8 mm. Impaired reaches, in general, are study streams with $\geq$35 percent of particles smaller than 8 mm. This threshold may be higher for certain types of streams, such as those with low gradient.

Cumulative Distribution		
Class	Reference	Study
2	6	20
4	6	23
8	7	32
16	12	57
32	18	86
64	23	94
128	45	97
256	87	99
512	98	99
1024	100	99
2048	100	99
4096	100	100

Histogram		
Class	Reference	Study
2	6	20
4	0	3
8	1	9
16	5	25
32	6	29
64	5	9
128	23	3
256	42	2
512	11	0
1024	3	0
2048	0	0
4096	0	2



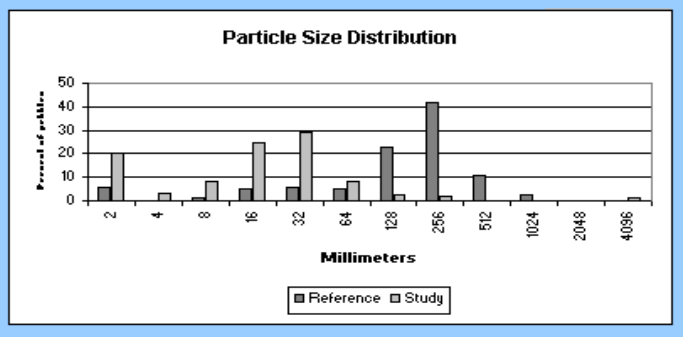


Figure 1. Example analysis of pebble count data.

LITERATURE CITED

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling. 1999. Rapid Bioassessment Protocols for Use in Wadeable Streams and Rivers: Periphyton, Benthic Macroinvertebrates and Fish, Second Edition. EPA 841-B-99-002. United States Environmental Protection Agency; Office of Water. Washington, D.C.

Bevenger, G. S., and R. M. King. 1995. A Pebble Count Procedure for Assessing Watershed Cumulative Effects. RM-RP-319 United States Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, CO.

Plafkin, J.L, M.T. Barbour, K.D. Porter, S.K. Gross, and R.M. Hughes. 1989. Rapid Bioassessment Protocols for Use in Streams and Rivers: Benthic Macroinvertebrates and Fish. EPA/440/4-89-001. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

Shull, D. R., and M. J. Lookenbill (editors). 2018. Water Quality Monitoring Protocols for Streams and Rivers. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

CHAPTER 5 ASSESSMENT DETERMINATION AND DELISTING METHODS

Prepared by:

Gary Walters and Molly Pulket
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2015

Edited by:

Gary Walters and Molly Pulket
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2017

ASSESSMENT DETERMINATIONS

To meet the objective of creating accurate and precise determinations using the methods detailed above, DEP's assessments are conducted on a segment-by-segment basis of the National Hydrologic Dataset (NHD) flowline layer in a DEP Geographic Information System (GIS) application. Unlike most states that assess whole watersheds probabilistically, DEP conducts a statewide census primarily using targeted monitoring to identify individual stream reaches as attaining or impaired. This results in more detailed and accurate assessments of the waterbody, significantly reduces the need to revisit sites, and allows DEP to focus resources on only those segments of a waterbody that are not meeting applicable WQS, discussed below.

Using the Methods

Independent Applicability

The assessment methods detailed in Chapters 2–4 constitute the current "decision rules" DEP uses when making assessments. These methods are understood to be independently applicable when making assessment determinations. This is based on USEPA guidance, which mandates that all assessment methods must be evaluated on a stand-alone basis (USEPA 2002). One exception to the independent applicability rule is with discrete and continuous physiochemical sampling methodologies for parameters that can be measured by both methods (e.g., pH, dissolved oxygen). Given that continuous datasets are simply more robust than discrete datasets, continuous datasets can be used to reassess or delist assessment determinations based on discrete datasets. Because of a continuous dataset's ability to better capture important daily and seasonal variations, extreme caution should be used in reassessing or delisting a continuous assessment using discrete data. If possible, continuous datasets are preferred for parameters where this technology is available.

Narrative Criteria

The assessment methods detailed in Chapters 2–4 may also be used to assess the narrative criteria found in 25 Pa. Code §93.6(a) and (b). For narrative criteria assessments, these methods may be used in a "weight of evidence" approach to come to a final assessment determination.

Outside Data

In addition to the data DEP collects, DEP readily accepts and values all data from outside agencies and the public for use in the making assessments. However, different data types and levels of quality assurance determines how exactly those data are used. DEP's tiered data acceptance strategies follow the same general tiered framework as described by the Chesapeake Bay Monitoring Cooperative's Prioritization Report

(Chesapeake Bay Monitoring Cooperative 2017). Tier 1 data is generally defined as educational or environmental screening data that has known quality and a study plan, but does not follow DEP or USEPA quality assurance plans. These data will not be used for assessment determination purposes, but can be used by DEP to highlight areas of interest for future monitoring efforts. Tier 2 data have clearly defined quality assurance plans and procedures, but may not have followed DEP monitoring protocols described in the Water Quality Monitoring Protocols for Streams and Rivers (Shull and Lookenbill 2018). These data may not be used for assessment determination purposes, but can be used for other purposes such as trend or performance analysis. Tier 3 data are assessment level data that have approved quality assurance plans, follow appropriate study designs, and follow DEP monitoring protocols (Shull and Lookenbill 2018). Individuals seeking to provide DEP with Tier 3 data should also be trained and audited by DEP staff before submitting data.

Some interstate surface waters of Pennsylvania have water quality regulation through compact commissions. These waters are comprised of the Ohio River and Delaware River mainstems. The Ohio River Valley Water Sanitation Commission (ORSANCO) and the Delaware River Basin Commission (DRBC) have established methodology, in consultation with DEP and other compact states, to assess the attainment of WQS in compliance with CWA Section 305(b) and provide those results to the states and USEPA. DEP reviews these data and results to make appropriate assessments for both Section 303(d) and 305(b) in the Integrated Report. These assessments apply to the protected uses of the Ohio River from the confluence of the Allegheny and Monongahela Rivers to the PA/WV state line and for the West Branch Delaware River at the PA/NY state line and the mainstem Delaware River from the confluence of the East and West Branches through the Delaware estuary to the PA/DE state line.

Sample Design Considerations

Thoughtful study design and execution are critical to assuring water quality sampling efforts provide the information necessary to make assessment decisions. More information on acceptable sampling design procedures are found in DEP monitoring protocols (Shull and Lookenbill 2018). For assessment determination purposes, DEP utilizes both targeted and probabilistic sampling designs. However, DEP believes the targeted “judgment-based” sampling design is the most suited method to assess WQS and uses. Targeting sampling not only focuses in on sources and causes of potential impairment, it also delimits the spatial effect of the impact. This translates into more accurate assessments. In addition, properly implemented targeted sampling provides information that is necessary if a TMDL is developed. Probabilistic sampling designs, can also be useful for assessment determination purposes, especially when waterbodies lack significant environmental stress or are rather homogeneous in land

use. In these cases, probabilistic sampling can provide accurate information without overextending resources. When a probabilistic sampling design is employed, statistical analysis is conducted to determine miles of attaining and impaired stream miles. The results are then translated into assessment units for the Integrated Report. If probabilistic results return a significant mix of assessment decisions, then the watershed may be revisited using a targeted sampling design to obtain more detailed information for assessments.

Requirements for Making Assessments

Assessments will be completed with data that has been collected using appropriate sampling design, see Monitoring Book (Shull and Lookenbill 2018). Sampling sites and locations are positioned to account for changes in water quality due to influences such as major tributaries, point and nonpoint source impacts, land use changes, soil characteristics, and geology. Additional samples are collected at the limits of these changes to effectively "bracket" potential sources of water quality differences. The minimum length of any assessment unit is typically ½ mile. Any assessment unit less than ½ mile may be considered a localized impact and likely will not be reported in the Integrated Report. There is no set maximum assessment unit length; however, the size is limited by the DEP GIS application to efficiently save and return results from the database. Approximately 55 segments of the NHD flowline is recommended as a maximum assessment unit length to avoid GIS application issues.

Decision Framework

DEP will implement the following framework when evaluating monitoring data in the use assessment decision process. The details of this appraisal process may vary from application to application based on the unique characteristics and contexts of each situation. However, DEP will follow this process as often as possible to maintain consistency in the use assessment decision process and so that interested stakeholders can clearly see how DEP evaluates data for assessments. The decision framework aims to document and communicate each step of the decision process in a clear, consistent manner addressing the study designs, data quality, data analysis, assumptions, uncertainties, and consequences associated with each use assessment decision. DEP attempts to be as concise as possible within this framework while not compromising adequate discussion of critical issues influencing the decisions.

(1) **Describe monitoring effort.** Describe the waterbody and the watershed, including basin size, land uses, geologies, and other characteristics. Discuss any germane history and context pertaining to the monitoring effort. To the extent possible, describe the motivations and intentions of the monitoring effort, including the individuals and organizations involved

as well as the intended use of the information collected. Clearly state study goals. Describe and map monitoring locations. Include any photographs.

(2) **Check data quality.** Evaluate any study plans and objectives, including sampling plan design details such as recordkeeping, data management, training, sampling techniques, and analytical methods. Check data for typos and other anomalies. Document non-detects and censored data.

(3) **Gather information on likely sources of variation.** At a minimum, this information will typically include characterization – and quantification where possible – of tributary locations, upstream discharges, geologies, and land uses. Potential sources of this information include stream gages, climatological records, and discharge monitoring reports. Include maps, figures, and diagrams as needed. Discuss relevant physical, chemical, and biological processes and other potential sources of variation for the parameter(s) of concern. Address context-specific considerations (e.g., dams).

(4) **Explore data.** Perform various graphical analyses (e.g., histograms, probability distribution functions, boxplots, time-series plots, scatterplots with likely sources of variation, LOWESS) to visually explore and illustrate data characteristics. Document summary statistics (e.g., minimum, maximum, mean, median, standard deviation).

(5) **Evaluate data representativeness.** Evaluate how representative samples are of unmonitored conditions, mindful of the sampling plan design (e.g., sample collection frequency, locations, timing, targeting) and the likely sources of variation with special attention to any critical sampling times and locations. Consider if the system is likely to be spatially well-mixed at monitoring location(s) and how quickly conditions are likely to change in time.

(6) **Describe the relevant standards.** Identify which criteria are being evaluated and the uses to which they apply. Describe how the parameters of concern impact the protected use (i.e., exposure pathways, detrimental effects) being assessed. Review the associated regulatory language including any relevant criterion rationale documentation.

(7) **Apply appropriate analytical procedures.** Select and apply appropriate analytical techniques, mindful of the sampling plan design, monitoring objectives, and the relevant criteria, parameters, and context. State and verify any assumptions associated with each analytical technique. Evaluate decision error rates, if applicable. For hypothesis tests, evaluate null hypothesis choice. Discuss the frequency, duration, and magnitude of any criteria violations.

(8) **Consider other sources of relevant use assessment information.** Additional sources of information may include: previous or concurrent monitoring efforts; data from water supply intakes; biological surveys; and discharge monitoring reports.

(9) **Evaluate all relevant lines of evidence.** Bring together the previous steps into a narrative that addresses contextual data interpretations, possible counter arguments, alternative decision choices, and decision consequences, including evaluation of decision error consequences. Explicitly address any policy ramifications if applicable.

(10) **Decide.** Decide what to do with the dataset and waterbody in question. At a minimum, each decision will include placing the waterbody in one of the Integrated Report categories.

Natural Conditions Exception

Natural quality is defined in 25 Pa. Code § 93.1 as "The water quality conditions that exist or that would reasonably be expected to exist in the absence of human related activity." In accordance with the provisions of Pennsylvania's WQS, waters that have naturally occurring pollutant concentrations, or "natural quality," that prevent the attainment of an established use will not be assessed as impaired, if it can be demonstrated that anthropogenic sources do not cause or contribute significantly to the non-attainment and the pollutant(s) of concern are generated by natural processes.

Reassessment of Previously Assessed Waters

DEP completed the first statewide ALU assessment of wadeable waters (SSWAP) in 2006 and began reassessment with new methods during 2006 in the eastern regions of the state. The primary focus of reassessment is the attaining waters from the first statewide ALU assessment. The current assessment methodology is more rigorous than the SSWAP method and, as a result, the reassessment of attained waters is to confirm that these waters are attaining ALU. The goal is to reassess all SSWAP ALU attaining waters by 2025. Reassessment of impaired waters is a lower priority unless

conditions have changed as a result of restoration or implementation of a TMDL or it is believed the water may have been listed in error. Reasons to reassess include confirmation of the original source and cause determination and collection of additional data necessary for TMDL development or alternative restoration plans. Following implementation of the TMDL targets and other restoration plans, reassessment should occur after sufficient time has passed to allow for recovery. In general, reassessment following implementation should occur five years after restoration activities have been completed, and if full restoration to WQS has not occurred, reassessment should occur at five-year intervals.

Other use assessments include Fish Consumption, Recreation and Potable Water Supply. Reassessment of these uses should occur within 10 years and identify any changes in status. Assessment of recreation use is ongoing and has only been completed on approximately 40% of the waters of Pennsylvania. The priority for assessment is on assessing the waters that have not yet been assessed as well as reassessing waters that were first assessed prior to 2008.

INTEGRATED REPORT CATEGORY ASSIGNMENT

Chapter 1 introduced and described the Integrated Report Categories. This section describes the assignment of a waterbody segment to one of the Categories based upon the results of the assessment. Categories 1 and 2 are for waters attaining protected uses. Waterbody segments that have been assessed and are attaining all uses are assigned to Category 1. Waterbody segments that have been assessed and are attaining at least one use are assigned to Category 2. Category 3 is reserved for waters that are not assessed for any uses due to insufficient information to complete an assessment.

Impaired waters are assigned to Category 4 or 5. Waters assigned to Category 4 are impaired for one or more uses; however, these waters do not require a TMDL to be developed. Category 4 is comprised of 3 subcategories: 1) Category 4a applies when a TMDL has been completed and approved by USEPA; 2) Category 4b applies when a use impairment caused by a point source pollutant is being addressed by the state through other pollution control requirements and a schedule of compliance. and 3) Category 4c applies when a use is impaired, but the impairment is not caused by a pollutant (i.e., Flow Alterations, Habitat Modification, Water/Flow Variability and Filling and Draining).

Waters assigned to Category 5 are impaired by pollutants for one or more uses and require the development of a TMDL. Category 5 has one subcategory, 5alt, that is comprised of waters that have been identified for water quality restoration through an alternative approach before a TMDL is completed.

DELISTING IMPAIRMENT CAUSES

When conditions improve in impaired waters it is possible to delist a cause or causes of impairment from the CWA § 303(d) list. In addition, if a cause of impairment is no longer appropriate, it can be removed despite the waterbody remaining impaired for other sources or causes. Any removal of a cause of impairment on the 303(d) list is subject to EPA review and approval and must come with reasoning and data to support the change.

Waters impaired by pollutants for one or more uses are listed on Category 5 of the Integrated Report. Impaired waters can also be placed on Category 4 if they do not require a TMDL. The Integrated Report Categories are discussed in detail in the Purpose section of the Introduction Chapter of this book. Generally, the term “delisting” describes the process of moving a waterbody from Category 4 or 5 (impaired waters) to Category 1 or 2 (attaining waters). A specific cause can also be delisted from a waterbody; however, the waterbody will remain on Category 4 or 5 due to another cause(s) of impairment. There are multiple reasons to delist a waterbody. A modified list of USEPA’s reasons are in Table 1 below.

Table 1. USEPA Delisting Reasons

	Delisted Reason	Type of Delisting
1	Discharge is in compliance	(4b); Attainment
2	Applicable WQS attained, due to new assessment method.	Attainment
3	Applicable WQS attained, due to change in WQS	Attainment
4	Applicable WQS attained, due to restoration activities.	Attainment
5	Applicable WQS attained; original basis for listing was incorrect.	Error; Attainment
6	Applicable WQS attained; reason for recovery unspecified	Attainment
7	Applicable WQS attained; based on new data	Attainment
8	Refinement of terminology of listing cause	Change; Attainment
9	WQS no longer applicable	Change; Attainment
10	Water determined to not be a water of the state	Change; Attainment
11	Data and/or information lacking to determine WQ status; original basis for listing was incorrect	Error

Requirements

Generally, it takes the same or greater level of data rigor used to make the impairment determination to delist. To remove a waterbody from Category 4b, documentation must be provided showing the facility is in compliance with their permit conditions and/or their discharge is no longer the cause of impairment. This documentation could include one year of Discharge Monitoring Reports (DMRs) with data showing the assessed use is attaining, or data showing there is a different cause for the impairment.

To justify reasons 2-7 in Table 1, an assessment must be conducted to show the waterbody is now meeting its use. The data requirements to demonstrate these improvements are found in Table 2. The applicable data and a detailed map displaying the waterbody must be provided to DEP. Appendix B contains an example map of a stream delisting and details the information that should be depicted on the map. For ALU assessments, the macroinvertebrate station(s) and the new attaining IBI score(s) must be displayed. Recreational use assessments should map the station(s) and display the attaining geometric mean(s). If an assessment is based on chemistry, the data showing attainment must be provided. Any other pertinent information or data to justify the delisting should also be provided.

Table 2. Data Requirements for Delisting 303(d) Waters

Assessed Use	Delisting Data Requirements
Aquatic Life - macroinvertebrate	Aquatic macroinvertebrate data, collected using DEP data collection protocols (Shull and Lookenbill 2018), that generates an IBI score above the attainment benchmark set by the sampling protocol. Multiple stations are required to bracket land use changes, nonpoint and point source influences, and any other influences that could affect water quality within the potential delisted waterbody.
Aquatic Life - chemistry	Macroinvertebrate sample results are preferred. Chemistry results must demonstrate that the applicable criterion is being met 99% of the time as set forth in 25 Pa. Code Chapters 93 and 96. For other criteria, the samples must be collected for analysis consistent with requirements outlined in Chapter 3 of this book.
Recreation	The geometric mean of all 5 samples must be below the criterion for fecal coliforms (200 cfu/100 ml) and no single sample above 400cfu/100 ml as described in the Bacteriological Assessment Method for recreational use section of this book.
Potable Water Supply	Sampling should target the critical period when criteria exceedances are expected. Samples must be collected at the point of withdrawal prior to the treatment process. Results must demonstrate that the applicable criterion is being met 99% of the time as set forth in 25 Pa. Code Chapters 93 and 96.
Fish Consumption	Fish tissue results showing the improved contamination level and the recommended fish advisory change.

LITERATURE CITED

Chesapeake Monitoring Cooperative. 2017. Prioritization Report: How volunteer and nontraditional monitoring can help fill data gaps in the Chesapeake Bay Watershed.

Shull, D. R., and M. J. Lookenbill (editors). 2018. Water Quality Monitoring Protocols for Streams and Rivers. Pennsylvania Department of Environmental Protection, Harrisburg, Pennsylvania.

USEPA. 2002. Consolidated assessment and listing methodology: Toward a compendium of best practices. U.S. Environmental Protection Agency, Office of Wetlands, Oceans, and Watersheds. Washington, D.C.

CHAPTER 6 SOURCE AND CAUSE DETERMINATION METHODS

GENERAL SOURCE AND CAUSE DETERMINATION METHODS

Prepared by:

Gary Walters
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2015

Edited by:

Gary Walters
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2017

INTRODUCTION

Once it is determined through the assessment decision process (Chapter 5) that one or more uses are impaired, the next steps are to identify the source and cause of the impairment. Section 303(d) of the CWA and 40 CFR § 130.7 requires listing those waters impaired by pollutants that will not achieve WQS after the application of technology-based effluent limitations, of more stringent effluent limitations required by state requirements and of other pollution control requirements (e.g., best management practices). This is an important component of the decision process as source and cause determinations distinguish waters that require the development of a TMDL versus waters that require some other method of restoration of the WQS.

Sources of impairment can be divided into two general categories, point and non-point sources. In general, point sources are discharges from pipes or discrete conveyances. A “point source discharge” is defined in 25 Pa. Code § 93.1 as a pollutant source regulated under the National Pollution Discharge Elimination System (NPDES). Section 93.1 defines a “nonpoint source” as a pollution source which is not a point source discharge. An example of nonpoint sources are those sources of runoff coming over the land surface.

As with sources, causes of impairment can be assigned to two general categories; pollutant and pollution. “Pollutant” is defined in 25 Pa. Code Section 92a.2 as a contaminant or other alteration of the physical, chemical, biological or radiological integrity of surface water that causes or has the potential to cause pollution as defined in section 1 of The Clean Streams Law (35 P.S. §681.1). Examples of pollutants are substances such as iron, pesticides, pathogens, or sediment that prevent the attainment of uses. For the purpose of listing waters pursuant to section 303(d) of the CWA, “pollution” is described as habitat modifications and impacts related to water volume and/or flow. Any water impaired by a pollutant, and listed in Category 5, requires the development of a TMDL. Waters impaired by pollution do not require a TMDL and may be restored through other restoration methods.

SOURCE AND CAUSE DETERMINATION METHOD

Source

The method to determine source of impairment relies heavily on a thorough reconnaissance and knowledge of the watershed that is being assessed. Prior to monitoring of the watershed, the investigator compiles all known point and nonpoint sources of pollution. Field reconnaissance should be conducted in addition to a desktop reconnaissance of aerial photography to identify any potential sources of pollution. This information is then used in conjunction with sampling locations and data to assign the

most probable source of the impairment. A full list of potential sources is provided in Appendix A.

Cause

Most causes of impairment can be determined in a similar manner as the source determination, thorough reconnaissance and knowledge of the watershed coupled with knowledge of ecological and biological responses to pollution and the applicable narrative and numeric water quality criteria. The causal identification process includes identifying all probable causes, evaluation of biological, physical and chemical data and the observed response in the stream. Many causes of impairment will be obvious to the investigator such as excess sediment causing siltation impairments or metals precipitate covering the stream bed. Chemical impairments are determined through the analysis and evaluation of discrete and continuous water chemistry data and the applicable water quality criteria. A full list of potential causes and their context is provided in Appendix A.

There are instances when cause determination will require additional monitoring following specific protocols and a more structured casual identification process. This process may rely on a weight of evidence approach from multiple lines of evidence to arrive at the cause of an observed impairment. This will typically be the case when interpreting the narrative criteria at §93.6(a) and (b).

ACID PRECIPITATION SOURCE AND CAUSE DETERMINATION METHOD

Prepared by:

Martin Friday
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2000

Edited by:

Dustin Shull
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2017

INTRODUCTION

Acid precipitation impairment is difficult to detect using only biological collection protocols and assessment methods, particularly when the impairment is due to episodic acidification. Small, forested, headwater streams with low alkalinity are generally unproductive. Low numbers of benthic macroinvertebrates with relatively low diversity are frequently observed in these types of streams. The collected organisms are also generally sensitive to organic pollution, so the benthic community will normally be dominated by taxa with low Hilsenhoff scores. Depending on the season and recent precipitation history, field water chemistry measurements will document the low alkalinity, but may fail to detect a low pH event. Assuming that no major component of the benthic community is missing (e.g. mayflies), the biological assessment method may lead to the potentially erroneous conclusion of no biological impairment.

Current biological assessment protocols may fail to identify acid precipitation impacts because they do not assess the fish community. A fish community may slowly decline as year classes are lost to episodic acidification and sensitive species are eliminated from a given stream reach, but this trend may go unnoticed if the benthos alone is used to detect biological impairment. Macroinvertebrates are better able to recolonize stream reaches than fish due to the shorter time between successive generations, and may not exhibit the same symptoms as fish communities when challenged by episodic acidification. Thus, a relatively healthy macroinvertebrate community may not indicate that a healthy fish community is present, and therefore may not give a complete indication of the stream's biological impairment due to acid precipitation.

Macroinvertebrate metrics provide only an indirect indication of potential acid precipitation impairment. When abundance and diversity are obviously low, community composition is abnormal (e.g., no mayflies), and field alkalinity and pH are both low (alkalinity <5 ppm; pH <5.0), macroinvertebrate assessment methods can support a decision of biological impairment due to acidification. When these conditions are not observed and acid impairment is suspected, a more detailed investigation may be warranted to conclusively identify an acid precipitation problem. Other evidence that may also trigger a detailed follow-up survey would include anecdotal information indicating a decline in a fishery; cessation of trout stocking by PFBC due to poor survival; and fisheries data documenting population changes and species loss over time.

ACID PRECIPITATION SOURCE AND CAUSE DETERMINATION

The best way to document acid precipitation impairment is to collect water samples during spring snowmelt or storm events that document conditions known to be lethal to fish. The most critical measurements are pH and dissolved aluminum. Low pH and high concentrations of dissolved aluminum have been linked to high fish mortality in studies of episodic acidification (Fiss and Carline 1993, Sharpe and Drohan 1999). Dissolved inorganic monomeric aluminum is the aluminum species most strongly correlated to fish mortality, but analysis for this form of aluminum is more complicated than for the more traditional “total dissolved aluminum” concentration. Total dissolved aluminum concentrations obtained via the standard method of field filtration through a 0.45 µm filter are only weakly correlated with lethal response in fish, and are of limited value for identifying impairment due to acidification. An alternate dissolved aluminum analysis that correlates well with inorganic monomeric aluminum concentrations and is useful for identifying acid impairment is one conducted on water samples filtered through a 0.1 µm filter (Van Sickle et al. 1996).

Follow-up sampling to detect acid impairment should be concentrated during storm events and periods of heavy snowmelt. Ideally, water samples should be collected during peak flows to characterize worst-case conditions. Grab samples collected during high flow events should be adequate for most follow-up surveys. A low flow sample may be collected for comparison, but is not necessary. Standard Analysis Code 910 (SAC 910) has been established for use when investigating potential acid precipitation problems. The analyses conducted as part of SAC 910 are listed in Table 1. The most important parameters for identifying acid precipitation impairment are pH and dissolved aluminum concentrations (with 0.1 µm filtration). Prior to shipping the sample to the lab, a 500 ml aliquot must be filtered through a 0.1 µm filter.

If the high flow sample documents stressful conditions (i.e. low pH and high dissolved aluminum levels), then some degree of biological impairment is likely. Elevated dissolved aluminum concentrations (>150 µg/L) and low pH (<5.8) can be lethal to brook trout, depending on duration of exposure. When a stream survey documents pH depression and dissolved aluminum levels above 150µg/L (after 0.1 µm filtration), it is appropriate to consider the stream to be biologically impaired due to acid precipitation. For 303d list reporting purposes, pH will be the cause of impairment. Acid precipitation is the source on impairment.

Table 1. Analyses included under the Standard Analysis Code for acid precipitation samples (SAC 910).

Test Description	Reporting Units
Specific Conductance	µmhos/cm
pH	pH units
Alkalinity total as $CaCO_3$	mg/L
Acidity, mineral as $CaCO_3$	mg/ L
Calcium, total	mg/ L
Magnesium, total	mg/ L
Chloride, total	mg/ L
Sulfate, total	mg/ L
Iron, total	µg/ L
Manganese, total by trace elements	µg/ L
Aluminum, total by trace elements	µg/ L
Aluminum, dissolved 0.1µm filter	µg/ L

LITERATURE CITED

Fiss, F. C., and R. F. Carline. 1993. Survival of brook trout embryos in three episodically acidified streams. Transaction of the American Fisheries Society 122:268–278.

Sharpe, W. E., and J. R. Drohan (editors). 1999. The Effects of Acidic Deposition on Aquatic Ecosystems in Pennsylvania. Proceedings of the 1998 PA Acidic Deposition Conference, Vol II. Environmental Research Institute, University Park, Pennsylvania.

Van Sickle, J. J., J. P. Baker, H. A. Simonin, B. P. Baldigo, W. A. Kretser, and W. E. Sharpe. 1996. Episodic acidification of small streams in the northeastern United States: fish mortality in field bioassays. Ecological Applications 6:408–421.

EUTROPHICATION CAUSE DETERMINATION METHOD FOR SMALL STREAMS (≤50 Mi^2 DRAINAGE AREA)

Prepared by:

Charles McGarrell
Pennsylvania Department of Environmental Protection
Office of Water Programs
Bureau of Clean Water
11th Floor: Rachel Carson State Office Building
Harrisburg, PA 17105

2018

INTRODUCTION

The USEPA describes nutrient pollution as one of America's most widespread, costly and challenging environmental problems. Within the context of nutrient pollution of streams, the term eutrophication refers to the process by which elevated nutrient levels (especially phosphorus and nitrogen) stimulate the growth of algae and/or aquatic plants, and alters the quantity and quality of organic matter available as food for aquatic organisms. In addition to modifying the trophic structure of stream ecosystems, eutrophication can alter physical habitat conditions, stimulate the growth of toxin-producing algae, and can produce large daily (diel) fluctuations in dissolved oxygen (DO) and pH that, in some cases, fall below or rise above levels protective of aquatic life.

Over the past several years, DEP staff have collected nutrient; benthic chlorophyll-a; continuously monitored DO, pH, and water temperature; and benthic macroinvertebrate community data from small streams statewide. The technical background behind the development of the ECD protocol can be found in McGarrell (2018). The conceptual model shown in Figure 1 illustrates the cause/response relationships linking nutrient enrichment to stream biological integrity that was used as a framework for developing this Eutrophication Cause Determination (ECD) protocol. The ECD protocol provides a method for quantitatively assessing the impact of nutrient enrichment on Pennsylvania's small streams (drainage area ≤ 50 mi^2) The intended use of the ECD protocol is for determining if eutrophication is a cause of ALU impairment, under the context of nutrient enrichment, after DEP's Wadeable Freestone Riffle-Run Stream Macroinvertebrate Assessment Method indicates impairment.

The ECD protocol uses a multiple lines of evidence approach for determining if eutrophication is a cause of ALU impairment. Stream ecosystem parameters used in the protocol include: diel DO swing characteristics, water quality criteria for DO and pH, benthic chlorophyll-a concentration, diel DO swing-diel pH swing relationships, and diel DO swing- diel water temperature swing relationships. A graphical summary of the ECD Protocol is shown in (Figure 2).

THE EUTROPHICATION CAUSE DETERMINATION (ECD) PROTOCOL

Data Collection and Analysis

Baseflow (non-storm event) water column total phosphorus (TP) and total nitrogen (TN) samples are to be collected for laboratory analysis when continuous data sondes are first deployed, during each subsequent data sonde maintenance event (approximately monthly), and when sondes are retrieved. Water column nutrient samples are to be collected and processed in accordance with Shull (2013).

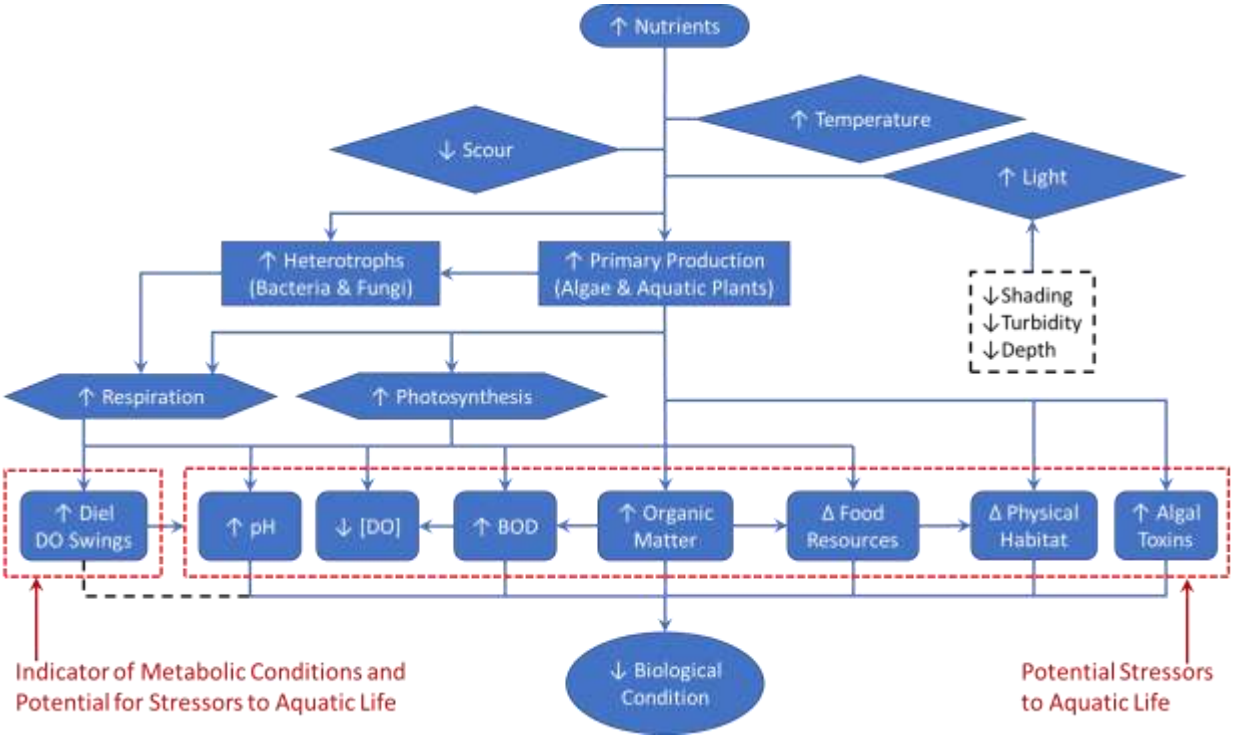


Figure 1. Conceptual model of how nutrient enrichment and eutrophication impact stream biological condition (modified from Heiskary and Bouchard (2015), Minnesota Eutrophication Criteria for Streams and Rivers).

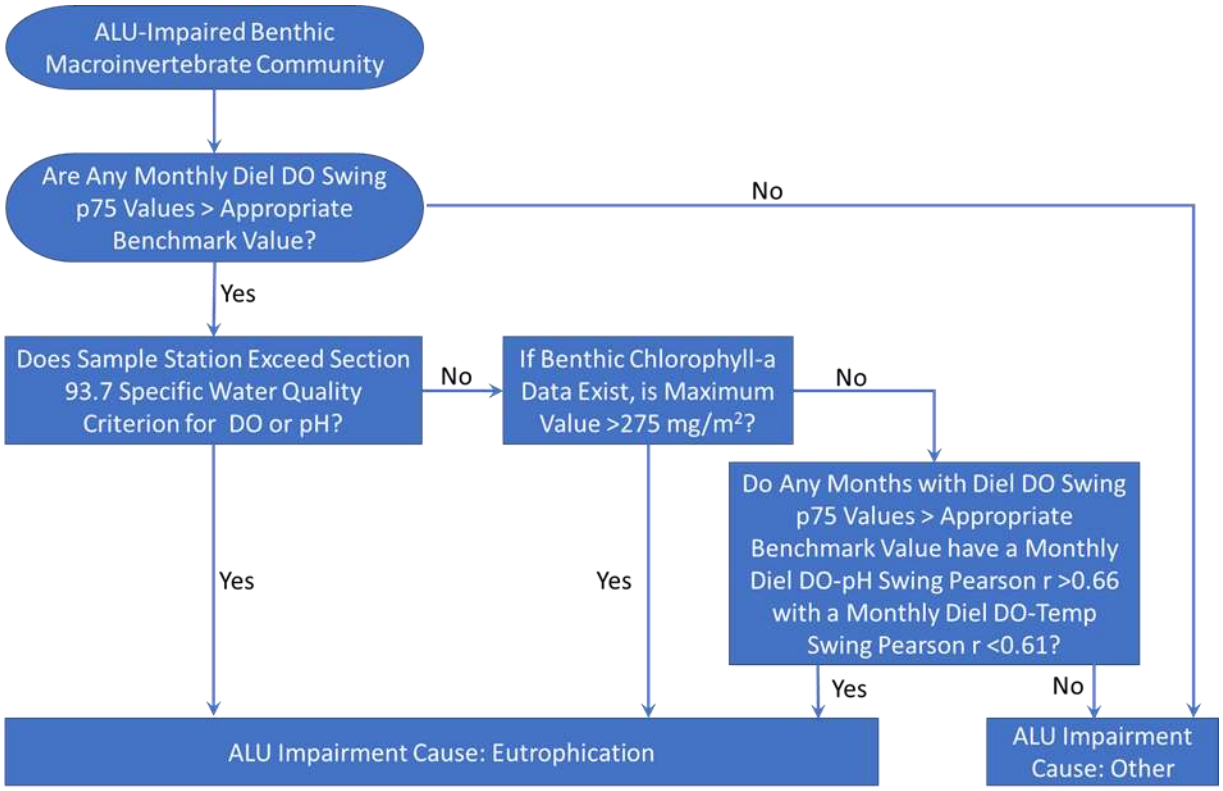


Figure 2. Graphical summary of the Eutrophication Cause Determination Protocol.

Photo-documenting or otherwise noting field observations of primary production levels (algal and/or aquatic macrophyte growth) at continuously monitored sample stations is an important part of the field data collection component of the ECD Protocol. Photographs that clearly show in-stream primary production levels should be taken on each sample station visit. At least one benthic periphyton sample should be collected at each sample station while the data sonde is deployed. Benthic periphyton samples are to be collected using DEP's Quantitative Benthic Epilithic (QBE) Periphyton Sampling Method (Butt 2017), and efforts should be made to collect samples when primary production rates appear to be relatively high, based on professional judgement and visual observations made during routine data sonde maintenance events.

Water column nutrient data and information pertaining to primary production levels can be very helpful when trying to ascertain the extent of nutrient enrichment at a specific reach of stream. In some cases, water column nutrient levels are excessively high and indicative of a nutrient-enriched system. However, some nutrient-enriched, highly productive stream reaches have very high diel DO swings that are strongly correlated with daily pH swings, but have very low water column phosphorus and nitrogen concentrations due to algal uptake of nutrients. In these cases, where elevated levels of primary production occur under seemingly low levels of nutrient enrichment, benthic chlorophyll-a concentration values and photo-documentation of excessive algal or aquatic macrophyte growth become even more important.

Continuously monitored DO, pH, and water temperature data are collected between March and October and are collected, graded, and approved for use in accordance with DEP's Continuous Physicochemical Data Collection Protocol (Hoger et al. 2017). Diel DO, pH, and water temperature swing values are calculated for days in which continuous data are collected over at least 75% of the day (e.g., a minimum of 36 readings at ½ hour intervals). Diel swing values are calculated as the difference between the maximum and minimum values recorded on a given day (Figure 3).

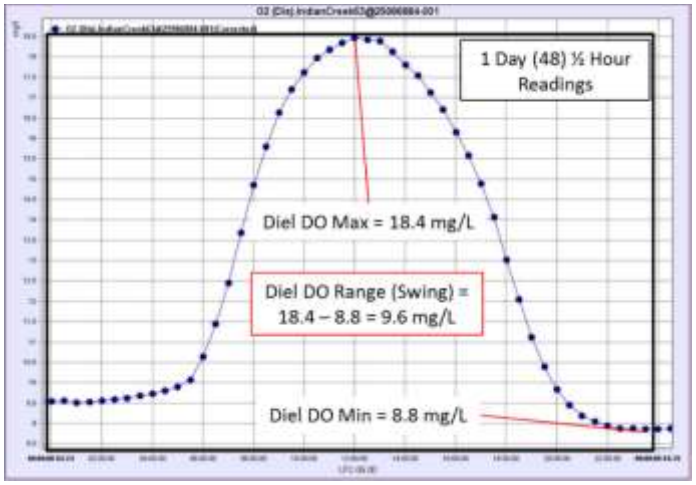


Figure 3. Graphical representation of the calculation of diel DO swing values from DO data monitored continuously over a period of 24 hours.

All useable diel DO swing values recorded within a given month are summarized using the 75th percentile (p75) value of the diel swing values recorded in that month. Diel DO swing p75 values are only generated for months that have usable diel DO swing values recorded for a minimum of 50% of days in that month. For example, if a sonde was deployed at Station X from March 1 to March 31, 2017, and yielded only 12 diel DO swing values, no p75 would be calculated for that month, because 12 days are less than 50% of the 31 days in March.

In addition to the requirement of having usable diel DO swing values recorded for a minimum of 50% of the days in a given month, a minimum of 15 pairs of diel DO-pH swing and diel DO-water temperature swing values are required for calculating monthly correlation values. Examples of how monthly diel DO swing p75 and correlation values are calculated are provided in Table 1 with results shown graphically in Figures 4 and 5.

Table 1. Example spreadsheet calculation of a monthly diel DO swing p75 value of 8.0 mg/L and monthly diel DO swing-diel pH swing and monthly diel DO swing-diel water temperature swing correlation coefficients of 0.95 and 0.14, respectively, from 31 days of data recorded at a small (drainage area ≤50 mi^2) ALU impaired stream in Physiographic Region A.

	A	B	C	D	E	F	G	H
1	Example Continuous Monitoring Data							
2	Date	Diel DO Swing (mg/L)	Diel pH Swing	Diel Water Temp Swing (C°)	Diel DO Swing p75 (mg/L)	Correlation Pairs (N)	Diel DO-pH Swing Correlation Coefficient r	Diel DO-Temp Swing Correlation Coefficient r
3	5/1/2013	7.1	1.5	4.8	8.0	31	0.95	0.14
4	5/2/2013	8.4	1.6	3.5				
5	5/3/2013	9.2	1.7	5.1	Formula in Cell E3		=PERCENTILE.INC(B3:B33,0.75)	
6	5/4/2013	9.2	1.6	1.7	Formula in Cell F3		=COUNT(B3:B33)	
7	5/5/2013	9.8	1.8	5.3	Formula in Cell G3		=CORREL(B3:B33,C3:C33)	
8	5/6/2013	9.1	1.6	4.3	Formula in Cell H3		=CORREL(B3:B33,D3:D33)	
9	5/7/2013	7.7	1.6	4.3				
10	5/8/2013	8.0	1.6	2.8				
11	5/9/2013	8.3	1.6	3.7				
12	5/10/2013	6.5	1.4	4.6				
13	5/11/2013	7.4	1.5	5.4				
14	5/12/2013	8.1	1.6	5.0				
15	5/13/2013	7.6	1.5	4.4				
16	5/14/2013	7.2	1.6	3.6				
17	5/15/2013	2.2	0.3	1.7				
18	5/16/2013	3.1	0.7	5.0				
19	5/17/2013	4.4	0.8	4.8				
20	5/18/2013	4.4	0.8	3.9				
21	5/19/2013	6.0	1.1	5.9				
22	5/20/2013	6.3	1.2	4.6				
23	5/21/2013	7.1	1.3	3.0				
24	5/22/2013	6.5	1.2	4.3				
25	5/23/2013	7.2	1.4	5.8				
26	5/24/2013	7.7	1.4	6.3				
27	5/25/2013	7.6	1.4	6.9				
28	5/26/2013	8.0	1.5	6.5				
29	5/27/2013	8.0	1.4	4.0				
30	5/28/2013	7.7	1.5	6.3				
31	5/29/2013	7.1	1.3	4.4				
32	5/30/2013	6.9	1.0	2.8				
33	5/31/2013	6.7	1.3	6.0				

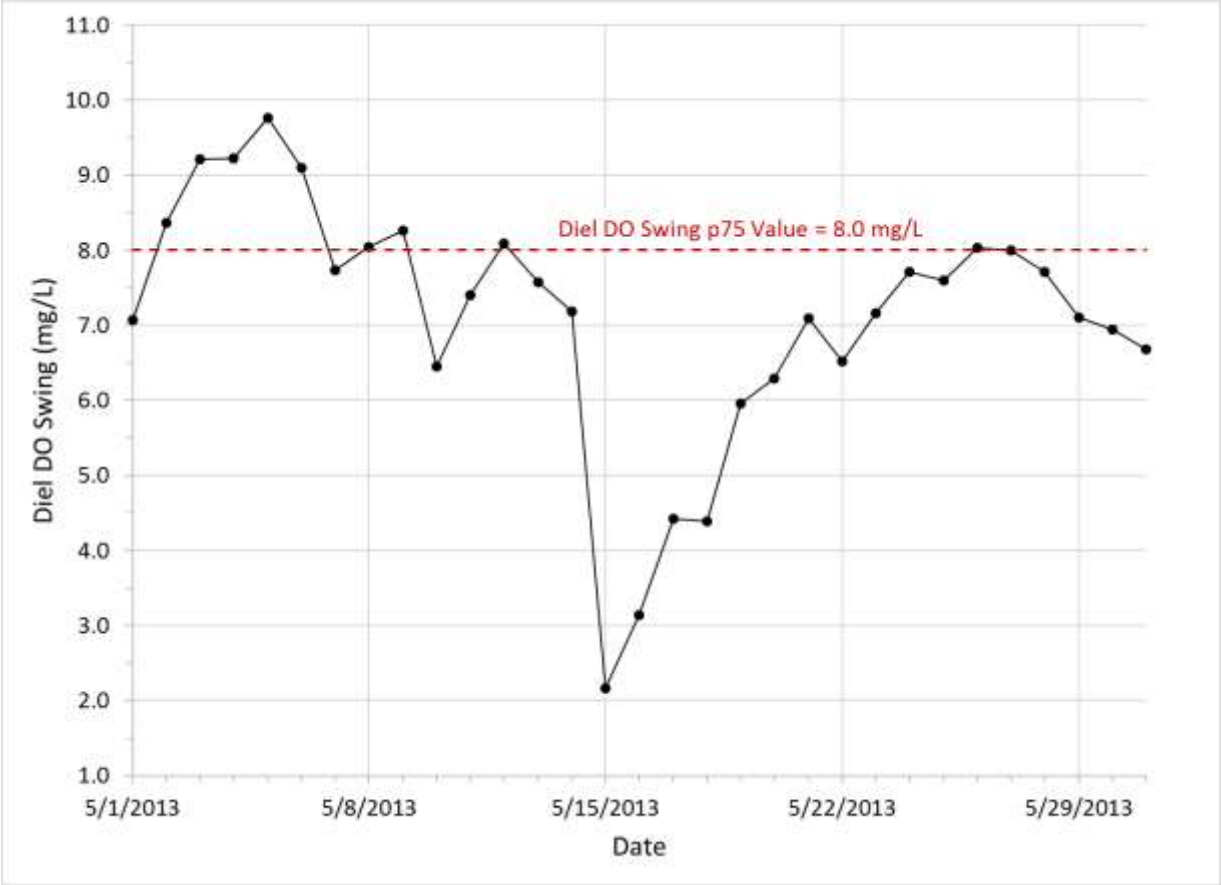
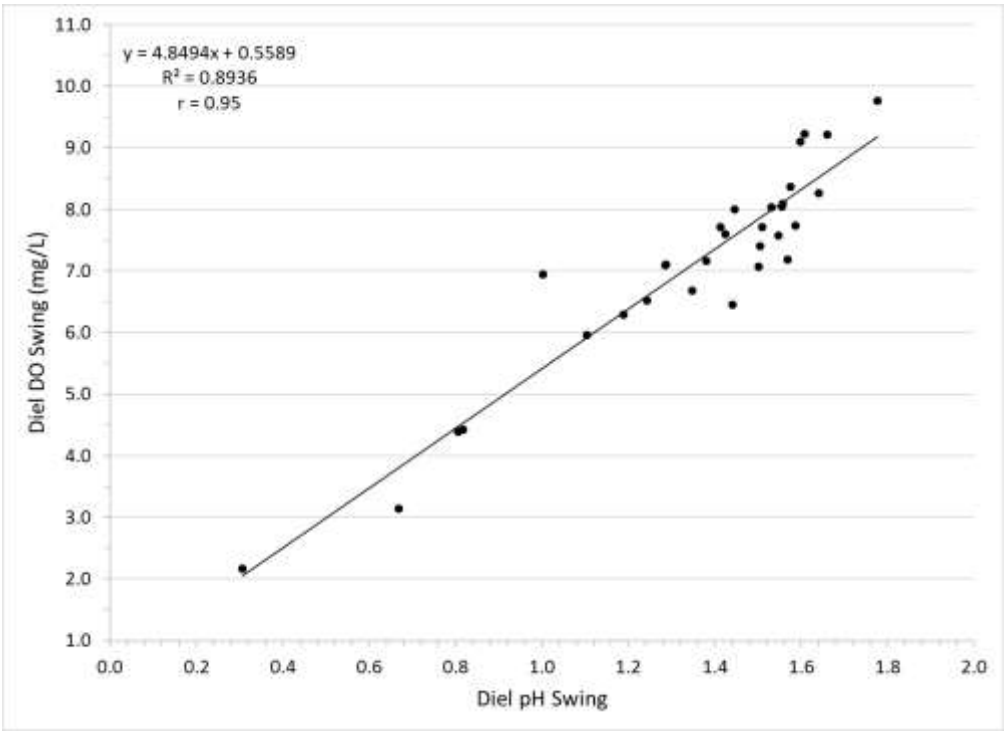
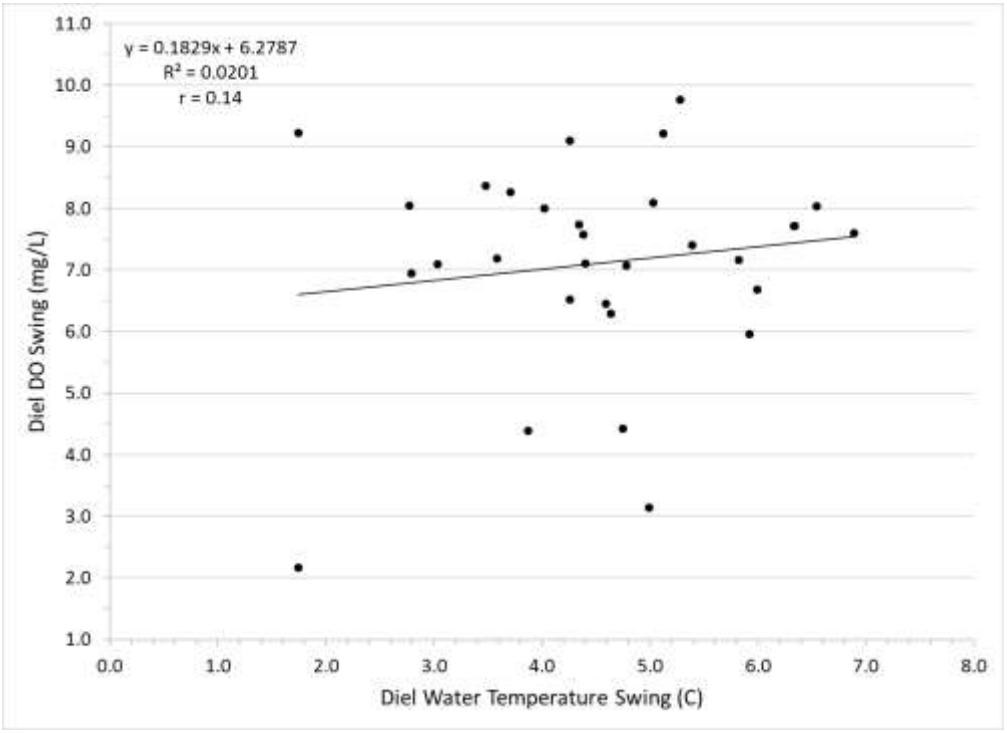


Figure 4. Graphical representation of data from Table 1 showing individual diel DO swing values and the monthly diel DO swing 75th percentile (p75) value of 8.0 mg/L.



(A)



(B)

Figure 5. Graphical representation of data from Table 1 showing (A) diel DO swing vs. diel pH swing values and corresponding monthly Pearson Correlation r-value of 0.95 and (B) diel DO swing vs. diel water temperature swing values and corresponding monthly Pearson Correlation r-value of 0.14.

Eutrophication Cause Determinations

The first step in the ECD Protocol is to determine if the ALU impaired stream is subject to excessive diel swings in DO. This is accomplished by comparing the monthly diel DO p75 values recorded at the ALU impaired stream to the benchmark values shown in Table 2. Separate diel DO swing benchmark values were developed within the context of 2-month sample periods and the Physiographic Regions shown in Figure 6.

If no monthly diel DO swing p75 values recorded at the ALU impaired stream exceed the appropriate Table 2 diel DO swing benchmark value the cause of ALU impairment is determined to be something other than eutrophication (Figure 2). If any monthly diel DO swing p75 value recorded at an ALU impaired stream segment exceeds the appropriate diel DO swing p75 benchmark value, eutrophication is identified as a cause of ALU impairment if:

1. The stream segment exceeds water quality criteria for DO or pH greater than 1% of the time, based on Hoger et al. (2017) (Figure 2), or

2. Any benthic periphyton sample collected in the stream segment has a chlorophyll-a concentration >275 mg/m^2 (Figure 2), or

3. Any monthly diel DO swing p75 that exceeds the appropriate diel DO swing p75 benchmark value has a monthly diel DO swing-diel pH swing Pearson correlation r-value >0.66 with a monthly diel DO swing-diel water temperature swing Pearson correlation r-value <0.61 (Figure 2).

Table 2. Eutrophication Cause Determination Protocol benchmark values.

Monthly Diel DO Swing p75 Benchmark Values (mg/L)	**Physiographic Region**	
Sample Period	**A**	**B**
March-April	2.8	1.5
May-June	1.7	1.4
July-August	1.8	1.3
September-October	2.0	1.5
Maximum Benthic Chlorophyll-a Value (mg/m^2)	275	
Monthly Correlation Benchmark Values	**Pearson Correlation Coefficient (r)**	
Monthly Diel DO Swing-Diel pH Swing	>0.66	
Monthly Diel DO Swing-Diel Water Temperature Swing	<0.61	

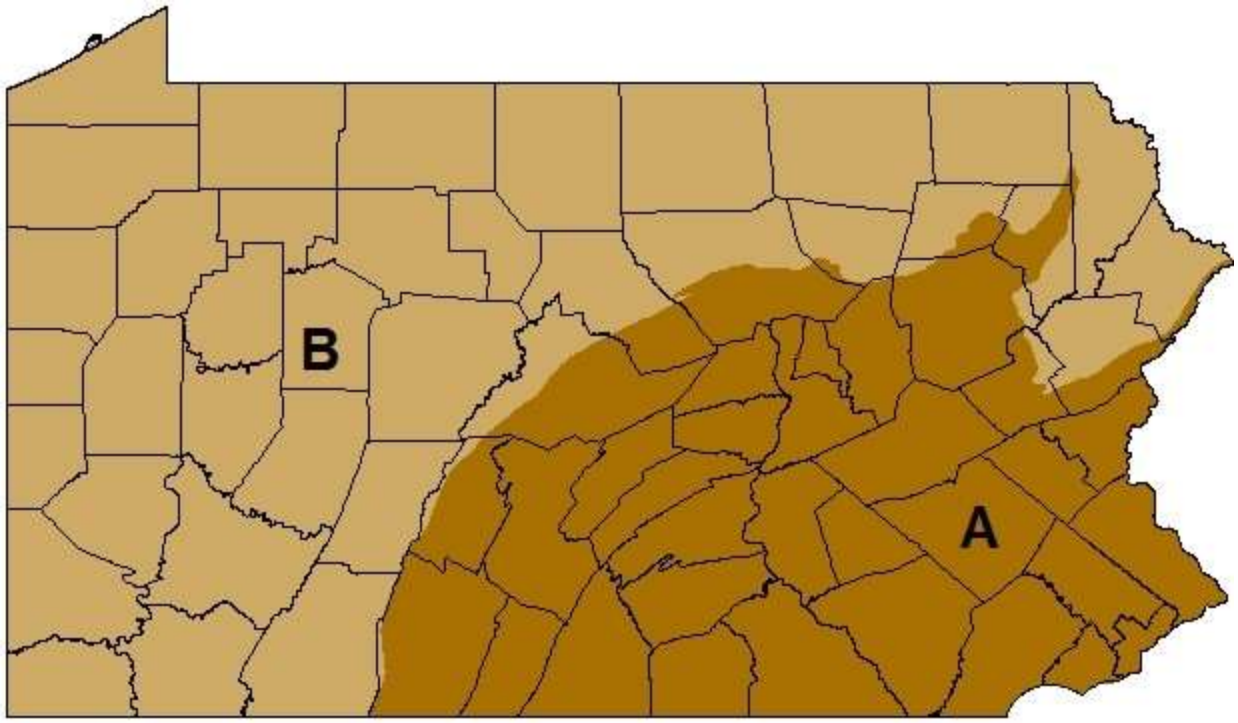


Figure 6. Eutrophication Cause Determination Protocol Physiographic Regions.

The following is an example application of the ECD Protocol to the data shown in Table 1. In this example, it is assumed that the stream segment meets water quality criteria for DO or pH and no benthic chlorophyll-a samples exceeded a concentration of 275 mg/m^2. Based on the ECD Protocol, eutrophication is identified as a cause of ALU impairment because the following conditions are met:

1. The monthly diel DO swing p75 value of 8.0 mg/L exceeds the benchmark value of 1.7 mg/L for Physiographic Region A streams during the May-June sample period, **AND**

2. The monthly diel DO-pH swing correlation r-value of 0.95 is >0.66, **AND**

3. The monthly diel DO-water temperature swing correlation r-value of 0.14 is <0.61.

In the example above, ECD Protocol results indicate the sample station has excessively high diel DO swings. Furthermore, the strong correlation between diel DO swings and diel pH swings, in conjunction with a weak correlation between diel DO swings and diel water temperature swings, indicates the excessive diel DO swings are related to stream metabolic processes (photosynthesis and respiration rates), not the water temperature conditions of the stream.

LITERATURE CITED

Butt, J. S. 2017. Periphyton data collection protocol. Chapter 3, pages 83-131. *In* Shull, D. R. and M. J. Lookenbill (editors). Water Quality Monitoring Protocols for Streams and Rivers. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania.

Heiskary, S. A. and R. W. Bouchard, Jr. 2015. Development of eutrophication criteria for Minnesota streams and rivers using multiple lines of evidence. Freshwater Science 34:574-592.

Hoger, M. S., D. R. Shull, and M. J. Lookenbill. 2017. Continuous Physicochemical Data Collection Protocol. Chapter 4, pages 22-87. *In* Shull, D. R. and M. J. Lookenbill (editors). Water Quality Monitoring Protocols for Streams and Rivers. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania.

McGarrell, C. 2018. Eutrophication Cause Determination Protocol: Technical Support Document. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania.

Shull, D. R. 2013. Discrete Water Chemistry Data Collection Protocol. Chapter 4, pages 8-21. *In* Shull, D. R. and M. J. Lookenbill (editors). Water Quality Monitoring Protocols for Streams and Rivers. Pennsylvania Department of Environmental Protection. Harrisburg, Pennsylvania.

APPENDIX A: SOURCES AND CAUSES

SOURCES

Below details the current list of available impairment sources allowable by the USEPA.

Source List (Available from USEPA)
ABOVE GROUND STORAGE TANK LEAKS (TANK FARMS)
ACCIDENTAL RELEASE/SPILL
ACID MINE DRAINAGE
AGRICULTURAL RETURN FLOWS
AGRICULTURAL WATER DIVERSION
AGRICULTURE
AIRPORTS
ANIMAL FEEDING OPERATIONS (NPS)
ANIMAL HOLDING/MANAGEMENT AREAS
ANIMAL SHOWS AND RACETRACKS
AQUACULTURE (NOT PERMITTED)
AQUACULTURE (PERMITTED)
ATMOSPHERIC DEPOSITION
ATMOSPHERIC DEPOSITION - ACIDITY
ATMOSPHERIC DEPOSITION - NITROGEN
ATMOSPHERIC DEPOSITION - TOXICS
AUCTION BARNS
BALLAST WATER RELEASES
BARGE CANAL IMPACTS
BASEFLOW DEPLETION FROM GROUNDWATER WITHDRAWALS
BROWNFIELD (NON-NPL) SITES
CARGO LOADING/UNLOADING
CERCLA NPL (SUPERFUND) SITES
CHANGES IN ORDINARY STRATIFICATION AND BOTTOM WATER HYPOXIA/ANOXIA
CHANGES IN TIDAL CIRCULATION/FLUSHING
CHANNEL EROSION/INCISION FROM UPSTREAM HYDROMODIFICATIONS
CHANNELIZATION
CHEMICAL LEAK/SPILL
COAL MINING
COAL MINING (SUBSURFACE)
COAL MINING DISCHARGES (PERMITTED)
COMBINED SEWER OVERFLOWS
COMMERCIAL DISTRICTS (INDUSTRIAL PARKS)
COMMERCIAL DISTRICTS (SHOPPING/OFFICE COMPLEXES)
COMMERCIAL HARBOR AND PORT ACTIVITIES
CONFINED ANIMAL FEEDING OPERATIONS - CAFOS (POINT SOURCE)
CONFINED ANIMAL FEEDING OPERATIONS (NPS)
CONSTRUCTION

CONSTRUCTION STORMWATER DISCHARGE (PERMITTED)
CONTAMINATED GROUNDWATER
CONTAMINATED SEDIMENTS
COOLING WATER INTAKE STRUCTURES (IMPINGEMENT OR ENTRAINMENT)
CRANBERRY PRODUCTION
CROP PRODUCTION (CROP LAND OR DRY LAND)
CROP PRODUCTION (IRRIGATED)
CROP PRODUCTION (NON-IRRIGATED)
CROP PRODUCTION WITH SUBSURFACE DRAINAGE
DAIRIES
DAM CONSTRUCTION (OTHER THAN UPSTREAM FLOOD CONTROL PROJECTS)
DAM OR IMPOUNDMENT
DEICING (STORAGE/APPLICATION)
DISCHARGES FROM BIOSOLIDS (SLUDGE) STORAGE, APPLICATION OR DISPOSAL
DISCHARGES FROM MUNICIPAL SEPARATE STORM SEWER SYSTEMS (MS4)
DISCHARGES FROM OFFSHORE OIL AND GAS EXPLORATION (PERMITTED)
DREDGE MINING
DREDGING (E.G., FOR NAVIGATION CHANNELS)
DROUGHT-RELATED IMPACTS
DRY WEATHER FLOWS WITH NPS POLLUTANTS
EROSION AND SEDIMENTATION
EROSION FROM DERELICT LAND (BARREN LAND)
FORCED DRAINAGE PUMPING
FOREST ROADS (ROAD CONSTRUCTION AND USE)
FRESHETS OR MAJOR FLOODING
GOLF COURSES
GRAZING IN RIPARIAN OR SHORELINE ZONES
GROUNDWATER LOADINGS
HABITAT MODIFICATION - OTHER THAN HYDROMODIFICATION
HARDROCK MINING DISCHARGES (PERMITTED)
HARVESTING/RESTORATION/RESIDUE MANAGEMENT
HEAP-LEACH EXTRACTION MINING
HIGHWAY/ROAD/BRIDGE RUNOFF (NON-CONSTRUCTION RELATED)
HIGHWAYS, ROADS, BRIDGES, INFRASTRUCTURE (NEW CONSTRUCTION)
HISTORIC BOTTOM DEPOSITS (NOT SEDIMENT)
HYDROSTRUCTURE IMPACTS ON FISH PASSAGE
ILLEGAL DUMPS OR OTHER INAPPROPRIATE WASTE DISPOSAL
ILLICIT CONNECTIONS/HOOK-UPS TO STORM SEWERS
IMPACTS FROM ABANDONED MINE LANDS (INACTIVE)
IMPACTS FROM GEOTHERMAL DEVELOPMENT
IMPACTS FROM HYDROSTRUCTURE FLOW REGULATION/MODIFICATION
IMPACTS FROM LAND APPLICATION OF WASTES
IMPACTS FROM RESORT AREAS

IMPERVIOUS SURFACE/PARKING LOT RUNOFF
INDUSTRIAL LAND TREATMENT
INDUSTRIAL POINT SOURCE DISCHARGE
INDUSTRIAL THERMAL DISCHARGES
INDUSTRIAL/COMMERCIAL SITE STORMWATER DISCHARGE (PERMITTED)
INTERNAL NUTRIENT RECYCLING
INTRODUCTION OF NON-NATIVE ORGANISMS (ACCIDENTAL OR INTENTIONAL)
LAKE FERTILIZATION
LANDFILLS
LEAKING UNDERGROUND STORAGE TANKS
LEGACY/HISTORICAL POLLUTANTS
LITTORAL/SHORE AREA MODIFICATIONS (NON-RIVERINE)
LIVESTOCK (GRAZING OR FEEDING OPERATIONS)
LOSS OF RIPARIAN HABITAT
LOW HEAD DAMS
LOW WATER CROSSING
MANAGED PASTURE GRAZING
MANURE LAGOONS
MANURE RUNOFF
MARINA BOAT CONSTRUCTION
MARINA BOAT MAINTENANCE
MARINA DREDGING OPERATIONS
MARINA FUELING OPERATIONS
MARINA RELATED SHORELINE HABITAT DEGRADATION
MARINA/BOATING PUMPOUT RELEASES
MARINA/BOATING SANITARY ON-VESSEL DISCHARGES
MARINAS AND RECREATIONAL BOATING
MILL TAILINGS
MINE TAILINGS
MINING
MOTORIZED WATERCRAFT
MOUNTAINTOP MINING
MUNICIPAL (URBANIZED HIGH DENSITY AREA)
MUNICIPAL POINT SOURCE DISCHARGES
MUNICIPAL POINT SOURCE IMPACTS FROM INADEQUATE INDUSTRIAL/COMMERCIAL PRETREATMENT
NATURAL CONDITIONS - WQS ANALYSES NEEDED
NATURAL SOURCES
NATURAL-BEAVER DAMS/LOG JAMS
NATURAL-DROUGHT
NATURAL-FLOOD
NATURALLY OCCURRING ORGANIC ACIDS
NATURAL-SNOWMELT
NON-METALS MINING DISCHARGES (PERMITTED)

NON-POINT SOURCE
NPS POLLUTION FROM MILITARY BASE FACILITIES (OTHER THAN PORT FACILITIES)
NPS POLLUTION FROM MILITARY PORT FACILITIES
OFF-ROAD VEHICLES
ON-SITE TREATMENT SYSTEMS (SEPTIC SYSTEMS AND SIMILAR DECENTRALIZED SYSTEMS)
OPEN PIT MINING
OTHER MARINA/BOATING ON-VESSEL DISCHARGES
OTHER RECREATIONAL POLLUTION SOURCES
OTHER SHIPPING RELEASES (WASTES AND DETRITUS)
OTHER SPILL RELATED IMPACTS
OTHER TURF MANAGEMENT
PACKAGE PLANT OR OTHER PERMITTED SMALL FLOWS DISCHARGES
PESTICIDE APPLICATION
PETROLEUM/NATURAL GAS ACTIVITIES
PETROLEUM/NATURAL GAS PRODUCTION ACTIVITIES (PERMITTED)
PIPELINE BREAKS
PLACER MINING
POINT SOURCE(S) - UNSPECIFIED
POLLUTANTS FROM PUBLIC BATHING AREAS
POST-DEVELOPMENT EROSION AND SEDIMENTATION
POTASH MINING
RANGELAND GRAZING
RCRA HAZARDOUS WASTE SITES
RECREATION AND TOURISM (NON-BOATING)
REDUCED FRESHWATER FLOWS
RELEASES FROM WASTE SITES OR DUMPS
REMOVAL OF RIPARIAN VEGETATION
RESIDENTIAL DISTRICTS
RUNOFF FROM FOREST/GRASSLAND/PARKLAND
RURAL (RESIDENTIAL AREAS)
SALT STORAGE SITES
SALTWATER INTRUSION
SAND/GRAVEL/ROCK MINING OR QUARRIES
SANITARY SEWER OVERFLOWS (COLLECTION SYSTEM FAILURES)
SEAFOOD PROCESSING OPERATIONS
SEDIMENT RESUSPENSION (CLEAN SEDIMENT)
SEDIMENT RESUSPENSION (CONTAMINATED SEDIMENT)
SEPTAGE DISPOSAL
SEWAGE DISCHARGES IN UNSEWERED AREAS
SHALLOW LAKE/RESERVOIR
SHIPBUILDING, REPAIRS, DRYDOCKING
SILVICULTURE ACTIVITIES
SILVICULTURE HARVESTING

SILVICULTURE, FIRE SUPPRESSION
SITE CLEARANCE (LAND DEVELOPMENT OR REDEVELOPMENT)
SOURCE UNKNOWN
SOURCES OUTSIDE STATE JURISDICTION OR BORDERS
SPECIALITY CROP PRODUCTION
SPILLS FROM TRUCKS OR TRAINS
STREAMBANK MODIFICATIONS/DESTABILIZATION
SUBSURFACE (HARDROCK) MINING
SURFACE MINING
SURFACE WATER DIVERSIONS
SURFACE WATER WITHDRAWALS
TOTAL RETENTION DOMESTIC SEWAGE LAGOONS
TRANSFER OF WATER FROM AN OUTSIDE WATERSHED
UIC WELLS (UNDERGROUND INJECTION CONTROL WELLS)
UNKNOWN POINT SOURCE
UNPERMITTED DISCHARGE (DOMESTIC WASTES)
UNPERMITTED DISCHARGE (INDUSTRIAL/COMMERCIAL WASTES)
UNRESTRICTED CATTLE ACCESS
UNSPECIFIED DOMESTIC WASTE
UNSPECIFIED LAND DISTURBANCE
UNSPECIFIED UNPAVED ROAD OR TRAIL
UNSPECIFIED URBAN STORMWATER
UPSTREAM SOURCE
URBAN RUNOFF/STORM SEWERS
WASTES FROM PETS
WATER DIVERSIONS
WATERFOWL
WATERSHED RUNOFF FOLLOWING FOREST FIRE
WET WEATHER DISCHARGES (NON-POINT SOURCE)
WET WEATHER DISCHARGES (POINT SOURCE AND COMBINATION OF STORMWATER, SSO OR CSO)
WETLAND DRAINAGE
WILDLIFE OTHER THAN WATERFOWL
WOODLOT SITE CLEARANCE
WOODLOT SITE MANAGEMENT
YARD MAINTENANCE

CAUSES

Below details the current list of available impairment causes allowable by the USEPA. The context that USEPA places each cause into is also provided for reference purposes; however, categories are not used as causes of impairment.

Cause List (Available from USEPA)	Cause Context
ALGAE	ALGAL GROWTH
ALGAL TOXINS	ALGAL GROWTH
BROWN TIDE	ALGAL GROWTH
CHLOROPHYLL-A	ALGAL GROWTH
CHLOROPHYLL-A - AQUATIC LIFE USE SUPPORT	ALGAL GROWTH
CHLOROPHYLL-A - PRIMARY CONTACT RECREATION	ALGAL GROWTH
CURLY-LEAF PONDWEED	ALGAL GROWTH
FANWORT	ALGAL GROWTH
HARMFUL ALGAL BLOOMS	ALGAL GROWTH
HYDRILLA	ALGAL GROWTH
SEA LETTUCE	ALGAL GROWTH
SUSPENDED ALGAE	ALGAL GROWTH
AMMONIA	AMMONIA
AMMONIA NITROGEN	AMMONIA
AMMONIA, TOTAL	AMMONIA
AMMONIA, UN-IONIZED	AMMONIA
ABNORMAL FISH DEFORMITIES, EROSIONS, LESIONS, TUMORS (DELTS)	BIOTOXINS
BIOTOXINS	BIOTOXINS
CYANOBACTERIA HEPATOTOXIC MICROCYSTINS	BIOTOXINS
CYANOBACTERIA HEPATOTOXIC NODULARINS	BIOTOXINS
CYANOBACTERIA NEUROTOXIC ANATOXINS	BIOTOXINS
CYANOBACTERIA NEUROTOXIC SAXITOXINS	BIOTOXINS
CAUSE UNKNOWN	CAUSE UNKNOWN
FISH KILL DUE TO THERMAL MODIFICATIONS	CAUSE UNKNOWN - FISH KILLS
FISH KILL(S)	CAUSE UNKNOWN - FISH KILLS
AQUATIC PLANT BIOASSESSMENTS	CAUSE UNKNOWN - IMPAIRED BIOTA
BENTHIC MACROINVERTEBRATES	CAUSE UNKNOWN - IMPAIRED BIOTA
BENTHIC MACROINVERTEBRATES BIOASSESSMENTS	CAUSE UNKNOWN - IMPAIRED BIOTA
BIOLOGICAL INTEGRITY	CAUSE UNKNOWN - IMPAIRED BIOTA
COMBINED BIOTA/HABITAT BIOASSESSMENTS	CAUSE UNKNOWN - IMPAIRED BIOTA
ESTUARINE BIOASSESSMENTS	CAUSE UNKNOWN - IMPAIRED BIOTA
FISH BIOASSESSMENTS	CAUSE UNKNOWN - IMPAIRED BIOTA
HABITAT ASSESSMENT	CAUSE UNKNOWN - IMPAIRED BIOTA
INDEX OF BIOLOGICAL INTEGRITY (IBI)	CAUSE UNKNOWN - IMPAIRED BIOTA
PERIPHYTON (AUFWUCHS) INDICATOR BIOASSESSMENTS	CAUSE UNKNOWN - IMPAIRED BIOTA
CHLORINE	CHLORINE

CHLORINE DIOXIDE	CHLORINE
CHLORINE, RESIDUAL (CHLORINE DEMAND)	CHLORINE
FREE CHLORINE	CHLORINE
2,3,7,8-TETRACHLORODIBENZOFURAN	DIOXINS
2,3,7,8-TETRACHLORODIBENZO-P-DIOXIN	DIOXINS
DIBENZOFURAN	DIOXINS
DIOXIN	DIOXINS
DIOXIN - FISH CONSUMPTION ADVISORY	DIOXINS
DIOXIN (INCLUDING 2,3,7,8-TCDD)	DIOXINS
DIOXIN IN FISH TISSUE	DIOXINS
FURAN COMPOUNDS	DIOXINS
CANCER RISK COMPOUNDS	FISH CONSUMPTION ADVISORY
COMMERCIAL FISHING BAN	FISH CONSUMPTION ADVISORY
FISH CONSUMPTION ADVISORY - DDE, DDD	FISH CONSUMPTION ADVISORY
FISH CONSUMPTION ADVISORY - DDE, DDT	FISH CONSUMPTION ADVISORY
HAZARD INDEX COMPOUNDS	FISH CONSUMPTION ADVISORY
HEXACHLOROBENZE - FISH CONSUMPTION ADVISORY	FISH CONSUMPTION ADVISORY
MIREX - FISH CONSUMPTION ADVISORY	FISH CONSUMPTION ADVISORY
SHELLFISHING BAN	FISH CONSUMPTION ADVISORY
ALTERATION IN STREAM-SIDE OR LITTORAL VEGETATIVE COVERS	HABITAT ALTERATIONS
FISH PASSAGE BARRIER	HABITAT ALTERATIONS
HABITAT ALTERATIONS	HABITAT ALTERATIONS
LOSS OF INSTREAM COVER	HABITAT ALTERATIONS
OTHER ANTHROPOGENIC SUBSTRATE ALTERATIONS	HABITAT ALTERATIONS
PHYSICAL SUBSTRATE HABITAT ALTERATIONS	HABITAT ALTERATIONS
DEWATERING	HYDROLOGIC ALTERATION
FLOW REGIME MODIFICATION	HYDROLOGIC ALTERATION
SALINITY CHANGE DUE TO CHANGE IN FLOW	HYDROLOGIC ALTERATION
STREAM MODIFICATION	HYDROLOGIC ALTERATION
TIDAL FLOW ALTERATION	HYDROLOGIC ALTERATION
WETLANDS DRAINAGE	HYDROLOGIC ALTERATION
WETLANDS DREDGED/FILLED	HYDROLOGIC ALTERATION
MERCURY	MERCURY
MERCURY - FISH CONSUMPTION ADVISORY	MERCURY
MERCURY IN FISH TISSUE	MERCURY
MERCURY IN SEDIMENT	MERCURY
MERCURY IN WATER COLUMN	MERCURY
MERCURY, DISSOLVED	MERCURY
MERCURY, TOTAL	MERCURY
METHYLMERCURY	MERCURY
ALUM IN SEDIMENT	METALS (OTHER THAN MERCURY)
ALUMINUM	METALS (OTHER THAN MERCURY)

ALUMINUM, DISSOLVED	METALS (OTHER THAN MERCURY)
ALUMINUM, TOTAL	METALS (OTHER THAN MERCURY)
ARSENIC	METALS (OTHER THAN MERCURY)
ARSENIC IN SEDIMENT	METALS (OTHER THAN MERCURY)
ARSENIC, TRIVALENT	METALS (OTHER THAN MERCURY)
BERYLLIUM	METALS (OTHER THAN MERCURY)
CADMIUM	METALS (OTHER THAN MERCURY)
CADMIUM IN SEDIMENT	METALS (OTHER THAN MERCURY)
CHROMIUM	METALS (OTHER THAN MERCURY)
CHROMIUM IN SEDIMENT	METALS (OTHER THAN MERCURY)
CHROMIUM, HEXAVALENT	METALS (OTHER THAN MERCURY)
CHROMIUM, TOTAL	METALS (OTHER THAN MERCURY)
CHROMIUM, TRIVALENT	METALS (OTHER THAN MERCURY)
COBALT	METALS (OTHER THAN MERCURY)
COPPER	METALS (OTHER THAN MERCURY)
COPPER IN SEDIMENT	METALS (OTHER THAN MERCURY)
COPPER, DISSOLVED	METALS (OTHER THAN MERCURY)
COPPER, TOTAL	METALS (OTHER THAN MERCURY)
GOLD	METALS (OTHER THAN MERCURY)
GOLD IN SEDIMENT	METALS (OTHER THAN MERCURY)
IRON	METALS (OTHER THAN MERCURY)
IRON TOTAL RECOVERABLE	METALS (OTHER THAN MERCURY)
LEAD	METALS (OTHER THAN MERCURY)
LEAD IN SEDIMENT	METALS (OTHER THAN MERCURY)
MANGANESE	METALS (OTHER THAN MERCURY)
METALS	METALS (OTHER THAN MERCURY)
MOLYBDENUM	METALS (OTHER THAN MERCURY)
NICKEL	METALS (OTHER THAN MERCURY)
NICKEL IN SEDIMENT	METALS (OTHER THAN MERCURY)
SELENIUM	METALS (OTHER THAN MERCURY)
SELENIUM IN FISH TISSUE	METALS (OTHER THAN MERCURY)
SELENIUM IN SEDIMENT	METALS (OTHER THAN MERCURY)
SELENIUM, DISSOLVED	METALS (OTHER THAN MERCURY)
SELENIUM, TOTAL	METALS (OTHER THAN MERCURY)
SILVER	METALS (OTHER THAN MERCURY)
SILVER IN SEDIMENT	METALS (OTHER THAN MERCURY)
STRONTIUM	METALS (OTHER THAN MERCURY)
THALLIUM	METALS (OTHER THAN MERCURY)
TITANIUM	METALS (OTHER THAN MERCURY)
VANADIUM	METALS (OTHER THAN MERCURY)
ZINC	METALS (OTHER THAN MERCURY)
ZINC IN FISH TISSUE	METALS (OTHER THAN MERCURY)
ZINC IN SEDIMENT	METALS (OTHER THAN MERCURY)

ZINC IN SHELLFISH	METALS (OTHER THAN MERCURY)
ZINC, CHRONIC	METALS (OTHER THAN MERCURY)
ZINC, DISSOLVED	METALS (OTHER THAN MERCURY)
ZINC, TOTAL	METALS (OTHER THAN MERCURY)
AQUATIC PLANTS (MACROPHYTES)	NOXIOUS AQUATIC PLANTS
NOXIOUS AQUATIC PLANTS	NOXIOUS AQUATIC PLANTS
NON-NATIVE AQUATIC PLANTS	NUISANCE EXOTIC SPECIES
NON-NATIVE FISH/SHELLFISH/ZOOPLANKTON	NUISANCE EXOTIC SPECIES
ZEBRA MUSSEL, DREISSENA POLYMORPH	NUISANCE EXOTIC SPECIES
NOXIOUS AQUATIC PLANTS NATIVE	NUISANCE NATIVE SPECIES
EUTROPHICATION	NUTRIENTS
NITRATE	NUTRIENTS
NITRATE/NITRITE (NITRITE + NITRATE AS N)	NUTRIENTS
NITRITE	NUTRIENTS
NITROGEN	NUTRIENTS
NITROGEN, AMMONIA	NUTRIENTS
NITROGEN, NITRATE	NUTRIENTS
NITROGEN, NITRITE	NUTRIENTS
NITROGEN, TOTAL	NUTRIENTS
NUTRIENTS	NUTRIENTS
PHOSPHATE	NUTRIENTS
PHOSPHORUS	NUTRIENTS
PHOSPHORUS, TOTAL	NUTRIENTS
TOTAL KJEHLDAHL NITROGEN (TKN)	NUTRIENTS
TROPHIC STATE INDEX (TSI)	NUTRIENTS
DIESEL FUEL	OIL AND GREASE
OIL AND GREASE	OIL AND GREASE
PETROLEUM HYDROCARBONS	OIL AND GREASE
RESIDUAL SURFACE AND SUB-SURFACE OIL/TAR BALLS/TAR MATS	OIL AND GREASE
BIOCHEMICAL OXYGEN DEMAND (BOD)	ORGANIC ENRICHMENT/OXYGEN DEPLETION
CHEMICAL OXYGEN DEMAND (COD)	ORGANIC ENRICHMENT/OXYGEN DEPLETION
DISSOLVED OXYGEN	ORGANIC ENRICHMENT/OXYGEN DEPLETION
ORGANIC ENRICHMENT	ORGANIC ENRICHMENT/OXYGEN DEPLETION
ORGANIC ENRICHMENT (SEWAGE) BIOLOGICAL INDICATORS	ORGANIC ENRICHMENT/OXYGEN DEPLETION
SEDIMENT OXYGEN DEMAND	ORGANIC ENRICHMENT/OXYGEN DEPLETION
TOTAL ORGANIC CARBON (TOC)	ORGANIC ENRICHMENT/OXYGEN DEPLETION
ATTAINING ASSESSMENT UNIT WITH PROTECTION PLAN	OTHER CAUSE
DEBRIS	OTHER CAUSE
DISSOLVED GAS SUPERSATURATION	OTHER CAUSE
NATURAL LIMITS	OTHER CAUSE
OSMOTIC PRESSURE	OTHER CAUSE
PAPER SLUDGE	OTHER CAUSE

POLLUTANTS IN URBAN STORMWATER	OTHER CAUSE
RESIDUES	OTHER CAUSE
SCUM/FOAM	OTHER CAUSE
SLUDGE	OTHER CAUSE
SODIUM ABSORPTION RATIO	OTHER CAUSE
SURFACTANTS	OTHER CAUSE
TOTAL DISSOLVED GAS	OTHER CAUSE
VISIBLE OIL AND SODIUM ADSORPTION RATIO (SAR)	OTHER CAUSE
BACTERIA (OYSTER WATERS)	PATHOGENS
ENTEROCOCCUS	PATHOGENS
ESCHERICHIA COLI (E. COLI)	PATHOGENS
FECAL COLIFORM	PATHOGENS
PATHOGENS	PATHOGENS
TOTAL COLIFORM	PATHOGENS
VIRUSES (ENTERIC)	PATHOGENS
1,2-DIBROMO-3-CHLOROPROPANE	PESTICIDES
1,2-DICHLOROETHANE	PESTICIDES
1,2-DICHLOROPROPANE	PESTICIDES
1,3-DICHLOROPROPENE	PESTICIDES
2,3-DICHLOROPROPENE	PESTICIDES
2,4,5-TP (SILVEX)	PESTICIDES
2,4,5-TRICHLOROPHENOL	PESTICIDES
2,4-DINITROPHENOL	PESTICIDES
2-METHYLNAPHTHALENE	PESTICIDES
4,4'-DDD	PESTICIDES
4,4'-DDE	PESTICIDES
4,4'-DDT	PESTICIDES
ACETAMIDE	PESTICIDES
ACROLEIN	PESTICIDES
ACRYLONITRILE	PESTICIDES
ALACHLOR	PESTICIDES
ALDICARB	PESTICIDES
ALDRIN	PESTICIDES
ALPHA-BHC	PESTICIDES
ALPHA-ENDOSULFAN (ENDOSULFAN 1)	PESTICIDES
AMETRYN	PESTICIDES
AMITROLE	PESTICIDES
ATRAZINE	PESTICIDES
BENTAZON	PESTICIDES
BETA-BHC	PESTICIDES
BETA-ENDOSULFAN (ENDOSULFAN 2)	PESTICIDES
BHC	PESTICIDES
BIFENTHRIN	PESTICIDES

BUTYLATE	PESTICIDES
CAPTAN	PESTICIDES
CARBARYL	PESTICIDES
CARBOFURAN	PESTICIDES
CARBOFURAN/FURADAN	PESTICIDES
CARBON DISULFIDE	PESTICIDES
CHLORAMBEN	PESTICIDES
CHLORDANE	PESTICIDES
CHLORDANE IN FISH TISSUE	PESTICIDES
CHLORDANE IN SEDIMENT	PESTICIDES
CHLORINATED PESTICIDES	PESTICIDES
CHLOROBENZILATE	PESTICIDES
CHLOROPHENOXY HERBICIDES (2,4,5,-TP)	PESTICIDES
CHLOROPHENOXY HERBICIDES (2,4-D)	PESTICIDES
CHLOROTHALONIL	PESTICIDES
CHLORPYRIFOS	PESTICIDES
CHRORDANE	PESTICIDES
CYANAZINE	PESTICIDES
CYCLOATE	PESTICIDES
CYPERMETHRIN	PESTICIDES
DACTHAL	PESTICIDES
DALAPON	PESTICIDES
DDD (DICHLORODIPHENYLDICHLOROETHANE)	PESTICIDES
DDD IN FISH TISSUE	PESTICIDES
DDE (DICHLORODIPHENYLDICHLOROETHYLENE)	PESTICIDES
DDE IN FISH TISSUE	PESTICIDES
DDT (DICHLORODIPHENYLTRICHLOROETHANE)	PESTICIDES
DDT IN FISH TISSUE	PESTICIDES
DDT IN SEDIMENT	PESTICIDES
DDT METABOLITES	PESTICIDES
DEHP (DI-SEC-OCTYL PHTHALATE)	PESTICIDES
DELTA-BHC	PESTICIDES
DEMETON	PESTICIDES
DIALLATE	PESTICIDES
DIAZINON	PESTICIDES
DIBUTYL PHTHALATE	PESTICIDES
DICHLOBENIL	PESTICIDES
DICHLORVOS	PESTICIDES
DICOFOL	PESTICIDES
DIELDRIN	PESTICIDES
DIELDRIN IN FISH TISSUE	PESTICIDES
DIELDRIN IN SEDIMENT	PESTICIDES
DIMETHOATE	PESTICIDES

DIMETHYL PHTHALATE	PESTICIDES
DIMETHYL THALATE	PESTICIDES
DINITRO-O-CRESOL	PESTICIDES
DINOSEB	PESTICIDES
DIPHENAMID	PESTICIDES
DIQUAT	PESTICIDES
DISULFOTON	PESTICIDES
DIURON	PESTICIDES
DYFONATE (FONOFOS OR FONOPHOS)	PESTICIDES
ELDRIN	PESTICIDES
ENDOSULFAN	PESTICIDES
ENDOSULFAN SULFATE	PESTICIDES
ENDOTHALL	PESTICIDES
ENDRIN	PESTICIDES
ENDRIN ALDEHYDE	PESTICIDES
ENDRIN IN SEDIMENT	PESTICIDES
EPTC	PESTICIDES
ETHELYNE DIBROMIDE	PESTICIDES
ETHOPROP	PESTICIDES
FIPRONIL	PESTICIDES
FLUOMETURON	PESTICIDES
FONOFOS	PESTICIDES
FORMALDEHYDE	PESTICIDES
GLYPHOSATE	PESTICIDES
GUTHION	PESTICIDES
HEPTACHLOR	PESTICIDES
HEPTACHLOR EPOXIDE	PESTICIDES
HEPTACHLOR EPOXIDE IN FISH TISSUE	PESTICIDES
HEXACHLOROBENZENE	PESTICIDES
HEXACHLOROCYCLOHEXANE	PESTICIDES
HEXACHLOROCYCLOHEXANE (HCH)	PESTICIDES
HEXACHLOROPHENE	PESTICIDES
HEXAZINONE	PESTICIDES
INDENO[1,2,3-CD]PYRENE	PESTICIDES
KEPONE	PESTICIDES
LAMDA-CYHALOTHRIN	PESTICIDES
LINDANE	PESTICIDES
LINURON	PESTICIDES
MALATHION	PESTICIDES
METHANOL	PESTICIDES
METHOXYCHLOR	PESTICIDES
METHYL BROMIDE	PESTICIDES
METHYL PARATHION	PESTICIDES

METOLACHLOR	PESTICIDES
METRIBUZIN	PESTICIDES
MIREX	PESTICIDES
MOLINAT (ODRAM)	PESTICIDES
MOLINATE	PESTICIDES
M-XYLENE	PESTICIDES
NAPHTHALENE	PESTICIDES
NAPROPAMIDE	PESTICIDES
NITROFEN	PESTICIDES
ORYZALIN	PESTICIDES
OXADIAZON	PESTICIDES
OXAMYL (VYDATE)	PESTICIDES
OXYFLUORFEN	PESTICIDES
P,P' DDD	PESTICIDES
PARATHION	PESTICIDES
PEBULATE	PESTICIDES
PERMETHRIN	PESTICIDES
PESTICIDES	PESTICIDES
PHOTOMIREX	PESTICIDES
PICLORAM	PESTICIDES
PROMETON	PESTICIDES
PROMETRYN	PESTICIDES
PRONAMIDE	PESTICIDES
PROPACHLOR	PESTICIDES
PROPARGITE	PESTICIDES
PROPAZINE	PESTICIDES
PROPOXUR	PESTICIDES
P-XYLENE	PESTICIDES
PYRETHROIDS	PESTICIDES
QUINTOZENE	PESTICIDES
SIMAZINE	PESTICIDES
SIMETRYN	PESTICIDES
TEBUTHIURON	PESTICIDES
TERBACIL	PESTICIDES
TERBUFOS	PESTICIDES
TETRACHLORVINPHOS	PESTICIDES
TOXAPHENE	PESTICIDES
TOXAPHENE IN FISH TISSUE	PESTICIDES
TOXAPHENE IN SEDIMENT	PESTICIDES
TRIALLATE	PESTICIDES
TRIBUTYLTIN	PESTICIDES
TRICHLORFON	PESTICIDES
TRIFLURALIN	PESTICIDES

VERNOLATE	PESTICIDES
XYLENE	PESTICIDES
ZINEB	PESTICIDES
ALKALINITY	PH/ACIDITY/CAUSTIC CONDITIONS
PH	PH/ACIDITY/CAUSTIC CONDITIONS
PH, HIGH	PH/ACIDITY/CAUSTIC CONDITIONS
PH, LOW	PH/ACIDITY/CAUSTIC CONDITIONS
PCBS - FISH CONSUMPTION ADVISORY	POLYCHLORINATED BIPHENYLS (PCBS)
PCBS IN FISH TISSUE	POLYCHLORINATED BIPHENYLS (PCBS)
PCBS IN MIGRATORY SPECIES	POLYCHLORINATED BIPHENYLS (PCBS)
PCBS IN SEDIMENT	POLYCHLORINATED BIPHENYLS (PCBS)
POLYCHLORINATED BIPHENYLS (PCBS)	POLYCHLORINATED BIPHENYLS (PCBS)
ALPHA PARTICLES	RADIATION
BARIUM	RADIATION
BETA PARTICLES AND PHOTON EMITTERS	RADIATION
CESIUM	RADIATION
RADIATION	RADIATION
RADIUM	RADIATION
TRITIUM	RADIATION
URANIUM	RADIATION
CHLORIDE	SALINITY/TOTAL DISSOLVED SOLIDS/CHLORIDES/SULFATES
SALINITY	SALINITY/TOTAL DISSOLVED SOLIDS/CHLORIDES/SULFATES
SALINITY/TOTAL DISSOLVED SOLIDS/CHLORIDES	SALINITY/TOTAL DISSOLVED SOLIDS/CHLORIDES/SULFATES
SODIUM	SALINITY/TOTAL DISSOLVED SOLIDS/CHLORIDES/SULFATES
SPECIFIC CONDUCTIVITY	SALINITY/TOTAL DISSOLVED SOLIDS/CHLORIDES/SULFATES
SULFATE	SALINITY/TOTAL DISSOLVED SOLIDS/CHLORIDES/SULFATES
SULFATE + CHLORIDE	SALINITY/TOTAL DISSOLVED SOLIDS/CHLORIDES/SULFATES
TOTAL DISSOLVED SOLIDS (TDS)	SALINITY/TOTAL DISSOLVED SOLIDS/CHLORIDES/SULFATES
COARSE SEDIMENT	SEDIMENT
FINE SEDIMENT	SEDIMENT
PARTICLE DISTRIBUTION (EMBEDDEDNESS)	SEDIMENT
SEDIMENT	SEDIMENT
SEDIMENTATION	SEDIMENT
SEDIMENTATION/SILTATION	SEDIMENT
SILICA	SEDIMENT
SILICATE	SEDIMENT
SILTATION	SEDIMENT
TOTAL SUSPENDED SEDIMENT	SEDIMENT
COLOR	TASTE, COLOR, AND ODOR
ODOR	TASTE, COLOR, AND ODOR
TASTE	TASTE, COLOR, AND ODOR

TEMPERATURE	TEMPERATURE
THERMAL MODIFICATIONS	TEMPERATURE
AMBIENT BIOASSAYS - ACUTE AQUATIC TOXICITY	TOTAL TOXICS
AMBIENT BIOASSAYS - CHRONIC AQUATIC TOXICITY	TOTAL TOXICS
CHRONIC TOXICITY	TOTAL TOXICS
SEDIMENT BIOASSAY	TOTAL TOXICS
TOXICITY	TOTAL TOXICS
WATER COLUMN BIOASSAY	TOTAL TOXICS
WHOLE EFFLUENT TOXICITY (WET)	TOTAL TOXICS
ANTIMONY	TOXIC INORGANICS
ASBESTOS	TOXIC INORGANICS
BORON	TOXIC INORGANICS
CYANIDE	TOXIC INORGANICS
FLUORIDE	TOXIC INORGANICS
HYDROGEN SULFIDE	TOXIC INORGANICS
PERCHLORATE	TOXIC INORGANICS
1,1,1,2-TETRACHLOROETHANE	TOXIC ORGANICS
1,1,1-TRICHLOROETHANE	TOXIC ORGANICS
1,1,2,2-TETRACHLOROETHANE	TOXIC ORGANICS
1,1,2-TRICHLOROETHANE	TOXIC ORGANICS
1,1-DICHLORO-1,2,2-TRIFLUOROETHANE	TOXIC ORGANICS
1,1-DICHLOROETHANE	TOXIC ORGANICS
1,1-DICHLOROETHYLENE	TOXIC ORGANICS
1,2,3,4-TETRACHLOROBENZENE	TOXIC ORGANICS
1,2,4,5-TETRACHLOROBENZENE	TOXIC ORGANICS
1,2,4-TRICHLOROBENZENE	TOXIC ORGANICS
1,2,4-TRIMETHYLBENZENE	TOXIC ORGANICS
1,2-BUTYLENE OXIDE	TOXIC ORGANICS
1,2-DIBROMOETHANE	TOXIC ORGANICS
1,2-DICHLOROBENZENE	TOXIC ORGANICS
1,2-DICHLOROETHYLENE	TOXIC ORGANICS
1,2-DIPHENYLHDRAZINE	TOXIC ORGANICS
1,2-DIPHENYLHYDRAZINE	TOXIC ORGANICS
1,2-PROPANEDIOL	TOXIC ORGANICS
1,3-BUTADIENE	TOXIC ORGANICS
1,3-DICHLOROBENZENE	TOXIC ORGANICS
1,4-DICHLOROBENZENE	TOXIC ORGANICS
1,4-DIOXANE	TOXIC ORGANICS
2,2'-DICHLORODIETHYL ETHER	TOXIC ORGANICS
2,2'-DICHLORODIISOPROPYL ETHER	TOXIC ORGANICS
2,4,6-TRICHLOROPHENOL	TOXIC ORGANICS
2,4-DIAMINOTOLUENE	TOXIC ORGANICS
2,4-DICHLOROPHENOL	TOXIC ORGANICS

2,4-DIMETHYLPHENOL	TOXIC ORGANICS
2,4-DINITROTOLUENE	TOXIC ORGANICS
2,5-DICHLOROPHENOL	TOXIC ORGANICS
2,6-DINITROTOLUENE	TOXIC ORGANICS
2-ACETYLAMINOFLUORENE	TOXIC ORGANICS
2-BUTANONE	TOXIC ORGANICS
2-CHLOROETHYL VINYL ETHER	TOXIC ORGANICS
2-CHLORONAPHTHALENE	TOXIC ORGANICS
2-CHLOROPHENOL	TOXIC ORGANICS
2-ETHOXYETHANOL	TOXIC ORGANICS
2-HEXANONE	TOXIC ORGANICS
2-METHOXYETHANOL	TOXIC ORGANICS
2-METHYLPHENOL	TOXIC ORGANICS
2-METHYLPYRIDINE	TOXIC ORGANICS
2-NITROPHENOL	TOXIC ORGANICS
3,3'-DICHLOROBENZIDINE	TOXIC ORGANICS
3,3'-DIMETHOXYBENZIDINE	TOXIC ORGANICS
3,3'-DIMETHYLBENZIDINE	TOXIC ORGANICS
3,4-DICHLOROPHENOL	TOXIC ORGANICS
3-CHLOROPHENOL	TOXIC ORGANICS
4,4-DICHLORO-2-BUTENE	TOXIC ORGANICS
4,4'-ISOPROPYLIDENEDIPHENOL	TOXIC ORGANICS
4,4'-METHYLENEBIS	TOXIC ORGANICS
4-AMINOBIPHENYL	TOXIC ORGANICS
4-BROMOPHENYLPHENYL ETHER	TOXIC ORGANICS
4-CHLORO-3-METHYLPHENOL (3-METHYL-4-CHLOROPHENOL)	TOXIC ORGANICS
4-CHLOROPHENOL	TOXIC ORGANICS
4-DIMETHYLAMINOAZOBENZENE	TOXIC ORGANICS
4-METHYL-2-PENTANONE (MIBK)	TOXIC ORGANICS
4-METHYLPHENOL	TOXIC ORGANICS
4-NITROPHENOL	TOXIC ORGANICS
5-NITRO-O-TOLUIDINE	TOXIC ORGANICS
ACENAPHTHENE	TOXIC ORGANICS
ACENAPHTHYLENE	TOXIC ORGANICS
ACETALDEHYDE	TOXIC ORGANICS
ACETOCHLOR	TOXIC ORGANICS
ACETONE	TOXIC ORGANICS
ACETONITRILE	TOXIC ORGANICS
ACRYLAMIDE	TOXIC ORGANICS
ALKYLBENZENE	TOXIC ORGANICS
ALLYL ALCOHOL	TOXIC ORGANICS
ALLYL CHLORIDE	TOXIC ORGANICS
ALPHA-BNC	TOXIC ORGANICS

ALPHA-NAPHTHYLAMINE	TOXIC ORGANICS
ANILINE	TOXIC ORGANICS
ANTHRACENE	TOXIC ORGANICS
BENTAZONE	TOXIC ORGANICS
BENZ[A]ANTHRACENE	TOXIC ORGANICS
BENZAL CHLORIDE	TOXIC ORGANICS
BENZENE	TOXIC ORGANICS
BENZIDINE	TOXIC ORGANICS
BENZO[A]ANTHRACENE	TOXIC ORGANICS
BENZO[A]PYRENE	TOXIC ORGANICS
BENZO[A]PYRENE (PAHS)	TOXIC ORGANICS
BENZO[B,K]FLUORANTHENES	TOXIC ORGANICS
BENZO[B]FLUORANTHENE	TOXIC ORGANICS
BENZO[B]FLUORENE	TOXIC ORGANICS
BENZO[G,H,I]PERYLENE	TOXIC ORGANICS
BENZO[K]FLUORANTHENE	TOXIC ORGANICS
BENZO[K]FLUORENE	TOXIC ORGANICS
BENZOFLUORANTHENES TOTAL (B+K+J)	TOXIC ORGANICS
BENZOIC ACID	TOXIC ORGANICS
BENZOPYRENE	TOXIC ORGANICS
BENZOYL CHLORIDE	TOXIC ORGANICS
BENZYL ALCOHOL	TOXIC ORGANICS
BENZYL CHLORIDE	TOXIC ORGANICS
BETA-NAPHTHYLAMINE	TOXIC ORGANICS
BIPHENYL	TOXIC ORGANICS
BIS(2 ETHYLHEXYL)PHTHALATE	TOXIC ORGANICS
BIS(2 ETHYLHEXYL)PHTHALATE AND PHENOL	TOXIC ORGANICS
BIS(2-CHLORO-1-METHYLETHYL)	TOXIC ORGANICS
BIS(2-CHLOROETHOXY)METHANE	TOXIC ORGANICS
BIS(2-CHLOROISOPROPYL) ETHER	TOXIC ORGANICS
BIS(2-ETHYLHEXYL) PHTHALATE	TOXIC ORGANICS
BIS(N-OCTYL) PHTHALATE	TOXIC ORGANICS
BIS-2-CHLOROETHYL ETHER	TOXIC ORGANICS
BISPHTHALATE	TOXIC ORGANICS
BROMACIL	TOXIC ORGANICS
BROMODICHLOROMETHANE	TOXIC ORGANICS
BROMOFORM	TOXIC ORGANICS
BUTYL BENZYL PHTHALATE	TOXIC ORGANICS
BUTYRALDEHYDE	TOXIC ORGANICS
CARBON TETRACHLORIDE	TOXIC ORGANICS
CESETHYLATRAZINE	TOXIC ORGANICS
CHLORAMINES	TOXIC ORGANICS
CHLORINATED BENZENES	TOXIC ORGANICS

CHLORINATED PHENOLS	TOXIC ORGANICS
CHLOROACETIC ACID	TOXIC ORGANICS
CHLOROBENZENE (MONO)	TOXIC ORGANICS
CHLORODIBROMOMETHANE	TOXIC ORGANICS
CHLORODIFLUOROMETHANE	TOXIC ORGANICS
CHLOROETHANE	TOXIC ORGANICS
CHLOROFORM	TOXIC ORGANICS
CHLOROMETHANE	TOXIC ORGANICS
CHLOROMETHYL METHYL ETHER	TOXIC ORGANICS
CHLOROPHENYL-4 PHENYL ETHER	TOXIC ORGANICS
CHLOROPRENE	TOXIC ORGANICS
CHRYSENE	TOXIC ORGANICS
CIS-1,2-DICHLOROETHYLENE	TOXIC ORGANICS
COAL ASH	TOXIC ORGANICS
COAL TAR	TOXIC ORGANICS
CONTAMINATED SEDIMENTS (BENZO[A]ANTHRACENE)	TOXIC ORGANICS
CONTAMINATED SEDIMENTS (BENZO[A]PYRENE)	TOXIC ORGANICS
CONTAMINATED SEDIMENTS (CHRYSENE)	TOXIC ORGANICS
CONTAMINATED SEDIMENTS (FLUORANTHENE)	TOXIC ORGANICS
CONTAMINATED SEDIMENTS (HYDROCARBONS)	TOXIC ORGANICS
CONTAMINATED SEDIMENTS (PAHS)	TOXIC ORGANICS
CONTAMINATED SEDIMENTS (PHENANTHRENE)	TOXIC ORGANICS
CONTAMINATED SEDIMENTS (PYRENE)	TOXIC ORGANICS
CREOSOTE	TOXIC ORGANICS
CRESOL (MIXED ISOMERS)	TOXIC ORGANICS
CUMENE	TOXIC ORGANICS
CYCLOHEXANAMINE, N-ETHYL-1-PHENYL-	TOXIC ORGANICS
CYCLOHEXANE	TOXIC ORGANICS
CYMENE	TOXIC ORGANICS
DEETHYLATRAZINE	TOXIC ORGANICS
DESETHYLATRAZINE	TOXIC ORGANICS
DESISOPROYLATRAZINE	TOXIC ORGANICS
DI(2-ETHYLHEXYL) ADIPATE	TOXIC ORGANICS
DI-2-ETHYLHEXYL PHTHALATE	TOXIC ORGANICS
DIAMINOTOLUENE (MIXED ISOMERS)	TOXIC ORGANICS
DIBENZ[A,H]ANTHRACENE	TOXIC ORGANICS
DIBROMOCHLOROMETHANE	TOXIC ORGANICS
DICHLOROBENZENE (MIXED ISOMERS)	TOXIC ORGANICS
DICHLOROBROMOMETHANE	TOXIC ORGANICS
DICHLORODIFLUOROMETHANE	TOXIC ORGANICS
DICHLOROETHANE	TOXIC ORGANICS
DICHLOROETHANE/POLYCYCLIC AROMATIC HYDROCARBONS	TOXIC ORGANICS
DICHLOROETHYLENE/1,1-DCE	TOXIC ORGANICS

DICHLOROETHYLENES	TOXIC ORGANICS
DICHLOROMETHANE	TOXIC ORGANICS
DIETHYL PHTHALATE	TOXIC ORGANICS
DI-N-BUTYL PHTHALATE	TOXIC ORGANICS
DINITROTOLUENE	TOXIC ORGANICS
DI-N-OCTYL PHTHALATE	TOXIC ORGANICS
DISINFECTION BY-PRODUCTS	TOXIC ORGANICS
DODECYLBENZENE	TOXIC ORGANICS
EPICHLOROHYDRIN	TOXIC ORGANICS
ETHER, BIS CHLOROMETHYL	TOXIC ORGANICS
ETHYLBENZENE	TOXIC ORGANICS
ETHYLENE	TOXIC ORGANICS
ETHYLENE GLYCOL	TOXIC ORGANICS
ETHYLENE OXIDE	TOXIC ORGANICS
ETHYLENE THIOUREA	TOXIC ORGANICS
FISH CONSUMPTION ADVISORY - PAHS	TOXIC ORGANICS
FLUORANTHENE	TOXIC ORGANICS
FLUORENE	TOXIC ORGANICS
FORMIC ACID	TOXIC ORGANICS
HALOMETHANES	TOXIC ORGANICS
HEXACHLOROBUTADIENE	TOXIC ORGANICS
HEXACHLOROCYCLOPENTADIENE	TOXIC ORGANICS
HEXACHLOROETHANE	TOXIC ORGANICS
HEXAMETHYLPHOSPHORAMIDE	TOXIC ORGANICS
HYDRAZINE	TOXIC ORGANICS
HYDROCARBONS	TOXIC ORGANICS
HYDROCARBONS - NON PRIORITY	TOXIC ORGANICS
HYDROCARBONS - PRIORITY ORGANICS	TOXIC ORGANICS
HYDROQUINONE	TOXIC ORGANICS
ISOBUTYRALDEHYDE	TOXIC ORGANICS
ISOPHORONE	TOXIC ORGANICS
ISOPROPANOL	TOXIC ORGANICS
ISOSAFROLE	TOXIC ORGANICS
MALEIC ANHYDRIDE	TOXIC ORGANICS
M-CRESOL	TOXIC ORGANICS
M-DICHLOROBENZENE	TOXIC ORGANICS
M-DINITROBENZENE	TOXIC ORGANICS
METHACRYLONITRILE	TOXIC ORGANICS
METHYL BLUE	TOXIC ORGANICS
METHYL CHLORIDE	TOXIC ORGANICS
METHYL ETHYL KETONE	TOXIC ORGANICS
METHYL HYDRAZINE	TOXIC ORGANICS
METHYL IODIDE	TOXIC ORGANICS

METHYL ISOBUTYL KETONE	TOXIC ORGANICS
METHYL METHACRYLATE	TOXIC ORGANICS
METHYL TERTIARY-BUTYL ETHER (MTBE)	TOXIC ORGANICS
METHYLENE BROMIDE	TOXIC ORGANICS
METHYLENE CHLORIDE	TOXIC ORGANICS
MTBE	TOXIC ORGANICS
N-BUTYL ALCOHOL	TOXIC ORGANICS
N-BUTYLBENZYLPHTHALATE	TOXIC ORGANICS
NITRILOTRIACETIC ACID	TOXIC ORGANICS
NITROBENZENE	TOXIC ORGANICS
NITRODIBUTYLAMINE,N	TOXIC ORGANICS
NITROGLYCERIN	TOXIC ORGANICS
NITROSODIETHYLAMINE,N	TOXIC ORGANICS
N-NITROSODIETHYLAMINE	TOXIC ORGANICS
N-NITROSODIMETHYLAMINE	TOXIC ORGANICS
N-NITROSO-DI-N-BUTYLAMINE	TOXIC ORGANICS
N-NITROSODIPHENYLAMINE	TOXIC ORGANICS
N-NITROSODIPROPYLAMINE	TOXIC ORGANICS
N-NITROSOMORPHOLINE	TOXIC ORGANICS
N-NITROSO-N-ETHYLUREA	TOXIC ORGANICS
N-NITROSO-N-METHYLUREA	TOXIC ORGANICS
N-NITROSOPIPERIDINE	TOXIC ORGANICS
N-NONYLBENZENE	TOXIC ORGANICS
NONPRIORITY ORGANICS/PAHS	TOXIC ORGANICS
NONYLPHENOL	TOXIC ORGANICS
O-CRESOL (2-METHYLPHENOL)	TOXIC ORGANICS
OCTACHLOROSTYRENE	TOXIC ORGANICS
OCTOCHLORONAPHTHALENE	TOXIC ORGANICS
O-DICHLOROBENZENE	TOXIC ORGANICS
OIL SPILL - PAHS	TOXIC ORGANICS
ORGANIC CHEMICALS	TOXIC ORGANICS
ORGANICS	TOXIC ORGANICS
OTHER ORGANICS	TOXIC ORGANICS
OTHER ORGANICS (FLORIDE)	TOXIC ORGANICS
OTHER TOXIC ORGANICS	TOXIC ORGANICS
O-TOLUIDINE	TOXIC ORGANICS
O-TOLUIDINE HYDROCHLORIDE	TOXIC ORGANICS
O-XYLENE	TOXIC ORGANICS
PAH1 - 2 & 3 RING POLYCYCLIC AROMATIC HYDROCARBONS	TOXIC ORGANICS
PAH2 - 4 RING POLYCYCLIC AROMATIC HYDROCARBONS	TOXIC ORGANICS
PAH3 - 5 & 6 RING POLYCYCLIC AROMATIC HYDROCARBONS	TOXIC ORGANICS
PARALDEHYDE	TOXIC ORGANICS
PCE	TOXIC ORGANICS

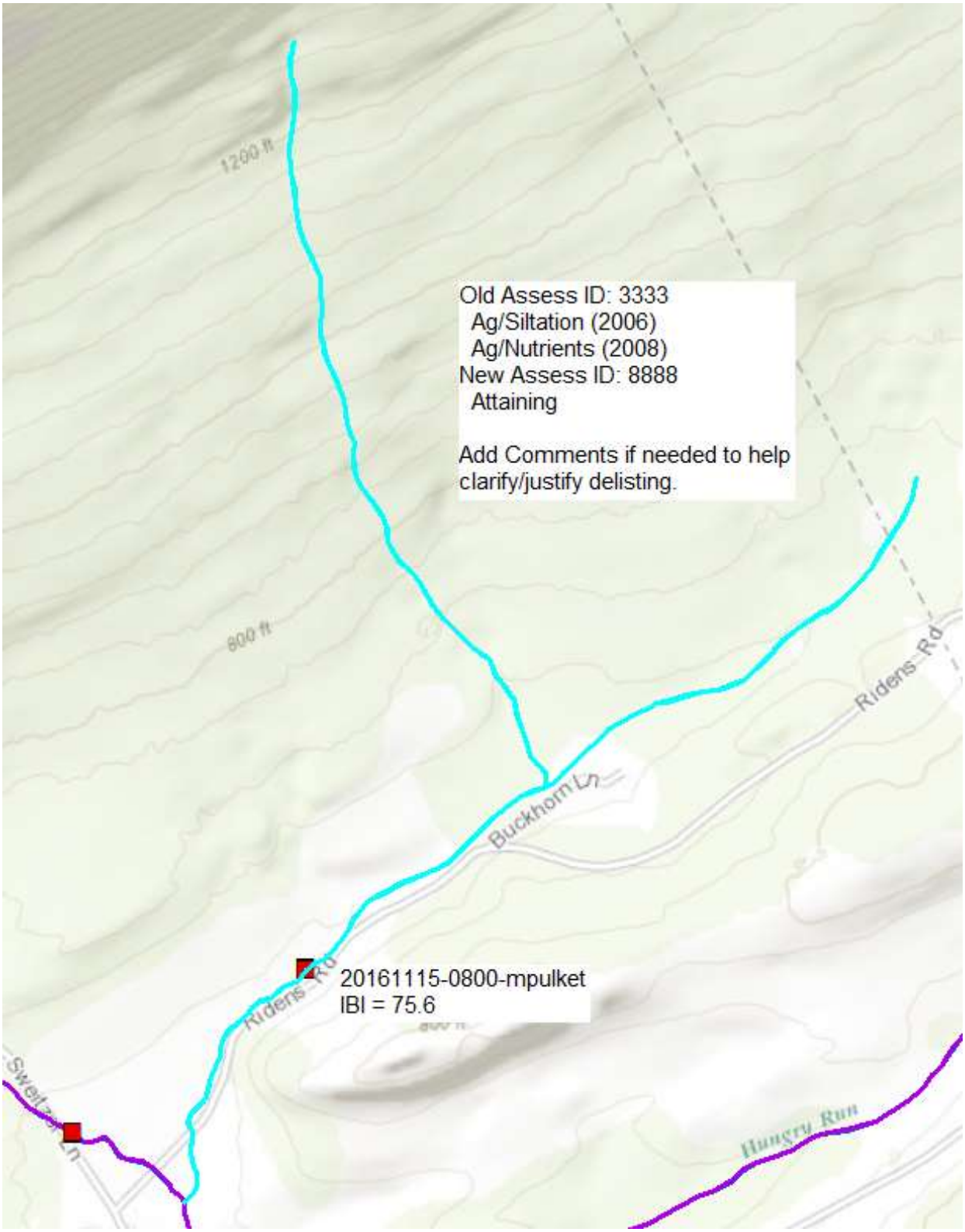
P-DICHLOROBENZENE	TOXIC ORGANICS
PENTACHLOROBENZENE	TOXIC ORGANICS
PENTACHLOROPHENOL (PCP)	TOXIC ORGANICS
PERFLUOROOCTANE SULFONATE (PFOS)	TOXIC ORGANICS
PERFLUOROOCTANE SULFONATE (PFOS) IN FISH TISSUE	TOXIC ORGANICS
PHENANTHRENE	TOXIC ORGANICS
PHENOL	TOXIC ORGANICS
PHTHALATE ESTERS	TOXIC ORGANICS
PHTHALIC ANHYDRIDE	TOXIC ORGANICS
PHTHLATE	TOXIC ORGANICS
PICRIC ACID	TOXIC ORGANICS
POLYBROMINATED BIPHENYLS	TOXIC ORGANICS
POLYCYCLIC AROMATIC HYDROCARBONS (PAHS) (AQUATIC ECOSYSTEMS)	TOXIC ORGANICS
PROPIONALDEHYDE	TOXIC ORGANICS
PROPYLENE GLYCOL	TOXIC ORGANICS
PROPYLENE OXIDE	TOXIC ORGANICS
PYRENE	TOXIC ORGANICS
PYRIDINE	TOXIC ORGANICS
QUINOLINE	TOXIC ORGANICS
QUINONE	TOXIC ORGANICS
RDX (HEXAHYDRO-1,3,5-TRINITRO-1,3,5-TRIAZINE)	TOXIC ORGANICS
SAFROLE	TOXIC ORGANICS
SEC-BUTYL ALCOHOL	TOXIC ORGANICS
STYRENE	TOXIC ORGANICS
TETRACHLOROETHYLENE	TOXIC ORGANICS
TETRACHLOROETHYLENE/PCE	TOXIC ORGANICS
THIOUREA	TOXIC ORGANICS
TOLUENE	TOXIC ORGANICS
TOTAL AROMATIC HYDROCARBONS	TOXIC ORGANICS
TOTAL BENZOFLUORANTHENES	TOXIC ORGANICS
TOTAL TRIHALOMETHANE (TTHM)	TOXIC ORGANICS
TRANS-1,2-DICHLOROETHENE	TOXIC ORGANICS
TRANS-1,2-DICHLOROETHYLENE	TOXIC ORGANICS
TRIBUTYLIN TBT (TRIBUTYLSTANNE)	TOXIC ORGANICS
TRICHLORINATED ETHANES	TOXIC ORGANICS
TRICHLOROETHYLENE (TCE)	TOXIC ORGANICS
TRICHLOROFLUOROMETHANE (CFC-11)	TOXIC ORGANICS
TRICLOPYR	TOXIC ORGANICS
TRIETHYLENE GLYCOL DICHLORIDE	TOXIC ORGANICS
VINYL ACETATE	TOXIC ORGANICS
VINYL BROMIDE	TOXIC ORGANICS
VINYL CHLORIDE	TOXIC ORGANICS

VINYLIDENE CHLORIDE	TOXIC ORGANICS
VOLATILE ORGANICS (VOCS)	TOXIC ORGANICS
TRASH	TRASH
TOTAL SUSPENDED SOLIDS (TSS)	TURBIDITY
TRANSPARENCY / CLARITY	TURBIDITY
TURBIDITY	TURBIDITY

APPENDIX B: DELISTING EXAMPLE

HUC: 02050305
Aquatic Life Use
Category 5 to 2

Test Creek Delisting



Required Information:

1. Title the map with the waterbody name.
2. Include the HUC, the use being assessed, Assessment ID with source(s) and cause(s) of impairment, new Assessment ID if available, and new status.
3. If any causes will remain, that must be made clear.
4. Highlight or clearly depict the stream segment(s) or lake being delisted.
5. Label all stations with GISkey (yyyymmdd-HHMM-collector; e.g., 20161115-0800-mpulket) or unique station identifier.
6. Include the IBI score, geometric mean, and/or chemistry data when applicable.
7. Use an appropriate basemap layer; this example uses World Topographic Map.
8. The following information is not required but is very useful for tracking delistings.
    a. The listing date of the Sources/Causes (in parentheses above),
    b. The current and new IR category
    c. Any comments or additional information to help clarify and justify the delisting.