Development of a Basin-wide Benthic Index of Biotic Integrity for Non-tidal Streams and Wadeable Rivers in the Chesapeake Bay Watershed:

Final Report to the Chesapeake Bay Program Non-Tidal Water Quality Workgroup

May 9, 2011

Claire Buchanan¹ Katherine Foreman² Jacqueline Johnson³ Adam Griggs¹





Chesapeake Bay Program A Watershed Partnership



¹ Interstate Commission on the Potomac River Basin, 51 Monroe St. PE-08, Rockville, MD 20850

- ² University of Maryland Center for Environmental Science, located at the Chesapeake Bay Program Office, 410 Severn Avenue, Suite 112, Annapolis, Maryland 21403
- ³ Interstate Commission on the Potomac River Basin, located at the Chesapeake Bay Program Office, 410 Severn Avenue, Suite 112, Annapolis, Maryland 21403

ICPRB Report 11-1

This report can be downloaded from the Publications tab of the Commission's website, <u>www.potomacriver.org</u>. To receive hard copies of the report, please write

Interstate Commission on the Potomac River Basin 51 Monroe St., PE-08 Rockville, MD 20850

or call 301-984-1908.

Disclaimer

The opinions expressed in this report are those of the authors and should not be construed as representing the opinions or policies of the U. S. Government, the U. S. Environmental Protection Agency, the several states, the signatories or Commissioners to the Interstate Commission on the Potomac River Basin, or the University of Maryland Center for Environmental Science. No official endorsement should be inferred.

Suggested citation for this report

Buchanan, C., K. Foreman, J. Johnson, and A. Griggs. 2011. Development of a Basin-wide Benthic Index of Biotic Integrity for Non-Tidal Streams and Wadeable Rivers in the Chesapeake Bay Watershed: Final Report to the Chesapeake Bay Program Non-Tidal Water Quality Workgroup. ICPRB Report 11-1. Report prepared for the US Environmental Protection Agency, Chesapeake Bay Program.

Acknowledgementsiv
Executive Summaryv
Introduction1
Scope of report2
Adjustments and assumptions2
Data Sources
Bioregions4
Kruskal-Wallis one-way analysis of variance6
Category and Regression Tree Analysis6
Coastal Plain
Non-Coastal Plain10
Candidate biological metric identification10
Reference and Degraded site selection12
Piedmont, Ridges, and Valleys12
Northern Appalachian Plateau and Uplands and North Central Appalachians 14
Impacts of stream order on biological metrics
Impacts of season on biological metrics
Influence of Karst geology on biological metrics
Standardization of count metrics
Biological metric testing and selection21
Statistical test for significance
Discrimination efficiencies
Metric selection
Scoring the biological metrics and calculating the index
Jackknife validation of non-Coastal Plain B-IBIs
Narrative Condition Ratings
Findings
Comparisons with state ratings
Mapping the Chessie B-IBI results40
Possible Next Steps
References44
Appendix A Feeding Guild, Habit, and Tolerance Assignments for Family-Level Taxa

Table of Contents

Appendix B Season and Strahler Order Effects on Biological Metrics

Appendix C Number of Family-Level Taxa in the Piedmont, Ridges, and Valleys Bioregions

Appendix D Discrimination Efficiencies of the 42 Candidate Metrics

Acknowledgements

This project was made possible through partial support provided by US Environmental Protection Agency grants CB-97339104, CB-97339105, CB-97339702 and I-98339410, and by the Interstate Commission on the Potomac River Basin (ICPRB) and the University of Maryland Center for Environmental Science (UMCES).

The Basin-wide Index of Biotic Integrity for benthic macroinvertebrates in Chesapeake non-tidal streams and wadeable rivers (Chessie B-IBI) is based on a Potomac River basin index produced by LeAnne Astin, formerly of ICPRB and now with the Fairfax County Stormwater Planning Division. The Chessie B-IBI in its current form was developed in collaboration with biologists from across the Chesapeake Bay watershed. An adhoc CBP workgroup created to guide development of the Chessie B-IBI consisted of benthic macroinvertebrate experts from the six states in the watershed (New York, Pennsylvania, Maryland, Virginia, West Virginia, and Delaware) as well as federal, academic, and River Basin Commission partners. The authors wish to give special thanks to the members of the adhoc workgroup for their diligence in providing technical guidance and feedback: A.J. Smith (NYDEC), Aimee Budd (VADEQ), Bill Richardson (US EPA Region 3), Brian Chalfant (PADEP), Charlie Poukish (MDE), Dan Boward (MD DNR), Ed Reilly (NYDEC), Ellen Dickey (DNREC), Greg Garman (VCU), Greg Pond (US EPA Region 3), Hassan Mirsajadi (DNREC), Jeff Bailey (WVDEP), Jen Hoffman (SRBC), John Wirts (WVDEP), Kevin McGonigal (SRBC), Maggie Passmore (US EPA Region 3), Mike Fritz (EPA-CBPO), Nita Sylvester (EPA-CBPO), Peter Tango (USGS-CBPO), Rick Hoffman (VADEQ), Rod Kime (PADEP), Ron Klauda (MD DNR), Scott Stranko (MD DNR), Tony Prochaska (MD DNR), and Wayne Davis (EPA). Other members of the Chesapeake Bay Program's Non-Tidal Water Quality Workgroup as well as the Indicator Workgroup provided input on final presentation of the results.

The Chessie B-IBI could not have been produced without the cooperation of the staff of the following monitoring programs who responded to data requests: Delaware Biological Monitoring Program; the US EPA EMAP, MAHA, and WSA programs; Fairfax Co. Stream Quality Assessment Program; Frederick Co. Watershed Management Program; Howard Co. Bio-Monitoring and Assessment Program; Loudoun Co. Stream Quality Assessment Program; Maryland Biological Stream Survey; Montgomery Co. Dept. of Environmental Protection; National Forest Service Stream Assessment; New York Routine Statewide Monitoring Program; Pennsylvania's Other Water Quality Assessments, Surface Water Monitoring and Unassessed Watersheds programs; the Pennsylvania USGS; Prince Georges Co. Programs and Planning Division; SRBC's Watershed Assessment and Protection (TMDLs) and Watershed Assessment Programs; USGS NAWQA program; Virginia's DEQ Benthic Monitoring and INteractive STream Assessment Resource (VCU) programs; and West Virginia Div. of Water and Waste Management.

The authors especially thank Maggie Passmore (US EPA Region 3) and Brian Chalfant (PA DEP) for their insightful comments and suggestions during the conceptual and developmental stages of this index; Mike Mallonee (ICPRB-CBPO) and Andrea Nagel (ICPRB) for their guidance and help in creating, populating, and updating the relational database that houses the primary data; Mary Ellen Ley (USGS-CBPO), CBP Quality Assurance Officer, for her help in resolving quality control and assurance issues as they arose; Jan Ducnuigeen (ICPRB) for delineating the New York and Pennsylvania watersheds above benthic sampling sites in the Northern Appalachian Plateau and Upland ecoregion; Greg Pond (US EPA Region 3) for providing the rarefaction computer programs; Fred Irani (USGS-CBPO) for his help in generating the maps of sampling sites and watershed condition; and Aaron Gorka (CRC-CBPO), Jeni Keisman (UMCES-CBPO), Peter Tango (USGS-CBPO) and Scott Phillips (USGS) for their technical support, encouragement and guidance.

Finally, the authors wish to acknowledge the support of the Interstate Commission on the Potomac River Basin (ICPRB), the University of Maryland Center for Environmental Science (UMCES), and the US EPA Chesapeake Bay Program (CBP) without which this project would not have been accomplished.

Executive Summary

The objective of the Chesapeake Bay basin-wide Benthic Index of Biological Integrity, or "Chessie B-IBI," is to evaluate macroinvertebrate community health in non-tidal streams and wadeable rivers in a uniform manner and in the context of the entire Chesapeake Bay basin. The index is one of several Chesapeake Bay Program analysis tools applied to multi-jurisdictional monitoring data for watershed assessments. The index is "reference-based" meaning sampling locations are evaluated according to how similar they are to sites with little or no measured anthropogenic disturbance. Data from multiple stream monitoring programs are incorporated into a common database structure and family-level metrics (indicators) are calculated from the data. The metrics are scored according to their similarity to Reference sites in each bioregion (areas with different geomorphologic traits). The scores of the metrics most responsive to anthropogenic disturbance are then averaged to obtain an index score. Index scores are rated on a 5-tiered scale: Excellent (67%-100%), Good (50%-67%), Fair (30%-50%), Poor (17%-30%), and Very Poor (0%-17%). This report describes the approach used in 2010 with the guidance of Chesapeake region biologists to refine an earlier version of the index.

The Chessie B-IBI method is designed so that results can be compared across all jurisdictional and geographical boundaries in the Chesapeake Bay basin. Variation in the Chessie B-IBI can be linked to land-based activities in individual watersheds. The best stream communities tend to be in forested areas with little land disturbance, low levels of pollution, and undisturbed in-stream and streamside habitat. The worst stream communities usually occur in watersheds with high levels of urbanization, agriculture and mining activities. Streams in these areas are compromised by extreme land disturbance, high levels of pollution, altered water flow, and poor quantity and quality of streamside vegetation. Efforts to lessen the multiple anthropogenic impacts and restore stream ecosystems should improve Chessie B-IBI scores in the affected watersheds. Efforts to protect undisturbed areas should maintain high Chessie B-IBI scores over time.

The Chessie B-IBI was applied to 10,833 stream sampling locations across the Chesapeake Bay basin. A subset of 7,886 locations were used to evaluate small (Hydrologic Unit Code 10, or HUC10) and moderate-sized (HUC8) watersheds for the 9-year period of 2000 – 2008 (**Figure i**). Monitoring programs selected this subset of locations randomly or systematically, so the data are presumably unbiased and can be combined and used to represent whole watershed status. Additional data from random/systematic sites will improve the map resolution, especially in Pennsylvania and New York. The index and thresholds used to rate the index scores are not static and can be improved with more monitoring data from Reference-quality sites. Next steps for development and analysis of the Chessie B-IBI are suggested. They include recommendations to a) calculate trends in the Chessie B-IBI over time with a subset of the data collected from fixed locations, and b) develop a methodology for using random/systematic sampling design data to assess whole watershed trends.

At this point in its development, the index is useful for regional evaluations and targeting to help managers and watershed groups focus stream restoration and protection efforts. It is not intended for state regulatory purposes and should not circumvent the indexes of stream health developed and used by state monitoring programs for regulatory purposes.



Figure i. Chessie B-IBI ratings of small (HUC10) and moderately sized (HUC8) watersheds in the Chesapeake Bay basin. HUC10 watersheds are overlain on HUC8 watersheds. Only data from randomly or systematically (grid) sampled sites are used to illustrate status. <u>Note</u>: in the 2000 - 2008 period, Pennsylvania and New York monitoring programs collected data from many sampling sites identified as "targeted" (such as sites below an pollutant source) and these data were not included here in order to avoid the presumed bias introduced by targeted site data. Watershed ratings in Pennsylvania and New York are therefore less certain because they are derived from fewer random/systematic sites (see Report Figure 18).

Introduction

Chesapeake Bay Program (CBP) partners collaborated in 2010 to refine a prototype benthic index of biotic integrity ("Chessie B-IBI") for regional assessments of the health of the streams and wadeable rivers in the Chesapeake Bay watershed (basin). Multiple federal, state, river basin, and local monitoring programs collect benthic macroinvertebrate samples in the 64,000 square mile Chesapeake Bay watershed. These programs use somewhat similar field methods and calculate a common suite of indicators from the data. The challenge is that each program scores and evaluates these indicators differently in order to assess its waters for regulatory purposes, which sometimes makes assessments of the data at a regional scale impossible. The purpose of the Chessie B-IBI is to evaluate non-tidal benthic macroinvertebrate community health in a uniform manner and in the context of the entire Chesapeake Bay basin.

In 2007 the CBP Non-Tidal Water Quality workgroup (NTWG) unsuccessfully attempted to combine on a single map the stream assessments made from the six Chesapeake watershed states' 303(d)/305(b) integrated reports. In 2008 the NTWG tried another approach and expanded to the entire Chesapeake Bay watershed an existing B-IBI originally developed for the multi-jurisdictional Potomac River basin (Astin 2006, 2007). The Potomac River index uniformly scores data collected by different monitoring programs in non-tidal streams and wadeable rivers. The first attempt to develop the Chesapeake Bay basin-wide B-IBI was documented in 2008 (Foreman et al. 2008). Results from this B-IBI were published in the 2008 and 2009 Chesapeake Bay Program's "Bay Barometer" (US EPA 2008 and 2009). After the B-IBI's initial release, CBP partners requested improvements to the Chessie B-IBI. An adhoc workgroup consisting of benthic macroinvertebrate experts drawn from the NTWG was formed to guide the improvements. The following high and moderate priority tasks identified by the workgroup were accomplished in 2010 and the results are described in this report:

High Priority

- Review biological metrics and consider additional candidate metrics
- Redo identification of Reference and Degraded sites for Piedmont, Ridges, Valley, and Northern Appalachian bioregions with additional NY and MD data included; use BCG (biological condition gradient) concepts to guide selection of habitat & water quality parameter thresholds for Reference and Degraded instead of relying strictly on percentiles of the Reference communities
- Re-test biometric sensitivity are Reference and Degraded values significantly different?
- Choose biometrics for inclusion in the B-IBI for each bioregion.
- Redo and re-test biometric thresholds for all bioregions. In particular, work with NYDEC and Coastal Plain states to improve the Northern Appalachian and Coastal Plain thresholds.
- Evaluate different scoring approaches for biometrics: discrete (1-3-5), continuous (0%-100%) or hybrid (apply a gradient between 1 and 5 thresholds); compare discrimination and classification efficiencies for each approach
- Select a scoring approach and redo jackknife validations

Moderate Priority

- Compare subsampling effects and usefulness of data rarefaction: 100 vs. 200 count at family level (can actually be done with original data set),
- Assess influence of catchment area and/or stream order on the B-IBI
- Assess influence of seasonality on B-IBI results
- Assess influence of limestone on B-IBI results and develop a limestone-specific B-IBI

Scope of report

This report documents the steps taken to perform the analytical tasks above and improve the prototype Chessie B-IBI. The data preparation and analysis were performed in 2010 and early 2011 by staff of the Interstate Commission on the Potomac River Basin (ICPRB) and University of Maryland Center for Environmental Science (UMCES) with guidance from the adhoc workgroup. The resulting Chessie B-IBI is based on bioregion-specific indices designed to produce comparable index scores that can be used for stream evaluations across the Chesapeake Bay watershed regardless of state boundaries. The coastal plain regions of the Chesapeake watershed were assessed using the Coastal Plain Macroinvertebrate Index (CPMI) which was a product of the EPA Mid-Atlantic Integrated Assessment Program (US EPA 1997, Maxted et al. 2000). New interstate indices were developed for the remaining, non-coastal plain geographic regions of the Chesapeake watershed. Chessie B-IBI scores are useful for regional watershed assessment purposes, but not necessarily for state regulatory purposes, i.e., 303(d)/305(b) integrated reports. Although numeric scores of the state biological indices appear to compare favorably to Chessie B-IBI numeric scores, state ratings (e.g., "not impaired," "somewhat impaired," "impaired") are not directly comparable to the five Chessie B-IBI ratings ("excellent," "good," "fair," "poor," and "very poor").

Some moderate or additional priority tasks were either partially completed or not attempted. They include:

- reevaluate the effects of targeted vs. random sites on B-IBI results
- compare state index scores and Chessie B-IBI scores directly
- assess the influence of blackwater streams on B-IBI results
- assess the effects of gear type on results (literature-based evaluation)
- investigate the utility of including state identified Reference sites
- consider alternative approaches to rating the Chessie B-IBI (4 categories instead of 5)
- validate the B-IBI with data not used in developing and calibrating the improved Chessie B-IBI

Adjustments and assumptions

As data from various Chesapeake stream monitoring programs were prepared and analyzed, the following standards and adjustments to the data—and the assumptions behind them—were carried forward from the initial effort:

- All taxonomic counts are "rolled up" to family-level. Some monitoring programs in the watershed perform genus-level counts of benthic macroinvertebrates while others perform family-level counts. Rolling up to the family-level avoids biasing biological metrics such as taxa richness that are directly affected by the level of taxonomic identification.
- Only monitoring results using a version of the US Environmental Protection Agency (US EPA) Rapid Bioassessment Protocols (RBP) were included in the analysis. It is assumed that the RBP methods for evaluating habitat features provide a common basis which holds even though individual monitoring programs have modified the original protocols (US EPA 1989).
- Only sampling sites from streams and wadeable rivers of Strahler stream order 1-4 (1:100,000 NHD layer) were used to develop the Chessie B-IBI. The River Continuum concept (Vannote et al. 1980) predicts biological community changes as streams and rivers increase in size.
- Sampling sites are classified into "bioregions" which are regions of similar topography, soils, geologic features, and vegetation. Regional classification is known to minimize or remove differences in stream biological communities caused by natural factors. Bioregions were created from aggregations of US EPA Level IV ecoregions (US EPA 2010a) and used to group the data prior to analysis.

Biological metrics and the placement of their scoring thresholds were carefully selected to
minimize the suspected or observed impacts on the Chessie BIBI results of different sampling
season and stream order. Differences in monitoring program sampling methodologies can
affect the resulting raw data; however they may not greatly influence the results of family-level
assessments when they have a common basis such as the US EPA RBP methods (e.g., Astin 2006)
and are unit-less (percent) or standardized (100-count subsample) metrics.

Data Sources

The ICPRB and Chesapeake Bay Program Office (CBPO) acquired historical and current benthic, physical, and chemical data for non-tidal streams and wadeable rivers from 23 federal, state, local, and river basin commission monitoring programs throughout the Chesapeake Bay basin (**Table 1**). Benthic macroinvertebrate, habitat, and water quality data for the 1992-2009 time period were obtained from the data providers, reprocessed, and quality checked and assured. A tabular summary of the states' collection, counting, and assessment methods is available in US EPA 2010b. Methods used by federal, county and other programs are described in various documents that can be obtained directly from the sources or from the Living Resources Data Manager/Analyst at the US EPA CBP office in Annapolis, MD. Methodological and data comparability issues were addressed and documented before the data were incorporated into a uniform relational database structure. A subset of the available data was used to develop the Chesapeake Bay basin-wide B-IBI. To be used, a sampling event had to have been processed

	Agency			Total	* B-IBI
Agency Program Name	Code	Start Date	End Date	Count	Count
Delaware Biological Monitoring Program	DNREC	3/20/2000	11/18/2008	1128	
Frederick County Watershed Management Program	FC-DPW	6/7/1999	9/18/2009	402	
Fairfax County Stream Quality Assessment Program	FC-SPS	4/18/1999	10/3/2008	484	
Howard Co Bio-Monitoring and Assessment Program	HC_DPW	3/7/2001	3/30/2009	266	219
Loudoun County Stream Quality Assessment Program	LC-DBD	3/27/2009	7/24/2009	500	197
Montgomery Co Dept. of Environmental Protection	MC-SPS	9/1/1989	4/13/2009	3938	
Maryland Biological Stream Survey	MDDNR	5/10/1994	8/25/2008	6502	681
New York Routine Statewide Monitoring Program	NYDEC	7/29/2002	8/7/2008	346	308
Pennsylvania Other Water Quality Assessments	PADEP	1/8/2006	2/5/2007	46	35
Pennsylvania Surface Water Monitoring Program	PADEP	4/13/2000	12/18/2008	750	465
Pennsylvania Unassessed Watersheds	PADEP	5/24/2000	12/4/2003	154	14
Pennsylvania USGS	PADEP	3/12/1999	10/24/2008	138	29
Prince Georges Co Programs and Planning Division	PGC-DER	6/23/1994	4/7/2008	584	
SRBC-Watershed Assessment and Protection-TMDL	SRBC	7/20/1994	12/5/2006	2059	1
SRBC-Watershed Assessment Program	SRBC	4/14/1986	8/15/2008	4321	395
EPA-EMAP Wadeable Streams Assessment	USEPA	4/27/1993	9/13/1996	334	246
EPA-Mid-Atlantic Highlands Assessment	USEPA	5/21/1997	9/14/1998	157	80
EPA-Wadeable Stream Assessment Program	USEPA	7/20/2004	11/10/2004	63	31
National Forest Service Stream Assessment	USFS	5/18/2000	5/8/2003	7	
USGS-National Water Quality Assessment Program	USGS	6/2/1993	8/27/2008	243	
Virginia DEQ Benthic Monitoring Program	VADEQ	5/20/1992	12/15/2008	3249	2033
INteractive STream Assessment Resource	VCU	1/1/1999	11/6/2008	1030	
West Virginia Div. of Water and Waste Management	WVDEP	8/19/1996	1/16/2008	2421	686

Table 1. Sampling events with benthic data included in the Chessie B-IBI database as of March 7, 2011. *Data used to develop the non-Coastal Plain Chessie B-IBI index; each sampling event included water quality measurements and stream habitat scores as well as macroinvertebrate taxa counts.

and incorporated into the Chessie B-IBI database before June 2010. After the index was developed, all available benthic sampling data between 2000 and 2008 were scored with the Chessie B-IBI.

Bioregions

Classification is a technique commonly used to create groups with similar characteristics that are distinct in meaningful ways from one another. The technique can reveal key relationships between biota and their environment by reducing "noise" in the data caused by confounding natural factors. The Chesapeake watershed stream data were classified by regional hydrogeomorphologic features into "bioregions" to reduce the confounding influence of climate, geography, soil, and other natural factors. Indexes specific to each bioregion could then be developed using those sites least degraded by anthropogenic factors (Reference sites) but still reflecting the influences of the natural factors.

It is well known that regional differences in topography, soils, geography, and vegetation will affect the structure and function of benthic macroinvertebrates stream communities (Feminella 2000, Hawkins et al. 2000, Kennan 1999). The US EPA has established ecoregions as a classification framework to group similar ecosystems into classes with components of relatively homogeneous quality and quantity (e.g., Olmernik 1995, Woods et al. 1999). To create a hydrogeomorphologic stream classification for the Potomac River basin, Astin (2006) aggregated US EPA Level IV ecoregions (see US EPA 2010a) into four "bioregions": Southeastern Coastal Plains, Piedmont, Valleys, and Ridges (**Table 2**). Astin found little or no ecoregion-related differences within the stream communities of each of her bioregions and the aggregation of Level IV ecoregions provided sufficient numbers of samples (n sizes) of biological data for analysis purposes. Note: several small, under-represented sections of the Level III Central Appalachians and Western Allegheny Plateau ecoregions that overlap the western boundary of the Potomac River basin were included in Astin's Ridges bioregion even though they are geologically distinct areas.

The four Potomac bioregions identified in Astin (2006) were retained for the Chessie B-IBI because they extend across most of the Chesapeake Bay basin. Pennsylvania's Level IV Reading Prong ecoregion was incorporated into the Piedmont bioregion. The Level III Mid-Atlantic Coastal Plain ecoregion which lies mostly outside of the Potomac coastal plain was added to the classification framework. Two additional bioregions were created to complete the framework in the northern Chesapeake Bay basin: the North Central Appalachians (NCA) and Northern Appalachian Plateau and Uplands (NAPU). These bioregions are effectively the Level III Northern Appalachian Plateau and Uplands (60) and North Central Appalachians (62) ecoregions, however, NAPU includes small pieces of the neighboring ecoregion 83 (Eastern Great Lakes Lowlands) in New York (**Table 2**). A map of the bioregions is presented in **Figure 1**. The hydrogeomorphic characteristics of each ecoregion are summarized in **Table 3**.

Two statistical tests were performed to assess the validity of the chosen classification framework: a Kruskal-Wallis one-way analysis of variance and a Category and Regression Tree Analysis. To control for the possible confounding influences of seasonality and anthropogenic disturbance, the tests were performed on data collected in Spring only (March – May) *and* from the least-degraded sites (see "Reference and Degraded site selection"). The Ridges, Valleys, Piedmont, NAPU and NCA bioregions were included in both analyses. The remaining two bioregions—Southeastern Coastal Plain and Mid-Atlantic Coastal Plain—were not included because an existing index of biotic integrity was applied to these data (see "Coastal Plain").

Table 2. Correspondence between the bioregions used to develop the Chessie B-IBI and the US EPA Level III and IVecoregion classifications. * Level IV ecoregions with limestone geology.

Bioregion	EPA	Level III Ecoregion	EPA I	Level IV Ecoregion
Northern	60	Northern Appalachian	60a	Glaciated Low Plateau
Appalachian		Plateau and Uplands	60b	Northeastern Uplands
Plateau and			60d	Finger Lakes Uplands and Gorges
opianas			60e	Glaciated Allegheny Hills
	83	Eastern Great Lakes and Hudson Lowlands	83f	Mohawk Valley
North Central	62	North Central	62a	Pocono High Plateau
Appalachians		Appalachians	62b	Low Poconos
			62c	Glaciated Allegheney High Plateau
			62d	Unglaciated Allegheney High Plateau
Middle Atlantic	63	Middle Atlantic Coastal	63b	Chesapeake-Pamlico Lowlands and Tidal Marshes
Coastal Plain		Plain	63c	Swamps and Peatlands
			63d	Virginian Barrier Islands and Coastal Marshes
			63e	Mid-Atlantic Flatwoods
			63f	Delmarva Uplands
Southeastern	65	Southeastern Plains	65n	Chesapeake Rolling Coastal Plain
Plains			65m	Rolling Coastal Plain
Ridges	66	Blue Ridge	66a	Northern Igneous Ridges
			66b	Northern Sedimentary and Metasedimentary Ridges
	67	Ridge and Valley	67d	Northern Dissected Ridges and Knobs
			67c	Northern Sandstone Ridges
			67i	Southern Dissected Ridges and Knobs
			67h	Southern Sandstone Ridges
	69	Central Appalachians	69a	Forested Hills and Mountains
			69b	Uplands and Valleys of Mixed Land Use
	70	Western Allegheny Plateau	70c	Pittsburgh Low Plateau
Piedmont	45	Piedmont	45c	Carolina Slate Belt
			45e	Northern Inner Piedmont
			45f	Northern Outer Piedmont
			45g	Triassic Basins
	58	Northeastern Highlands	58h	Reading Prong
	64	Northern Piedmont	64d	Piedmont Limestone/Dolomite Lowlands*
			64c	Piedmont Uplands
			64b	Trap Rock and Conglomerate Uplands
			64a	Triassic Lowlands
Valleys	67	Ridge and Valley	67e	Anthracite Subregion
			67a	Northern Limestone/Dolomite Valleys*
			67b	Northern Shale Valleys
			67f	Southern Limestone/Dolomite Valleys & Low Rolling Hills*
	1		67g	Southern Shale Valleys



Figure 1. Bioregion classification framework for the B-IBI.

Kruskal-Wallis one-way analysis of variance

The Kruskal Wallis (KW) test is a nonparametric method for testing equality of population medians among groups. The analysis tested for differences in the medians of the candidate biological metric values in each bioregion. If there are significant differences in metric values, the classification framework can be considered appropriate. Results from the KW test show most biological metrics (see "Candidate biological metric identification" for a complete list of the metrics) have significantly different values in the Piedmont, Ridges, and Valleys bioregions (p<0.05). Only 4 of the 42 metrics showed no significant difference between bioregions. These four metrics are % Limestone, % Net Caddisfly, % Swimmer, and % Trichoptera. A similar KW analysis was conducted for the Northern Appalachian Plateau and Uplands and the North Central Appalachians and showed similar results. These results support the use of bioregion classification for indicator development.

Category and Regression Tree Analysis

The Category and Regression Tree (CART) analysis was performed to further investigate bioregions as a classification framework and also to determine if variance in the data could be better explained by individual variables such as latitude or karst geology. CART is a non-parametric decision tree technique that splits the data into increasing similar groups. These groups are formed by a collection of rules that classify data based on which independent variables differentiate observations in the dependent variable (in this case, the biological data) the best. The algorithm begins by splitting the data into two groups according to which independent parameter reduces variability the most (first node) and then continues to split the data at subsequent nodes with lesser classification strength. The tree stops splitting when the program determined no further gain in classification strength can be made by categorizing the data. The resulting CART can be used to determine which variables are responsible for the majority of the variance in the data and therefore which are better to use for a classification scheme.

Six independent variables were included in the analysis: latitude, elevation class, Strahler stream order (1-4), Level IV ecoregions, limestone geology (presence/absence), and hydrogeomorphic regions (HGMR). A hydrogeomorphic GIS layer (USGS 2008) was used to determine which sampling sites were located directly on limestone geology. A CART analysis was conducted on each of 15 dependent variables which were all raw count metrics randomly standardized to a sample count of 100. Again,

Bioregion			Climate/Avg.			Stream
(Level III ecoregion)	Physiography	Crestal Elevations	Growing Season	Geology	Soils	Gradient
Northern Appalachian	rolling hills, open	1,300 to 2,000 ft	continental, cool	glaciated plateau, till partly	stony	low and
Plateau and Uplands	valleys, and low		summers and cold	covers nonresistant, nearly		moderate
(primarily 60)	mountains		winters	horizontal to slightly deformed		
. , .			100-160 days	shale, siltstone, sandstone		
North Central	plateau surfaces, high	1,800-2,600 ft	continental, cool	partially glaciated elevated	derived from sandstone,	low, moderate
Appalachians	hills, and low		summers and cold	plateau, more rugged and	shale, and till	and steep
(62)	mountains		winters	elevated than (60), horizontally		
			100-165 days	bedded sandstone, shale,		
				siltstone, conglomerate, and coal		
Ridges	mountainous, well	500-5,700 ft	climate varies	resistant and deformed	stony, sometimes sandy,	steep, cool,
(66, parts of 67 and	dissected ridges that		significantly	metavolcanic, igneous,	some shale barrens	trellised with
others)	are folded and faulted		120 -180 days	sedimentary, and		many riffles,
				metasedimentary rock,		actively down-
				anthracite coal		cutting
Valleys	rolling valleys and low	variable	climate varies	shale or limestone; sinkholes,	diverse, base-rich soils	low, warm; low
(parts of 67)	hills, some valleys very		significantly	underground streams, and other		stream density
	broad		130-180 days	karst features common		in limestone
						valleys
Piedmont	(64) low rounded hills,	(64) 325-1,300 ft	(64) humid	(64) unglaciated metamorphic,	(64) deep, well-developed	(64, 45) low and
(64, 45)	irregular plains, and	(45) 200-1,000 ft	continental, cold	igneous, and sedimentary rocks,	fertile soils	moderate with
	open valleys	with monadnocks	winters, hot	has some limestone-dolomite	(45) clay-rich, acidc, and	falls, islands,
	(45) low rounded hills	reaching 2,000 ft	summers	lowlands	relatively low in base	and rapids in
	and ridges, irregular		170-210 days	(45) unglaciated, deeply	saturation; cultivation has	Fall-line zone
	plains, shallow valleys,		(45) humid, warm	weathered, deformed	caused significant soil loss	
	and scattered		temperate climate	metamorphic rocks intruded by		
	monadnocks		170-210 days	igneous material; some		
				sedimentary rocks		
Middle Atlantic	low, nearly flat plain,	0-100 ft	temperate, ocean-	unglaciated, unconsolidated	poorly drained alluvial sand	very low, often
Coastal Plain	with many swampy or		modifed	sediments	and silt, estuarine sand and	poorly incised
(63)	marshy areas		175-255 days		silt, saline marsh deposits,	and lack a
					and marine sand, silt, and	defined channel
					clay	
Southeastern	low, irregular or rolling	0-300 ft	temperate, ocean-	unglaciated, unconsolidated	sand, silt, and clay	low, sandy
Plains	plains		modifed	sediments		bottom
(65)			160-225 days			

Table 3. Geomorphological descriptions of bioregions. Details based on Level III and IV ecoregion descriptions in Wood et al. 1999.

only data from least-disturbed sites (see "Reference and Degraded site selection") and Spring were used in order to minimize the effects of anthropogenic impacts and the potential effect of seasonality on the results. The selected biological metrics represent a mixture of taxonomic and functional groups that should indicate if communities vary significantly according to one or more of the independent variables.

The CART results are summarized in **Table 4**, which shows the first and second splits in the classification and regression tree. "X" denotes the first split in the tree (i.e. the strongest discriminator) and * denotes the two categories in the second splits of the regression tree (i.e. the second strongest discriminators). All of the biological metrics have Level IV ecoregion in at least one of the first two splits in the regression tree. Eleven of the 15 metrics classified most strongly on Level IV ecoregions. Filterer and Plecoptera counts had elevation as the strongest discriminator and Scraper and Swimmer counts had latitude as the strongest discriminator, but Level IV ecoregion was the second strongest discriminator for all four. These results confirm that some common stream biological metrics respond strongly to ecoregion. They support the use of a classification framework such as the US EPA ecoregions which is based on biotic and abiotic features that collectively affect or reflect differences in ecosystem quality and integrity (Omernik 1995). These features include geology, physiography, climate, soils, and hydrology (**Table 3**). The bioregion classification framework incorporates all of these elements and is thus an appropriate classification tool for Chesapeake Bay basin non-tidal streams and small rivers.

Coastal Plain

The US EPA Level III Mid-Atlantic Coastal Plain and Southeastern Coastal Plain ecoregions include all coastal plain areas within the Chesapeake Bay Watershed (US EPA 2010a). The Mid-Atlantic Coastal Plain ecoregion encompasses the eastern shores of Maryland and Virginia and a portion of Delaware.

Table 4. Category and Regression Tree Analysis results. "X" denotes the highest discriminating category while * denotes the second most discriminating categories for classifying 15 biological metrics. The data were limited to the following in order to remove potential confounding factors: 1) Only Reference sites, 2) only sample results standardized to n=100 (all samples with n<100 count removed), and 3) only Spring samples.

Metric	US EPA			Hydrogeo-		
(count per 100-	Level IV			morphic	Strahler	Karst
count sample)	Ecoregion	Elevation	Latitude	regions	order	geology
Habit Groups						
Climber	Х	*		*		
Swimmer	*	*	х			
Clinger	Х		*	*		
FeedingGroups						
Shredder	Х	*			*	
Collector	Х	*				
Gatherer	Х	*	*			
Filterer	*	Х	*			
Scraper	*		х	*		
Composition						
Chironomid	Х		*			*
Trichoptera	Х		*			
Plecoptera	*	Х	*			
Ephemeroptera	Х	*		*		
EPT	X,*	*				
EPT (no Hydropsychidae)	X,*	*				
Diptera	X,*					

The Southeastern Coastal Plain ecoregion includes the Maryland and Virginia western shores (**Figure 1**). These two ecoregions are unique from others in the Chesapeake Bay watershed in that they have flat terrain, unconsolidated sediments, nutrient-rich soils and low-gradient "swamp streams" (e.g., Wood et al. 1999, Maxted et al. 2000).

Due to these differences, biological monitoring programs typically use benthic macroinvertebrate monitoring and assessment protocols specific to the coastal plain area. Current assessment protocols used by Virginia and Delaware agencies are based on methods developed by the interstate Mid-Atlantic Coastal Streams Workgroup (US EPA 1997, Maxted et al. 2000). The index was intended to rate stream benthic macroinvertebrate condition across the mid-Atlantic seaboard of the United States. It was created with a multi-jurisdictional data set from several coastal plain ecoregions and rigorously tested by scientists across six states where it showed high classification efficiencies for degraded sites (86%). Maryland Department of Natural Resources (MD DNR) has developed two other coastal plain indexes. The older version (Stribling et al. 1998) was applied to all Chesapeake coastal plain sites and used in the 2008 CBP report. The newer version (Southerland et al. 2005) was tested but not used in the 2009 CBP report. It proved accurate in scoring benthic macroinvertebrate communities in Maryland but tended to over-estimate benthic macroinvertebrate condition when applied to Virginia and Delaware sites.

For the 2010 analysis, a version of the Coastal Plain Macroinvertebrate Index (CPMI) developed by the Mid-Atlantic Coastal Streams Workgroup and adapted to family-level taxonomic identifications (M. Passmore US EPA, personal communication; A. Budd VADEQ, personal communication) was tested on the Chesapeake region data. Coastal Plain samples collected by 10 agencies and organizations are presently included in the Chessie B-IBI database. Samples were collected in Spring (Mar-May), Summer (June-September) and/or Fall/Winter (October-February) depending on the program. The original CPMI is based on 100-count samples of taxa identified to the genus level. A CPMI protocol for scoring taxa counts at the family level was established to accommodate biological monitoring programs that perform family level identifications such as the Virginia Department of Environmental Quality. The CPMI family level thresholds for scoring the component biological metrics in Middle Atlantic Coastal Plain Ecoregion (63N) and Southeastern Coastal Plain Ecoregion (65) were applicable to the entire Chesapeake coastal plain data set after high count samples had been standardized to 100 counts.

The CPMI modified to family-level metrics was selected for scoring Chesapeake coastal plain regions for the 2010 CBP report. The existence of an interstate index applicable to all Chesapeake coastal plain watersheds removed the need to independently develop a coastal plain version of the Chessie B-IBI. Additional analysis could potentially refine and improve the index; however time did not allow for this in the current effort. Before applying the CPMI to the available data, samples with high total counts were standardized to 100 counts per sample (see "Standardization of count metrics" for details).

The following five metrics comprise the modified CPMI: Taxa Richness, Ephemeroptera-Plecoptera-Trichoptera (EPT) Taxa Richness, % Ephemeroptera, Hilsenhoff Biotic Index (FBI), and % Clinger. Each metric is scored on a scale of 0-6, with separate scoring thresholds for each of the two coastal plain Level IV ecoregions. Metrics are scored 0, 2, 4, or 6 depending on their value relative to percentiles in reference communities. The Maxted et al. (2000) CPMI score is calculated by summing the five component metric scores, resulting in an index scale of 0-30. The same approach is used for the family level CPMI but the metric scoring thresholds are adjusted to accommodate family level data. Metric scoring thresholds for the ecoregions found in the Chesapeake Bay watershed are shown in **Tables 5a** and **5b**. For CBP purposes, the Mid-Atlantic Coastal Plain and Southeastern Coastal Plain ecoregions were adopted as Chessie B-IBI bioregions. The CPMI scores were expressed as a percent of the maximum score of 30 to make them directly comparable to non-Coastal Plain index values which are on a 0% - 100% scale. See the section "Narrative condition ratings" for details on how index scores for the Coastal Plain bioregion were qualitatively rated.

	Scoring Thresholds				
Metric	6	4	2	0	
Taxa Richness	>17	<=17 And >= 11	<11 And >= 6	<6	
EPT Taxa Richness	>3	3	2	<2	
%Ephemeroptera	>=37	<37 And >= 25	<25 And >= 12	<12	
FBI	<= 5.1	> 5.1 And <= 6.1	> 6.1 And <= 7	>7	
%Clinger	> 21	<= 21 And >= 14	< 14 And >= 7	<7	

 Table 5a. CPMI metric thresholds for the Middle Atlantic Coastal Plain Bioregion (Ecoregion 63N).

 Data are standardized to 100 counts per sample.

Table 5b. CPMI metric thresholds for the Southeastern Coastal Plain Bioregion (Ecoregion 65).Data are standardized to 100 counts per sample.

		Scoring	Thresholds	
Metric	6	4	2	0
Taxa Richness	> 18	<= 18 And >= 12	< 12 And >= 6	<6
EPT Taxa Richness	>= 7	< 7 And >= 5	< 5 And >= 3	<3
%Ephemeroptera	>= 25	< 25 And >= 16	< 16 And >= 8	<8
FBI	<= 5.7	> 5.7 And <= 6.5	> 6.5 And <= 7.2	>7.2
%Clinger	> 27	<= 27 And >= 18	< 18 And >= 9	<9

Non-Coastal Plain

A reference-based approach was used to develop bioregion-specific indexes for the non-coastal plain Chesapeake Bay watershed. The approach consisted of the following steps, described in detailed sections below:

- develop a list of candidate biological metrics
- identify Reference and Degraded sites using physical and chemical stream characteristics
- evaluate the influence of season, stream order, and limestone geology on the candidate biological metrics
- standardize count-based metrics to account for differences in sample enumeration
- performed a series of tests to evaluate metric suitability
- select best performing (most discriminating) biological metrics
- test and select a metric scoring approach
- validate the final scored results
- determine a narrative rating for the Chessie B-IBI scores

Candidate biological metric identification

A list of biological metrics was compiled from the following sources and considered for the non-coastal plain indexes: the Potomac River Basin-wide B-IBI (Astin 2006, 2007), NTWG workgroup member recommendations, and a composite list of all biological metrics used by the major sampling organizations in the watershed. A total of 127 metrics were suggested. This list was reduced to 42 (**Table 6**) using the following criteria to exclude metrics:

- if the metric was genus-based and not applicable to family level assessments
- if the metric was strongly affected by sample size (includes all non-normalized abundance metrics such as EPT abundance and count-based metrics such as number of EPT taxa).
- if there was a large amount of redundancy between similar metrics (the known best performer was retained and the other(s) discarded, e.g., variations of EPT taxa metrics).

Finally, if a metric was affected by sample size but is known to be responsive to anthropogenic disturbance and is commonly used in indexes elsewhere in the watershed, it was standardized to a

Table 6. Forty-two metrics considered for index development in non-coastal plain bioregions. Type: T, tolerance; R, richness; H, habit; FG, feeding guild; C, composition. Metrics that have been standardized to a 100-count sample have the suffix "100."

Metric	Description	Туре
ASPT_MOD	Average tolerance score per taxon modified to family level	Т
BECK_100	Becks Index, sample standardized to a total count of 100	Т
EPHEMEROPTERA_TAXA_CNT_100	Ephemeroptera family count, sample standardized to a total count of 100	R
EPT_TAXA_COUNT_NO_TOL_100	EPT families excluding tolerants, sample standardized to total count of 100	R <i>,</i> T
FBI	Family level Hilsenhoff Biotic Index	Т
GOLD	1 minus proportional # of gastropods, oligochaetes, Diptera individuals	С
LOG10_SEL_EPTD	Log 10 of selected Ephemeroptera, Plecoptera, Trichoptera and Diptera taxa	Т
MARGALEFS	Margalef's Index	R
PCT_AMPHIPOD	Percent of individuals that are amphipods	С
PCT_CHIRONOMID	Percent of individuals that are chironomids	С
PCT_CLIMB	Percent of individuals that are climbers	Н
PCT_CLING	Percent of individuals that are clingers	Н
PCT CLINGER TAXA	Percent of family level taxa that are clingers	Н
PCT_COLLECT	Percent of individuals that are collectors	FG
PCT_CORBICULA	Percent of individuals that are Corbicula clams	С
PCT_DIPTERA	Percent of individuals that are Diptera	С
PCT DOM1	Percent of individuals that belong to the most common family level taxa	т
PCT DOM2	Percent of individuals that belong to the 2 most common family level taxa	Т
PCT DOM3	Percent of individuals that belong to the 3 most common family level taxa	Т
PCT EPHEMEROPTERA	Percent of individuals that are Ephemeroptera	С
PCT EPT	Percent of individuals that are Ephemeroptera, Plecoptera, Trichoptera (EPT)	С
PCT EPT TAXA RICH	Percent of family level taxa that are EPT	R
PCT_FILTERERS	Percent of individuals that are filterers	FG
PCT GATHER	Percent of individuals that are gatherers	FG
PCT LIMESTONE	Percent of individuals that are isopods, amphipods, and Ephemeralla	С
PCT_NET_CADDISFLY	Percent net-spinning caddisflies	С
PCT NON INSECT	Percent taxa that are not insects	С
PCT OLIGOCHAETA	Percent Oligochaeta	С
PCT_PLECOPTERA	Percent Plecoptera taxa	С
PCT SCRAPER	Percent scraper	FG
PCT SENSITIVE	Percent sensitive taxa (family level tolerance value \leq 3)	Т
PCT SHREDDER	Percent shredders	FG
PCT_SWIMMER	Percent swimmers	н
PCT TOLERANT	Percent tolerant taxa (family level tolerance value \geq 7)	т
PCT TRICHOPTERA	Percent Trichoptera taxa	С
PCT TRICHOPTERA NO TOL	Percent Trichoptera taxa (excluding Hydropsychidae)	т
PCT URBAN INTOL	Percent of urban intolerants	т
SENSITIVE TAXA COUNT 100	Sensitive family-level taxa count, sample standardized to total count of 100	т
SIMPSON_DIVERSITY	Simpson Diversity index	R
SW	Shannon Wiener Index	R
TAXA RICH 100	Taxa richness, sample standardized to total count of 100	R
TOLERANT_TAXA_COUNT_100	Tolerant family-level taxa count, sample standardized to total count of 100	Т

sample count of 100 and retained (see "Standardization of count-based metrics" below for details). Given this final criteria, the following 6 count-based metrics were standardized and retained for analysis: Taxa Richness, Beck's Index, EPT Taxa count no tolerant, Ephemperoptera taxa count, Sensitive taxa count, and Tolerant taxa count. The standardized metrics are indicated with the suffix "100" in **Table 6**. Overall, 13 composition, 5 feeding guild, 4 habit, 7 richness, and 13 tolerance metrics were retained for analysis. All metrics that were used in the 2008 and 2009 version of the Chessie B-IBI are included in this list. **Appendix A** lists the feeding guild, habit, and tolerance attributes assigned to each family-level taxon and used in the metric calculations. *Note: this list has been updated since the 2008/2009 Chessie B-IBI calculations were done*.

Reference and Degraded site selection

Physical and chemical parameters of Chesapeake Bay basin streams and wadeable rivers were used to create a gradient of stream conditions ranging from "least disturbed" to "most disturbed" by anthropogenic impacts. Habitats classified as least-disturbed are called "Reference"; those classified as most disturbed are called "Degraded." Ideally, metric values of the biological communities in the opposing stream habitat types are widely separated and thus useful for establishing biological scales with which to measure community responses to anthropogenic disturbance. Reference and Degraded sampling events¹ in the Piedmont, Ridges, and Valleys were identified using an approach based on stream habitat and water quality parameters; those in the NAPU and NCA bioregions were identified using a combination of watershed land use, water quality, and stream habitat parameters.

Piedmont, Ridges, and Valleys

Reference and Degraded sites were determined using six habitat and two water quality metrics originally identified in Astin (2006). The six habitat metrics are either directly measured by all the Chesapeake Bay basin states overlapping these bioregions or made up of different components that are measured by the states and can be equated. Most basin monitoring programs have adopted habitat evaluation procedures from the US EPA Rapid Bioassessment Protocols for scoring stream habitat (Plafkin et al. 1989, Barbour et al. 1999), hence their results tend to be comparable. The metrics are:

Habitat

- Bank Stability (BANK_STAB): stream bank erosion or potential for erosion
- Channel Alteration (CHAN_ALT): changes to the natural stream channel (e.g., straightening, dredging, artificial armor, bar development)
- Riparian zone (RIP_ZONE): the extent of naturally vegetated stream edge and riparian zone (if riparian zone is given as measurement rather than score, it is converted to a score from 0 20)
- Embeddedness (EMBED, previously SUB_QUAL, or Substrate Quality): the extent to which rocks (gravel, cobble, and boulders) and snags are covered or sunken into the silt, sand, or mud of the stream bottom (if embeddedness is a % rather than a score, it is converted to a 0 20 score)
- Habitat Heterogeneity (HAB_HETERO): habitat diversity, e.g., riffle frequency, pool/glide/eddy quality
- Instream Condition (INSTR_COND): the abundance of suitable substrate for benthic macroinvertebrates and/or cover for fish

Water quality

- pH (pH)
- Specific Conductivity (COND)

¹ Changing environmental conditions can cause the classification of a repeatedly sampled site to differ over time, therefore sampling events (dates) rather than sites were classified as Reference, Degraded, or neither (Mixed).

The pH and conductivity values were obtained directly from the collecting agencies. Scores for BANK_STAB, CHAN_ALT, RIP_ZONE, and EMBED were obtained directly from the collecting agencies and converted to a 0-20 scale if necessary. Scores for HAB_HETERO and INSTR_COND were calculated and/or converted from state-specific metrics to standard metrics for the purposes of comparability with the following equations:

HAB_HETERO = RIFF or POOL or (RIFF + POOL)/2 INSTR_COND = EPI_SUB or COVER or (EPI_SUB + COVER)/2

where RIFF is riffle frequency; POOL is the overall mix of four basic pool types: large-shallow, large-deep, small-shallow, and small-deep (see for example the RBP pool variability evaluation in Barbour et al. 1999); EPI_SUB is the relative quantity and quality of natural structures in the stream suitable for epifauna, e.g., cobble, gravel, woody debris; and COVER is the relative quantity and quality of natural structures in the stream suitable for fish, e.g., root wads, snags, bank undercuts, fallen logs, large boulders.

Astin (2006) identified Reference or Degraded events in the Potomac River basin using literaturesupported thresholds for water quality parameters and fixed percentiles that delineated the most and least degraded ends of each habitat metric's distribution in a bioregion. The identical approach was

used to classify Reference and Degraded events in earlier iterations of the Chessie B-IBI (Foreman et al 2008, CBPO 2008 and 2009). A sampling event was classified as Reference if all habitat scores and water quality measurements met or exceeded the Reference criteria. If one or more of the water quality metrics or three or more of the habitat metrics met the Degraded criteria, the event was classified as Degraded. All other locations were classified as Mixed.

In the current analysis, the pH and conductivity thresholds for separating Reference and Degraded events remained unchanged in the Piedmont and Valleys bioregions (**Table 7**). The pH thresholds for Degraded conditions in the Ridges bioregion were broadened to pH<5 and pH>9.5 to compensate for a generally weaker buffering capacity in those streams (e.g., Wood et al. 1999). Instead of fixed percentiles of the metric's score distributions, specific scores from the customary 0-20 habitat metric scoring scale were used to classify each of the six habitat metrics. Three different sets of thresholds were explored and compared to the percentile-based thresholds:

Option 1: score is \leq 10 for Degraded and \geq 18 for Reference

Option 2: score is \leq 5 for Degraded and \geq 18 for Reference

Table 7. Classification criteria for Reference andDegraded events in the Piedmont, Ridges andValleys bioregions. Habitat parameters are on a 0-20 scale; pH, PSU; COND, umhos/cm.

Reference events

(All criteria must be met for an event to qualify as Reference)

Reference)			
Parameter	Piedmont	Ridges	Valleys
BANK_STAB	<u>></u> 16	<u>></u> 16	<u>></u> 16
CHAN_ALT	<u>></u> 16	<u>></u> 16	<u>></u> 16
HAB_HETERO	<u>></u> 16	<u>></u> 16	<u>></u> 16
INSTR_COND	<u>></u> 16	<u>></u> 16	<u>></u> 16
RIP_ZONE	<u>></u> 16	<u>></u> 16	<u>></u> 16
EMBED	<u>></u> 16	<u>></u> 16	<u>></u> 16
рН	6-9	6-9	6-9
COND	<u><</u> 500	<u><</u> 500	<u><</u> 500

Degraded events

(Criteria for pH or COND or at least three habitat metrics must be met for an event to qualify as Degraded)

must be met for an event to quality as begraded)					
Parameter	Piedmont	Ridges	Valleys		
BANK_STAB	<u><</u> 5	<u><</u> 5	<u><</u> 5		
CHAN_ALT	<u><</u> 5	<u><</u> 5	<u><</u> 5		
HAB_HETERO	<u><</u> 5	<u><</u> 5	<u><</u> 5		
INSTR_COND	<u><</u> 5	<u><</u> 5	<u><</u> 5		
RIP_ZONE	<u><</u> 5	<u><</u> 5	<u><</u> 5		
EMBED	<u><</u> 5	<u><</u> 5	<u><</u> 5		
рН	<6 or >9	<5 or >9.5	<6 or >9		
COND	>1000	>1000	>1000		



Figure 2. Reference and Degraded events in the Piedmont, Ridges, and Valleys bioregions

Option 3: score is ≤ 5 for Degraded and ≥ 16 for Reference

Option 3 was selected because a) candidate biological metrics showed a strong degree of separation between Reference and Degraded, and b) the option classified the data into Reference, Mixed, and Degraded pools of sufficient size for analysis (n>20) (**Table 7**).

Reference and Degraded events identified in Option 3 were broadly distributed across the Chesapeake Bay basin (**Figure 2**). The Piedmont had 222 Reference and 76 Degraded events, Ridges had 154 Reference and 42 Degraded events, and Valleys had 57 Reference and 92 Degraded events.

Northern Appalachian Plateau and Uplands and North Central Appalachians

Adequate habitat data were not available in Northern Appalachian Plateau and Uplands (NAPU) and the North Central Appalachians (NCA) bioregions for identifying Reference and degraded conditions. Therefore, alternative approaches to identifying Reference and degraded sites were employed.

During the 2000 to 2008 period, NYDEP collected only one of the six habitat metrics of interest (embeddedness) in the NAPU bioregion while PADEP, SRBC and US EPA monitored all the parameters. The NAPU bioregion consists mostly of sites in New York. An alternate method based on land use was used to identify Reference and Degraded sites in conjunction with available water quality and habitat data. The watershed above each sampling location was delineated with the Multi-watershed Delineation GIS tool (http://hydrology.neng.usu.edu/mwdtool/) and %forest cover, %imperviousness, and number of NPDES permits were calculated from layers provided by the Chesapeake Bay Program. Watersheds with >78% forest cover and <1% impervious surface show minimal anthropogenic disruption in flow regime and are thus considered reference quality with respect to flow (C. Buchanan, ICPRB, personal communication); an absence of National Pollutant Discharge Elimination System (NPDES) permits suggests a low level of point source pollution. These watershed criteria were used with pH, conductivity, and embeddedness to classify Reference streams in the NAPU bioregion (Table 8). The Reference group pH boundaries were 6-9 PSU. The conductivity threshold was 500 umhos/cm. The criterion for embeddedness was relaxed to 6 in order to increase the number of sampling events in the Reference data pool. Land use features could not be used in NCA to identify Degraded sites because none of the delineated watersheds had land uses significant enough to potentially degrade stream benthic communities. A site was classified as Degraded if EMBED < 10 or pH <5.5 or pH >9 or COND > 500 umhos/cm (Table 8).

The land use information was also used in the NCA bioregion to support classification of Reference events. The NCA bioregion had the fewest number of sampling events associated with a full

complement of water quality, habitat and benthic macroinvertebrate data. Reference criteria for habitat metric scores were eased from 16 to 15 in order to increase the number of sampling events in the Reference data pool. Only pH, conductivity and embeddedness criteria were applied to identify Degraded.

After the NCA and NAPU criteria were applied to the available data, there were 28 Reference and 20 Degraded events in the NAPU and 26 Reference and 40 Degraded events in the NCA. **Table 8** lists the various criteria for identifying Reference and Degraded events in the NAPU and NCA bioregions. Figure 3 shows the distribution of Reference and Degraded events in both bioregions. The easing of the habitat criteria and the scarcity of both Reference and Degraded sites weakens the ability of the biological metrics to discriminate between good and poor quality streams. Future data collection should correct this problem and strengthen the classification efficiency of the overall index. The New York Department of Environmental Protection recently added more habitat parameters to its stream monitoring programs and this will improve Reference and Degraded site identification.

Table 9 summarizes the number of Referenceand Degraded sites in all bioregion.

Impacts of stream order on biological metrics

The relative size of the stream from which a sample is collected can potentially affect benthic metrics and ultimately a regional B-IBI score. The most desirable metrics for the Chesapeake regional B-IBI would be metrics with little stream order variation since the regional B-IBI is intended for 1st through 4th order running waters, or streams and wadeable rivers. The 42 candidate metrics were compared across Strahler stream order to test for these effects (**Appendix B**). Only sampling events from Reference conditions were used to minimize confounding effects of anthropogenic factors.

Table 8. Classification criteria for Reference andDegraded events in the North Central Appalachian (NCA)and Northern Appalachian Plateau & Upland (NAPU)bioregions.

Reference events

(All criteria must meet criteria for an event to qualify as Reference)

Parameter	NCA	NAPU
BANK_STAB	<u>></u> 15	
CHAN_ALT	<u>></u> 15	
HAB_HETERO	<u>></u> 15	
INSTR_COND	<u>></u> 15	
RIP_ZONE	<u>></u> 15	
EMBED	<u>></u> 15	<u>></u> 6
Watershed forested area	>78%	>78%
Watershed impervious surface	<1%	<1%
Watershed NPDES permits	none	none
рН	6-9	6-9
COND (umhos/cm)	<u><</u> 500	<u><</u> 500

Degraded events

(NCA: criteria for pH or COND or at least three habitat metrics must be met for an event to qualify as Degraded; NAPU: criteria for EMBED and [pH or COND] must be met for an event to qualify as Degraded)

Parameter	NCA	NAPU
BANK_STAB	<u><</u> 5	
CHAN_ALT	<u><</u> 5	
HAB_HETERO	<u><</u> 5	
INSTR_COND	<u><</u> 5	
RIP_ZONE	<u><</u> 5	
EMBED	<u><</u> 5	<u><</u> 15
рН	<5.5 or >9	<5.5 or >9
COND (umhos/cm)	>500	>500

Table 9. Number of Reference and Degradedsites in each bioregion

Bioregion	# REF	# DEG
Piedmont	222	76
Ridges	154	42
Valleys	57	92
NAPU	28	20
NCA	26	40



Figure 3. Reference and Degraded sites in the Northern Appalachian Plateau and Uplands and the North Central Appalachians.

Just one of the 42 biological metrics, %Shredder, showed a consistent stream order effect in all five bioregions, with percents highest in 1st order streams. As a whole, the stream order comparisons indicate that as the size of the Reference data pool increases, the number of metrics showing significant differences (p<0.05) across stream order also increases (Figure 4). Specifically, the NCA and NAPU bioregions have the smallest total numbers (26 and 28, respectively) and the fewest biological metrics showing significant differences across stream order whereas the Piedmont has the largest total numbers (222) and the most metrics showing significant differences.² Stream order effects-subtle or obvious-can be found in most stream metrics but may not be statistically

significant until sample number is large. Careful scoring of the metrics can minimize stream order effects.

An example is shown in **Figure 5** where a significant stream order effect on %EPT is minimized by placing a metric scoring threshold at 48% (i.e., values below 48% score low and values above 48% score higher). The metric %EPT is a strong indicator of anthropogenic impacts in the Piedmont bioregion.



Figure 4. The number of biological metrics in each bioregion that show significant differences across Strahler stream order (left panel) and across season (right panel), *versus* total number of samples in the bioregion. Total number of biological metrics tested is 42. Only Reference condition samples were included in order to minimize the effects of anthropogenic factors.

² This trend can be expected if variability due to stream order is fairly consistent between bioregions. This is because measures of variability or error around a mean or median are a decreasing function of n. For example, the standard error of a mean is the standard deviation of the observations divided by the square root of n. Similarly in non-parametric tests like the Kruskal-Wallis which compares the medians of two or more group, the proportion of each group's ranked observations greater than the overall median is divided by the group's n. Increasing n sizes ultimately result in tighter confidence intervals around each group's mean or median if variance around the mean or median is relatively stable. This appears to be the case with stream order effects on the 42 candidate biological metrics.

Impacts of season on biological metrics

The season in which a sample is collected can also potentially affect benthic metrics and ultimately a regional B-IBI. The most desirable metrics for the Chesapeake regional B-IBI would be metrics with little seasonal variation given the variety of sampling seasons represented in the basin's monitoring programs. The 42 candidate metrics were compared across season to test for these effects (Appendix B). Only sampling events from Reference conditions were used in order to minimize confounding effects of anthropogenic factors. The results did not reveal a strong relationship between bioregion n size and count of metrics with significant seasonal differences, although there is a slight increase with increasing n (Figure 4). Overall, the western-most bioregions with elevated, generally steeper terrains (NCA, Ridges) had more biological metrics with significant seasonal differences. Of the remaining bioregions, NAPU with the smallest n size had the fewest metrics





showing significant differences. Five of the 42 metrics (%Diptera, %Gatherers, %Net Caddisfly, %Trichoptera, Gold index) showed seasonal differences in 4 of the 5 bioregions; none showed consistent seasonal-effects in all bioregions. Values of %Diptera and the Gold Index were usually highest in Spring; those of %Net Caddisfly were usually highest in Summer; %Gatherers showed a preference for Spring.

The results suggest that while season does influence some biological metrics significantly, other environmental factors may have more influence on most of the metrics. Again, judicial placement of metric scoring thresholds can minimize the influence of seasonal effects. An example is shown in **Figure 6** where a significant seasonal effect on %Diptera is minimized by placing a threshold at 11.7% (i.e., values below 11.7% score higher with decreasing % and values above 11.7% score low).

Influence of Karst geology on biological metrics

Limestone or "karst" geology with its many springs and close ties to groundwater has a profound effect on stream networks in the



Figure 6. Distributions of %Diptera in All Reference (REF) and All Degraded (DEG) sites of the Piedmont bioregion and in Reference streams during Spring, Summer and Fall/Winter. Frequencies are calculated as a percent of each group's n size.

Table 10. Comparison of medians in the "Valleys No Limestone" and "ValleysLimestone" groups in Reference conditions. Wilcoxon test: **, p < 0.01; *, $0.01 \le p$ ≤ 0.05 . The median and inter-quartile range (IQR) of each group is shown.

Metric	No Limestone Median (IOR)	Limestone Median (IOR)	n
	(n = 26)	(n = 31)	P
ASPT MOD	4.1 (3.8-4.6)	4.1 (3.5-4.4)	
BECK 100	9 (7-10.8)	9 (6-12)	
EPHEMEROPTERA TAXA CNT 100	4 (3-4)	4 (3.25-4)	
EPT TAXA COUNT NO TOL 100	8 (7-9)	8 (6-9)	
FBI	4.3 (3.5-5)	4.0 (3.6-4.5)	
GOLD	0.8 (0.6-0.9)	0.8 (0.7-0.9)	
LOG10 SEL EPTD	1.3 (1.1-1.5)	1.3 (1.1-1.5)	
MARGALEFS	3.3 (2.7-3.7)	3.0 (2.5-3.4)	
PCT_AMPHIPOD	0 (0-0)	0 (0-0)	
PCT_CHRONOMIDAE	8.9 (4.5-23.4)	5.3 (1.9-11.7)	
PCT CLIMB	3.4 (1.7-8.2)	2.4 (0.7-7.0)	
PCT CLING	61.7 (45.3-82.5)	76.4 (60.6-87.2)	**
PCT CLINGER TAXA	63.3 (53.5-69.1)	70 (63.6-75)	*
PCT_COLLECT	64.1 (54.8-74.1)	68.5 (56.5-78.8)	
PCT_CORBICULA	0 (0-0)	0 (0-0)	
PCT DIPTERA	13 (8.6-33)	9.2 (4.6-22.9)	
PCT_DOM1	32.2 (26.6-41.8)	29.1 (24.8-36.7)	
PCT DOM2	52.5 (47.3-62.9)	48.4 (41.4-56.2)	
PCT DOM3	68.2 (57.8-71.7)	63.9 (56.2-70)	
PCT_EPHEMEROPTERA	21.6 (16.3-31.8)	38.5 (24.9-54.9)	**
PCT EPT	58.1 (34.8-68.7)	68.3 (54.9-75.6)	
PCT EPT TAXA RICH	53.6 (50-59.7)	57.1 (50.8-64.2)	
PCT_FILTERERS	21 (8.9-30.8)	24 (11.8-32.7)	
PCT GATHER	43 (27.3-51.6)	36.2 (25.2-56.3)	
PCT LIMESTONE	3.1 (1.5-8.9)	5.7 (3.1-21.7)	=.055
PCT NET CADDISFLY	10.1 (4.3-19.6)	17.2 (3.7-24)	
PCT NON INSECT	2.4 (1-11.9)	2.6 (0-8.8)	
PCT OLIGOCHAETA	0.4 (0-1)	0 (0-1)	
PCT PLECOPTERA	5.4 (1.9-25.5)	3.4 (0.3-6.2)	*
PCT SCRAPER	10.6 (4.2-22.1)	17.5 (11.6-28.1)	
PCT SENSITIVE	32.1 (22.2-57.4)	46.6 (36.7-54.7)	
PCT SHREDDER	5.1 (0.9-13.1)	3.2 (0.7-8.5)	
PCT_SWIMMER	10.1 (2.6-16.6)	10.9 (6-19.4)	
PCT_TOLERANT	12.8 (8.1-25.7)	5.8 (3.5-14.7)	*
PCT_TRICHOPTERA	12 (6.4-20.9)	18.2 (5.5-28.5)	
PCT_TRICHOPTERA_NO_TOL	3.7 (1.3-9.3)	4.7 (1.3-12.5)	
PCT_URBAN_INTOL	82.8 (75.5-87.3)	85.7 (81.1-91.2)	
SENSITIVE_TAXA_COUNT 100	6.5 (5-8)	6 (5-9)	
SIMPSON_DIVERSITY	0.81 (0.77-0.84)	0.83 (0.80-0.86)	
SW	2.1 (1.9-2.2)	2.1 (1.9-2.3)	
TAXA_RICH 100	15 (13-17)	15 (12.3-16)	
-	• • •	. ,	

Valleys bioregion. The bioregion CART analysis (Table 3) initially suggested there is little effect of geology (carbonate and mite) on the 42 candidate gical metrics. It became ent later in the analysis that geology could potentially t a few of the biological rics in the Valleys bioregion le 10). In particular, the es of %Clinger, %Clinger , and %Ephemeroptera and ibly %Limestone individuals higher in Valleys regions laying karst geology and es of %Plecoptera and lerants were higher in eys regions overlaying shale other bedrock types .05). e were sufficient sample bers in the Valleys bioregion vide it into a Limestone eys-L) group and a No stone (Valleys-NL) group and the influence of karst ogy through the entire B-IBI elopment process. The USGS

"Preliminary Map of Potentially Karstic Carbonate Rocks in the Central and Southern Appalachian States" (USGS 2008) was used to assign each sampling event to one of the two groups based on the presence or absence limestone bedrock at the site. The same habitat and water quality criteria used in the Valleys bioregion to identify Reference and Degraded events were applied to the two groups.

Parallel steps were taken with each group and with the groups combined to select discriminating metrics, score them, and combine them into an index. The results are presented in the relevant report sections below.

Standardization of count metrics

Metrics such as Taxa Richness and the Becks Index that are computed directly from a sample's raw count increase in value up to an asymptote as the number of individuals counted per sample increases. Thus, count metrics from monitoring programs with different sample counting protocols should not be directly compared if the asymptote has not been reached. In the Chessie B-IBI database, sample counts range from 1 to 1480. Counts ~100 and ~200 are most frequent, reflecting the commonly used 100- and 200-count protocols, respectively (**Figure 7**). The effect of sample count on family-level taxa richness in Chesapeake Bay basin non-Coastal Plain bioregions can be seen in **Figure 8**. Taxa richness increases curvi-linearly as sample count increases, and it is evident that asymptotes have not been achieved in most of the different bioregion-site types. As with species- and genus-level taxa, the rarer or uncommon families are more likely to be found with a 200-count protocol compared to a 100-count protocol. **Figure 8** also suggests that Degraded sites for the most part have significantly fewer family-level taxa per sample than Reference sites for comparable sample counts. **Figure 9** confirms this finding and shows Degraded sites to have between 1 and 6 fewer family-level taxa than Reference sites when sample count is randomly reduced to a sample total count of 100.

To explore sample count effects on the candidate biological metrics, samples with counts greater than 180 organisms were randomly thinned to 100 organisms using a Utah State University computer program developed for this purpose (<u>http://cnr.usu.edu/wmc/htm/predictive-</u><u>models/usingandbuildingmodels</u>). The candidate metrics were calculated from these "standardized" samples and compared to the same metrics calculated from the actual samples. Comparisons were restricted to Reference sites to avoid anthropogenic effects confounding the results. Data from just the Ridges bioregion were used in this exercise because only this bioregion had a large number of Reference samples with total counts greater than 180 (n=68). **Figure 10** shows that as one would expect, "proportional" metrics, or those with raw counts of individuals expressed as a percent of the total



Figure 7. Frequency distribution of total counts in samples from the Piedmont (n=1,829), Ridges (n=1,016), Valley (n=1,520), North Central Appalachians (n=283), and Northern Appalachian Plateau & Upland (n=772) bioregions.



Figure 8. Curvi-linear (power) relationships between sample total count and the number of family-level taxa per sample for Reference (REF) and Degraded (DEG) site types (r^2 indicated). Solid line, significant relationship (p<0.05); dashed line, non-significant relationship.

sample count (n_i/N), are not affected by sample count for the most part. A few proportional metrics do appear affected by sample count. Specifically, metrics for rare families or functional groups with very low counts (e.g., %Climber) are forced into an angular pattern when standardized to 100 counts. %EPT Taxa Richness, or the count of EPT family taxa divided by the count of all family-level taxa, can involve relatively low numbers and the results of standardized samples are slightly offset from those calculated from the entire sample. The Shannon Wiener diversity index derived from the proportions of each



Figure 9. Family-level taxa richness in samples standardized to a total sample count of 100. Site types are Degraded = DEG, Mixed = MIX, and Reference = REF (see text for details). Samples are grouped by bioregion (Piedmont, Ridges, Valleys, North Central Appalachians, and Northern Appalachian Plateau & Uplands). Box and whisker plots show 5th, 25th, 75th and 95th; median values are indicated.

family-level taxa³ also shows a slight downward shift in standardized samples. The offset, or bias, that is created in certain richness or diversity metrics when samples are standardized to 100 individuals is likely due to an increased probability of finding rarer families in samples with higher counts.

In the interest of representing the benthic community to each sample's fullest potential, the nonstandardized version of most of the metrics were used as candidates for the purpose of developing a Chessie B-IBI. The exceptions were family-level Taxa Richness, Sensitive Taxa Count, Tolerant Taxa Count, Becks Index, Ephemeroptera Taxa Count, and EPT Taxa Count (No Hydropsychidae), all of which were standardized prior to analysis and metric selection.

The statistical method "rarefaction" was used to investigate how family-level taxa richness changes as the *total* number of counted individuals increases in the Ridges, Valleys and Piedmont bioregions (see details in **Appendix C**). The resulting "taxa-accumulation" curves in each bioregion are not necessarily the same. While taxa richness in *individual samples* tends to be higher at Reference sites, the cumulative number of family-level taxa found at a given sampling intensity (total number counted) in the Piedmont and Valleys appears to be higher in Degraded samples rather than in Reference samples. In the Ridges, the Reference and Degraded taxa-accumulation curves appear to overlap closely. Projections of the curves to their asymptotes indicate the asymptotes are reached somewhere between 100 and 120 family-level taxa, suggesting that at least 100-120 family-level taxa per site type (Reference, Degraded) can be expected in each of the three bioregions with extensive sampling.

Biological metric testing and selection

Biological metrics to comprise the Chessie B-IBI were selected from the pool of 42 candidates in an iterative process. With few exceptions, the selected metrics showed significant statistical differences between Reference and Degraded sites *and* correctly identified both Reference and Degraded sites with

³ Shannon Wiener Index: H' = $\sum_{i=1}^{S} (p_i \ln p_i)$, where S is total number of taxa in a sample and p_i is the proportion of the *i*th species.



Figure 10. Cumulative frequency distributions of metrics calculated from standardized (100 count) and non-standardized Reference samples in the Ridges bioregion (n=68). Dark blue line, non-standardized samples where total n > 180; light blue line, standardized samples where total n = 100.

a high degree of certainty. The selected metrics also represented a range of metric types (composition, richness, habit, feeding guild, and tolerance). The following sections describe the process used to evaluate each biological metric's eligibility for the Chessie B-IBI.

Statistical test for significance

A Kruskal-Wallis one-way analysis of variance (KW) was performed on each biological metric in each bioregion to test for significant differences between the Reference and Degraded data pools (**Table 11**). Again, n sizes of the Reference and Degraded data pools affected the statistical results. The Piedmont, Ridges, and Valleys (combined) with the largest data pools had the most number of metrics exhibiting statistically significant differences; the NCA and NAPU bioregions with the smallest data pools had the least (**Table 11**). Also, Reference and Degraded sites in the NAPU bioregion were identified with limited habitat data, resulting in weaker differences between Reference and Degraded conditions.

The KW test identifies an initial tier of metrics responsive to anthropogenic impacts. The results are sufficiently influenced by n size as to be unreliable measures of a metric's ability to discern anthropogenic impacts. More biological metrics would likely exhibit significant KW tests in the NCA and NAPU bioregions if more Reference samples were incorporated into the analysis.

Discrimination efficiencies

Discrimination efficiencies (DEs) measure the responsiveness of the metrics to anthropogenic impacts in a different way. A metric's discrimination efficiency is a quantitative measure of its ability to correctly identify Reference or Degraded sites. A middle (M) percentile and a tail (T) percentile of the distribution of a metric's values in its Reference pool are used to classify all of the metric's values into three categories: "furthest from the Degraded condition," "somewhat removed from the Degraded condition," and "most like the Degraded condition." These three categories were assigned the numeric values of 5, 3, and 1, respectively (Figure 11). For each metric, the percentage of samples in the Reference data pool scoring 3 or 5 is the metric's DE for Reference sites and the percentage of samples in the Degraded data pool scoring 1 is the metric's DE for Degraded sites. A simple screeing or trial and

Table 11. Results from the Kruskal-Wallis test for significance between Reference (REF) and Degraded (DEG) sites in the Piedmont, Ridges, Valleys (combined), NCA, and NAPU. "X" denotes statistically significant at p<0.05. Highlighted metrics are significant in all ecoregions.

Metric		Piedmont	Ridges	Valleys	NCA	NAPU	
	REF n =	222	154	57	26	28	
	DEG n =	76	42	92	40	20	
ASPT MOD		х	Х	Х	Х		
BECK100		Х	Х	Х	Х	Х	
EPHEMEROPTERA TAX	A CNT100	х	Х	Х	Х		
EPT TAXA COUNT NO	TOL100	Х	Х	Х	Х	Х	
FBI	_	х	Х	Х			
GOLD		х		Х			
LOG10 SEL EPTD		х	Х	Х	х		
MARGALEFS		Х	Х	Х	Х	Х	
PCT AMPHIPOD		х		Х			
		х	Х	Х			
PCT_CLIMB		х					
PCT_CLING		х		х			
PCT CLINGER TAXA		х	х	х			
PCT_COLLECT		X		X			
PCT_CORBICULA		X					
PCT_DIPTERA		x	х	х			
PCT_DOM1		x	X	x	х		
PCT_DOM2		x	x	x	x		
		x	x	x	x		
		x	x	x	x		
DCT EDT		x	x	x	~		
DCT EDT TAXA RICH		X	X	x	x		
		×	~	Λ	~		
		×		v	v	v	
		~	v	~	v	~	
		v	^	v	^		
PCT_NET_CADDISFLT		×		× v			
		×		~			
		×	v	v		v	
PCI_PLECOPTERA		×	~	X	v	~	
PCI_SCRAPER		X	X	X	X		
PCI_SENSITIVE		X	X	X			
PCI_SHREDDER		N.	v	V	V		
PCI_SWIMINER		X	X	X	х		
PCI_IOLERANI		X	Х	X			
PCT_TRICHOPTERA		X		Х			
PCT_TRICHOPTERA_NO	HYDRO	X	X	Х		Х	
PCT_URBAN_INTOL		Х	Х	Х			
SENSITIVE_TAXA_COUN	IT100	Х	Х	Х	Х	Х	
SIMPSON_DIVERSITY		Х	Х	Х	Х		
SW		Х	Х	Х	Х		
TAXA_RICH100		Х	Х	Х	Х	Х	
TOLERANT_TAXA_COUI	NT100	х	Х				
# metrics with significar differences betw REF ar	nt nd DEG	40	29	35	19	8	

error approach (Cattel 1966) was used to adjust the T percentile (in increments of 5 %iles) with the goal of first making the Reference and Degraded DEs as similar as possible (within 10% of each other) and then as high as possible. Strong discrimination is typically found when the T percentile is between the 10th and 45th%ile of the Reference distribution for metrics decreasing in value with degradation and between the 70th and 95th%ile for metrics increasing with degradation. The best performing metrics are considered those with DEs greater than 70% for both Reference and Degraded indicating more than 70% of both the Reference and Degraded data pools are correctly classified when the T percentile is applied.



Figure 11. Three categories (1, 3, 5) are created from the distribution of a hypothetical metric's values in its Reference (REF) data pool. In this example, the metric increases in value with environmental improvement. Category 5 is delineated by the Ref distribution's M percentile (usually the 50th) and comprises metric values that are furthest from those in the Degraded (DEG) pool. Category 1 is delineated by the REF distribution's T percentile and comprises metric values that are most like those in the DEG pool. The T percentile is used to calculate discrimination efficiencies.

The average DE for each metric in each bioregion is shown in Table 12. Further detail is provided in Appendix D. There are many metrics with high DEs in the Piedmont, Ridges, and Valleys (combined, Valleys-L, Valleys-NL), and fewer metrics with high DEs in NCA and NAPU. Again, NCA and NAPU metric performance is impacted by the limited ability to identify Reference and Degraded in these regions and/or their smaller n sizes. Additional biological metrics will likely exhibit high DEs when more Reference samples are incorporated into the analysis in the NCA and NAPU bioregions. Percent Corbicula, % Amphipod, and to some extent % Oligochaeta did not lend themselves to the DE calculations because their counts were mostly zeros. The Reference and Degraded DEs of several other metrics could never be "balanced" to within 10% of each other for a variety of reasons. With one exception in the NAPU, these two groups of metrics were not considered for the Chessie B-IBI when they occurred.

Metric selection

In reviewing the KW test and discrimination efficiency results for each bioregion's candidate biological metrics, four criteria were applied before a final set of metrics was selected for the Chessie B-IBI. 1) There had to be little

redundancy in the metrics comprising the Chessie B-IBI. Pearson correlation coefficients were used to determine if similar or related metrics were redundant and responding in nearly identical ways. If r>0.90, the metrics were considered to be too closely correlated to each other and only one could be used in the Chessie B-IBI. Every attempt was made to reduce redundancies and no combination of metrics chosen had r values >0.90. 2) If two or more similarly functioning metrics performed equally well, preference was given to the more commonly used metrics (e.g., EPT, Hilsenhoff Biotic Index). 3) The set of selected metrics had to represent several of the metric types: taxa richness, composition, habit, tolerance, and feeding guild. 4) A metric's combined DE should preferably be greater than 70%, and should be no lower than 65%. (The 70% criterion was relaxed to 65% in the Valleys-NL, NCA, and NAPU bioregions in order to have a sufficient large pool of candidate metrics from which to choose.)

Various combinations of the most qualified metrics were tested in each bioregion until the classification efficiency was maximized. The classification efficiency (CE) is analogous to the metric DEs but it expresses the success of the combined metrics, or index, in identifying Reference and Degraded sites. For purposes of calculating the CE, the numeric categories 1, 3, or 5 (see **Figure 11**) associated with the index's metric values for each sample are averaged. The averages range from 1 to 5 due to the 1-3-5 classification approach applied to the individual metrics. A threshold of 3 is used to classify the average score of each sample as correctly identifying Reference or Degraded sites. The percent of samples in the Reference data pool with an average score equal to or greater than 3 is the classification efficiency for Reference sites and the percent of samples in the Degraded data pool with an average score less than 3 is the classification efficiency for Degraded sites. **Table 13** summarizes the biological metrics chosen for each bioregion, the T (and M) percentiles associated with their maximized Reference and Degraded DEs, and their combined CEs. The selection process in each bioregion is discussed in more detail below.

Table 12. Average discrimination efficiency (DE) of biological metrics for Reference and Degraded sites. (a), individual DEs for Reference and Degraded could not be brought to within 10% of each other ("balanced"), usually because the Reference and Degraded distributions overlap too much; (b), many metric values were zero (0) so the calculated percentiles were not useful. Highlighted DEs indicate the metrics that were eventually selected for the Chessie B-IBI in each bioregion.

			Valleys	Valleys			
Metric	Piedmont	Ridges	(combined)	NL	Valleys L	NCA	NAPU
ASPT_MOD	82.3%	72.3%	78.1%	71.2%	85.7%	58.3%	(a)
BECKS_100	78.0%	85.3%	73.1%	74.0%	78.7%	67.1%	68.9%
EPHEMEROPTERA_TAXA_CNT_100	74.0%	85.6%	(a)	(a)	(a)	75.3%	(a)
EPT_TAXA_COUNT_NO_TOL_100	68.7%	87.2%	74.2%	75.0%	83.5%	77.9%	64.6%
FBI	90.0%	68.4%	77.2%	67.3%	81.9%	54.4%	68.9%
GOLD	87.7%	57.3%	61.8%	64.4%	60.6%	51.3%	59.6%
LOG10_SEL_EPTD	74.9%	75.6%	72.0%	66.3%	81.9%	76.0%	65.4%
MARGALEFS	63.4%	81.8%	72.6%	69.2%	80.3%	73.4%	74.3%
PCT_AMPHIPOD	(b)	(b)	(b)	(b)	(b)	(b)	(b)
PCT_CHIRONOMIDAE	89.6%	57.3%	62.3%	65.4%	60.6%	(a)	60.4%
PCT_CLIMB	(a)	56.2%	64.0%	(a)	(a)	(a)	(b)
PCT_CLING	(a)	(a)	66.2%	(a)	72.1%	(a)	(a)
PCT_CLINGER_TAXA	(a)	64.9%	71.5%	(a)	77.8%	65.8%	(a)
PCT_COLLECT	81.2%	53.6%	68.2%	67.3%	73.3%	59.5%	62.9%
PCT_CORBICULA	(b)	(b)	(b)	(b)	(b)	(b)	(b)
PCT_DIPTERA	87.1%	54.8%	65.6%	65.4%	60.6%	50.0%	60.4%
PCT_DOM1	72.3%	76.6%	68.2%	61.5%	80.3%	70.9%	49.3%
PCT_DOM2	78.0%	78.0%	69.9%	61.5%	82.8%	69.6%	54.3%
PCT_DOM3	76.7%	79.2%	72.0%	63.5%	84.1%	66.4%	51.8%
PCT_EPHEMEROPTERA	81.3%	83.0%	76.8%	71.2%	84.4%	69.0%	57.9%
PCT_EPT	83.1%	58.4%	71.5%	66.3%	75.8%	(a)	64.6%
PCT_EPT_TAXA_RICH	(a)	75.8%	77.3%	74.0%	80.7%	70.9%	(a)
PCT_FILTERERS	74.2%	53.6%	61.7%	56.7%	64.8%	49.4%	51.8%
PCT_GATHER	86.4%	46.4%	62.4%	57.7%	68.0%	(a)	75.0%
PCT_LIMESTONE	(a)	68.4%	55.5%	55.8%	52.4%	77.8%	(b)
PCT_NET_CADDISFLY	73.6%	47.6%	58.6%	60.6%	59.4%	49.4%	48.2%
PCT_NON_INSECT	65.8%	50.0%	60.8%	58.7%	66.0%	(b)	(a)
PCT_OLIGOCHAETA	(b)	(a)	(b)	53.8%	(b)	(b)	(a)
PCT_PLECOPTERA	76.2%	(a)	68.8%	67.3%	74.6%	(a)	76.4%(a)
PCT_SCRAPER	76.7%	81.8%	72.5%	65.4%	78.3%	71.5%	(a)
PCT_SENSITIVE	89.1%	68.4%	75.2%	66.3%	81.6%	54.4%	66.4%
PCT_SHREDDER	50.7%	(a)	59.2%	60.6%	56.2%	62.0%	51.8%
PCT_SWIMMER	71.6%	78.0%	67.2%	60.6%	75.8%	65.9%	60.4%
PCT_TOLERANT	90.2%	57.3%	74.1%	69.2%	79.1%	56.9%	59.6%
PCT_TRICHOPTERA	73.6%	50.2%	59.7%	63.5%	(a)	48.8%	(a)
PCT_TRICHOPTERA_NO_TOL	66.0%	61.0%	64.6%	63.5%	(a)	(a)	65.4%
PCT_URBAN_INTOL	66.7%	(a)	71.0%	66.3%	(a)	45.6%	60.4%
SENSITIVE_TAXA_COUNT_100	80.0%	84.7%	(a)	74.0%	81.9%	69.6%	71.4%
SIMPSON_DIVERSITY	77.4%	79.0%	70.4%	62.5%	82.8%	71.5%	54.3%
SW	76.7%	82.8%	72.5%	71.2%	80.3%	73.4%	60.4%
TAXA_RICH_100	65.7%	81.4%	72.5%	71.2%	76.6%	74.7%	73.2%
TOLERANT_TAXA_COUNT_100	74.2%	57.6%	(a)	(a)	(a)	(a)	(a)

Matrics by Biorogian	Change w/	Stat.	DE	Tuno	REF	REF T %ile	REF M %ilo	REF M %ile	CE REF n
Diadmont	uegru.	Din.	DE	Type	1 76112	value	IVI 7611E	value	DEG II
	POS	v	v	т	>0 00	151	<0.60	3 63	01 7%
%Collectors	POS	x	x	FG	>0.90	71 02	<u><</u> 0.00	52 71	BFE n = 222
%Dintera	POS	x	x	ru C	>0.85	11 71	<u><</u> 0.50	6.60	DEG n = 76
%EPT	NEG	x	x	C C	<0.05	48 12	<u><0.70</u> >0.50	72 24	DEG II = 70
Shannon Wiener	NEG	x	x	R	<0.13	1 92	>0.50	2.24	
Vallevs - All	NLO	~	~	n	<u>-</u> 0.20	1.52	<u>-</u> 0.50	2.17	
Becks Index 100	NEG	х	х	т	<0.30	6.9	>0.50	9.0	81.9%
%Ephemeroptera	NEG	x	x	Ċ	<0.25	18.97	>0.50	29.33	RFF n = 57
%EPT Taxa Richness	NEG	x	x	R	<0.25	50.00	>0.50	55.56	DFG n = 92
%Scrapers	NEG	x	x	FG	<0.25	7.02	>0.50	14.41	
Shannon Wiener	NEG	x	x	R	<0.25	1.88	>0.50	2.09	
Valleys - No Limestone (NLV	()	~	~			2.00	<u>_</u> 0.00	2.00	
ASPT MOD	, POS	х	х	т	>0.70	4.4	<0.50	4.1	79.8%
%Ephemeroptera	NEG	x	X	C	<0.25	16.26	>0.50	21.57	REF n = 26
%EPT Taxa Richness	NEG	X	Х	R	<0.25	50.00	>0.50	53.59	DEG n = 52
%Scrapers	NEG	x	65%	FG	< 0.35	5.78	>0.50	10.56	
Shannon Wiener	NEG	X	X	R	< 0.30	1.97	>0.50	2.07	
Valleys - Limestone (LV)					_		-		
ASPT MOD	POS	х	х	т	>0.85	4.45	<0.60	4.2	91.1%
~ %Ephemeroptera	NEG	х	х	С	<u>-</u> 0.15	20.83	<u>></u> 0.50	38.51	REF n = 31
%EPT Taxa Richness	NEG	х	х	R	<0.20	50.00	>0.50	57.14	DEG n = 40
%Scrapers	NEG	х	х	FG		11.54	>0.50	17.54	
Shannon Wiener	NEG	х	х	R	<0.20	1.81	>0.50	2.14	
Ridges					-		-		
Becks Index 100	NEG	х	Х	т	<u><</u> 0.25	10	<u>></u> 0.50	13	85.5%
% Ephemeroptera	NEG	х	х	т	<u><</u> 0.15	13.51	<u>></u> 0.50	26.83	REF n = 154
%Scrapers	NEG	х	Х	FG	<u><</u> 0.15	3.37	<u>></u> 0.50	11.06	DEG n = 42
%Swimmers	NEG	х	х	н	<u><</u> 0.20	4.59	<u>></u> 0.50	10.73	
Shannon Wiener	NEG	х	х	С	<u><</u> 0.20	1.99	<u>></u> 0.50	2.26	
Northern Appalachian Platea	au (NAPU)								
FBI	POS		Х	т	<u>></u> 0.70	4.88	<u><</u> 0.50	4.57	86.1%
%Gatherers	POS	х	Х	FG	<u>></u> 0.75	56.72	<u><</u> 0.50	49.33	REF n = 28
%Plecoptera	NEG	х	Х	С	<u><</u> 0.30	1.96	<u>></u> 0.50	3.71	DEG n = 20
Taxa Richness 100	NEG	х	х	R	<u><</u> 0.30	13.4	<u>></u> 0.50	15	
%Trichoptera No Hydropsy.	NEG	х	65%	т	<u><</u> 0.40	3.08	<u>></u> 0.50	4.26	
North Central Appalachians	(NCA)								
EPT Taxa Count No Tol. 100	NEG	х	Х	R, T	<u><</u> 0.25	9.0	<u>></u> 0.50	11.0	77.8%
%Scrapers	NEG	Х	Х	FG	<u><</u> 0.25	8.45	<u>></u> 0.50	13.30	REF n = 26
Taxa Richness 100	NEG	Х	Х	R	<u><</u> 0.20	15	<u>></u> 0.50	17	DEG n = 40
%Ephemeroptera	NEG	Х	69%	С	<u><</u> 0.25	24.26	<u>></u> 0.50	40.22	
Shannon Wiener	NEG	х	Х	R	<u><</u> 0.25	2.19	<u>></u> 0.50	2.38	

Table 13. Biological metrics selected for the Chessie B-IBI in each bioregion. See text for details.

Piedmont The Family Level Hilsenhoff Biotic Index (FBI), %Collector, % Diptera, % EPT, and the Shannon Wiener Index were the final metrics selected for the Piedmont bioregion. These biological metrics represent community composition, feeding guild, richness, and tolerance features. No habit metric performed well enough to merit final selection into the index. Except for % Diptera, the chosen metrics were not significantly affected by season. Due to the large n size of the Piedmont's Reference data pool, most Piedmont's metrics show significant differences across stream order, therefore choosing biological metrics with significant stream order effects could not be avoided. However, the differences were not large despite their statistical significance (see **Appendix B**) and T %iles were identified that could minimize the stream order effects. The Piedmont-specific index correctly identified both Reference and Degraded conditions 91.7% of the time, or better than 9 in 10 samples.

Ridges Beck's Index 100, % Ephemeroptera, % Scraper, % Swimmer, and the Shannon Wiener Index were the biological metrics selected for the Ridges bioregion. These metrics represent a combination of community composition, feeding guild, habit, and tolerance. The % Ephemeroptera and % Scraper metrics were significantly affected by season with lowest values in Summer; the other three metrics were not affected. Only the Shannon Wiener Index showed significant differences across stream order, but the differences were not large (**Appendix B**). The Ridges-specific index correctly identified both Reference and Degraded conditions about 85.5% of the time.

Valleys Beck's Index 100, % Ephemeroptera, % EPT Taxa Richness, % Scraper, and the Shannon Wiener Index were the biological metrics selected for the Valleys-All bioregion (Limestone and Non-Limestone groups combined). They represent a combination of community composition, feeding guild, richness, and tolerance features of the stream communities in this bioregion. None of the metrics chosen were strongly affected by season or stream order (p<0.01), although several showed weak responses (0.01<p<0.05) (**Appendix B**). Identical metrics were selected for the Limestone and No Limestone groups of the Valleys bioregion except Beck's Index 100 was replaced with ASPT_MOD. ASPT_MOD is similar to the Beck's Index in that it is another tolerance metric. Using the ASPT_MOD instead of Beck's Index slightly increased the CEs of the two groups. The Valleys-All bioregion had a CE of 81.9%; the No Limestone group had a CE of 79.8%; the Limestone group had a CE of 91.1%. (Although two Valleys groups had relatively high CEs, the final decision was to use the Valleys-All index. See below for details.)

North Central Appalachians EPT Taxa Count No Tolerants 100, % Ephemeroptera, % Scraper, Shannon Wiener Index, and Taxa Richness 100 count were the biological metrics selected for the NCA bioregion. These metrics represent a combination of composition, feeding guild, richness, and tolerance features of the community (EPT Taxa Count No Tolerants 100 can be categorized as both a richness and tolerance metric). The % Ephemeroptera metric was slightly weaker than the others with a combined DE of 69%, however when combined with the other metrics it appeared to help the overall performance of the index. The CE of the overall index in this bioregion was 77.8%, and the lowest of all the non-Coastal Plain bioregions. The strength of the individual metrics and the CE of the overall index for this bioregion would be improved if the classification criteria for Reference and Degraded sites in this region were further refined. This is a small bioregion and small n sizes were a limiting factor in identifying Reference and Degraded communities.

Northern Appalachian Plateau and Uplands Family Level Hilsenhoff Index, % Gatherer, % Plecoptera, % Trichoptera no Hydropsychidae, and Taxa Richness 100 were selected for the NAPU bioregion. These metrics represent a combination of composition, feeding guild, richness, and tolerance features. The typical EPT metrics did not perform well in this bioregion. % Trichoptera no Hydropsychidae was a weaker performing metric than the others (combined DE of 65.4%), however when combined with the

other metrics it appeared to help the overall performance (CE) of the index. Family Level Hilsenhoff Biotic Index did not show a statistically significant difference between Reference and Degraded sites (**Table 11**) but it was retained because it worked well in the index when combined with other metrics. The CE of the overall index for this bioregion was 86.1%. The strength of the individual metrics and the CE could potentially improve if the classification criteria for Reference and Degraded sites in this region were further refined. This region currently lacks stream habitat and water quality data to adequate classify stream condition and confidently identify Reference and Degraded sites.

Overall, the biological metrics chosen for the Chessie B-IBI index in the Piedmont, Ridges, Valleys, NCA, and NAPU were top performers meeting most of the selection criteria. The metrics represent a diversity of biological community features and functions, discriminate well between Reference and Degraded conditions, and are not redundant with other metrics in the index.

Scoring the biological metrics and calculating the index

Metric values at each sampling location are individually scored and the scores averaged to obtain a final index, or Chessie B-IBI, score of stream condition. Monitoring programs in the Chesapeake Bay watershed currently use a variety of approaches for scoring the raw values of individual biological metrics before they are combined in an overall index. A consistent metric scoring approach was needed for non-Coastal Plain bioregions where the index was being developed from the raw data. Four metric scoring options were investigated to determine their effectiveness in differentiating between good and poor quality stream communities. All the options use thresholds selected from each metric's distribution of values to score the metric. Three options rely on distributions of the metric values in Reference communities to establish scoring thresholds; one relies on the distributions in all available samples. The options are described below. Thresholds not in parentheses apply to metrics that decrease in value when a site is disturbed such as %EPT and Taxa Richness; those in parentheses apply to metrics that increase in value with disturbance such as FBI and %Collector. In all options, the scores of the component metrics are averaged to obtain a final, index score for each sampling event.

Option 1) 1-3-5 Categorical Scoring Approach

This approach uses the T%ile (T'%ile) and M%ile of the Reference distribution to divide the raw values of each metric into three categories scored 1, 3, and 5. The T%ile is a low percentile in the tail of the Reference distribution – often the 25%ile – for metrics that decrease with disturbance. A hypothetical example is shown in **Figure 11**. (The T'%ile is a high percentile – often the 75%ile – for metrics that increase with disturbance.) Values for T%ile (T'%ile) are set so that the discrimination efficiencies (DE) of both Reference and Degraded are within 10% of each other. The M%ile is the 50th percentile for most metrics.⁴ Values above the M%ile of the Reference distribution (below M%ile for metrics that increase in value with disturbance) are most different from those in Degraded and receive a score of 5. Values below T%ile (above T'%ile) are the most similar to Degraded values and receive a score 1. Remaining values receive a score of 3. This index score is based on a scale ranging from 1 to 5. It is easily converted to a 0% - 100% scale by subtracting 1 and multiplying the result by 25. This approach is identical to the one used above to calculated DEs.

⁴ In several bioregions, separation of raw metric values in Reference and Degraded communities was such that classification efficiencies (CEs) of a few metrics increased when the M%ile threshold was set so that more or less than half of the Reference values scored 5 (e.g., FBI, %Collector, %Diptera).

Option 2) T%ile – M%ile Gradient Scoring Approach

This approach uses the same M and T (T') percentiles described above but creates a linear scale between the M and T(T') metric values. For metrics that decrease in value with stream disturbance, values less than or equal to T%ile of the Reference are assigned a score of 0% and values greater than or equal to M%ile of the Reference are assigned a score of 100%. Metric values between the T%ile and M%ile are scored proportionally from 0% to 100% according to the equation:

Score =
$$\frac{X - X_T}{X_M - X_T} * 100$$

where X is the metric value, X_M is the M%ile value, and X_T is the T%ile value. The reverse is done for metrics that increase with stream disturbance. Metric values less than the M%ile of the Reference are score 100%, values greater than the T'%ile of the Reference are scored 0%, and values between the M%ile and T'%ile are scored proportionally according to the equation:

$$\text{Score} = \frac{X_{T'} - X}{X_{T'} - X_M} * 100$$

Option 3) 0%-100% Reference Gradient Scoring Approach

This approach uses the 5th and 95th percentiles of each metric's raw values in the Reference community samples in a given bioregion to create a linear scale. For metrics that decrease in value with stream disturbance, the 5th%ile of the Reference is assigned a score of 0% and the 95%ile of the Reference is assigned a score of 100%. Metric values less than the 5%ile and greater than the 95%ile are scored 0% and 100%, respectively. Metric values between the 5th%ile and 95th%ile are scored proportionally from 0% to 100% according to the equation:

Score =
$$\frac{X - X_5}{X_{95} - X_5} * 100$$

where X is the metric value, X_{95} is the 95th%ile value, and X_5 is the 5%ile value. The reverse is done for metrics that <u>increase</u> with stream disturbance. Metric values less than the 5th%ile of the Reference are score 100%, values greater than the 95th%ile of the Reference are scored 0%, and values between the 5th%ile and 95th%ile are scored proportionally according to the equation:

Score =
$$\frac{X_{95} - X}{X_{95} - X_5} * 100$$

Option 4) 0-100% ALL Gradient Scoring Approach

This approach is basically the same as (3) above but it uses the 5th and 95th percentiles of metric values in <u>all</u> the available samples in a given bioregion to establish the 0% - 100% scale for scoring.

Examples from the Piedmont of the metric scoring scales are given in **Figure 12** for the Hilsenhoff Family Biotic Index (FBI) and %Diptera which increase with disturbance and %EPT and the Shannon-Wiener index which decrease with disturbance. Evident in the graphs is the sharper demarcation between high and low scores in Options 1 and 2. Both use the same M percentile from the middle of the distribution

of Reference raw values to assign the highest possible score to half or more of the Reference group. Both also use the same T (or T') percentile in the tail of the Reference distribution to assign the lowest possible score to roughly a quarter of the Reference group. In contrast, Option 3 assigns the highest possible score to only 5% and the lowest possible score to only 5% of each metric's raw values in the Reference group. Option 4, whose thresholds are strongly influenced by the relative frequencies of Reference and Degraded sites in each Chesapeake Bay bioregion, assigns the highest possible score to anywhere from 3% to 32% of each metric in the Reference group and rarely assigns the lowest possible score to this group.

The distributions of the Chessie B-IBI index scores which result from the four scoring options applied to the identical data (i.e., data subset where Reference and Degraded sites are identified with physical and water quality parameters) are shown as box-and-whisker plots in **Figures 13**, **14**, and **15**. The component metrics of each index rarely all score 100% (or 5) in a given sample, but the likelihood of an individual metric scoring 100% (or 5) is greatest in Options 1 and 2. Hence, distributions of the index's Degraded, Mixed, and Reference groups in Options 1 and 2 are higher on the 0% - 100% scale than those of Option 3 in each bioregion, even though scoring criteria in all three options are derived from the same group of Reference samples. In Option 4, the positions of the Degraded, Mixed, and Reference distributions on the 0%-100% scale depend entirely on the proportions of low, moderate, and high quality streams in the entire pool of samples.



Figure 12. Metric scoring scales for Hilsenhoff Family Biotic Index (FBI), %Diptera, %EPT, and Shannon-Wiener index derived from the four different scoring approaches, in the Piedmont. For the purpose of directly comparing Option 1 examples with the others, its 1, 3, and 5 scores were converted to 0%, 50%, and 100%, respectively. X-axis, all possible metric values found in the database; Y-axis, metric score. Brown, 1-3-5 categorical method (1); blue, T%ile – M%ile gradient method (2); green, 0%-100% REF gradient method (3); and red, 0%-100% ALL gradient method (4). See text for details.
As expected, Options 1 and 2 produce very similar index results because they use the same T(T') and M percentiles (top two rows in Figures 13-15). Both methods clearly separate Reference and Degraded index scores in all bioregions, and the Reference and Degraded groups are approximately balanced around 50%, or the middle of the scoring scale, with roughly ³/₄ or more of each Reference group scoring above 50% and more than $\frac{4}{5}$ of each Degraded group scoring below 50%. Reference medians are between 40 and 79 percentage points higher than their corresponding Degraded medians, and Reference 25th%ile and Degraded 75th%ile never overlap. In Option 1, the median values of the six bioregions occur in a narrow range between 60% and 70% and average 67.8%. The medians are slightly more variable in Option 2, ranging between 59.4% - 79.3% and averaging 66.2%. Reference 25th%ile values in both options straddle the 50% score and are roughly as variable as the medians across the bioregions. They average 54.3% in Option 1 and 52.1% in Option 2. This consistency in the Reference index values of the median and 25th%ile statistics is to be expected given that half (and sometimes more) of the values of each component metric scored 100% (or 5) in the Reference group in Options 1 and 2. This acts to stabilize the resulting index and anchor it to the upper end of the scoring scale. Variability in the index scores of the Reference group is introduced by the proportion of each component metric that scores 3 or 1 (Option 1) or by the distribution of scores between 1% and 99% (Option 2) in each bioregion.

Separation between Reference and Degraded indexes is not as clear in the Option 3 scale where each metric is scored according to a linear gradient between the 5th and 95th percentiles of the Reference group. Reference medians range from 39% to 62% with an average of 52.7%, so they are roughly equivalent to the 25% iles of the index distributions in Options 1 and 2. The Reference 25th% ile and Degraded 75th% ile actually overlap in Valleys No Limestone. Reference group medians in Option 3 are between 18 and 56 percentage points higher than their corresponding Degraded group medians and hence closer to Degraded scores than the Reference medians in Options 1 and 2.

Consistent differences between Reference and Degraded do not emerge in Option 4 where the distributions are heavily influenced by the relatively number of high, medium and low quality streams in each of the six non-Coastal Plain bioregions/karst groups. Reference group medians range from 43% to 84% with an average of 61.8%. As in Option 3, Option 4 medians are lower than the medians generated by Options 1 and 2 on the same Reference group. Option 4 Reference medians are anywhere from 14 to 54 percentage points higher than their corresponding Degraded group medians.

The vulnerability of the Option 4 scoring approach to proportions of low, medium and high quality streams in the entire bioregion and the low values of the Reference medians in Option 3 make these scoring options problematic. In Options 1 and 2, the consistency of the Reference group 50th and 25th percentiles across the six non-Coastal Plain bioregions/groups and the clear separation of Reference and Degraded index scores make these the preferred scoring options. Several state representatives in the Non-Tidal Water Quality Workgroup expressed a desire to have the metrics scored on a gradient scale rather than a categorical scale. In consideration of this, the T-M Reference Gradient scoring method, or Option 2, was used to score the biological metrics comprising the Chessie B-IBI.

As intended, the Reference group medians and percentiles across the six non-Coastal Plain bioregions/groups are very comparable in Option 2 (and in Option 1 as well) despite some variability related to small n size (**Figure 13-15**). The confounding influence of climate, geography, soil, and other natural factors have for the most part been minimized. This permits index scores from different non-Coastal Plain bioregions to be combined and makes watershed comparisons "seamless."



Figure 13. B-IBI scores derived from the four metric scoring options, Piedmont and Ridge bioregions. Box-andwhisker plots indicate distributions of Chessie B-IBI scores in Degraded (DEG), Mixed (MIX), and Reference (REF) site types. MIX are sites that do not classify as either REF or DEG. Median, **•**; box, interquartile range; whiskers, 5th%ile and 95th%ile. *Note:* these scoring methods comparisons were done before the feeding guild/habit/tolerance assignments (Appendix A) and metric scoring thresholds (Appendix D) were finalized, so values in these graphs deviate slightly from the final values.



Figure 14. B-IBI scores derived from the four metric scoring options, Valleys bioregion (with and without limestone). See Figure 13 heading for details.



Figure 15. B-IBI scores derived from the four metric scoring options, North Central Appalachians and the Northern Appalachian Plateau and Uplands. See Figure 13 heading for details.

Jackknife validation of non-Coastal Plain B-IBIs

A problem frequently encountered in developing indices is the lack of additional data for independent validation (Seegert, 2000). In the development of this index, there were 26, 28, and 57 Reference sites in the NCA, the NAPU, and the Valley bioregions, respectively, and data could not be reasonably withheld for validation purposes. Although the Piedmont had 222 Reference sites and the Ridges 154, a model cross validation technique known as "jackknife with replacement" (Snedecor & Cochran 1989) was applied to all bioregions in order to make error estimates comparable. Jackknife validation is thought to produce conservative estimates of classification rates among the currently accepted cross validation techniques (Olden et al. 2002).

In each bioregion, the Reference data pool had 10% of its records randomly removed at a time.⁵ The values of the Reference percentiles used to score each metric were recomputed from the truncated data set. The records were returned to the data set and the process repeated 100 times or until each combination of Reference samples had been withheld once. The resulting sets of simulated metric scoring criteria were each used to rescore all available data. The root mean squared error, total error, and bias were calculated on the differences in scores between the original and simulated results for each of the B-IBI metrics and for the overall B-IBI score. The total error estimates generated by the simulations provide a measure of the variability in individual B-IBI scores that propagate from uncertainty in the metric scoring thresholds due to sampling.

Total error estimates of the B-IBI scores range from 1.8% to 10.2% (**Figure 16**). The Piedmont, Ridges, and Valleys-All bioregions with 222, 154, and 57 Reference sites, respectively, have error estimates about 5% or less. The NCA and NAPU bioregions with only 26 and 28 Reference sites, respectively, have error estimates of more than 8%. When the Valleys bioregion was divided into Limestone (n=31) and No Limestone (n=26) groups, error estimates are 7.5% and 8.0%, respectively. Overall, the larger the pool of Reference sites, the smaller the estimate of B-IBI total error. Error estimates for the individual B-IBI

metrics range from 1.03% - 27.5% in the non-Coastal Plain bioregions and as high as 35.4% in the two Valleys bioregion groups when Valleys is split (Table 14). In each bioregion or group, the error estimate of the total B-IBI score is roughly half of the average of the component metrics error estimates. This suggests that, as expected, error inherent in one metric's ability to identify a Reference site tends to be outweighed by correct identifications made by the index's other metrics. Metrics standardized to 100 counts per sample (i.e., Becks 100, Taxa Richness 100, EPT Taxa Count No Tolerants 100) and those involving small numbers (e.g., %Trichoptera No Hydropsychidae)





⁵ This jackknife validation is more rigorous than the one applied in the initial development of the Chessie B-IBI where just one sample was withheld from the Reference data pool (Foreman et al. 2008).

Table 14. Jackknife validation results with 10% of Reference data pool withheld. The prefix "100" indicates themetric was standardized to 100 counts per sample. The Valley bioregion was analyzed as one unit (All) as wellas two subunits (Limestone, No Limestone). Sample count bins: "100-cnt," <150 counts per sample; "200-cnt</td>and higher," \geq 150 counts per sample.

Bioregion or Group	Metric	Total Erro
North Central Appalachian (NCA)	%Ephemeroptera	7.66%
REF total n = 26	%Scraper	6.25%
Jackknife n = 23	EPT Taxa Cnt No Tolerants 100	27.51%
(0% 100-cnt, 100% 200-cnt and higher)	Taxa Richness 100	15.10%
	Shannon Wiener	15.81%
	B-IBI Score	10.20%
Northern App. Plateau & Upland (NAPU)	%Gatherer	10.31%
REF total n = 28	%Plecoptera	23.229
Jackknife n = 25	%Trichoptera No Hydropsych.	18.34%
(77% 100-cnt, 8% 200-cnt and higher)	Taxa Rich 100	22.00%
	Hilsenhoff Family Biotic Index	14.28%
	B-IBI Score	8.56%
Piedmont	%Collect	5.88%
REF total n = 222	%Diptera	6.14%
Jackknife n = 200	%EPT	3.79%
(79% 100-cnt, 21% 200-cnt and higher)	Hilsenhoff Family Biotic Index	23.62%
	Shannon Wiener	4.48%
	B-IBI Score	4.78%
Ridges	%Ephemeroptera	2.17%
REF total n = 154	%Scraper	4.21%
Jackknife n = 139	%Swimmer	3.20%
(51% 100-cnt, 49% 200-cnt and higher)	Beck Index 100	1.03%
	Shannon Wiener	5.37%
	B-IBI Score	1.76%
Valley (All)	Beck Index 100	17.44%
REF total n = 57	%Ephemeroptera	8.10%
Jackknife n = 51	%EPT Taxa Richness	6.38%
(60% 100-cnt, 40% 200-cnt and higher)	%Scraper	10.26%
· · · · · · · · · · · · · · · · · · ·	Shannon Wiener	10.29%
	B-IBI Score	5.06%
Valley – Limestone Group	%Ephemeroptera	13.25%
REF total n = 31	%EPT Taxa Richness	9.63%
Jackknife n = 28	%Scraper	9.86%
(77% 100-cnt, 23% 200-cnt and higher)	ASPT_Mod	33.39%
<u> </u>	Shannon Wiener	6.92%
	B-IBI Score	7.46%
Valley - No Limestone Group	%Ephemeroptera	9.04%
REF total n = 26	%EPT Taxa Richness	20.05%
Jackknife n = 23	%Scraper	14.79%
(38% 100-cnt, 62% 200-cnt and higher)	ASPT_Mod	35.42%
<u> </u>	Shannon Wiener	7.09%
	B-IBI Score	7.96%

have higher error estimates even when they have good discrimination efficiencies. Error estimates of the component metrics are also higher when the Reference data pool contains few sites. For example, metric error estimates in the data-rich Piedmont and Ridges bioregions are with one exception below 6.2% while those in the two Appalachian bioregions, Valleys bioregion, and two Valleys groupings (Limestone, No Limestone) are all above 6.2%.

There are several possible causes of high total error in individual metrics and the overall index. The largest factor appears to be the number of Reference sites. The bigger the pool of Reference samples the better the precision of the index and many of its component biological metrics. The dashed line relating index total error and number of Reference sites in **Figure 16** suggests that 50 Reference samples is a threshold above which the index total error rate levels off at 5% or less. Another source of error, or more properly, uncertainty may be the influence of subsample count. Comparing just the two bioregions with the largest Reference pools, Ridges has a lower error estimate than Piedmont. About half of the Ridges samples have sample counts of 200 or more while most of the Piedmont samples are 100-count samples (**Table 14**). A third factor possibly influencing total error estimates could be the less reliable method used to classify Reference and Degraded streams in the NAPU bioregion. Total error estimates of all the component metrics in the NAPU bioregion were high, above 10%.

The Limestone Valley group has a natural stressor - limestone - that is not factored in when sites are classified as Reference or Degraded, perhaps explaining some of this group's high error estimate. Despite significant discrimination abilities in some of its metrics, the Valleys bioregion should probably not be divided into Limestone and No Limestone groups until more Reference sites are identified and used to refine the scoring criteria. This will facilitate lower total error estimates.

The jackknife validation results indicate total error estimates for the Chessie B-IBI scores will improve as the number of Reference samples increases up to about 50 samples. Incorporating more Reference sites in the NCA and NAPU analyses should theoretically improve total error estimates in these bioregions. More accurate identification of Reference and Degraded sites using stream habitat scores and water quality conditions can also be expected to improve error estimates in the NAPU bioregion.

Narrative Condition Ratings

Narrative condition ratings are used to communicate numerical monitoring results to a broad audience in a succinct, descriptive manner. Five rating categories provide good resolution on the quality of stream conditions. Previously, narrative condition ratings for Chessie B-IBI scores were based on the means of values representing key percentiles in four bioregion-specific Reference and Degraded communities, i.e., the means of the 50th, 25th, and 10th percentiles of the Reference communities across bioregions, and the mean of the 50th percentile of the Degraded communities across bioregions (Foreman et al. 2008). If the original B-IBI scale of 7 - 35 is converted to the 0%-100% scale used for the refined Chessie B-IBI, these thresholds would assign an "Excellent" condition to an index score of \geq 71.4%; "Good" to 57.1% - <71.4%; "Fair" to 42.9% - <57.1%; "Poor" to 25% - <42.9%; and "Very Poor" to <25%. The same thresholds would apply to the Coastal Plain B-IBI scores which were previously calculated using an MDDNR method (Stribling et al. 1998).

In the current effort to improve the Chessie B-IBI the following changes were made:

 more stringent habitat and water quality criteria for Reference and Degraded sites were used in the Piedmont, Ridges and Valleys bioregions

- two new bioregions—NCA and NAPU—were created from the original "Northern Appalachian" bioregion and new habitat and water quality criteria for Reference and Degraded were established for these bioregions
- new combinations of five instead of seven metrics were incorporated into the B-IBI in all bioregions
- the metric scoring thresholds were adjusted to balance the Reference and Degraded discrimination efficiencies
- a different (gradient) protocol was used to score the metrics

These changes warrant adjusting the thresholds used to assign narrative condition ratings. As before, we decided *a priori* to use a single set of thresholds to assign narrative ratings to the B-IBI scores across all bioregions.

The means of the 75th, 25th and 5th percentiles of the B-IBI Reference values in non-Coastal Plain bioregions closely align with the rating thresholds currently applied by Virginia to its Coastal Plain familylevel CPMI scores (**Table 15**). Specifically, the mean of the Reference 75th%iles in the non-Coastal Plain bioregions (84.5%) is approximately equal to the lower boundary of the "Not Impaired" condition rating for the CPMI (80.0%). Thus, about half of the Chessie B-IBI scores that are *least like* Degraded scores are equivalent to the "Not Impaired" CPMI scores. The mean of the Reference 25th%iles in non-Coastal Plain bioregions (47.4%) is approximately equal to the boundary separating the "Slightly Impaired" and "Moderately Impaired" categories (51.0%). In the Chessie B-IBI, this range is where the tails of the Degraded and Reference distributions overlap the most. Finally, the mean of the Reference 5th%ile (12.1%) is close to the upper boundary of the "Severely Impaired" category for the CPMI (17.0%). Chessie B-IBI values in this low range are most like those in Degraded conditions. The good alignment between the Reference 5th, 25th, and75th percentiles and the CPMI rating thresholds indicates the Chessie B-IBI and CPMI scales are representing similar Reference communities. It allows us to develop a common set of stream narrative ratings for all Chesapeake bioregions.

We examined several ways of creating a 5-category rating scale for the non-Coastal Plain B-IBI and matching it to the existing CPMI 4-category rating scale. In consideration of the slight shift in index scores resulting from the new T-M Gradient metric scoring approach (compare Options 1 and 2 results in **Figures 13-15**), we felt the best choice for the non-Coastal Plain bioregions was to rate index scores for

Table 15. The Chessie B-IBI scores at selected percentiles of the index's Reference distribution in each non-Coastal Plain bioregion. Scores are compared to the rating criteria currently applied to the family-level Coastal Plain Macroinvertebrate Index in Coastal Plain bioregions (obtained from A. Budd of VADEQ and converted to a 0%-100% scale).

Bioregion	5 th %ile	10 th %ile	25 th %ile	50 th %ile	75 th %ile	95 th %ile
Non-Coastal Plain Bioregions						
Valleys - All	24.5	33.1	43.7	60.0	80.0	98.8
Ridges	6.3	22.2	49.2	69.5	91.2	100
Piedmont	20.2	40.4	60	76.8	88.2	100
North Central Appalachians	6.7	16.1	36.0	70.1	84.0	100
Northern Appalachian Plateau & Uplands	2.6	24.6	47.9	60.4	78.9	96.9
Average	12.1	27.3	47.4	67.4	84.5	99.1
Coastal Plain Bioregions						
Southeastern Coastal Plain and Mid-Atlantic Coastal Plain	17.0%		51.0%		80.0%	
Impairment Ratings Severe	\rightarrow \leftarrow	Moderate	\rightarrow \leftarrow	Slight	\rightarrow \leftarrow	None

Excellent, Good, and Fair using the same *percentiles* of the index's Reference distributions that were chosen in the earlier Chessie B-IBI (Foreman et al. 2008). Specifically, the mean of the 50th%iles in **Table 15** establishes the lower boundary for Excellent, the mean of the 25th%iles is the lower boundary for Good, and the mean of the 10th%iles is the lower boundary for Fair. With these rating thresholds, approximately ¾ of Reference sites in the non-Coastal Plain will score Good or Excellent and less than a tenth will score Poor or Very Poor. This framework recognizes that benthic communities at sites with good habitat and water quality can still score poorly due to other stressors. In the interest of simplicity, we rounded the lower boundary of Excellent to 67%, of Good to 50%, and of Fair to 30% which places them more evenly on the 0%-100% scale. We decided to use the upper boundary of the CPMI Impaired category, which is 17% on the 0%-100% scale, as the threshold for separating the Chessie B-IBI Poor and Very Poor categories. The 17% threshold is close to the mean of the Reference 5th%iles (12.1%) in the non-Coastal Plain and divides roughly in half the scale below the Chessie B-IBI Fair category. **Table 16**

lists the narrative condition ratings applied to the Chessie B-IBI scores for all bioregions in the Chesapeake Bay basin.

A comparison of the *narrative condition ratings* produced using the old and new Chessie B-IBI methods and applied to the data set used in Foreman et al. 2008 (n=13,448) shows good agreement across the basin despite the method changes. Actual index scores are lower when they are calculated with the new method. However, because the same percentiles (50th, 25th, 10th) of the Reference group are used to create the Excellent, Good and Fair rating categories in both the new and old methods, the ratings agree exactly or very closely about 87% of the time. They disagree strongly less than 0.8% of the time. Table 16. CBP narrative conditionratings applied to B-IBI indexscores in all bioregions of theChesapeake Bay basin.

Rating	B-IBI Score
Excellent:	<u>></u> 67%
Good:	50% - <67%
Fair:	30% - <50%
Poor:	17% - <30%
Very Poor:	<17%

Findings

Chessie B-IBI scores were calculated for over 14,000 stream and wadeable river sampling events obtained from the different monitoring programs across the Chesapeake Bay watershed during the years 2000-2008. The results were averaged for sampling events collected at different times at the same location. This averaging was done to ensure that a location sampled multiple times is not given more weight than other locations. In a few cases many years of data from multiple organizations were averaged. In most other cases, only one event from one organization represented the average score at that location. A total of 10,833 sampling locations throughout the Chesapeake Bay watershed received Chessie B-IBI scores: 1,538 (14.2%) rated Excellent, 1,375 (12.7%) rated Good, 1,944 (18.0%) rated Fair, 1,609 (14.8%) rated Poor, and 4,367 (40.3%) rated Very Poor. The preponderance of Poor and Very Poor ratings is due in part to the intense sampling currently done in the heavily urbanized counties in the Washington, DC region. Results are also weighted by the relatively large Maryland data set.

Comparisons with state ratings

Chessie B-IBI qualitative ratings were compared to state qualitative ratings when possible in order to determine if the ratings were similar. An exact match was not expected given the different number of rating categories applied by the states and the different index thresholds used to delineate the rating categories (US EPA 2010b). Where state ratings were available, it was found that the Chessie B-IBI ratings for the same locations generally agree with the states' ratings. The Chessie B-IBI agrees most often with the state's ratings in the "best" and "worst" categories. There is variability in how

comparable the Chessie B-IBI and state categories are between the "worst" and "best" rankings. General agreement in scores was usually over 80%. Only a few locations were rated as excellent or good by the Chessie B-IBI where the states considered benthic conditions as poor, and *vice versa*. In most cases the percent mismatch was less than 10%. Regardless of the rating discrepancies, the standardized procedures used to derive Chessie B-IBI make it more useful than the various state-specific indexes for an interstate, Chesapeake Bay basin-wide evaluation of stream health and can lead to a better understanding of benthic and habitat issues in the basin. A more appropriate analysis of state and Chessie B-IBI results in the future would be a direct comparison of the Chessie B-IBI index scores and the state index scores, all placed on a 0% - 100% scale. This analysis removes the artificially imposed confines of the qualitative ratings.

Mapping the Chessie B-IBI results

More than half of the 10,833 stream and wadeable river locations across the Chesapeake basin are rated Poor or Very Poor by the Chessie B-IBI, but maps of sampling locations color-coded to reflect the ratings can sometime give a biased impression. In the mapping process, dots indicating the sampling locations overlap in intensely sampled areas. A reader will falsely conclude that most of the locations scored high if the green "Excellent" and "Good" ratings are plotted on top of the other ratings (**Figure 17**). Conversely, maps with the red "Poor" and "Very Poor" ratings plotted on top will give the opposite impression (not shown). Although index ratings at each location are informative at a fine scale, the information usually needed by natural resource agencies for communication purposes is the overall condition of streams on a watershed basis. In order to enhance interpretation and usefulness of the Chessie B-IBI results, Chessie B-IBI index scores were averaged across the smallest feasible watershed size. At this time, the smallest feasible watershed size proves to be the USGS Hydrologic Unit Category (HUC) 8 in sparsely sampled areas of the Chesapeake Bay basin and the smaller HUC10 in more intensely sampled areas.

The Chessie B-IBI scores used to represent watershed conditions were limited to those from randomly sampled or systematically sampled (grid) sites that were not originally selected by the monitoring programs to target areas of known degradation or high quality (**Figure 18**). Included in this group are locations that were randomly/systematically selected and sampled just once, and those that were randomly/systematically selected and then repeatedly sampled in the 2000-2008 period. (Chessie B-IBI scores from the latter type are first averaged before being grouped with other location results.) By using randomly/systematically selected locations, Chessie B-IBI scores can be averaged across a watershed without introducing the presumed bias that would be associated with targeted sampling designs. A subset of 7,886 locations in the Chesapeake Bay basin met the random/systematically sampling design requirement. Interestingly, the results for this subset of locations were very similar to those for all 10,833 locations: 13.9% versus 14.2% rated Excellent, 13.0% versus 12.7% rated Good, 18.5% versus 17.9% rated Fair, 15.2% versus 14.9% rated Poor, and 39.4% versus 40.3% rated Very Poor. A closer look at the selection process and purpose for sampling the "targeted" locations is warranted to determine if they in fact bias the overall watershed results.

A minimum n size was considered important before averaging Chessie B-IBI scores across HUC8 and HUC10 watersheds, so the following rules were developed:

- if a HUC8 has n<10 locations it does not receive a rating
- if a HUC8 has n>= 10 locations the B-IBI scores are averaged to generate a HUC8 score, which is then rating according Chessie B-IBI rating thresholds
- if a HUC10 has n<5 locations it does not receive a rating

- if a HUC10 has n>= 5 locations the B-IBI scores are averaged to generate a HUC10 score, which is then rating according to Chessie B-IBI rating thresholds
- HUC10 watersheds are mapped *on top of* HUC8s watersheds in order to provide the finer resolution of the results

Figure i in the Executive Summary illustrates the result of this mapping procedure. It is easily apparent from the map where random/systematic sampling designs have and have not been used extensively. There are fewer sites identified as random or systematic in New York and Pennsylvania compared to other states (see Report Figure 18). Chessie B-IBI ratings in New York and Pennsylvania are therefore less certain because they are derived from fewer data. Random/systematic sampling was instituted earlier or intensively in West Virginia, Delaware, Virginia, and Maryland and HUC10 coverage in those states is complete for the 2000-2008 period. Future additions of random/systematic sampled data to the database will improve HUC10 watershed characterizations in more of the Chesapeake Bay basin. If analyses of the targeted location results show that certain types of these locations do not significantly bias the overall watershed results, their inclusion would allow more HUC10 watershed evaluations.

Figure i shows a clear link between the Chessie B-IBI scores and land-based activities in individual watersheds. The poorest stream indexes occur in highly urbanized watersheds such as those in the Baltimore-Washington D.C. metropolitan region. Stream health is compromised in urban areas by extreme land disturbance and an abundance of paved surfaces. These stressors result in high levels of pollution, altered stream flow, and poor quantity and quality of streamside vegetation. Lower scores in the Chesapeake Bay basin are also present in areas with intense agricultural activity such as the lower Eastern Shore and south central Pennsylvania. Excess nutrients and sediment compromise stream health in these areas. The Upper West Branch of the Susquehanna River in Pennsylvania appears to be compromised by mining activity which causes habitat alterations and toxic plumes that negatively impact benthic stream populations. The highest Chessie B-IBI scores are typically found in minimally disturbed watersheds with low levels of pollution and stable in-stream and streamside habitats. These watersheds tend to be clustered in forested areas along the western side of the Chesapeake Bay basin.

Trend analysis of Chessie B-IBI scores at fixed sample locations that have been repeatedly sampled for a long time period is possible now. The methodology for assessing whole watershed trends using the randomly sampled data does not exist and an effort should be made for its development. A technical issue impeding development of this methodology is the fact that many programs monitor benthic macroinvertebrates on a rotating cycle, with data from one year representing a two- to 10-year period. Given the nature of the data, it is expected that change in watershed condition over time will be represented in a minimum of 5 year increments. It is also expected that the time frame for pooling the Chessie B-IBI index scores to represent the condition of streams will change as this methodology is developed. The time frame used to portray status in this report was the 9-year period of 2000-2008.

Possible Next Steps

Future work to improve upon the Chessie B-IBI could include the following tasks:

- explore the effects of targeted vs. random/systematic sites on Chessie B-IBI watershed results
- develop a limestone-specific scoring approach for affected regions in the Valleys and possibly the Piedmont (more limestone data would need to be incorporated in order to do this)



Figure 17. Chessie B-IBI ratings for all 10,833 locations, 2000-2008. Dots representing Excellent and Good locations overlay and obscure those representing Fair, Poor and Very Poor locations in densely sampled areas.



Figure 18. Chessie B-IBI ratings for 7,803 randomly selected locations, 2000-2008. Dots representing Excellent and Good locations overlay and obscure those representing Fair, Poor and Very Poor locations in densely sampled areas.

- improve the Chessie B-IBI performance in the North Central Appalachians and the North Appalachian Plateau and Uplands regions by acquiring new habitat and water quality data and better identifying Reference and Degraded sites
- test the rigorousness of the Chessie B-IBI in the coastal plain bioregions by identifying Reference and Degraded sites in those bioregions and testing the discrimination efficiencies of the familylevel biometrics
- compare in more detail the several Coastal Plain indexes that have been developed and evaluate if and how they differ when calculated from the same data set
- calculate trends in the Chessie B-IBI over time with a subset of the data collected from fixed locations
- develop a methodology for using random/systematic sampling design data to assess whole watershed trends
- validate the Chessie B-IBI with new data as it is added to the database
- determine the influence of blackwater systems on Chessie B-IBI scores and potentially having a blackwater-specific scoring approach for affected regions

References

- Astin, L. E. 2007. Developing biological indicators from diverse data: The Potomac Basin-wide Index of Benthic Integrity (B-IBI). Ecological Indicators, 7: 895-908.
- Astin, L.E. 2006. Data synthesis and bioindicator development for nontidal streams in the interstate Potomac River basin, USA. Ecological Indicators, 6: 664-685.
- 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. Available online at: http://www.epa.gov/owow/monitoring/rbp/download.html
- Buchanan, C. Personal communication. September 2010. Interstate Commission on the Potomac River Basin. cbuchan@icprb.org
- Budd, A. Personal communication. October 2010. Virginia Department of Environmental Quality. <u>Aimee.Budd@deq.virginia.gov</u>
- Burton, J. and J. Gerritsen. 2003. A Stream Condition Index for Virginia Non-Coastal Streams. Report prepared by Tetra Tech for Virginia Department of Environmental Quality.
- Cattel, R.B., 1966. The meaning and strategic use of factor analysis. In: Cattel, R.B. (Ed.), Handbook of Multivariate Experimental Psychology. Rand McNally, Chicago, pp. 143-174.
- Colwell, R. K. 2009. EstimateS: Statistical estimation of species richness and shared species from samples. Version 8.2. User's Guide and application published at: <u>http://purl.oclc.org/estimates</u>.
- Feminella, J. W. 2000. Correspondence between stream macroinvertebrate assemblages and 4 ecoregions of the southeastern USA. J. N. Am. Benthol. Soc. 19(3):442-461.
- Foreman, K., C. Buchanan, and A. Nagel. 2008. Development of ecosystem health indexes for non-tidal wadeable streams and rivers in the Chesapeake Bay basin. Report to the Chesapeake Bay Program Non-Tidal Water Quality Workgroup. December 5, 2008.
- Hawkins, C. P., R. H. Norris, J. Gerritsen, R. M. Hughes, S. K. Jackson, R. K. Johnson, and R. J. Stevenson. 2000. Evaluation of the use of landscape classifications for the prediction of freshwater biota:

synthesis and recommendations. J. N. Am. Benthol. Soc. 19(3):541-556.

- Kennan, J. G. 1999. Relation of macroinvertebrate community impairment to catchment characteristics in New Jersey streams. JAWRA 35(4):939-955.
- Maxted, J., M. Barbour, J. Gerritsen, V. Poretti, N. Primrose, A. Silvia, D. Penrose, and R. Renfrow. 2000. Assessment framework for mid-Atlantic coastal plain streams using benthic macroinvertebrates. Journal of American Benthological Society, 19 (1): 128-144.
- Merritt, R. W., K. W. Cummins, and M. B. Berg. 2008. An Introduction to the Aquatic Insects of North America, Fourth Edition. Kendall Hunt Publishing Co., Iowa, pp. 1158.
- Olden, J. D., D. A. Jackson, and P. R. Peres-Neto. 2002. Predictive Models of Fish Species Distributions: A Note on Proper Validation and Chance Predictions. Transactions of the American Fisheries Society, 131, 329–336.
- Omernik, J.M. 1987. Ecoregions of the conterminous United States (map supplement): Annals of the Association of American Geographers 77(1):118-125, map scale 1:7,500,000.
- Omernik, J.M. 1995. Ecoregions: a spatial framework for environmental management, in Davis, W.S. and Simon, T.P. (eds.), Biological assessment and criteria, tools for water resource planning and decision making. Lewis Publishers, Boca Raton, Florida. p. 31-47.
- Passmore, M. Personal communication. 2011. US Environmental Protection Agency, Region III, Wheeling, WV. <u>passmore.Margaret@epamail.epa.gov</u>
- 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. US Environmental Protection Agency EPA/444/4-89-001.
- Seegert, G. 2000. The Development, Use, and Misuse of Biocriteria With an Emphasis on the Index of Biotic Integrity. Environmental Science and Policy 3:S51-S58
- Snedecor, G. W. and W. G. Cochran. 1989. Statistical Methods, Eighth Edition, Iowa State University Press.
- Southerland, M., G. M. Rogers, M. J. Kline, R. P. Morgan, D. M Boward, P. F. Kazyak, R. J. Klauda, and S.
 A. Stranko. 2005. New biological indicators to better assess the condition of Maryland streams. Maryland Department of Natural Resources. DNR-12-0305-0100.
- Stribling J.B., B.K. Jessup, J.S. White, D. Boward, and M.Hurd. 1998. Development of a Benthic Index of Biotic Integrity for Maryland Streams. Prepared by Tetra Tech, Inc., Owings Mills, MD and Maryland Department of Natural Resources, Monitoring and Non-tidal Assessment Program. CBWP-MANTA-EA-98-3.
- US Environmental Protection Agency. 1997. "Field and laboratory methods for macroinvertebrate and habitat assessment of low gradient nontidal streams"; Mid-Atlantic Coastal Streams Workgroup, Environmental Services Division, Region 3, Wheeling, WV; 23 pages with appendices.
- US Environmental Protection Agency. 2008. Bay Barometer. A Health and Restoration Assessment of the Chesapeake Bay and Watershed in 2008. EPA-903-R-09-001 March 2009, Chesapeake Bay Program Office, Annapolis, MD
- US Environmental Protection Agency. 2009. Bay Barometer. A Health and Restoration Assessment of the Chesapeake Bay and Watershed in 2008. Available online at: http://www.chesapeakebay.net/indicatorshome.aspx?menuitem=14871

- US Environmental Protection Agency. 2010a. Western Ecology Division Ecoregion Maps and GIS Resources. US EPA Level IV ecoregion. Available online at: (http://www.epa.gov/wed/pages/ecoregions.htm)
- US Environmental Protection Agency. 2010b. Wadeable Streams Macroinvertebrate Methods Comparison [Microsoft Excel Book]. May 27, 2010. U.S. EPA Region III, Wheeling, WV. Contact: Kelly Krock, 304-234-0242, krock.kelly@epa.gov.
- US Geological Survey. 2008. "Preliminary Map of Potentially Karstic Carbonate Rocks in the Central and Southern Appalachian States."
- Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell, and C. E. Cushing. 1980. The River Continuum Concept. Can. J. Fish. Aquat. Sci. 37:130-137.
- Wood, A.J., J.M. Olmernik, and D.D. Brown. 1999. Level III and IV ecoregions of Delaware, Maryland, Pennsylvania, Virginia, and West Virginia. U.S. Environmental Protection Agency technical document, Corvallis, Oregon. 61pp. Available online at: <u>http://www.epa.gov/wed/pages/ecoregions/reg3_eco.htm</u>

APPENDIX A

Feeding Guild, Habit, and Tolerance Assignments for Family-Level Taxa

The family-level taxa listed in **Table A-1** have been reported in the multi-jurisdictional agency data. Tolerance values, functional feeding groups, and habit characteristics were assigned from a combination of sources, including the jurisdictional agency taxa lookup tables and Merritt et al. 2008. Feeding guild and habit characteristics do not typically change across geographic regions, but family-level tolerance values can be region-specific.

A taxonomic group may have a lower tolerance value at the limits of its geographic range than at its interior, as it is at the limits of its ecological niche. For this reason, the jurisdictional agencies within the Chesapeake basin maintain their own taxa tolerance assignments. For the purposes of this Chesapeake Basin Index, a single tolerance value was needed to calculate metrics similarly across the basin. Family-level tolerance values were obtained from the state jurisdictional agencies and were compared for their similarities and differences. An average tolerance value was calculated for each taxon and formed the basis for consideration by the state agencies. If the averaged tolerance value did not differ from the state agencies' values by more than 1 unit the average was used. For those taxa that deviate more than 1 unit, a single tolerance value was reached by consensus of representatives of the state agencies during a conference call held on May 8, 2009. The tolerance values in **Table A-1** represent either a consensus or an average of the state values.

Family-level feeding guild and habit assignments were drawn primarily from Merritt et al. 2008. Note: if a particular genus was the only/dominant representative of a family in the Chesapeake Bay basin and its feeding guild or habit assignment differed from the generalized family assignment, the genus-level assignment was made the family-level assignment. Exclusion taxa are those reported in the data that are either a) not known to occur in this region, or b) accidentally collected terrestrial or marine taxa that should not be included for the assessment of flowing waters.

Table A-1. Family-level attribute assignments. Feeding guild: CF, collector-filterer; CG, collector-gatherer; PR, predator; SC, scraper; SH, shredder. Habit: BU, burrower; CB, climber; CN, clinger; SK, skater; SP, sprawler; SW, swimmer. GOLD index taxa: D, Diptera; G, Gastropoda; O, Oligochaeta. EPTD Taxa: E, Ephemeroptera; P, Plecoptera; T, Trichoptera; D, Diptera.

	Taxon							GOLD		
	Serial	Feeding			Beck	EPTD	Urban	Index	Non-	Excl.
Family	Number	Guild	Habit	Tolerance	Class	Таха	Intolerant	Таха	Insect	Таха
ACARIFORMES	0690740	PR		6	3				Х	
ACARINA	0082754	PR	CN	7	3				Х	
AEOLOSOMATIDAE	0068423	CF	BU	8	3				Х	
AESHNIDAE	0101596	PR	СВ	3	2		Х			
AGRIIDAE	0181180									
AMELETIDAE	0568544	SC	SW,CN	1	1	Е	Х			
AMPHARETIDAE	0067718	CG	BU	10	3				Х	Х
AMPHIPODA	0093294	CG	SP	7	3				Х	
ANCYLIDAE	0076568	SC	CB	7	3			G	Х	
ANISOPTERA	0101594	PR	CB,SP	6	3					
ANNELIDA	0064357	CG	BU	10	3				Х	
ANTHURIDAE	0092144	CG	SP	8	3				Х	
APATANIIDAE	0598182	SC	CN,CB,SP	1	1	Т				

	Taxon							GOLD		
	Serial	Feeding			Beck	EPTD	Urban	Index	Non-	Excl.
Family	Number	Guild	Habit	Tolerance	Class	Таха	Intolerant	Таха	Insect	Таха
ARACHNIDA	0082708	PR	CN	5	3				Х	
ARRENURIDAE	0082862	PR	CN	6	3				Х	
ASELLIDAE	0092657	CG	SP	8	3		Х		Х	
ASTACIDAE	0097324									Х
ATHERICIDAE	0130928	PR	SP,BU	2	2	D	Х	D		
ATURIDAE	0082973	PR	CN	3	2				Х	
BAETIDAE	0100755	CG	SW	5	3	Е	Х			
BAETISCIDAE	0101493	CG	SP	3	2	Е				
BELOSTOMATIDAE	0103683	PR	CB,SW	9	3					
BITHYNIIDAE	0070745	SC	CB	8	3			G	Х	
BIVALVIA	0079118	CF	BU	8	3				Х	
BLEPHARICERIDAE	0121227	SC	CN	0	1	D		D		
BRACHYCENTRIDAE	0116905	SH	CN	1	1	Т	Х			
BRACONIDAE	0152867									Х
BRANCHIOBDELLIDA	0069168	CG		5	3			0	Х	
BRANCHIOBDELLIDAE	0069169	CG	CN	6	3			0	Х	
BRANCHIURA	0089403	PR	BU	10	3				Х	
CAENIDAE	0101467	CG	SP	7	3	Е	Х			
CALAMOCERATIDAE	0116529	SH	SP	3	2	Т	х			
	0085258	CG	SW		_				х	
CALOPTERYGIDAE	0102043	PR	CB	5	3					
	0097336	CG	SP	5	3		X		х	
	0069209	PR	BU	6	3		~	0	X	
	0102643	SH		2	2	P	x	0	X	
	0102045	DR	014,01	2	2		Λ			x
	0103234	1 1		-	2	D		П		X
	0122075	DR	SD	6	2	<u>ס</u>	x	D		Λ
	0125886	PR	SP	7	3		Λ	D		
	0123880	6	BII	7	2	D	v			
	0127917			/ 1	1	D D		D		
	0103202	сп СП			2	Г	^			
	0114309				2				v	
	0083652			0	2					
	0083170			10	2				^ V	
	0142200	CG	БО	10	2			D	^	v
	0143299	חח	CD	0	2	U		D		Χ
	0102077			<u>ہ</u>	3					
	0109216	58	BU	5	3					
	0099237		SP,SK	/	3				V	
	0085257	CG	SW	8	3				X	
CORBICULIDAE	0081381	CF	BU	6	3				Х	
	0102026	РК	BO	3	2		X			
CORDULIIDAE	0102020	PR	SP	5	3		X			
CORIXIDAE	0103364	PR	SW	7	3					
COROPHIIDAE	0093584	CG	SP						Х	
CORYDALIDAE	0115023	PR	CN,CB	3	2		Х			
COSMOPTERIGIDAE	0693962	SH	BU							
COSMOPTERYGIDAE	0117915	SH	BU							
COSSIDAE	0117906									Х
CRAMBIDAE	0693963	SH	CB	5	3					
CRANGONYCTIDAE	0095080	CG	SP	5	3		Х		Х	
CULICIDAE	0125930	CG	SW	8	3	D		D		
CURCULIONIDAE	0114666	SH	CN	5	3					

	Taxon							GOLD		
	Serial	Feeding			Beck	EPTD	Urban	Index	Non-	Excl.
Family	Number	Guild	Habit	Tolerance	Class	Таха	Intolerant	Таха	Insect	Таха
CYNIPIDAE	0154031									Х
DAPHNIIDAE	0083872	CF	SW	8	3				Х	
DECAPODA	0095599	SH	SP	6	3				Х	
DIPSEUDOPSIDAE	0598168	CF	BU	5	3	Т				
DIPTERA	0118831	CG	SP	5	3	D		D		
DIXIDAE	0125809	CG	SW,CL	2	2	D		D		
DOLICHOPODIDAE	0136824	PR	SP,BU	4	2	D		D		
DORYLAIMIDAE	0062983	PR	SP						Х	
DREISSENIDAE	0081330	CF	BU	5	3				Х	
DRYOPIDAE	0113999	SC	CN,CB	5	3					
DUGESIIDAE	0054552	PR	SP	7	3				Х	
DYTISCIDAE	0111963	PR	SW	5	3					
ELMIDAE	0114093	CG	CN	5	3		Х			
EMPIDIDAE	0135830	PR	CN,SP,BU	6	3	D		D		
ENCHYTRAEIDAE	0068510	CG	BU	10	3			0	Х	
ENTOMOBRYIDAE	0099643	CG	SP	7	3					
EPHEMERELLIDAE	0101232	CG	CN	2	2	Е	Х			
EPHEMERIDAE	0101525	CG	BU	4	2	Е	Х			
EPHEMEROIDEA	0101508	CG	BU	3	2	Е				
EPHEMEROPTERA	0100502	CG	CN,SP,SW	3	2	Е				
EPHYDRIDAE	0146893	CG	BU	7	3	D		D		
ERIRHINIDAE	0616734									Х
ERPOBDELLIDAE	0069438	PR	SP	9	3				Х	
EUHOLOGNATHA	0609847	SH	SP	2	2	Р				
GAMMARIDAE	0093745	SH	SP	5	3				Х	
GASTROPODA	0069459	SC	СВ	6	3			G	Х	
GELASTOCORIDAE	0103768	PR	SP	7	3					
GERRIDAE	0103801	PR	SK	7	3					
GERROIDEA	0676832	PR	SK	7	3					
GLOSSIPHONIIDAE	0069357	PR	SP	8	3				Х	
GLOSSOSCOLECIDAE	0069080	CG	BU	10	3			0	Х	
GLOSSOSOMATIDAE	0117120	SC	CN	1	1	Т	Х			
GOERIDAE	0115934	SC	CN	2	2	Т				
GOMPHIDAE	0101664	PR	BU	3	2		Х			
GORDIIDAE	0064227	SC	SP	4	2				Х	
GORDIOIDEA	0699878	PR	BU	4	2				Х	
GYRINIDAE	0112653	PR	SW	5	3					
HAEMOPIDAE	0568839	PR	SP	8	3				Х	
HALIPLIDAE	0111857	SH	СВ	5	3					
HAPLOTAXIDA	0068498	CG	BU	10	3			0	Х	
HAPLOTAXIDAE	0068504	CG	BU	8	3			0	Х	
HEBRIDAE	0103964	PR	СВ	6	3					
HELICOPSYCHIDAE	0117015	SC	CN	3	2	Т				
HELODIDAE	0113923	CG	CB,SP	6	3					
HELOPHORIDAE	0193642	PR	SW	5	3					
HEMERODROMIINI	0136304	PR	SP	7.6	3	D		D		
HEMIPTERA	0103359	PR	СВ	7	3					
HEPTAGENIIDAE	0100504	SC	CN	3	2	Е	Х			
HEXAMITIDAE	0553107				0					Х
HIRUDINEA	0069290	PR	SP	8	3				Х	
HIRUDINIDAE	0069407	PR	SP	8	3				Х	
HYALELLIDAE	0094022	SH	SP	7	3				Х	

	Taxon							GOLD		
	Serial	Feeding			Beck	EPTD	Urban	Index	Non-	Excl.
Family	Number	Guild	Habit	Tolerance	Class	Таха	Intolerant	Таха	Insect	Таха
HYDRACARINA	BAY0431	PR	SP	5.5	3				Х	
HYDRACHNIDAE	0083122	PR	CN	5	3				Х	
HYDRAENIDAE	0112756	PR	CN	6	3					
HYDRIDAE	0050844	PR	CN	4	2				Х	
HYDROBIIDAE	0070493	SC	СВ	8	3			G	Х	
HYDROCHIDAE	0722226	SH	SW	5	3					
HYDRODROMIDAE	0083224	PR	CN	5	3				Х	
HYDROIDA	0048740	PR	CN						Х	Х
HYDROMETRIDAE	0103938	PR	SK	7	3					
HYDROPHILIDAE	0112811	PR	SW	5	3					
HYDROPSYCHIDAE	0115398	CF	CN	5	3	Т	Х			
HYDROPTILIDAE	0115629	SC	CN	4	2	Т	Х			
HYDROZETIDAE	0553091	SC							Х	
HYDRYPHANTIDAE	0083212	PR	CN	5	3				Х	
HYGROBATIDAE	0083281	PR	CN	7	3				X	
HYMENOPTERA	0152741		0.1		0					х
	0099917	ſG	SP	7	3					Λ
	0092564		51	,	5				x	x
	0609506	CE	SW/ CN	2	2	F	x		Χ	Λ
	0003300			2	2	L	Λ		v	
	BAY0422			0	2				~	
	000024F	60		7	2					
	0099245		SP,SK	/ C	3				v	
	0083499	PK	CN	6	3				X	v
	0113835				2					X
	0083033	PR	CN	/	3				X	
	011/232	SH	CB	6	3	_				
LEPIDOSTOMATIDAE	0116/93	SH	CB,SP,CN	1	1		Х			
LEPTOCERIDAE	0116547	CG	CB,SP,CN	4	2					
LEPTOHYPHIDAE	0568545	CG	CN,SP	5	3	E				
LEPTOPHLEBIIDAE	0101095	CG	SW,CL	3	2	E	Х			
LESTIDAE	0102058	PR	CB,SW	9	3					
LEUCTRIDAE	0102840	SH	SP,CN	0	1	Р	Х			
LIBELLULIDAE	0101797	PR	SP	9	3		Х			
LIMNEPHILIDAE	0115933	SH	CB,SP,CN	4	2	Т				
LIMNESIIDAE	0083050	PR	CN	6	3				Х	
LIMNICHIDAE	0114030			5	3					Х
LIMNOCHARIDAE	0083145	PR	CN						Х	
LUMBRICIDAE	0069165	CG	BU	10	3			0	Х	
LUMBRICINA	0069069	CG	BU	10	3			0	Х	
LUMBRICULIDA	0068439	CG	BU	10	3			0	Х	
LUMBRICULIDAE	0068440	CG	BU	9	3			0	Х	
LYMNAEIDAE	0076483	SC	СВ	6	3			G	Х	
MACROMIIDAE	0102019	PR	SP	3	2					
MACROVELIIDAE	0103990	PR								Х
MARGARITIFERIDAE	0079914	CF	BU	5	3				Х	
MASENIIDAE	0056919	PR	SP	4	2				х	
MEGALOPTERA	0115000	PR	CN.CB	5	3					
MEGASCOLECIDAE	0069166		BU	3	5			0	х	х
MELITIDAE	0093746								~	x
MELYRIDAF	0113860	PR								X
	0102052	DR	SK	6	z					Λ
	0101079	(G	S\\/	2 2	2	F	Y			
	01010/0	20	J V V	2	2	L .	^			

	Taxon							GOLD		
	Serial	Feeding			Beck	EPTD	Urban	Index	Non-	Excl.
Family	Number	Guild	Habit	Tolerance	Class	Таха	Intolerant	Таха	Insect	Таха
MICROTURBELLARIA	0053965								Х	Х
MIDEOPSIDAE	0083476	PR	CN	6	3				Х	
MOLANNIDAE	0116473	SC	SP,CN	6	3	Т				
MOLLUSCA	0069458	CF	СВ	5	3				Х	
MUSCIDAE	0150025	PR	SP	7	3	D		D		
MUSCOMORPHA	0131750					D		D		
MYCETOPHILIDAE	0121636					D		D		Х
MYIDAE	0081688	CG	BU						Х	
MYMARIDAE	0153659									Х
NAIDIDAE	0068854	CG	BU	9	3			0	Х	
NAUCORIDAE	0103613	PR	CN,SW	5	3					
NEMATA	0563956			5	3				Х	Х
NEMATODA	0059490	CG	BU	9	3				Х	
NEMATOMORPHA	0064183	PR	BU	8	3				Х	
NEMERTEA	0057411	PR	CN	6	3				X	
NEMOURIDAE	0102517	SH	SP CN	2	2	Р	X			
	0101460	00 00	SP CN	3	2	F	Λ			
	0103747	PR	CR	7	2	-				
	0065870	0	BU	, 10	3				Y	
	0117219	с <u>о</u>		10	2				~	
	0117518			5	2					
	0112606		DU,CD	3	2					
	0103557	PR	SVV	7	3	D		D		
	0121292	SC		6	3	U		D		
	0103787	PR	CB	8	3					
	0101593	PR	CB	6	3	-				
ODONTOCERIDAE	0116496	SC	SP	0	1		Х	_		
OLIGOCHAETA	0068422	CG	BO	10	3			0	Х	
OLIGONEURIIDAE	0101029	CF	BU	2	2	E				
ONYCHIURIDAE	0099546	CG	SP	9	3					
ORTHOCLADIINAE	0128457	CG	CN	6	3	D		D		
OSTRACODA	0084195	CG	SW	8	3				Х	
OXIDAE	0083239	PR	CN	5	3				Х	
PALAEMONIDAE	0096213		SP	5	3				Х	
PANNOTA	0609505	CG	CN	4	2	Е				
PARNIDAE	0114000									
PELECORHYNCHIDAE	0130914	PR	SP,BU	5	3	D		D		
PELECYPODA	0079119	CF	BU	8	3				Х	
PELTOPERLIDAE	0102488	SH	CN,SP	2	2	Р	Х			
PERLIDAE	0102914	PR	CN	2	2	Р	Х			
PERLODIDAE	0102994	PR	CN,SP	2	2	Р	Х			
PETALURIDAE	0101659	PR	BU	4	2					
PHILOPOTAMIDAE	0115257	CF	CN	3	2	Т	Х			
PHORIDAE	0138921	CG	BU	6	3	D		D		
PHRYGANEIDAE	0115867	SH	CB	4	2	Т	Х			
PHYSIDAE	0076676	SC	CB	8	3			G	х	
PIONIDAE	0083330	PR	CN	6	3			-	X	
PISCICOLIDAE	0069296	PR	SP	7	3				X	
PISIDIIDAE	0081388	CF	BU	, 8	3				X	
	0054502	PR	SP	2 2	2				X	
	0034302	sr.	CB	7	2			G	Y	
	0053063	DR	SD	//	2			U	л У	
	0102467	сц	CN		2	P			Λ	
TELCOFTENA	0102407	211	CIN	2	2	r				

	Taxon							GOLD		
	Serial	Feeding			Beck	EPTD	Urban	Index	Non-	Excl.
Family	Number	Guild	Habit	Tolerance	Class	Таха	Intolerant	Таха	Insect	Таха
PLEUROCERIDAE	0071541	SC	CB	6	3			G	Х	
PODOCOPA	0084409	CG	SW	8	3				Х	
PODURIDAE	0099239	CG	SK	7	3					
POLYCENTROPODIDAE	0117043	PR	CN	6	3	Т	Х			
POLYDESMIDA	0154416								Х	Х
POLYMITARCYIDAE	0101569	CG	BU	2	2	Е	Х			
PORTUNIDAE	0098689	CG	SW	6	3				Х	
POTAMANTHIDAE	0101509	CF	BU	4	2	Е	Х			
PRODIAMESINAE	0128437					D		D		
PROSTIGMATA	0082770	PR	CN	6	3				Х	
PROTONEURIDAE	0102159	PR	CB	9	3					Х
PSEPHENIDAE	0114069	SC	CN	4	2		Х			
PSYCHODIDAE	0125351	CG	BU	10	3	D		D		
PSYCHOMYIIDAE	0115334	CG	CN	2	2	Т				
PTERONARCYIDAE	0102470	SH	CN,SP	1	1	Р	Х			
PTILODACTYLIDAE	0114265	SH	CN	4	2					
PTYCHOPTERIDAE	0125763	CG	BU	7	3	D		D		
PYRALIDAE	0117641	SH	СВ	5	3					
RHYACOPHILIDAE	0115096	PR	CN	1	1	Т	Х			
SABELLIDAE	0068076	CG	BU	10	3				Х	
SALDIDAE	0104063	PR	СВ	8	3					
SARCOPHAGIDAE	0151815	CG	BU			D		D		Х
SCATHOPHAGIDAE	0678096	SH	BU	7	3	D		D		
SCIARIDAE	0122702			6	3	D		D		Х
SCIOMYZIDAE	0144653	PR	BU	9	3	D		D		
SCIRTIDAE	0113924	SC	CB,SP	6	3					
SERICOSTOMATIDAE	0116982	SH	SP	2	2	Т	Х			
SIALIDAE	0115001	PR	BU.CN	4	2		Х			
SIMULIIDAE	0126640	CF	CN	6	3	D	Х	D		
SIPHLONURIDAE	0100951	CG	SW,CL	7	3	Е				
SISYRIDAE	0115085	PR	CB	3	2					
SMINTHURIDAE	0100258	CG	SK	7	3					
SPARGANOPHILIDAE	0068499	CG	BU	10	3			0	Х	
SPERCHONIDAE	0083005	PR	CN	8	3				Х	
SPERCHONTIDAE	BAY0452	PR	CL	8	3				Х	
SPHAERIIDAE	0081389	CF	BU	7.5	3				х	
SPHAERIUSIDAE	0112737	SC	CL							Х
SPHAERIUSIDAE	0678385	CF	BU	8	3					Х
SPHAEROMATIDAE	0092283									Х
SPIONIDAE	0066781		BU	10	3				Х	Х
SPONGILLIDAE	0047691	CF	-	3	2				Х	
STAPHYLINIDAE	0113265	PR	CN	8	3					
STRATIOMYIDAE	0130150	CG	SP	8	3	D	Х	D		
SYRPHIDAE	0139621	CG	BU	10	3	D		D		
TABANIDAE	0130934	PR	SP.BU	6	3	D	Х	D		
TAENIOPTERYGIDAE	0102788	SH	SP.CN	2	2	P	X	-		
TALITRIDAE	0095032	CG	SP	- 8	3				Х	
TANYDERIDAE	0125799	CG	SP	7	3	D		D		
TETRASTEMMATIDAF	0057556	PR	CN	. 7	3	-		-	Х	
THAUMALEIDAF	0126624	SC	CN	,	2	D		D		
TIPULIDAE	0118840	SH	BU	4	2	D	Х	 D		
TORRENTICOLIDAE	0083249	PR	CN	6	3				Х	

	Taxon							GOLD		
	Serial	Feeding			Beck	EPTD	Urban	Index	Non-	Excl.
Family	Number	Guild	Habit	Tolerance	Class	Таха	Intolerant	Таха	Insect	Таха
TORTRICIDAE	0117856	SH	BU,CB	5	3					
TRICHOCERIDAE	0125556					D		D		Х
TRICHOPTERA	0115095	SC	CN	4	2	Т				
TRICLADIDA	0054468	PR	SP	7	3				Х	
TRICORYTHIDAE	0101404	CG	SP	4	2	Е				
TROMBIDIFORMES	0082769	PR	CN	6	3				Х	
TUBIFICIDAE	0068585	CG	BU	10	3			0	Х	
TURBELLARIA	0053964	PR	SP	4	2				Х	
UENOIDAE	0568757	SC	CN	3	2	Т	Х			
UNIONICOLIDAE	0083072	PR	CN						Х	Х
UNIONIDAE	0079913	CF	BU	4	2				Х	
VALVATIDAE	0070345	SC	CB	4	2			G	Х	
VELIIDAE	0103885	PR	SK	6	3					
VENEROIDA	0080384	CF	BU	7	3				Х	
VIVIPARIDAE	0070304	SC	СВ	6	3		Х	G	х	
XANTHIDAE	0098748									Х
ZYGOPTERA	0102042	PR	СВ	9	3					

APPENDIX B

Season and Strahler Order Effects on Biological Metrics

The median values of the biological metrics are compared across season and across Strahler order using the Kruskal-Wallis test, a nonparametric test that determines if the medians of samples taken from two or more populations are statistically different. Only Reference (REF) condition samples were used in the analysis in order to remove or minimize the confounding influence of anthropogenic factors. Results for the 42 candidate metrics are shown in the following tables for the five bioregions outside the Coastal Plain. **, p < 0.01; *, $0.01 \le p \le 0.05$. The superscript notation (¹) indicates the metrics ultimately selected for the Chessie B-IBI in each bioregion. Grayed values indicate groups with n fewer than 10. *All metrics are calculated from family level counts*. Metric descriptions are given here:

Metric	Description
ASPT_MOD	Average tolerance score per taxon modified to family level
BECK_100	Becks Index, sample randomly standardized to a total count of 100
EPHEMEROPTERA_TAXA_CNT_100	Ephemeroptera count, sample randomly standardized to a total count of 100
EPT_TAXA_COUNT_NO_TOL_100	EPT families excluding tolerants, sample randomly standardized to total count of 100
FBI	Family level Hilsenhoff Biotic Index
GOLD	1 minus proportional abundance of gastropods, oligochaetes and Diptera individuals
LOG10_SEL_EPTD	Log 10 of selected Ephemeroptera, Plecoptera, Trichoptera and Diptera taxa
MARGALEFS	Margalef's Index
PCT_AMPHIPOD	Percent of individuals that are amphipods
PCT_CHIRONOMID	Percent of individuals that are chironomids
PCT_CLIMB	Percent of individuals that are climbers
PCT_CLING	Percent of individuals that are clingers
PCT_CLINGER_TAXA	Percent of family level taxa that are clingers
PCT_COLLECT	Percent of individuals that are collectors
PCT_CORBICULA	Percent of individuals that are Corbicula clams
PCT_DIPTERA	Percent of individuals that are Diptera
PCT_DOM1	Percent of individuals that belong to the most common family level taxa
PCT_DOM2	Percent of individuals that belong to the two most common family level taxa
PCT_DOM3	Percent of individuals that belong to the three most common family level taxa
PCT_EPHEMEROPTERA	Percent of individuals that are Ephemeroptera
PCT_EPT	Percent of individuals that are Ephemeroptera, Plecoptera and Trichoptera (EPT)
PCT_EPT_TAXA_RICH	Percent of family level taxa that are EPT
PCT_FILTERERS	Percent of individuals that are filterers
PCT_GATHER	Percent of individuals that are gatherers
PCT_LIMESTONE	Percent of individuals that are isopods, amphipods, and Ephemeralla
PCT_NET_CADDISFLY	Percent net-spinning caddisflies
PCT_NON_INSECT	Percent taxa that are not insects
PCT_OLIGOCHAETA	Percent Oligochaeta
PCT_PLECOPTERA	Percent Plecoptera taxa
PCT_SCRAPER	Percent scraper
PCT_SENSITIVE	Percent sensitive taxa (family level tolerance value < 3)
PCT_SHREDDER	Percent shredders
PCT_SWIMMER	Percent swimmers
PCT_TOLERANT	Percent tolerant taxa (family level tolerance value \geq 7)
PCT_TRICHOPTERA	Percent Trichoptera taxa
PCT_TRICHOPTERA_NO_TOL	Percent Trichoptera taxa (excluding Hydropsychidae)
PCT_URBAN_INTOL	Percent of urban intolerants
SENSITIVE_TAXA_COUNT_100	Sensitive family-level taxa count, sample randomly standardized to total count of 100
SIMPSON_DIVERSITY	Simpson Diversity index
SW	Shannon Wiener Index
TAXA_RICH_100	Taxa richness, sample randomly standardized to total count of 100
TOLERANT_TAXA_COUNT_100	Tolerant family-level taxa count, sample randomly standardized to total count of 100

PIEDMONT (REF n = 222)

	Season				Strahler Order						
Parameter	Spr	Sum	Fall/Win	р	1st	2nd	3rd	4th	р		
Count	90	62	70		61	93	25	43			
ASPT_MOD	4.1	4.1	3.95		3.5	4.1	4.5	4.3	**		
BECK_100	9	8	9		11	8	7	9	**		
EPHEMEROPTERA_TAXA_CNT_100	3	3	3	*	3	3	3	3			
EPT_TAXA_COUNT_NO_TOL_100	7	6.5	6		8	6	5	6	**		
FBI ¹	3.39	3.59	3.48		3.03	3.46	3.82	3.54	**		
GOLD	0.898	0.941	0.952	**	0.916	0.949	0.909	0.941			
LOG10_SEL_EPTD	1.332	1.216	1.312	*	1.301	1.342	1.362	1.255			
MARGALEFS	3.177	2.906	2.840		3.083	2.826	3.238	3.267	*		
PCT_AMPHIPOD	0	0	0		0.0	0.0	0.0	0.0			
PCT_CHIRONOMID	2.0	1.0	1.0	*	0.9	1.7	2.8	1.9	*		
PCT_CLIMB	8.5	11.2	11.2		5.3	9.5	19.4	14.0	**		
PCT_CLING	82.2	86.5	86.2		88.8	87.5	79.4	81.6	**		
PCT_CLINGER_TAXA	60.2	57.7	62.8		71.4	62.5	50.0	52.9	**		
PCT_COLLECT ¹	56.6	49.6	50.8		40.2	56.6	49.1	56.2	**		
PCT_CORBICULA	0.0	0.0	0.0	*	0.0	0.0	2.1	1.7	**		
PCT_DIPTERA ¹	7.1	2.4	3.6	**	4.8	4.3	4.7	3.1			
PCT_DOM1	26.9	26.1	26.3		25.0	30.3	23.8	23.5	**		
PCT_DOM2	45.5	45.8	46.2		44.0	49.2	45.5	41.1	**		
PCT_DOM3	59.9	60.1	60.9		56.4	63.4	60.0	54.8	*		
PCT_EPHEMEROPTERA	41.2	31.3	33.5	**	28.7	37.2	36.6	39.1	*		
PCT_EPT ¹	71.1	73.1	74.5		80.4	74.8	52.0	65.9	**		
PCT_EPT_TAXA_RICH	43.8	45.5	43.8		57.1	45.5	30.8	40.0	**		
PCT_FILTERERS	20.5	33.5	37.5	**	20.7	33.6	29.1	32.4	**		
PCT_GATHER	28.9	12.5	10.9	**	15.3	15.4	20.2	18.5			
PCT_LIMESTONE	13.3	2.0	2.4	**	4.2	3.5	5.7	4.8			
PCT_NET_CADDISFLY	7.6	20.7	25.2	**	10.7	20.9	12.1	15.5	**		
PCT_NON_INSECT	2.8	2.1	2.0		1.0	1.9	10.7	3.9	**		
PCT_OLIGOCHAETA	0.0	0.0	0.0	**	0.0	0.0	0.0	0.0	**		
PCT_PLECOPTERA	7.7	9.4	6.8		25.0	6.0	0.9	6.7	**		
PCT_SCRAPER	19.2	17.3	20.8		15.6	20.6	23.6	17.1	**		
PCT_SENSITIVE	65.3	60.1	64.5		75.0	65.5	53.7	62.0	**		
PCT_SHREDDER	2.2	1.8	3.1	*	13.7	1.5	1.9	2.6	**		
PCT_SWIMMER	7.4	12.5	12.4	*	7.3	11.7	6.9	17.0	**		
PCT_TOLERANT	3.6	3.1	2.6		1.9	3.2	6.3	4.7	**		
PCT_TRICHOPTERA	10.0	21.2	26.0	**	16.7	21.6	12.1	15.5	*		
PCT_TRICHOPTERA_NO_TOL	2.8	4.2	7.1	**	8.2	5.8	0.0	2.8	**		
PCT_URBAN_INTOL	76.7	75.0	81.5		90.0	81.3	64.7	70.0	**		
SENSITIVE_TAXA_COUNT_100	7	7	7		9	7	6	7	**		
SIMPSON_DIVERSITY	0.853	0.847	0.843		0.865	0.832	0.850	0.864	**		
SW ¹	2.268	2.157	2.164		2.302	2.104	2.185	2.321	**		
TAXA_RICH_100	15	14	14		15	14	16	16	*		
TOLERANT_TAXA_COUNT_100	2	2	1		1	2	3	2	**		

RIDGE	S
(REF n = 1	154)

		Season Strahler Order							
Parameter	Spr	Sum	Fall/Win	р	1st	2nd	3rd	4th	р
count	59	55	40		75	60	12	7	
ASPT_MOD	3.5	3.6	3.3		3.4	3.6	3.35	4.2	
BECKS_100 ¹	12.5	12	14		14	12	11.5	8	
EPHEMEROPTERA_TAXA_CNT_100	4	3	4		4	4	4	3	
EPT TAXA COUNT NO TOL_100	10	10	11		11	10	10.5	7.5	
FBI	3.80	3.79	3.39	**	3.56	3.65	3.84	4.36	
GOLD	0.779	0.836	0.867	**	0.813	0.849	0.813	0.689	
LOG10_SEL_EPTD	1.398	1.342	1.415		1.362	1.415	1.567	1.301	
MARGALEFS	3.349	3.161	3.498		3.472	3.179	3.341	2.659	*
PCT_AMPHIPOD	0	0	0	**	0.0	0.0	0.0	0.0	
PCT_CHIRONOMID	12.1	10.4	5.4		9.9	8.1	13.6	4.5	
PCT_CLIMB	0.9	0.9	2.3	**	1.0	1.4	0.7	0.9	
PCT_CLING	64.2	64.4	77.3		72.7	68.3	69.7	79.4	
PCT_CLINGER_TAXA	71.4	68.4	73.0		69.6	71.2	76.3	73.3	
PCT_COLLECT	60.4	56.6	47.5	*	52.4	56.4	72.0	75.3	*
PCT_CORBICULA	0.0	0.0	0.0		0.0	0.0	0.0	0.0	*
PCT_DIPTERA	18.6	15.7	10.0	**	16.7	14.2	17.5	7.8	
PCT_DOM1	27.1	26.9	23.4	*	23.6	28.5	26.3	26.4	
PCT_DOM2	47.9	45.9	39.8	*	43.1	48.2	45.8	48.5	
PCT_DOM3	58.8	59.8	52.5	*	53.6	60.2	58.2	63.2	*
PCT_EPHEMEROPTERA ¹	33.6	19.4	28.5	**	26.4	29.1	37.2	41.2	
PCT_EPT	70.5	71.8	76.8		72.1	72.0	74.8	60.7	
PCT_EPT_TAXA_RICH	66.7	65.0	66.7		66.7	63.9	71.7	56.3	*
PCT_FILTERERS	13.9	17.4	22.6	**	17.3	16.8	16.1	23.3	
PCT_GATHER	44.0	30.8	27.2	**	32.4	35.7	50.1	50.0	*
PCT_LIMESTONE	10.1	3.4	2.9	**	5.1	7.7	4.8	4.9	
PCT_NET_CADDISFLY	8.8	16.0	15.2	**	14.2	9.8	10.1	14.3	
PCT_NON_INSECT	0.9	1.7	0.5	*	1.7	0.9	0.2	2.1	*
PCT_OLIGOCHAETA	0.0	0.0	0.0		0.0	0.0	0.0	0.0	
PCT_PLECOPTERA	17.5	20.6	20.2		22.6	16.2	17.4	5.5	*
PCT_SCRAPER ¹	10.8	8.8	15.0	*	11.6	10.2	7.4	18.1	
PCT_SENSITIVE	56.2	51.6	64.3	**	56.5	55.9	54.4	45.1	
PCT_SHREDDER	13.2	21.1	18.7		20.8	13.2	11.5	2.7	*
PCT_SWIMMER ¹	11.3	10.4	10.3		10.1	10.9	15.7	16.6	
PCT_TOLERANT	12.5	11.3	5.6	**	10.9	9.1	13.4	5.8	
PCT_TRICHOPTERA	11.0	20.6	22.8	**	17.8	14.9	13.2	14.8	
PCT_TRICHOPTERA_NO_TOL	5.1	7.1	9.3	*	7.7	6.0	6.2	3.6	
PCT_URBAN_INTOL	89.5	88.9	93.9	**	90.9	89.5	90.2	80.0	
SENSITIVE_TAXA_COUNT_100	9	9	10		9	8.75	8	6.5	
SIMPSON_DIVERSITY	0.853	0.847	0.872		0.873	0.842	0.858	0.824	*
SW ¹	2.224	2.250	2.313		2.321	2.208	2.260	2.003	*
TAXA_RICH_100	16	16	16.5		16	16	15.5	13	
TOLERANT_TAXA_COUNT_100	1	1	1		1	1	1	1	

VAI		.E	YS	
(REF	n	=	57)	

		Season			Strahler Order				
Parameter	Spr	Sum	Fall/Win	р	1st	2nd	3rd	4th	р
Count	22	20	15		4	19	16	18	
ASPT_MOD	4.3	4	4.1		3.65	3.6	4	4.4	**
BECKS_100 ¹	9	10	7		12.5	12	9	7	
EPHEMEROPTERA_TAXA_CNT_100	4	4	3.5		4	4	4	4	
EPT TAXA COUNT NO TOL_100	8	8.5	7		10.5	9	8	6.25	
FBI	4.47	3.97	3.89		3.85	4.02	3.935	4.42	
GOLD	0.664	0.855	0.899	**	0.693	0.857	0.807	0.819	
LOG10_SEL_EPTD	1.255	1.267	1.301		1.438	1.342	1.266	1.216	
MARGALEFS	3.267	3.270	2.745		3.399	3.182	3.033	2.903	
PCT_AMPHIPOD	0	0	0		0.3	0.0	0.0	0.0	**
PCT_CHIRONOMID	13.0	8.0	3.9	**	22.3	5.8	9.3	5.2	
PCT_CLIMB	2.4	1.1	6.3		0.8	0.9	2.9	6.8	**
PCT_CLING	62.0	65.1	87.4	**	58.6	66.9	77.7	71.0	
PCT_CLINGER_TAXA	66.7	62.2	74.2		59.7	66.7	71.0	63.3	
PCT_COLLECT	73.1	62.7	68.5		60.9	68.0	72.2	65.0	
PCT_CORBICULA	0.0	0.0	0.0		0.0	0.0	0.0	0.0	**
PCT_DIPTERA	25.3	10.9	4.9	**	30.7	10.2	16.5	8.8	
PCT_DOM1	32.9	29.7	29.5		24.8	32.0	30.1	31.3	
PCT_DOM2	50.8	48.9	57.0		43.5	55.5	50.1	50.3	
PCT_DOM3	63.7	65.0	65.8		64.0	69.4	61.9	64.6	
PCT_EPHEMEROPTERA ¹	24.4	27.2	38.0		19.3	25.3	31.2	33.0	
PCT_EPT	50.8	66.3	71.6	**	64.7	68.3	64.1	59.0	
PCT_EPT_TAXA_RICH ¹	56.7	53.6	58.3		57.3	60.0	55.2	50.8	*
PCT_FILTERERS	12.4	27.4	32.0	**	16.8	23.2	21.7	25.4	
PCT_GATHER	53.5	34.8	25.5	**	40.0	35.9	45.7	35.5	
PCT_LIMESTONE	10.4	2.5	4.4	**	1.4	5.0	8.0	4.1	
PCT_NET_CADDISFLY	3.7	20.1	18.2	**	12.8	12.4	7.2	13.4	
PCT_NON_INSECT	5.0	1.2	2.9		1.1	1.4	2.8	9.7	
PCT_OLIGOCHAETA	1.0	0.0	0.0	**	0.0	0.0	0.4	0.0	
PCT_PLECOPTERA	3.5	8.0	0.9		25.6	5.5	4.9	1.6	**
PCT_SCRAPER ¹	12.4	10.2	24.9	*	4.4	12.8	11.9	23.5	*
PCT_SENSITIVE	35.8	47.3	47.4		51.0	45.7	50.0	35.8	
PCT_SHREDDER	3.4	6.0	2.7		21.5	4.3	6.7	0.3	**
PCT_SWIMMER	8.5	16.9	6.1	*	14.4	11.2	9.8	9.4	
PCT_TOLERANT	16.6	8.8	5.2	**	19.0	5.8	10.1	8.6	
PCT_TRICHOPTERA	7.5	21.6	28.7	**	14.3	18.3	10.1	16.2	
PCT_TRICHOPTERA_NO_TOL	3.9	3.6	9.1		3.7	4.5	5.1	4.1	
PCT_URBAN_INTOL	82.8	85.7	83.3		89.3	90.0	83.3	76.6	**
SENSITIVE_TAXA_COUNT_100	6	7	6		9	8	6	6	
SIMPSON_DIVERSITY	0.814	0.837	0.805		0.856	0.806	0.832	0.820	
SW ¹	2.131	2.162	1.988		2.252	2.064	2.164	2.016	
TAXA_RICH_100	15.5	15	13		17	15	14.5	15	
TOLERANT_TAXA_COUNT_100	2	2	2		1	2	2	2.5	

NORTH CENTRAL APPALACHIANS

(REF n = 26)

		Season							
Parameter	Spr	Sum	Fall/Win	р	1st	2nd	3rd	4th	р
Count	8	7	11		4	4	8	10	
ASPT_MOD	3.65	3.6	3.1	**	3.55	3.25	3.15	3.65	
BECK_100	13.5	12	15		15	13.5	15.5	12.5	
EPHEMEROPTERA_TAXA_CNT_100	4	3	4		4	4	4	3	
EPT_TAXA_COUNT_NO_TOL_100 ¹	10.5	9	14		13	11	12	10	
FBI	3.82	4.68	3.29	*	4.03	3.69	3.32	3.82	
GOLD	0.717	0.671	0.904	**	0.638	0.751	0.917	0.836	
LOG10_SEL_EPTD	1.632	1.398	1.863		1.700	1.706	1.851	1.585	
MARGALEFS	3.520	2.947	3.853		3.975	3.511	3.873	3.419	
PCT_AMPHIPOD	0	0	0		0.0	0.0	0.0	0.0	
PCT_CHIRONOMID	11.5	31.1	6.4	*	32.3	11.5	5.6	9.7	
PCT_CLIMB	0.5	1.8	2.0		1.5	0.4	1.5	1.3	
PCT_CLING	73.7	53.3	77.6	*	57.2	66.2	70.0	77.9	
PCT_CLINGER_TAXA	71.1	70.6	73.9		69.1	71.8	74.5	72.9	
PCT_COLLECT	58.4	74.7	53.3	*	56.3	62.1	65.7	66.4	
PCT_CORBICULA	0.0	0.0	0.0		0.0	0.0	0.0	0.0	
PCT_DIPTERA	28.3	32.9	9.6	**	36.2	24.9	8.3	16.2	
PCT_DOM1	22.0	31.1	20.1	*	31.3	21.9	20.6	28.1	
PCT_DOM2	40.3	54.4	36.3	*	45.0	37.8	36.9	51.4	
PCT_DOM3	52.9	64.5	50.9	*	53.6	49.0	51.2	63.6	
PCT_EPHEMEROPTERA ¹	38.5	26.1	46.7		24.9	39.6	46.2	30.2	*
PCT_EPT	60.8	56.1	82.4	**	59.1	69.1	83.2	60.8	
PCT_EPT_TAXA_RICH	63.5	58.8	70.0	**	64.0	65.8	68.0	61.9	
PCT_FILTERERS	18.5	18.6	17.3		3.7	9.6	18.8	20.6	**
PCT_GATHER	44.1	60.7	37.2	**	52.3	46.5	45.3	40.8	
PCT_LIMESTONE	9.2	1.0	12.6		7.7	9.5	15.7	7.3	
PCT_NET_CADDISFLY	8.1	17.4	12.0		3.8	6.5	13.5	17.3	*
PCT_NON_INSECT	0.0	0.0	0.0		0.2	0.0	0.0	0.0	
PCT_OLIGOCHAETA	0.0	0.0	0.0		0.0	0.0	0.0	0.0	
PCT_PLECOPTERA	12.3	8.6	16.5		27.9	19.3	15.7	8.7	
PCT_SCRAPER ¹	9.9	9.6	15.4	*	7.8	14.5	12.6	11.8	
PCT_SENSITIVE	46.2	31.9	69.0	*	49.9	58.0	67.5	44.9	
PCT_SHREDDER	12.6	6.7	10.2		26.2	11.8	8.3	5.1	*
PCT_SWIMMER	11.5	12.3	16.0		7.4	16.9	21.8	9.5	*
PCT_TOLERANT	12.0	31.1	6.4	*	30.3	11.5	6.0	9.7	
PCT_TRICHOPTERA	9.1	17.4	16.8		7.0	8.4	17.3	20.9	*
PCT_TRICHOPTERA_NO_TOL	3.2	5.5	5.5		5.7	4.1	5.2	5.1	
PCT_URBAN_INTOL	90.6	90.9	95.8		82.7	92.5	94.9	95.1	*
SENSITIVE_TAXA_COUNT_100	9	8	12		10.5	10	11	9	
SIMPSON_DIVERSITY	0.879	0.803	0.881	*	0.840	0.883	0.882	0.843	
SW ¹	2.426	2.082	2.467	*	2.383	2.443	2.475	2.235	
TAXA_RICH_100 ¹	17	15	18		18	16	18.5	16	
TOLERANT_TAXA_COUNT_100	1.5	1	1		1.5	1	1	1	

NORTHERN APPALACHIAN PLATEAU & UPLANDS

(REF n = 28)

		Season			Strahler Order				
Parameter	Spr	Sum	Fall/Win	р	1st	2nd	3rd	4th	р
count	5	22	1		3	7	13	5	
ASPT_MOD	3.9	4.4	3.8		3.9	4.4	4.2	3.7	
BECKS_100	14	8	12		14	8	9	10	
EPHEMEROPTERA_TAXA_CNT_100	4	4	4		5	4	3	5	
EPT_TAXA_COUNT_NO_TOL_100	10	7	10		10	7	7	8	
FBI ¹	4.49	4.60	4.06		4.42	4.52	4.89	4.58	
GOLD	0.738	0.774	0.778		0.738	0.791	0.768	0.790	
LOG10_SEL_EPTD	1.644	1.109	1.230	*	1.613	1.255	0.954	1.204	
MARGALEFS	3.551	2.903	3.482		4.084	3.787	2.881	2.908	*
PCT_AMPHIPOD	0	0	0		0.0	0.0	0.0	0.0	
PCT_CHIRONOMID	20.0	17.1	12.1		22.6	17.2	17.0	17.0	
PCT_CLIMB	0.3	0.0	0.0	*	1.1	0.3	0.0	0.0	**
PCT_CLING	57.0	65.5	53.5		52.4	56.0	67.0	76.2	
PCT_CLINGER_TAXA	66.7	59.4	64.7		68.2	57.9	60.0	70.6	
PCT_COLLECT	56.8	77.9	83.8		56.8	75.5	83.0	83.9	
PCT_CORBICULA	0.0	0.0	0.0		0.0	0.0	0.0	0.0	
PCT_DIPTERA	25.8	21.5	19.2		25.8	20.6	22.0	21.0	
PCT_DOM1	38.0	26.6	20.2		20.4	28.0	29.9	24.0	
PCT_DOM2	62.1	46.9	34.3		36.0	51.0	48.8	47.0	
PCT_DOM3	72.9	60.5	60.6		47.7	62.3	61.2	59.0	
PCT_EPHEMEROPTERA	30.5	23.0	48.5		30.5	33.3	16.0	20.3	
PCT_EPT	68.1	54.9	64.7		68.1	55.9	43.0	66.0	
PCT_EPT_TAXA_RICH	62.5	47.9	58.8	*	59.1	46.7	52.9	52.9	
PCT_FILTERERS	8.1	28.9	8.1		8.1	19.2	22.0	51.0	
PCT_GATHER ¹	48.7	47.5	75.8		48.7	50.0	53.6	43.0	
PCT_LIMESTONE	3.0	0.3	20.2		0.3	0.7	0.0	2.9	
PCT_NET_CADDISFLY	1.0	20.0	6.1	**	1.1	16.7	16.0	23.0	
PCT_NON_INSECT	0.4	1.5	3.0		0.4	1.0	2.0	0.0	
PCT_OLIGOCHAETA	0.0	0.0	3.0		0.0	0.0	1.0	0.0	
PCT_PLECOPTERA ¹	17.4	2.0	9.1	*	33.7	2.0	7.0	2.9	
PCT_SCRAPER	3.1	4.5	1.0		6.1	14.7	2.0	4.9	
PCT_SENSITIVE	48.7	30.0	45.5		48.7	33.0	21.0	30.0	
PCT_SHREDDER	17.4	2.9	8.1		35.4	3.0	4.7	1.5	**
PCT_SWIMMER	18.0	14.3	27.3		19.0	20.0	14.0	8.1	
PCT_TOLERANT	20.8	17.0	14.1		24.0	19.3	17.0	16.0	
PCT_TRICHOPTERA	2.9	21.2	7.1	**	3.9	19.2	17.0	23.0	
PCT_TRICHOPTERA_NO_HYDRO ¹	2.7	5.6	1.0		2.9	10.6	3.1	8.1	
PCT_URBAN_INTOL	76.2	85.2	88.2		70.8	85.7	84.6	93.8	
SENSITIVE_TAXA_COUNT_100	10	5	9		10	6	5	7	
SIMPSON_DIVERSITY	0.772	0.850	0.883		0.892	0.855	0.837	0.848	
SW	1.964	2.167	2.381		2.536	2.196	2.109	2.198	
TAXA_RICH_100 ¹	15	15	17		19	15	14	15	
TOLERANT_TAXA_COUNT_100	2	2	2		2	2	3	1	

APPENDIX C

Number of Family-Level Taxa in the Piedmont, Ridges, and Valleys Bioregions

The statistical method "rarefaction" was used to investigate how family-level taxa richness changes as the *total* number of counted individuals increases in the Ridges, Valleys and Piedmont bioregions. Sample counts of family-level taxa were pooled by bioregion and site type (Reference, Degraded) and each pool analyzed separately. The software program EstimateS 8.2.0 (Colwell 2009) was used to generate rarefaction - also called taxa accumulation - curves. Using the "sampling without replacement" option, EstimateS does the following:

"selects a single sample at random, computes the richness estimators (and diversity indexes, if requested) based on that sample, selects a second sample, re-computes the estimators using the pooled data from both samples, selects a third, re-computes, and so on until all samples in the matrix are included. Samples are added to the analysis in random order, without replacement (each sample is selected exactly once). Each distinct randomization accumulates the samples in a different order, but all samples are included in each randomization. The final value for the averaged, random-order species accumulation curve therefore matches, precisely, the total number of observed species."

In our analysis, the rarefaction curves were based on 100 randomizations. The curves developed for each bioregion-site type group are based on all of the family-level taxa actually identified in that group, or the known <u>pool</u> of taxa.

The rarefaction results show that family-level taxa accumulation curves in the Piedmont, Ridges, and Valleys bioregions⁶ are not identical, and the Reference (REF) and Degraded (DEG) curves are not positioned consistently with respect to each other (**Figure C-1**). While taxa richness in *individual samples* is generally higher at Reference sites, the overall number of family-level taxa found in all of the Degraded sites appears to be higher than those found in all of the Reference sites for a given sampling intensity (total number of individuals counted) in the Piedmont and Valleys. The trajectories of all of the accumulation curves seem to be converging on an asymptote somewhere between 100 and 120 family-level taxa, which suggests there are at least 100-120 family-level taxonomic groups per site type (Reference, Degraded) in each of the three bioregions. Considering there have been approximately 250 family-level taxa identified so far in all the bioregions (**Appendix A**), there is likely significant overlap in the taxa found in Reference and Degraded conditions.

At the 100-count level on the taxa accumulation curves, the number of taxa expected in the pooled samples is still increasing sharply in both Reference and Degraded samples, indicating the likelihood of finding rarer family-level taxa is still high with additional sample counts. The slopes of the accumulation curves begin to change at approximately 5,000-count in the Piedmont and Ridges bioregions and at a lower cumulative count in the Valleys. This result suggests cumulative counts of more than 5,000 are needed to find the rarest families, and a 200- or higher count protocol will be substantially more effective at finding rare or endangered taxa than a 100-count protocol.

⁶ Data from Reference and Degraded sites in the NCA and NAPU bioregions were insufficient to generate taxa accumulation curves.



Figure C-1. Rarefaction (taxa accumulation) curves for the Piedmont, Ridges and Valleys bioregions, developed with the EstimateS software package (Colwell 2009).

APPENDIX D

Discrimination Efficiencies of the 42 Candidate Metrics

The ability of each metric to correctly identify Reference or Degraded stream conditions is expressed as its Discrimination Efficiency (DE). A middle (M) percentile and a tail (T) percentile of the distribution of a metric's values in its Reference data pool are used to classify all of that metric's values into three categories: "furthest from the Degraded condition," "somewhat removed from the Degraded condition," and "most like the Degraded condition." These three categories are assigned the numeric values of 5, 3, and 1, respectively. For the purpose of calculating a metric's DEs, the percentage of samples in the Reference data pool scoring 3 or 5 is the metric's DE for Reference sites and the percentage of samples in the Degraded data pool scoring 1 is the metric's DE for Degraded sites. The average of these two DEs is the metric's overall DE.

The T percentiles can be selected to favor the correct identification of Reference sites or Degraded sites in each bioregion, or they can be selected to correctly identify Reference and Degraded equally well. The latter approach was used in developing the Chessie B-IBI because the intent of the index is to represent the status of all sites. After adjustment of the T %ile brought a metric's Reference and Degraded DEs to within 10% of each other, the T %ile was further adjusted (in 5%ile increments) to make the average of the Reference and Degraded DEs as high as possible. Strong discrimination is typically found when the T percentile is between the 10th and 45th%ile of the Reference distribution for metrics decreasing in value with degradation and between the 70th and 95th%ile for metrics increasing with degradation. The best performing metrics are considered those with DEs greater than 70% for both Reference and Degraded indicating more than 70% of both the Reference and Degraded data pools are correctly classified when the T percentile is applied.

The metric values associated with M and T percentiles of the Reference data pool, the Reference (REF) and Degraded (DEG) DEs, and the average or overall DEs are given in the following tables. Also included are the expected direction of change with degradation (POS, increase in metric value with degradation; NEG, decrease in the metric value with degradation) and the metric type (T, tolerance; R, richness; H, habit; FG, feeding guild; C, composition). Metrics that have been standardized to a 100-count sample have the suffix "100." The notation (a) indicates individual DEs for Reference and Degraded could not be brought to within 10% of each other because of the particular characteristics of the metric's Reference and Degraded distributions, and (b) indicates too many metric values were zero (0) and so the calculated percentiles were not useful.

			M %ile	T %ile			
Direction	Туре	Metric	Value	Value	DEG DE	REF DE	DE
POS	Т	ASPT_MOD	4.2	4.6	78.9%	85.6%	82.3%
NEG	Т	BECKS_100	9	7	76.3%	79.7%	78.0%
NEG	R	EPHEMEROPTERA_TAXA_CNT_100	3	3	71.1%	77.0%	74.0%
NEG	R,T	EPT_TAXA_COUNT_NO_TOL_100	7	6	67.1%	70.3%	68.7%
POS	Т	FBI	3.63	4.54	89.5%	90.5%	90.0%
NEG	С	GOLD	0.919	0.804	90.8%	84.7%	87.7%
NEG	Т	LOG10_SEL_EPTD	1.301	1.088	75.0%	74.8%	74.9%
NEG	R	MARGALEFS	3.06	2.70	61.8%	64.9%	63.4%
POS	С	PCT_AMPHIPOD	0.00	0.52	34.2%	89.6%	(a)
POS	С	PCT_CHIRONOMID	3.25	6.31	89.5%	89.6%	89.6%
NEG	Н	PCT_CLIMB	10.00	10.00	89.5%	50.9%	(b)
NEG	Н	PCT_CLING	85.64	83.83	94.7%	45.0%	(b)
NEG	Н	PCT_CLINGER_TAXA	61.33	57.48	80.3%	45.0%	(b)
POS	FG	PCT_COLLECT	52.71	71.02	77.6%	84.7%	81.2%
NEG	С	PCT_CORBICULA	0.00	0.00	0.0%	69.4%	(a)
POS	С	PCT_DIPTERA	6.60	11.71	89.5%	84.7%	87.1%
POS	т	PCT_DOM1	26.24	34.54	69.7%	74.8%	72.3%
POS	Т	PCT_DOM2	45.92	56.45	76.3%	79.7%	78.0%
POS	т	PCT_DOM3	63.14	69.89	73.7%	79.7%	76.7%
NEG	С	PCT_EPHEMEROPTERA	35.10	21.94	82.9%	79.7%	81.3%
NEG	С	PCT_EPT	72.24	48.12	81.6%	84.7%	83.1%
NEG	R	PCT_EPT_TAXA_RICH	43.75	41.18	61.8%	61.3%	(b)
NEG	FG	PCT_FILTERERS	30.14	18.62	73.7%	74.8%	74.2%
POS	FG	PCT_GATHER	16.92	39.20	88.2%	84.7%	86.4%
NEG	С	PCT_LIMESTONE	3.99	2.732	55.3%	59.9%	(b)
NEG	С	PCT_NET_CADDISFLY	14.68	7.43	72.4%	74.8%	73.6%
POS	С	PCT_NON_INSECT	2.26	5.70	61.8%	69.8%	65.8%
POS	С	PCT_OLIGOCHAETA	0.00	0.00	73.7%	80.2%	(a)
NEG	С	PCT_PLECOPTERA	7.55	2.80	77.6%	74.8%	76.2%
NEG	FG	PCT_SCRAPER	19.16	10.62	73.7%	79.7%	76.7%
NEG	Т	PCT_SENSITIVE	59.95	42.76	93.4%	84.7%	89.1%
NEG	FG	PCT_SHREDDER	3.856	2.08	46.1%	55.4%	50.7%
NEG	н	PCT_SWIMMER	10.45	4.36	68.4%	74.8%	71.6%
POS	Т	PCT_TOLERANT	3.23	12.30	90.8%	89.6%	90.2%
NEG	С	PCT_TRICHOPTERA	16.61	9.60	72.4%	74.8%	73.6%
NEG	т	PCT_TRICHOPTERA_NO_TOL	4.28	2.41	67.1%	64.9%	66.0%
NEG	т	PCT_URBAN_INTOL	78.18	70	63.2%	70.3%	66.7%
NEG	Т	SENSITIVE_TAXA_COUNT_100	7	6	81.6%	78.4%	80.0%
NEG	R	SIMPSON_DIVERSITY	0.849	0.796	75.0%	79.7%	77.4%
NEG	R	SW	2.17	1.92	73.7%	79.7%	76.7%
NEG	R	TAXA_RICH_100	15	14	69.7%	61.7%	65.7%
POS	Т	TOLERANT_TAXA_COUNT_100	2	2	77.6%	70.7%	74.2%

PIEDMONT (REF n = 222, DEG n = 76)

			M %ile	T %ile			
Direction	Туре	Metric	Value	Value	DEG DE	REF DE	DE
POS	т	ASPT_MOD	3.5	3.7	73.8%	70.8%	72.3%
NEG	т	BECKS_100	13	10	88.1%	82.5%	85.3%
NEG	R	EPHEMEROPTERA_TAXA_CNT_100	4	3	81.0%	90.3%	85.6%
NEG	R,T	EPT_TAXA_COUNT_NO_TOL_100	10	8	88.1%	86.4%	87.2%
POS	т	FBI	3.64	4.05	66.7%	70.1%	68.4%
NEG	С	GOLD	0.827	0.793	54.8%	59.7%	57.3%
NEG	т	LOG10_SEL_EPTD	1.380	1.085	71.4%	79.9%	75.6%
NEG	R	MARGALEFS	3.320	2.625	78.6%	85.1%	81.8%
POS	С	PCT_AMPHIPOD	0.00	0.00	23.8%	84.4%	(b)
POS	С	PCT_CHIRONOMID	9.29	11.80	54.8%	59.7%	57.3%
NEG	н	PCT_CLIMB	1.45	0.96	57.1%	55.2%	56.2%
NEG	н	PCT_CLING	70.82	62.21	40.5%	55.2%	(a)
NEG	н	PCT_CLINGER_TAXA	71.08	68.00	64.3%	65.6%	64.9%
POS	FG	PCT_COLLECT	54.93	54.93	57.1%	50.0%	53.6%
NEG	С	PCT_CORBICULA	0.00	0.00	0.0%	100.0%	(b)
POS	С	PCT_DIPTERA	17.91	15.55	59.5%	50.0%	54.8%
POS	т	PCT_DOM1	26.38	33.27	78.6%	74.7%	76.6%
POS	т	PCT_DOM2	48.60	57.19	76.2%	79.9%	78.0%
POS	т	PCT_DOM3	60.47	69.13	78.6%	79.9%	79.2%
NEG	С	PCT_EPHEMEROPTERA	26.83	13.51	81.0%	85.1%	83.0%
NEG	С	PCT_EPT	71.94	68.42	57.1%	59.7%	58.4%
NEG	R	PCT_EPT_TAXA_RICH	66.67	57.89	76.2%	75.3%	75.8%
NEG	FG	PCT_FILTERERS	17.25	17.25	57.1%	50.0%	53.6%
POS	FG	PCT_GATHER	34.44	34.44	42.9%	50.0%	46.4%
NEG	С	PCT_LIMESTONE	5.35	1.94	66.7%	70.1%	68.4%
NEG	С	PCT_NET_CADDISFLY	11.07	11.07	45.2%	50.0%	47.6%
POS	С	PCT_NON_INSECT	1.02	1.02	50.0%	50.0%	50.0%
POS	С	PCT_OLIGOCHAETA	0.00	0.45	40.5%	64.9%	(a)
NEG	С	PCT_PLECOPTERA	18.99	10.96	64.3%	74.7%	(a)
NEG	FG	PCT_SCRAPER	11.06	3.37	78.6%	85.1%	81.8%
NEG	т	PCT_SENSITIVE	56.11	46.44	66.7%	70.1%	68.4%
NEG	FG	PCT_SHREDDER	16.89	9.80	54.8%	70.1%	(a)
NEG	н	PCT_SWIMMER	10.73	4.59	76.2%	79.9%	78.0%
POS	т	PCT_TOLERANT	10.77	12.48	54.8%	59.7%	57.3%
NEG	С	PCT_TRICHOPTERA	15.86	14.34	45.2%	55.2%	50.2%
NEG	т	PCT_TRICHOPTERA_NO_TOL	6.90	4.60	57.1%	64.9%	61.0%
NEG	т	PCT URBAN INTOL	90.00	85.87	57.1%	74.7%	(a)
NEG	т	SENSITIVE_TAXA_COUNT_100	9	7	85.7%	83.8%	84.7%
NEG	R	SIMPSON_DIVERSITY	0.853	0.814	83.3%	74.7%	79.0%
NEG	R	SW	2.261	1.990	85.7%	79.9%	82.8%
NEG	R	TAXA_RICH_100	16	13.25	88.1%	74.7%	81.4%
POS	т	TOLERANT_TAXA_COUNT 100	1	1	54.8%	60.4%	57.6%

RIDGES (REF n = 154, DEG n = 42)

			M %ile	T %ile			
Direction	Туре	Metric	Value	Value	DEG DE	REF DE	DE
POS	Т	ASPT_MOD	4.1	4.4	77.2%	78.9%	78.1%
NEG	Т	BECKS_100	9	6.9	76.1%	70.2%	73.1%
NEG	R	EPHEMEROPTERA_TAXA_CNT_100	4	3	66.3%	87.7%	(a)
NEG	R,T	EPT_TAXA_COUNT_NO_TOL_100	8	6.4	78.3%	70.2%	74.2%
POS	т	FBI	4.10	4.70	77.2%	77.2%	77.2%
NEG	С	GOLD	0.827	0.675	58.7%	64.9%	61.8%
NEG	т	LOG10_SEL_EPTD	1.279	1.114	68.5%	75.4%	72.0%
NEG	R	MARGALEFS	3.156	2.697	75.0%	70.2%	72.6%
POS	С	PCT_AMPHIPOD	0.00	0.00	21.7%	94.7%	(b)
POS	С	PCT_CHIRONOMID	7.53	10.78	59.8%	64.9%	62.3%
NEG	н	PCT_CLIMB	2.98	1.46	63.0%	64.9%	64.0%
NEG	н	PCT_CLING	67.80	60.08	67.4%	64.9%	66.2%
NEG	н	PCT_CLINGER_TAXA	66.67	61.74	72.8%	70.2%	71.5%
POS	FG	PCT_COLLECT	66.67	74.18	66.3%	70.2%	68.2%
NEG	С	PCT_CORBICULA	0.00	0.00	0.0%	100.0%	(b)
POS	С	PCT_DIPTERA	11.29	17.12	66.3%	64.9%	65.6%
POS	т	PCT_DOM1	30.00	36.51	66.3%	70.2%	68.2%
POS	т	PCT_DOM2	50.28	57.42	69.6%	70.2%	69.9%
POS	т	PCT_DOM3	65.45	70.50	73.9%	70.2%	72.0%
NEG	С	PCT_EPHEMEROPTERA	29.33	18.97	78.3%	75.4%	76.8%
NEG	С	PCT_EPT	64.94	50.94	72.8%	70.2%	71.5%
NEG	R	PCT_EPT_TAXA_RICH	55.56	50.00	73.9%	80.7%	77.3%
NEG	FG	PCT_FILTERERS	23.20	20.00	62.0%	61.4%	61.7%
POS	FG	PCT_GATHER	37.72	45.52	65.2%	59.6%	62.4%
NEG	С	PCT_LIMESTONE	4.39	3.92	56.5%	54.4%	55.5%
NEG	С	PCT_NET_CADDISFLY	12.08	7.63	57.6%	59.6%	58.6%
POS	С	PCT_NON_INSECT	2.63	4.46	62.0%	59.6%	60.8%
POS	С	PCT_OLIGOCHAETA	0.00	0.44	52.2%	54.4%	(b)
NEG	С	PCT_PLECOPTERA	3.51	1.43	67.4%	70.2%	68.8%
NEG	FG	PCT_SCRAPER	14.41	7.02	69.6%	75.4%	72.5%
NEG	т	PCT_SENSITIVE	45.71	31.22	75.0%	75.4%	75.2%
NEG	FG	PCT_SHREDDER	6.02	2.72	58.7%	59.6%	59.2%
NEG	н	PCT_SWIMMER	10.74	6.19	69.6%	64.9%	67.2%
POS	т	PCT_TOLERANT	9.14	17.60	72.8%	75.4%	74.1%
NEG	С	PCT_TRICHOPTERA	15.65	10.61	59.8%	59.6%	59.7%
NEG	т	PCT_TRICHOPTERA_NO_TOL	4.72	2.64	69.6%	59.6%	64.6%
NEG	т	PCT_URBAN_INTOL	84.62	81.25	71.7%	70.2%	71.0%
NEG	т	SENSITIVE_TAXA_COUNT_100	6	5.4	82.6%	70.2%	(a)
NEG	R	SIMPSON_DIVERSITY	0.822	0.797	70.7%	70.2%	70.4%
NEG	R	SW	2.086	1.878	69.6%	75.4%	72.5%
NEG	R	TAXA_RICH_100	15	13	69.6%	75.4%	72.5%
POS	т	TOLERANT_TAXA_COUNT_100	2	2	54.3%	73.7%	(a)

VALLEYS ALL (REF n = 57, DEG n = 92)

VALLEYS LIMESTONE

(REF n = 31, DEG n = 40)

			M %ile	T %ile			
Direction	Туре	Metric	Value	Value	DEG DE	REF DE	DE
POS	т	ASPT_MOD	4.2	4.5	87.5%	83.9%	85.7%
NEG	т	BECK_100	9	6	80.0%	77.4%	78.7%
NEG	R	EPHEMEROPTERA_TAXA_CNT_100	4	3.25	85.0%	74.2%	(a)
NEG	R,T	EPT_TAXA_COUNT_NO_TOL_100	8	6	80.0%	87.1%	83.5%
POS	т	FBI	4.020	4.675	80.0%	83.9%	81.9%
NEG	С	GOLD	0.842	0.810	60.0%	61.3%	60.6%
NEG	т	LOG10_SEL_EPTD	1.255	1.079	80.0%	83.9%	81.9%
NEG	R	MARGALEFS	2.982	2.515	80.0%	80.6%	80.3%
POS	С	PCT_AMPIPOD	0.00	0.00	35.0%	96.8%	(b)
POS	С	PCT_CHIRONOMID	5.30	7.02	60.0%	61.3%	60.6%
NEG	н	PCT_CLIMB	2.44	1.60	75.0%	64.5%	(a)
NEG	Н	PCT_CLING	76.35	60.65	70.0%	74.2%	72.1%
NEG	Н	PCT_CLINGER_TAXA	70.00	57.14	75.0%	80.6%	77.8%
POS	FG	PCT_COLLECT	68.52	78.77	72.5%	74.2%	73.3%
NEG	С	PCT_CORBICULA	0.00	0.00	0.0%	100.0%	(b)
POS	С	PCT_DIPTERA	9.23	11.82	60.0%	61.3%	60.6%
POS	т	PCT_DOM1	29.13	37.72	80.0%	80.6%	80.3%
POS	т	PCT_DOM2	48.37	57.28	85.0%	80.6%	82.8%
POS	т	PCT_DOM3	63.85	70.59	87.5%	80.6%	84.1%
NEG	С	PCT_EPHEMEROPTERA	38.51	20.83	85.0%	83.9%	84.4%
NEG	С	PCT_EPT	68.29	54.91	77.5%	74.2%	75.8%
NEG	R	PCT_EPT_TAXA_RICH	57.14	50.00	77.5%	83.9%	80.7%
NEG	FG	PCT_FILTERERS	23.96	19.09	65.0%	64.5%	64.8%
POS	FG	PCT_GATHER	36.22	50.00	65.0%	71.0%	68.0%
NEG	С	PCT_LIMESTONE	5.69	5.27	50.0%	54.8%	52.4%
NEG	С	PCT_NET_CADDISFLY	17.19	11.02	57.5%	61.3%	59.4%
POS	С	PCT_NON_INSECT	2.63	5.34	67.5%	64.5%	66.0%
POS	С	PCT_OLIGOCHAETA	0.00	0.00	55.0%	58.1%	(b)
NEG	С	PCT_PLECOPTERA	3.37	0.33	75.0%	74.2%	74.6%
NEG	FG	PCT_SCRAPER	17.54	11.54	82.5%	74.2%	78.3%
NEG	т	PCT_SENSITIVE	46.58	33.61	82.5%	80.6%	81.6%
NEG	FG	PCT_SHREDDER	3.15	2.85	57.5%	54.8%	56.2%
NEG	н	PCT_SWIMMER	10.92	6.01	77.5%	74.2%	75.8%
POS	т	PCT_TOLERANT	5.83	16.36	77.5%	80.6%	79.1%
NEG	С	PCT_TRICHOPTERA	18.18	6.80	45.0%	71.0%	(a)
NEG	т	PCT_TRICHOPTERA_NO_TOL	4.73	2.41	75.0%	64.5%	(a)
NEG	т	PCT_URBAN_INTOL	85.71	82.84	77.5%	64.5%	(a)
NEG	т	SENSITIVE_TAXA_COUNT_100	6	5	80.0%	83.9%	81.9%
NEG	R	SIMPSON_DIVERSITY	0.827	0.793	85.0%	80.6%	82.8%
NEG	R	SW	2.138	1.811	80.0%	80.6%	80.3%
NEG	R	TAXA_RICH_100	14	12	72.5%	80.6%	76.6%
POS	Т	TOLERANT_TAXA_COUNT_100	2	2	52.5%	77.4%	(a)
VALLEYS NO LIMESTONE

(REF n = 26, DEG n = 52)

			M %ile	T %ile			
Direction	Туре	Metric	Value	Value	DEG DE	REF DE	DE
POS	Т	ASPT_MOD	4.1	4.4	69.2%	73.1%	71.2%
NEG	Т	BECK_100	9	7	71.2%	76.9%	74.0%
NEG	R	EPHEMEROPTERA_TAXA_CNT_100	4	3	57.7%	80.8%	(a)
NEG	R,T	EPT_TAXA_COUNT_NO_TOL_100	8	7	73.1%	76.9%	75.0%
POS	Т	FBI	4.335	4.910	65.4%	69.2%	67.3%
NEG	С	GOLD	0.751	0.619	63.5%	65.4%	64.4%
NEG	Т	LOG10_SEL_EPTD	1.301	1.176	67.3%	65.4%	66.3%
NEG	R	MARGALEFS	3.271	2.798	69.2%	69.2%	69.2%
POS	С	PCT_AMPIPOD	0.00	0.40	11.5%	92.3%	(b)
POS	С	PCT_CHIRONOMID	8.94	12.91	65.4%	65.4%	65.4%
NEG	н	PCT_CLIMB	3.44	2.82	67.3%	53.8%	(a)
NEG	н	PCT_CLING	61.70	59.98	65.4%	53.8%	(a)
NEG	Н	PCT_CLINGER_TAXA	63.25	54.20	53.8%	69.2%	(a)
POS	FG	PCT_COLLECT	64.10	72.02	65.4%	69.2%	67.3%
NEG	С	PCT_CORBICULA	0.00	0.00	0.0%	100.0%	(b)
POS	С	PCT_DIPTERA	13.00	28.01	65.4%	65.4%	65.4%
POS	Т	PCT_DOM1	32.22	35.35	57.7%	65.4%	61.5%
POS	т	PCT_DOM2	52.51	59.57	57.7%	65.4%	61.5%
POS	Т	PCT_DOM3	68.24	70.95	61.5%	65.4%	63.5%
NEG	С	PCT_EPHEMEROPTERA	21.57	16.26	69.2%	73.1%	71.2%
NEG	С	PCT_EPT	58.08	46.04	67.3%	65.4%	66.3%
NEG	R	PCT_EPT_TAXA_RICH	53.59	50.00	71.2%	76.9%	74.0%
NEG	FG	PCT_FILTERERS	21.04	20.15	59.6%	53.8%	56.7%
POS	FG	PCT_GATHER	42.95	49.14	53.8%	61.5%	57.7%
NEG	С	PCT_LIMESTONE	3.09	2.98	57.7%	53.8%	55.8%
NEG	С	PCT_NET_CADDISFLY	10.15	7.14	59.6%	61.5%	60.6%
POS	С	PCT_NON_INSECT	2.41	4.61	55.8%	61.5%	58.7%
POS	С	PCT_OLIGOCHAETA	0.41	0.47	53.8%	53.8%	53.8%
NEG	С	PCT_PLECOPTERA	5.42	2.04	65.4%	69.2%	67.3%
NEG	FG	PCT_SCRAPER	10.56	5.78	65.4%	65.4%	65.4%
NEG	т	PCT_SENSITIVE	32.13	25.27	63.5%	69.2%	66.3%
NEG	FG	PCT_SHREDDER	5.13	2.90	59.6%	61.5%	60.6%
NEG	н	PCT_SWIMMER	10.11	5.94	59.6%	61.5%	60.6%
POS	т	PCT_TOLERANT	12.78	22.92	69.2%	69.2%	69.2%
NEG	С	PCT_TRICHOPTERA	11.99	10.22	65.4%	61.5%	63.5%
NEG	т	PCT_TRICHOPTERA_NO_TOL	3.69	2.66	65.4%	61.5%	63.5%
NEG	т	PCT_URBAN_INTOL	82.84	80.68	67.3%	65.4%	66.3%
NEG	т	SENSITIVE_TAXA_COUNT_100	6.5	6	78.8%	69.2%	74.0%
NEG	R	SIMPSON_DIVERSITY	0.812	0.792	59.6%	65.4%	62.5%
NEG	R	SW	2.072	1.973	73.1%	69.2%	71.2%
NEG	R	TAXA_RICH_100	15	13.5	73.1%	69.2%	71.2%
POS	Т	TOLERANT_TAXA_COUNT_100	2	3	36.5%	84.6%	(a)

NORTH CENTRAL APPALACHIANS

(REF n = 26, DEG n = 40)

			M %ile	T %ile			
Direction	Туре	Metric	Value	Value	DEG DE	REF DE	DE
POS	Т	ASPT_MOD	3.5	3.5	55.0%	61.5%	58.3%
NEG	Т	BECK_100	14.5	11	65.0%	69.2%	67.1%
NEG	R	EPHEMEROPTERA_TAXA_CNT_100	4	4	77.5%	73.1%	75.3%
NEG	R,T	EPT_TAXA_COUNT_NO_TOL_100	11	9	75.0%	80.8%	77.9%
POS	Т	FBI	3.740	3.830	55.0%	53.8%	54.4%
NEG	С	GOLD	0.814	0.814	52.5%	50.0%	51.3%
NEG	т	LOG10_SEL_EPTD	1.724	1.580	75.0%	76.9%	76.0%
NEG	R	MARGALEFS	3.565	3.394	77.5%	69.2%	73.4%
POS	С	PCT_AMPHIPOD	0.00	0.00	0.0%	96.2%	(b)
POS	С	PCT_CHIRONOMID	9.00	15.71	45.0%	61.5%	(a)
NEG	н	PCT_CLIMB	1.05	0.95	35.0%	53.8%	(a)
NEG	н	PCT_CLING	69.94	65.25	22.5%	65.4%	(a)
NEG	н	PCT_CLINGER_TAXA	71.72	70.83	70.0%	61.5%	65.8%
POS	FG	PCT_COLLECT	63.66	66.95	57.5%	61.5%	59.5%
NEG	С	PCT_CORBICULA	0.00	0.00	0.0%	100.0%	(b)
POS	С	PCT_DIPTERA	18.34	18.34	50.0%	50.0%	50.0%
POS	т	PCT_DOM1	24.93	28.03	72.5%	69.2%	70.9%
POS	т	PCT_DOM2	42.02	48.73	70.0%	69.2%	69.6%
POS	т	PCT_DOM3	55.44	61.28	67.5%	65.4%	66.4%
NEG	С	PCT_EPHEMEROPTERA	40.22	24.26	65.0%	73.1%	69.0%
NEG	С	PCT_EPT	69.45	66.67	35.0%	53.8%	(a)
NEG	R	PCT_EPT_TAXA_RICH	65.11	62.20	72.5%	69.2%	70.9%
NEG	FG	PCT_FILTERERS	17.29	16.71	45.0%	53.8%	49.4%
POS	FG	PCT_GATHER	44.12	58.45	17.5%	73.1%	(a)
NEG	С	PCT_LIMESTONE	10.54	5.77	82.5%	73.1%	77.8%
NEG	С	PCT_NET_CADDISFLY	11.43	11.20	45.0%	53.8%	49.4%
POS	С	PCT_NON_INSECT	0.00	0.16	30.0%	69.2%	(b)
POS	С	PCT_OLIGOCHAETA	0.00	0.00	12.5%	80.8%	(b)
NEG	С	PCT_PLECOPTERA	14.11	15.32	60.0%	46.2%	(a)
NEG	FG	PCT_SCRAPER	13.30	8.45	70.0%	73.1%	71.5%
NEG	т	PCT_SENSITIVE	50.76	47.50	55.0%	53.8%	54.4%
NEG	FG	PCT_SHREDDER	9.30	7.73	62.5%	61.5%	62.0%
NEG	н	PCT_SWIMMER	12.99	9.46	62.5%	69.2%	65.9%
POS	т	PCT TOLERANT	9.68	11.08	60.0%	53.8%	56.9%
NEG	С	PCT TRICHOPTERA	15.48	15.48	47.5%	50.0%	48.8%
NEG	т	PCT TRICHOPTERA NO TOL	5.33	4.22	40.0%	65.4%	(a)
NEG	т	PCT URBAN INTOL	93.49	94.36	45.0%	46.2%	45.6%
NEG	т	SENSITIVE TAXA COUNT 100	10.5	8.5	70.0%	69.2%	69.6%
NEG	R	SIMPSON_DIVERSITY	0.867	0.825	70.0%	73.1%	71.5%
NEG	R	SW	2.376	2.197	77.5%	69.2%	73.4%
NEG	R	TAXA RICH 100	17	15	72.5%	76.9%	74.7%
POS	т	 TOLERANT_TAXA_COUNT_100	1	1	20.0%	76.9%	(a)

NORTHERN APPALACHIAN PLATEAU & UPLANDS

(REF n = 28, DEG n = 20)

			M %ile	T %ile			
Direction	Туре	Metric	Value	Value	DEG DE	REF DE	DE
POS	Т	ASPT_MOD	4.2	4.5	60.0%	71.4%	(a)
NEG	т	BECK_100	9	6.2	70.0%	67.9%	68.9%
NEG	R	EPHEMEROPTERA_TAXA_CNT_100	4	3.8	35.0%	60.7%	(a)
NEG	R <i>,</i> T	EPT_TAXA_COUNT_NO_TOL_100	7.5	7	65.0%	64.3%	64.6%
POS	Т	FBI	4.575	4.881	70.0%	67.9%	68.9%
NEG	С	GOLD	0.773	0.715	55.0%	64.3%	59.6%
NEG	т	LOG10_SEL_EPTD	1.176	1.099	70.0%	60.7%	65.4%
NEG	R	MARGALEFS	3.040	2.823	70.0%	78.6%	74.3%
POS	С	PCT_AMPHIPOD	0.00	0.00	15.0%	92.9%	(b)
POS	С	PCT_CHIRONOMID	17.08	19.20	60.0%	60.7%	60.4%
NEG	Н	PCT_CLIMB	0.00	0.00	0.0%	100.0%	(b)
NEG	Н	PCT_CLING	61.89	55.45	25.0%	64.3%	(a)
NEG	Н	PCT_CLINGER_TAXA	64.29	62.26	40.0%	53.6%	(a)
POS	FG	PCT_COLLECT	77.30	83.17	65.0%	60.7%	62.9%
NEG	С	PCT_CORBICULA	0.00	0.00	0.0%	100.0%	(b)
POS	С	PCT_DIPTERA	21.50	26.05	60.0%	60.7%	60.4%
POS	т	PCT_DOM1	27.57	29.65	45.0%	53.6%	49.3%
POS	т	PCT_DOM2	47.63	48.75	55.0%	53.6%	54.3%
POS	т	PCT_DOM3	61.12	62.09	50.0%	53.6%	51.8%
NEG	С	PCT_EPHEMEROPTERA	25.00	22.46	55.0%	60.7%	57.9%
NEG	С	PCT_EPT	56.94	53.00	65.0%	64.3%	64.6%
NEG	R	PCT_EPT_TAXA_RICH	51.47	50.00	40.0%	57.1%	(a)
NEG	FG	PCT_FILTERERS	23.71	22.06	50.0%	53.6%	51.8%
POS	FG	PCT_GATHER	49.33	56.72	75.0%	75.0%	75.0%
NEG	С	PCT_LIMESTONE	1.53	0.00	0.0%	100.0%	(a)
NEG	С	PCT_NET_CADDISFLY	19.05	19.20	50.0%	46.4%	48.2%
POS	С	PCT_NON_INSECT	1.19	1.91	40.0%	53.6%	(a)
POS	С	PCT_OLIGOCHAETA	0.00	0.48	25.0%	64.3%	(b)
NEG	С	PCT_PLECOPTERA	3.71	1.96	85.0%	67.9%	76.4%(a)
NEG	FG	PCT_SCRAPER	3.55	3.01	35.0%	53.6%	(a)
NEG	т	PCT SENSITIVE	30.50	24.76	65.0%	67.9%	66.4%
NEG	FG	PCT_SHREDDER	3.00	2.95	50.0%	53.6%	51.8%
NEG	н	PCT_SWIMMER	15.85	13.00	60.0%	60.7%	60.4%
POS	т	PCT TOLERANT	17.70	22.15	55.0%	64.3%	59.6%
NEG	С	PCT TRICHOPTERA	20.00	15.73	25.0%	64.3%	(a)
NEG	т	PCT TRICHOPTERA NO TOL	4.26	3.08	70.0%	60.7%	65.4%
NEG	т	PCT URBAN INTOL	85.17	81.65	60.0%	60.7%	60.4%
NEG	т	SENSITIVE TAXA COUNT 100	7	4.1	75.0%	67.9%	71.4%
NEG	R	SIMPSON DIVERSITY	0.849	0.839	55.0%	53.6%	54.3%
NEG	R	 SW	2.150	2.113	60.0%	60.7%	60.4%
NEG	R	TAXA RICH 100	15	13.4	70.0%	78.6%	74.3%
POS	т	 TOLERANT_TAXA_COUNT_100	2	2	45.0%	67.9%	(a)