VIRGINIA WATER RESOURCES RESEARCH CENTER

Development of
Aquatic Life Use Assessment Protocols
for
Class VII Waters in Virginia

**2014 Report of the Academic Advisory Committee
for
Virginia Department of Environmental Quality**

SPECIAL REPORT

**VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY
BLACKSBURG, VIRGINIA**

**SR55-2014
October 2014**

Stephen Schoenholtz, Director

DEVELOPMENT OF
AQUATIC LIFE USE ASSESSMENT PROTOCOLS
FOR
CLASS VII WATERS IN VIRGINIA

2014 Report of the Academic Advisory Committee for
Virginia Department of Environmental Quality

by:

Andrew L. Garey
Member, Academic Advisory Committee
Lab & Research Specialist II, Aquatic Ecology Laboratory
Department of Biology and VCU Rice Center for Environmental Sciences
Virginia Commonwealth University

Greg C. Garman
Member, Academic Advisory Committee
Associate Professor, Department of Biology
Director, Center for Environmental Studies
Virginia Commonwealth University

Leonard A. Smock
Member, Academic Advisory Committee
Professor, Department of Biology
Director, VCU Rice Center for Environmental Sciences
Virginia Commonwealth University

Edited by:
Jane L. Walker

Members of the 2014 Academic Advisory Committee to the Virginia Department of Environmental Quality

Stephen H. Schoenholtz, Chair
Virginia Water Resources Research Center /
    Department of Forest Resources and
    Environmental Conservation
Virginia Tech

E. Fred Benfield
Department of Biology
Virginia Tech

Paul Bukaveckas
Department of Biology / Center for
    Environmental Studies / Rice Center for
    Environmental Life Sciences
Virginia Commonwealth University

Andrew L. Garey
Department of Biology / Rice Center for
    Environmental Life Sciences
Virginia Commonwealth University

Gregory C. Garman
Department of Biology / Center for
    Environmental Studies
Virginia Commonwealth University

Carl Hershner
Department of Biology / Center for Coastal
    Resources Management
Virginia Institute of Marine Science
The College of William and Mary

Golde I. Holtzman
Department of Statistics
Virginia Tech

Wu-Seng Lung
Department of Civil and Environmental
    Engineering
The University of Virginia

Kevin J. McGuire
Virginia Water Resources Research Center /
    Department of Forest Resources and
    Environmental Conservation
Virginia Tech

Leonard A. Shabman
Resources for the Future

Eric P. Smith
Department of Statistics
Virginia Tech

Leonard A. Smock
Department of Biology / Rice Center for
    Environmental Life Sciences
Virginia Commonwealth University

Kurt Stephenson
Department of Agricultural and Applied
    Economics
Virginia Tech

Jane L. Walker
Virginia Water Resources Research Center
Virginia Tech

Gene Yagow
Department of Biological Systems
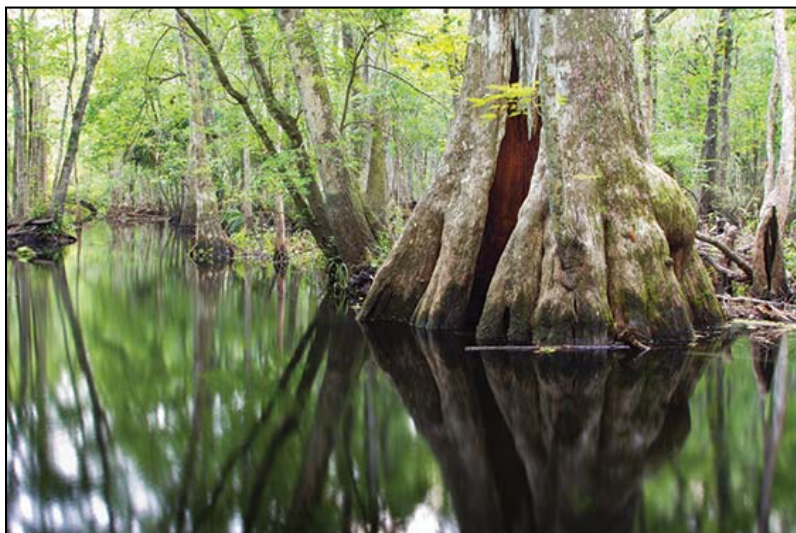   Engineering
Virginia Tech

Carl E. Zipper
Department of Crop and Soil Environmental
    Sciences
Virginia Tech

DEVELOPMENT OF
AQUATIC LIFE USE ASSESSMENT PROTOCOLS
FOR
CLASS VII WATERS IN VIRGINIA

**2014 Report of the Academic Advisory Committee for
Virginia Department of Environmental Quality**

Table of Contents

Introduction

Many of the streams and swamps that flow on Virginia's Coastal Plain physiographic province experience low dissolved oxygen and pH values due to naturally occurring conditions. The organisms that are endemic to these blackwater systems are typically physiologically and behaviorally adapted to these conditions. In contrast, aquatic taxa with geographically broader distributions lack specific adaptations to blackwater conditions, are uncommon in ecologically healthy blackwater systems, and may be indicators of impairment in these streams. As a consequence, it may be possible to use the taxonomic and functional composition of biotic assemblages in Coastal Plain streams and swamps of Virginia to accurately assess the status of these Class VII blackwaters and to evaluate the level of ecological health or impairment using biological criteria. The Academic Advisory Committee (AAC) to the Virginia Department of Environmental Quality (DEQ) has been tasked since 2011 with developing such a bioassessment tool for Virginia Class VII waters.

A previous report of the AAC described the initial development of the conceptual basis of a biology-based stream-assessment tool, termed the Blackwater Condition Index (BCI), for use in Class VII blackwater systems (Garman et al. 2012). That report also provided an initial working protocol for the use of field-based, non-biological criteria to identify blackwater systems as distinct from other Class VII waters in the region. A subsequent AAC report then presented work that further developed and tested a protocol that employs the best combination of field- and GIS-based habitat evaluation metrics to quantitatively separate Class VII blackwater streams from other types of Class VII waters as well as from Coastal Plain streams not appropriate for Class VII designation (Garman et al. 2013). That report also detailed areas of needed data and further validation of analytical tools developed during previous studies.

Following completion of the 2012 and 2013 reports (Garman et al. 2012, 2013), the AAC was tasked to continue development of the BCI. The primary objectives of this study thus were to: 1) further develop the blackwater habitat protocol (BHP), which is to be used to classify stream and swamp classes in a manner that is relevant to the fish and/or macroinvertebrate assemblages of interest; and 2) develop biological assemblage metrics that are able to distinguish between anthropogenically stressed and reference (relatively unimpaired) sites as indicated by a series of non-biological reference filters for each natural stream class, as determined by the BHP.

Based on ongoing discussions among AAC members and DEQ biologists, a multi-year study approach has been agreed upon, wherein project activities and reporting to be completed by June 2014 would include an initial inventory of existing data appropriate for the study, refinement of the prototype BHP, and refinement of the study design for continued development of the BCI. Activities conducted to fulfill these goals and conclusions with respect to each are organized according to the major Class VII tasks for the AAC during fiscal year 2014.

Task 1: Identification of Appropriate Study Sites and Assembly of Pertinent Datasets Needed for Data Analysis in the Development of a Bioassessment Protocol for Blackwater Streams

Virginia Commonwealth University (VCU) and DEQ personnel held office meetings on November 8, 2013; January 24, 2014; and April 11, 2014 to refine the previously developed draft strategy pertaining to swamp habitat assessment, discuss long-range plans for development of the BCI, and plan field activities. During these meetings, the group assembled a database that contained an inventory of available data from Virginia Coastal Plain stream and swamp sites that had been sampled by DEQ and/or VCU. Data imported into the BCI database were from the VCU INSTAR (INteractive STream Assessment Resource) database (http://instar.vcu.edu/; collection dates from 2004-present) and from the DEQ EDAS (Ecological Data Application System) database (collection dates from 2004-present). Macroinvertebrate data in EDAS exist primarily at the family level before 2009 and were therefore less useful for this study.

The database contains information on biological, physicochemical, and GIS-derived landscape data; designations of *Reference*, *Stressed,* or *Unknown* pertaining to the impairment status of each site; and designations of *Class-VII Blackwater*, *Class VII Non-blackwater,* and *Not Class-VII* based on each system's apparent natural typology. Both impairment and typology classifications were made based on cursory reviews of available data and the best professional judgment of biologists and have not yet been statistically validated. These preliminary classifications were necessary to make judgments regarding the dataset's breadth (distribution of data among site types) and depth (amount of biological, physicochemical, and landscape data available for each site).

After initial site visits and reviews of the available data, VCU and DEQ biologists agreed to an initial 2 x 3-factor blocked design based on natural system typology and anthropogenic impairment status for BHP validation and BCI construction. Therefore, each site was assigned to one of six groups, based on its preliminary designations: 1) Reference, Class VII, Blackwater (RBW7); 2) Reference, Class VII, Non-blackwater (R7); 3) Reference, Not Class VII (RX); 4) Stressed, Class VII, Blackwater (SBW7); 5) Stressed, Class VII, Non-blackwater (S7); or 6) Stressed, Not Class VII (SX) (Fig. 1). This blocked study design allows for the development of an analytical framework appropriate for the two major study objectives. Stream typology groups may be changed following further analysis of BHP classification metrics with biological assemblage data (discussed below under Task 3), but the study objectives and conceptual framework will remain as discussed here.

Each of the six data types must be included in the final BHP construction process. Therefore, the database was inventoried to determine whether sites contain appropriate data for inclusion and to determine where data gaps exist. The initial database contains only data where DEQ and/or VCU personnel are sufficiently familiar with the study sites to render professional judgment regarding site typology and impairment status. Data from additional sites exist in both DEQ and VCU datasets that may be appropriate for the project, but further GIS analysis, data review, and site reconnaissance are required to determine the appropriateness of including these data, as well as for determining the typology and impairment status of these sites. Whereas these additional

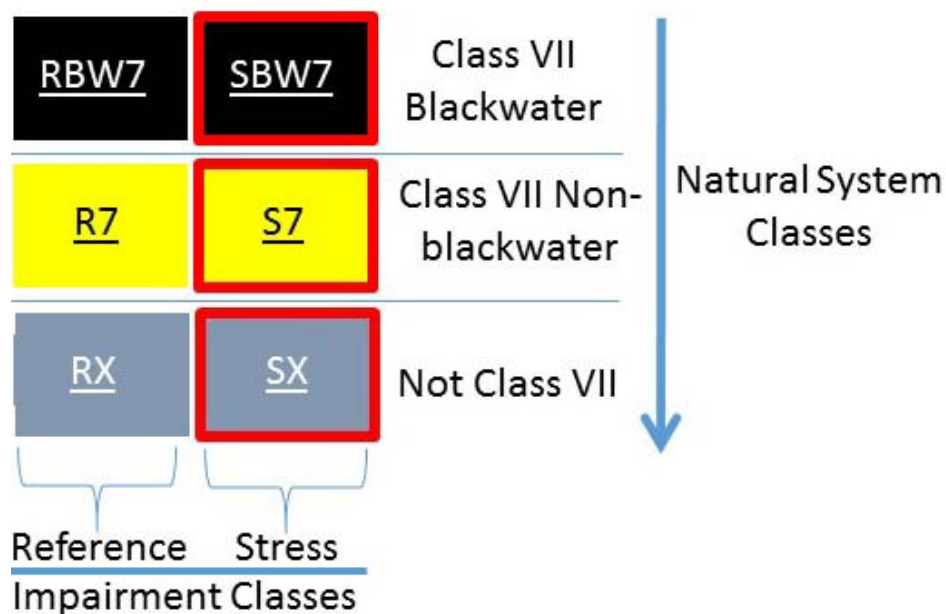data were not included in this initial inventory, they are discussed at the conclusion of this section.

Figure 1: Schematic of 2 x 3 factor blocked study design for grouping sites based on impairment status and natural typology classification. Designation classes: Reference, Class VII, Blackwater (RBW7); Reference, Class VII, Non-blackwater (R7); Reference, Not Class VII (RX); Stressed, Class VII, Blackwater (SBW7); Stressed, Class VII, Non-blackwater (S7); and Stressed, Not Class VII (SX).

1.1 Current Blackwater Condition Index Database

The current database contains 176 records representing 147 stream sites (records indicating separate sampling events at the same study site are currently allowed). Fields in the database contain information regarding the presence or absence of biological assemblage (fish and invertebrate) data, a suite of physicochemical and habitat parameters that were scored on site (including whether the current BHP was evaluated), and Geographic Information Systems (GIS) data that have been collected at the sites. DEQ and VCU biologists agreed to an initial suite of reference-filter criteria that will be used to determine each site's impairment status in a manner independent of the biological assemblage data that will be used for BCI construction (Table 1; see section 3.1 for further discussion on reference filters). The use of these reference filters will allow for verification or change of the preliminary impairment-status designations discussed above, as well as statistical evaluation of potential BCI metrics. Biological assemblage data were also evaluated based on the completeness of each dataset. Fish data were given evaluation codes of 0 (no data), 1 (data not at species level or presence/absence data only), or 2 (species-level, absolute abundance data). Macroinvertebrate data were given evaluation codes of 0 (no data), 1 (less than 200 individuals in subsamples or data are at the family or coarser taxonomic level), or 2 (genus or species-level data with at least 200-individual subsamples). For the sake of

brevity, further details regarding the database are not included here. A spreadsheet of the raw inventory data, as well as explanations of each included field, can be provided upon request.

Table 1: Reference filters for determination of site-impairment status; filters are subject to change based on data availability.

Parameter	Reference Threshold	Stressed Threshold
Physicochemistry		
Specific Conductance	<150 µS/cm	>350 µS/cm
Total Nitrogen	<1.5 mg/L	>3 mg/L
Total Phosphorus	<0.05 mg/L	>0.1 mg/L
pH	<6.5	>7.5
Other	No other measured parameters indicate site should be 303d listed	Other chemical stressors present that are likely to affect community
Land Cover		
GIS Land Use/Land Cover	>70 percent forested land cover in watershed	<50 percent forested land cover
Intact Riparian Vegetation	>50 m from both banks	<10 m, either bank, or <25 m from both banks
General Site Characteristics		
Point Sources/Others	No NPDES sites within watershed	NA*
Site Reconnaissance Land Use/Land Cover	No extensive development in the watershed that is likely to impact the system	NA*
Visible System Impairment	No visible signs of direct alteration to the water body (e.g., dams, weirs, levees, artificial channelization)	NA*

NPDES = National Pollutant Discharge Elimination System
NA* = Not applicable; these filters are not typically used to designate systems as stressed.

1.2 Initial Database Inventory

Study sites included in the BHP and BCI development process will include only those that contain datasets pertaining to both fish and macroinvertebrate assemblages, reference filters (specific conductance, total nitrogen, total phosphorus, pH, GIS-derived watershed land cover, and site-reconnaissance information), and BHP metrics. The cost associated with GIS land-cover evaluations is low relative to the field-collected biological data and water chemistry data. The BHP evaluation and site reconnaissance have only been conducted at a small number of study sites and will most probably be repeated at many sites to ensure that the most current set of observed conditions possible are available. Therefore, this inventory focused on fish- and macroinvertebrate-assemblage data and water chemistry data (i.e., specific conductance, total nitrogen, total phosphorus, and pH). To be counted as including fish or macroinvertebrate data, sites had to have been given a rank of 2 for each parameter (i.e., rank 1 data were considered of insufficient quality and were not included). To be counted as including water chemistry data, sites had to include at least one measurement of all four assessed chemical parameters. The majority of sites for which water chemistry data have been collected include multiple measurements of each parameter. Therefore, after further assessment of the dataset, the criterion for inclusion of sites may be made more stringent in the future to include only those wherein multiple physicochemical surveys have been conducted.

The database inventory consisted of a rudimentary estimation of the costs and time expenditures associated with further sampling activities at pre-established sites in the database, such that each contained all of the three major field data types (fish, macroinvertebrate, and water chemistry). When multiple surveys were conducted at the same site, records were combined to include all data types that had been collected. Only one record for each site was allowed for the inventory.

A computer program was developed in R (R Core Team 2013) to rank site records according to the relative costs and time expenditures associated with collection of the data already recorded at each site, as well as to track incidences where missing data occurred. Macroinvertebrate collections were given a rank of 3 (highest cost associated with collection); fish collections received a rank of 2; and water chemistry surveys were given a rank of 1 (lowest costs). Sites with multiple types of data were given higher ranks than sites with only one data type and were ranked in order of priority with respect to the types of data present. The ranking system for sites containing multiple data types was as follows:
- sites with all three data types: 7;
- sites with fish and macroinvertebrate data: 6;
- sites with macroinvertebrate and water chemistry data: 5; and
- sites with fish and water chemistry data: 4.

The ranked inventory system employed here allowed for evaluation of the relative time and cost expenditures already incurred at each site, determined the types of data needed for inclusion of sites in this project, and most importantly, indicated where future field investigations can be conducted in the most time- and cost-efficient manner. When additional replicate field sites are needed, highest-ranked sites should be visited first to maximize the use of previously acquired data and thus minimize additional costs to the state.

After ranking each site based on pre-existing data, we organized the sites according to the six preliminary designation groups. Of the 147 unique site records evaluated, biologists were confident in giving preliminary reference/stressed designations to 74. The remaining 73 sites were given typology designations, but biologists agreed that even preliminary impairment designations could not be made without further evaluation. Therefore, in order to take the most conservative approach to the inventory process, these sites were not included in the assessment of needed data, although they were given ranks based on data present because it is likely that many will be included in the study pending future evaluation.

Of the 74 sites with assigned designation classes, four were previously surveyed for all data types (rank: 7), including three likely reference sites (two sites in class R7 and one in class RX) and one stressed site (class SBW7) (Tables 2 and 3). Six of the site records included synoptic fish and macroinvertebrate data (two each in classes RBW7 and SX and one each in classes RX and SBW7). Comparatively more site records contained both macroinvertebrate and water chemistry data (38 of 74 records). The majority of these were sites sampled by DEQ personnel for validation of a new genus-level version of the Coastal Plain Macroinvertebrate Index. No site records contained data for fish and water chemistry but not macroinvertebrates, and none contained only macroinvertebrate data. Four records contained only fish data, and 16 site records contained only water chemistry data. Six of the 74 sites records with designation classes contained no biological or water chemistry data. These sites were the subject of field reconnaissance in preparation for other VCU and DEQ studies but were considered inappropriate for study purposes and therefore not sampled. All six sites were considered Class VII and impaired (two in class SBW7 and four in class S7). These sites were included in the dataset because their locations are known and they are accessible so that collection of data at these sites would be more efficient than establishment of entirely new sites.

Table 2: Inventory results showing the number of reference sites where the different types of data have been collected.

	All data (7)*	Fish and Macroinvertebrates (6)*	Macroinvertebrates and Chemistry (5)*	Fish and Chemistry (4)*	Macroinvertebrates Only (3)*	Fish Only (2)*	Chemistry Only (1)*
Blackwater Class VII (RBW7)	0	2	0	0	0	0	4
Non-BW Class VII (R7)	2	0	2	0	0	1	1
Not Class VII (RX)	1	1	15	0	0	1	1

* Columns are ordered according to the value ranking of existing data (1-7). Sites where all data were collected are considered of highest value (rank: 7), and sites for where only water chemistry was measured are of lowest value (rank: 1).

Table 3: Inventory results showing the number of stressed sites where the different types of data have been collected.

	All data (7)*	Fish and Macroinvertebrates (6)*	Macroinvertebrates and Chemistry (5)*	Fish and Chemistry (4)*	Macroinvertebrates Only (3)*	Fish Only (2)*	Chemistry Only (1)*
Blackwater Class VII (SBW7)	1	1	2	0	0	1	9
Non-BW Class VII (S7)	0	0	2	0	0	1	1
Not Class VII (SX)	0	2	17	0	0	0	0

* Columns are ordered according to the value ranking of existing data (1-7). Sites where all data were collected are considered of highest value (rank: 7), and sites for where only water chemistry was measured are of lowest value (rank: 1).

1.3 Further Data Requirements

There are no stringent, peer-reviewed standards with respect to the number of sites needed for the development of a statistically valid biological assessment index. The additional data acquisition required to provide synoptic fish, macroinvertebrate, and water chemistry data at 10 sites per class (total: 60 study sites) is the maximum amount feasible over the next one- to two-year study period. Therefore, the assessment of additional data required for the initial development of the BCI was conducted with the goal of achieving ten sites in each of the six designation classes.

A total of 98 additional surveys of either fish, macroinvertebrates, or water chemistry are required to achieve the 10-site per class study goal. It is important to note that the number of new surveys required of a given data type cannot be determined simply by subtracting the number of surveys of that type from the total number of sites needed for each class (ten). For example, for stressed blackwater Class VII sites (SBW7, Table 3), a total of eleven water chemistry surveys have been conducted, of which nine were chemistry-only surveys and two were surveys where water chemistry and biological data were both collected. However, under the most cost-efficient strategy for SBW7, two additional water chemistry surveys should be conducted because two sites within that class include biological data but not water chemistry data, and biological data are given priority.

A total of 47 fish surveys, 28 macroinvertebrate surveys, and 23 water chemistry surveys are needed to meet the study goal. Water chemistry will be collected at new study sites along with macroinvertebrate and fish data; therefore, the additional effort requirements to achieve the study goal would be 75 new field collection sites (for 47 fish and 28 macroinvertebrate surveys). The actual number of collections required will most probably be lower than 75 (and lower than 47, 28, and 23 for fish, macroinvertebrate, and water chemistry, respectively) because additional DEQ and VCU data exist that require more in-depth evaluation to determine their appropriateness for inclusion (discussed below).

The greatest need among all classes is for fish assemblage data (except for class RBW7 for which eight collections of both macroinvertebrates and fish are needed). Fish data are needed for seven to nine sites for all classes (Fig. 2). A relatively high need for macroinvertebrate data also exists. Eight macroinvertebrate collections are needed for classes RBW7 and S7, and six macroinvertebrate collections are needed for R7 and SBW7. The data requirement for macroinvertebrates has been fulfilled for non-Class VII systems (classes RX and SX). Requirements for additional water chemistry data were lowest overall among the three data types, ranging from one to seven additional sites needed per class (Fig. 2).
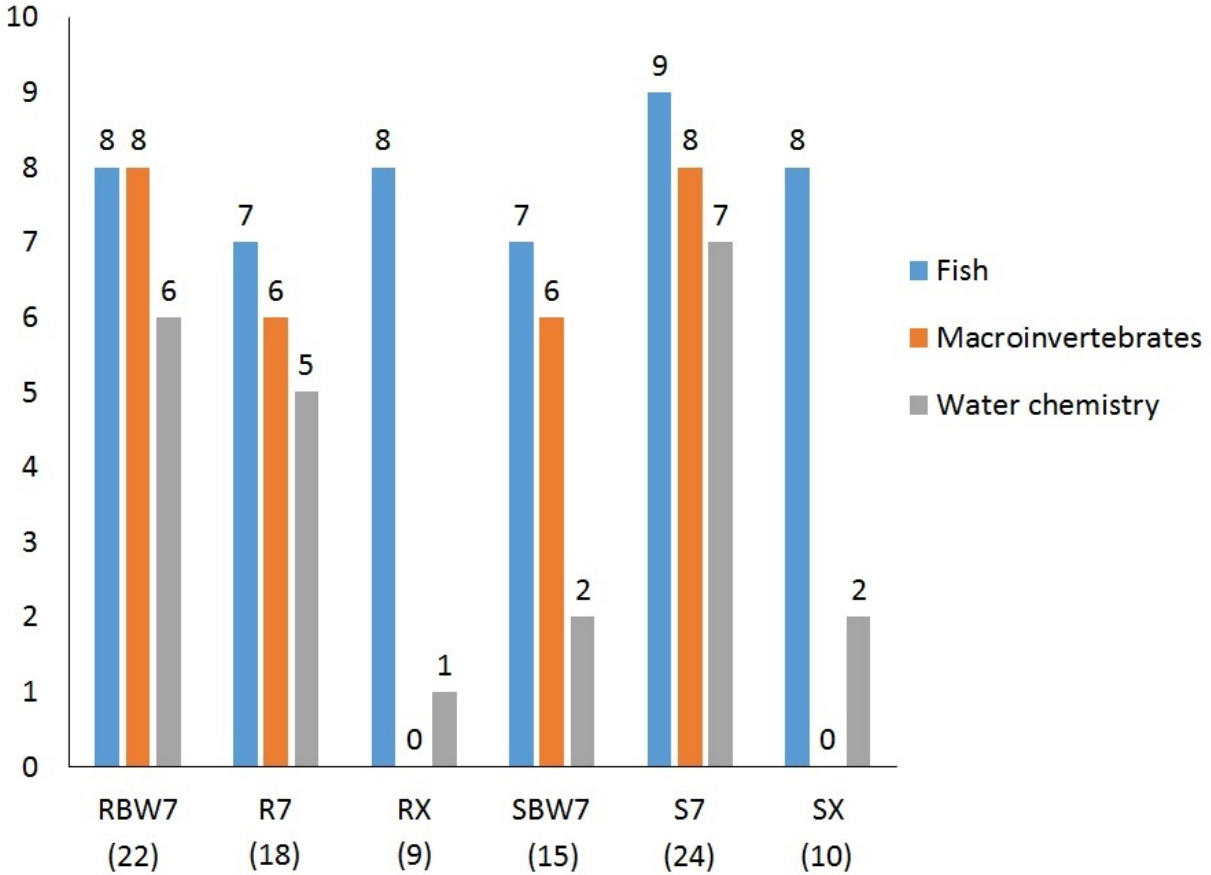
Figure 2: Number of stream surveys needed at study sites for the six designation classes to fulfill the goal of 10 sites per class. Designation classes: Reference, Class VII, Blackwater (RBW7); Reference, Class VII, Non-blackwater (R7); Reference, Not Class VII (RX); Stressed, Class VII, Blackwater (SBW7); Stressed, Class VII, Non-blackwater (S7); Stressed, Not Class VII (SX). Numbers below designation classes indicate the total number of surveys of all data types that need to be collected. In many cases surveys of different data types can be collected synoptically (i.e., numbers listed are greater than the number of separate site visits required per class).

1.4   Additional Pre-existing Data

The additional data requirements discussed in section 1.3 represent a highly conservative evaluation of the additional data needs of this investigation. There exist several sources of additional DEQ and VCU data that will serve to fill some of the existing data gaps. First, as discussed, 73 of the 147 site records were inventoried but not included in the available data totals because biologists could not make confident judgment regarding their impairment status. Given that much of the southern Coastal Plain where the sites are located is used for intensive agriculture, it is probable that several of these sites will be placed in the stressed category based on further evaluation of land cover and physicochemical data. These 73 site records contain a considerable amount of needed data, most especially with regard to fish assemblages (Fig. 3).

As such, after the ongoing data inventory process on these unknown sites is complete, the amount of data required via new collections at additional stressed sites is likely to be reduced.

Addition of records into the BCI database is ongoing. DEQ's Roger Stewart searched the DEQ Comprehensive Environmental Data System on April 21, 2014 and found a total of 688 sample collections in the Coastal Plain portions of the James River and Chowan River basins since 2005 for the four water chemistry parameters required for this investigation. During the same time period and in the same region, VCU collected 87 synoptic fish and macroinvertebrate samples that have not yet been included in the database because additional information is needed for site classification. Further evaluation is needed to determine the degree to which records from the two databases represent spatially synoptic surveys; however, these unevaluated sources of data have the potential to further reduce the need for additional field surveys.
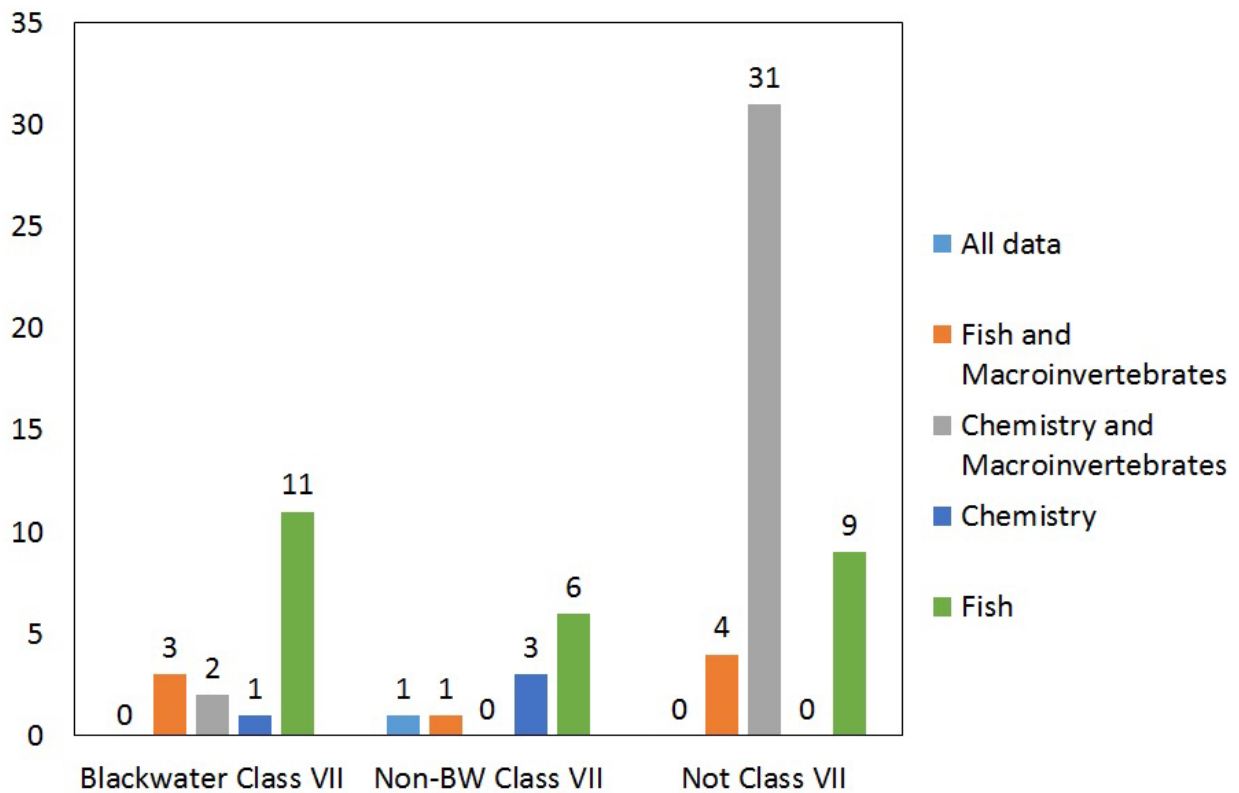

Figure 3: Available pre-existing data at study sites with unknown impairment status. No site records contained macroinvertebrate data only or fish and water chemistry only.

Task 2: Refinement of the Blackwater Habitat Protocol (BHP)

DEQ and VCU biologists conducted field evaluations of the BHP on July 31, 2013; February 20, 2014; February 27, 2014; and April 14, 2014. A total of 11 sites were evaluated. Four biologists conducted the evaluations on July 31, 2013 and February 20 and 27, 2014 (Warren Smigo, Bill Shanabruch and Tony Silvia from DEQ and Drew Garey from VCU), and three of the four participated on April 14, 2014. Each biologist independently conducted a separate habitat assessment. Following completion of the assessment at each site, participants compared and discussed their scoring process (although scores were not changed). This procedure helped identify sources of ambiguity in the metric descriptions and helped to identify differences in professional opinion that generate variability among investigators with respect to site classification. Evaluations began using the eight-metric BHP data sheet included in the 2013 report of the AAC, and the protocol was progressively altered based on input from biologists regarding areas of inherent ambiguity in the metric descriptions (see Appendices A-C for current BHP and notes regarding changes since 2013).

Mean disagreements for each site were determined by calculating all possible pairwise disagreements among investigators, as percentages of the maximum possible pairwise disagreement in BHP scores (24 points), and then averaging these values. Mean disagreements for each metric were calculated as average point differences among investigators. It is important to note that, like all distance measures, pairwise disagreements are non-independent (i.e., it is impossible for three investigators to all disagree with one another by 100 percent). Therefore, the maximum possible mean site disagreement is 66.67%, and the maximum possible mean metric disagreement is 2 points.

In order to determine if perceived disagreement among investigators was affected by the scoring method used, mean percentage disagreements for each site were also calculated using raw metric values (e.g., meters of wetland width and percent canopy cover). Large changes in percentage disagreement between raw values and scores would indicate that changing the scoring system (e.g., via use of a 10-point rather than a 3-point scale) could strongly affect site classification.

2.1 Overall Blackwater Habitat Protocol Site Scores

Eight of eleven sites visited for the BHP evaluation were designated by all investigators as blackwater, Class VII systems, and two sites were classified by all investigators as non-Class VII streams (Table 4). At the remaining site (Three Creek), investigators disagreed as to the best typology classification: three investigators designated the site as Class VII, blackwater, and one investigator designated the site as non-Class VII. Two of the three investigators designating Three Creek as Class VII indicated that the site was marginal between Class VII and non-Class VII.

Using the preliminary benchmark of 16 out of 24 points as a guideline for assigning Class VII status, results of the BHP scoring agreed with best professional judgment for five of eight sites originally designated as Class VII blackwater (mean scores: 16.0 to 20.8, Table 4, Fig. 4). The BHP protocol also correctly designated both non-Class VII streams (mean scores for Grays

Creek and Fox Branch: 10.8 and 5.8 respectively). At the three remaining sites originally designated as Class VII, blackwater (Assamoosick Swamp, Warwick Swamp, and Raccoon Creek), mean scores ranged between 15.0 and 15.8 – slightly below the 16-point threshold, although maximum scores at all three sites met or exceeded 16 (Table 4, Fig. 4). Field observations recorded at these three sites indicated that, although investigators agreed that they should be designated as Class VII, each had characteristics of both Class VII swamp systems as well as systems typically considered non-Class VII. Assamoosick Swamp and Warwick Swamp were described as large systems with extensive bordering wetlands typical of blackwater swamps but also exhibiting a main-channel area with a definitive flow path and high water velocity relative to most swamp systems. Field records showed that Raccoon Creek is bordered by extensive wetlands and exhibits generally sluggish flow, but it also has areas of moderate-to-rapid flow and a well-developed channel system along approximately 20 percent of the reach.

Table 4: Mean, minimum, and maximum scores among investigators at 11 Blackwater Habitat Protocol test sites.

	Designation Class*	Mean Score	Minimum Score	Maximum Score
Moores Swamp	SBW7	18.3	17	19
Cypress Swamp	RBW7	18.5	17	20
Otterdam Swamp	RBW7	20.8	20	22
German Swamp	RBW7	16.7	16	17
Seacorrie Swamp	RBW7	16.0	15	17
Grays Creek	SX	10.8	9	12
Fox Branch	SX	5.8	5	7
Assamoosick Swamp	RBW7	15.5	13	18
Warwick Swamp	RBW7	15.8	15	16
Raccoon Creek	SBW7	15.0	14	17
Three Creek	R?	13.0	12	14

* SBW7 = Stressed, Class VII, Blackwater; RBW7 = Reference, Class VII, Blackwater; SX = Stressed, Not Class VII; R? indicates that investigators disagreed as to the correct typology classification.
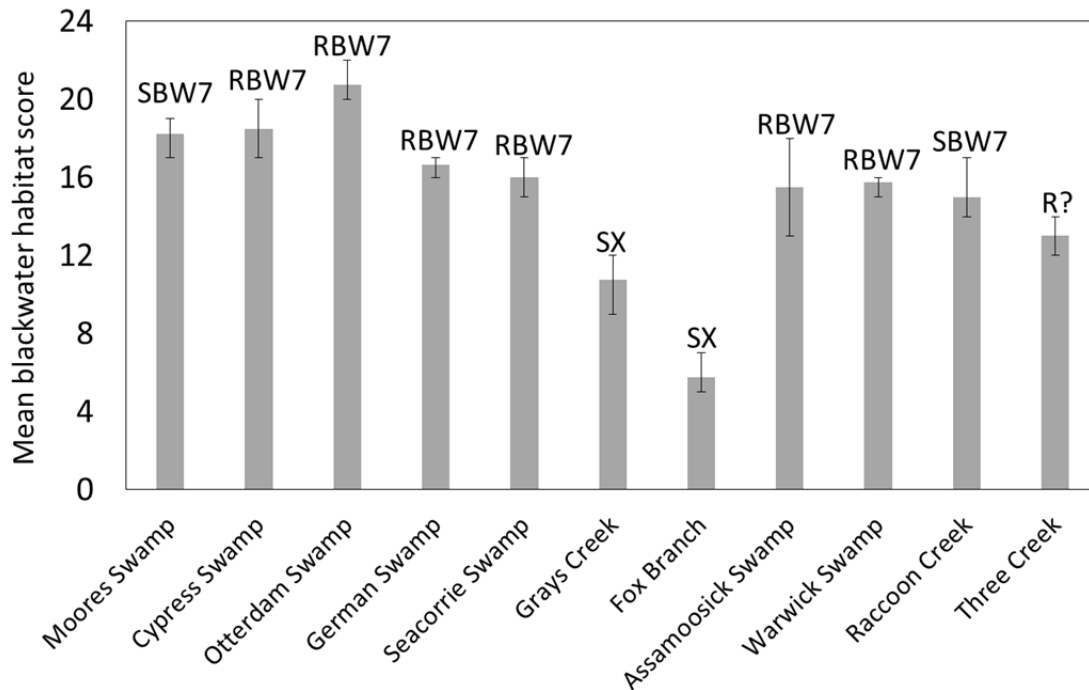
Figure 4: Mean Blackwater Habitat Protocol scores at 11 test sites. Error bars indicate range of scores among investigators. Designation classes: SBW7 = Stressed, Class VII, Blackwater; RBW7 = Reference, Class VII, Blackwater; SX = Stressed, Not Class VII; R? indicates that investigators disagreed as to the correct typology classification.

2.2 Blackwater Habitat Protocol: Disagreement among Metrics

Disagreements among investigators for each metric were generally low, with mean disagreements ranging from 0.05 points to 0.48 points (out of a maximum possible 2 points of disagreement). Forest-type scores exhibited the lowest amount of disagreement (mean: 0.05 points, Table 5, Fig. 5). The channel development, canopy cover, benthic organic matter, wetland width, and flow metrics all exhibited mean disagreements that exceeded 0.30 points, indicating that disagreements with respect to these metrics have the greatest potential to affect overall classification scores. To increase consistency in the interpretation of these metrics, several changes were made to the BHP field sheet. These changes are discussed below and listed in Appendix C.

The flow score exhibited the highest overall disagreement (mean: 0.48 points, Table 5, Fig. 5), although disagreement for this metric is expected to be greatly decreased in future investigations. After several field evaluations, it was discovered that investigators had different interpretations regarding the wording of the lowest-weighted flow category, which was originally described as "No perceptible flow" (Appendix A). Some investigators interpreted this description to mean no flow perceptible via visual observation of the water surface, whereas others interpreted it to mean that the current would not move fine benthic sediment downstream as it was disturbed and suspended in the water column. To increase consistency of interpretation, the category description was changed to "No perceptible flow/flow extremely sluggish", and the criterion

description on the reverse side of the habitat sheet was amended to indicate that "no perceptible flow" means no flow that is visible from a distance of 3 meters from the water surface.

Wetland width was highly variable along most reaches, which likely contributed to the disagreement among investigators for this metric (mean: 0.44 points, Table 5, Fig. 5). For future investigations, it is recommended that a laser rangefinder be used to measure wetland width at three or more equidistant points along the 100-m reach, and that the mean of these measurements be used for scoring.

Channel development was also interpreted differently by investigators. Some considered a channel to be a cross-sectional zone where flow was clearly greatest, whereas others considered it to be a defined bed/bank flow path. To avoid redundancy with the flow metric, it was agreed that the latter definition would be used, and the protocol category descriptions were changed accordingly (see metric 1 in Appendices A, B, and C).

The scoring system for the metric pertaining to benthic organic matter may require alteration. System benthos can be highly dominated by particulate organic matter and still not receive the highest score (3), which only occurs at 96% coverage or more. Therefore subtle variations in the estimation of the percent coverage of organic matter on the stream bottom can affect the scoring for this metric. Further investigation regarding the degree of variation in actual organic matter coverage among sites, as well as the biological relevance of this metric with respect to its relationships with fish and macroinvertebrate assemblages is needed before a sound decision can be made regarding the best scoring of this metric.

Investigators agreed that the canopy-cover metric was the most difficult to score. This estimation is most difficult in the cold season before leaf-out. It is expected that this metric will exhibit the most variation between Class VII blackwater and Class VII non-blackwater systems, but further evaluation is also needed to determine the best utility for this metric.

Table 5: Mean pairwise score disagreement among investigators for each of the eight Blackwater Habitat Protocol metrics.

Metric*	Mean Disagreement (points)	Std. Error Disagreement (points)
Forest type	0.05	0.04
Fl. plain elev.	0.15	0.08
Vegetation	0.27	0.12
Channel dev.	0.32	0.14
Canopy	0.36	0.09
Organic matter	0.44	0.11
Wetland width	0.44	0.12
Flow	0.48	0.15

*Metric names are abbreviated for space. See Appendices A and B for full metric names and descriptions.
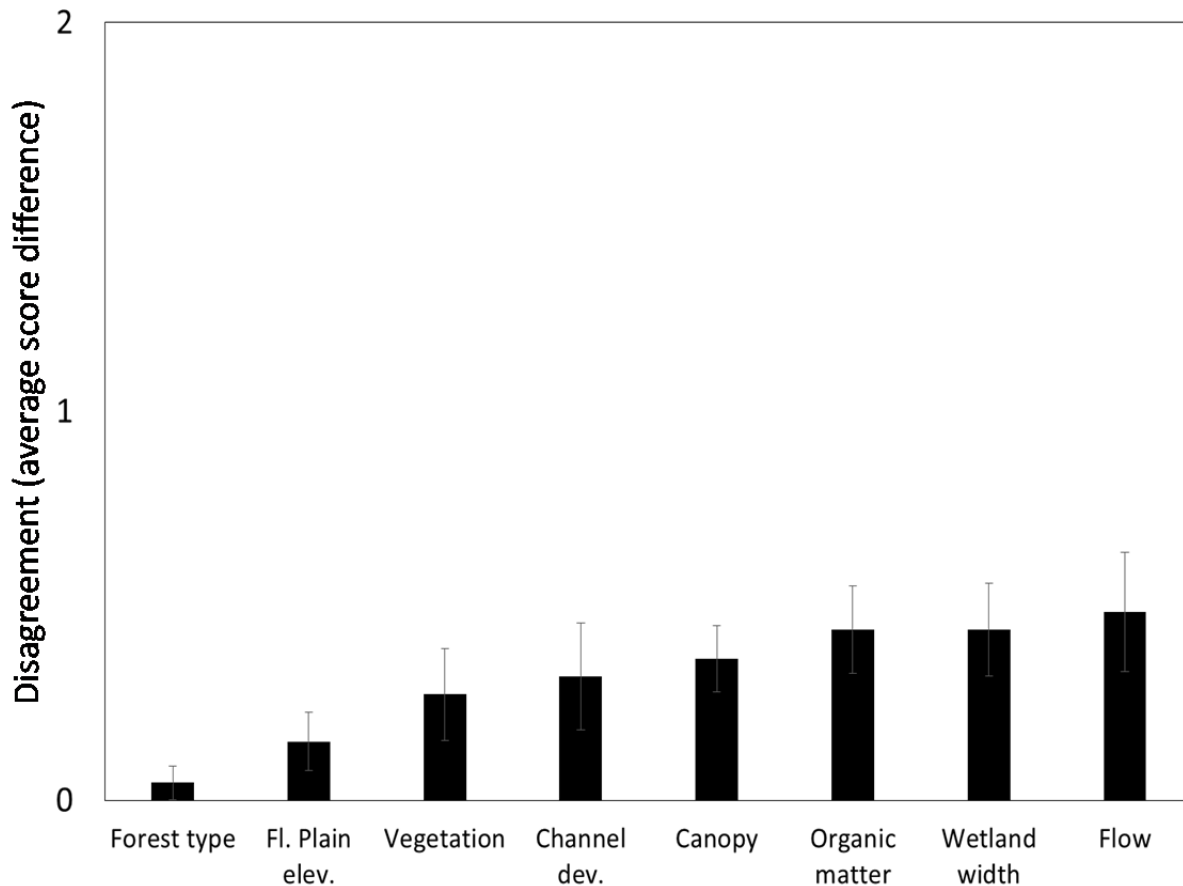
Figure 5: Mean disagreement among investigators for the eight Blackwater Habitat Protocol metrics. Error bars: +/- one standard error. Metric names are abbreviated for space. See Appendices A and B for full metric names and descriptions.

2.3 Blackwater Habitat Protocol: Disagreement among Sites

Eight of 11 sites evaluated exhibited less than ten percent mean pairwise disagreement among investigators. Three sites exhibited greater than 10 percent disagreement (Moores Swamp, Assamoosick Swamp, and Fox Branch), and one site exhibited greater than 20 percent disagreement (Fox Branch, mean disagreement: 23.6, Table 6, Fig. 6A).

Each of the three sites where mean disagreement was greater than 10 percent included relatively wide and densely forested wetland systems that were difficult to view in their entirety without a great deal of effort. In addition, high variability in system morphology and flow was noted at all three sites. Fox Branch exhibited high disagreements among investigators with respect to submerged and emergent vegetation (mean: 1.17 points), channel development (mean: 1.17 points), and wetland width (mean: 1.33 points) and also exhibited relatively high disagreement among investigators with respect to flow scores (mean: 0.5 points). Moores Swamp exhibited high disagreements among investigators for flood plain elevation (mean: 0.7 points), benthic organic matter (mean: 1.0 points), and flow scores (mean: 0.7 points). Assamoosick Swamp

exhibited high disagreement for canopy cover (mean: 0.7 points) and flow score (mean: 1.8 points).

When comparisons were made among investigator evaluations using raw values rather than metric scores, effects on mean percentage disagreements were variable among sites, reducing the mean percentage disagreement at some sites by as much five percent (Raccoon Creek) and increasing disagreement at other sites by as much as 44 percent (Seacorrie Swamp, Fig. 6B). These differences between scores and raw metric values indicate that changing the manner in which metrics are scored will affect site classification. Furthermore, because the differences between scores and raw values were variable among sites, changing the scoring system may change the perceived ordering of study sites along the wetland to stream continuum.

Table 6: Mean pairwise score disagreement among investigators for each of the 11 Blackwater Habitat Protocol evaluation sites.

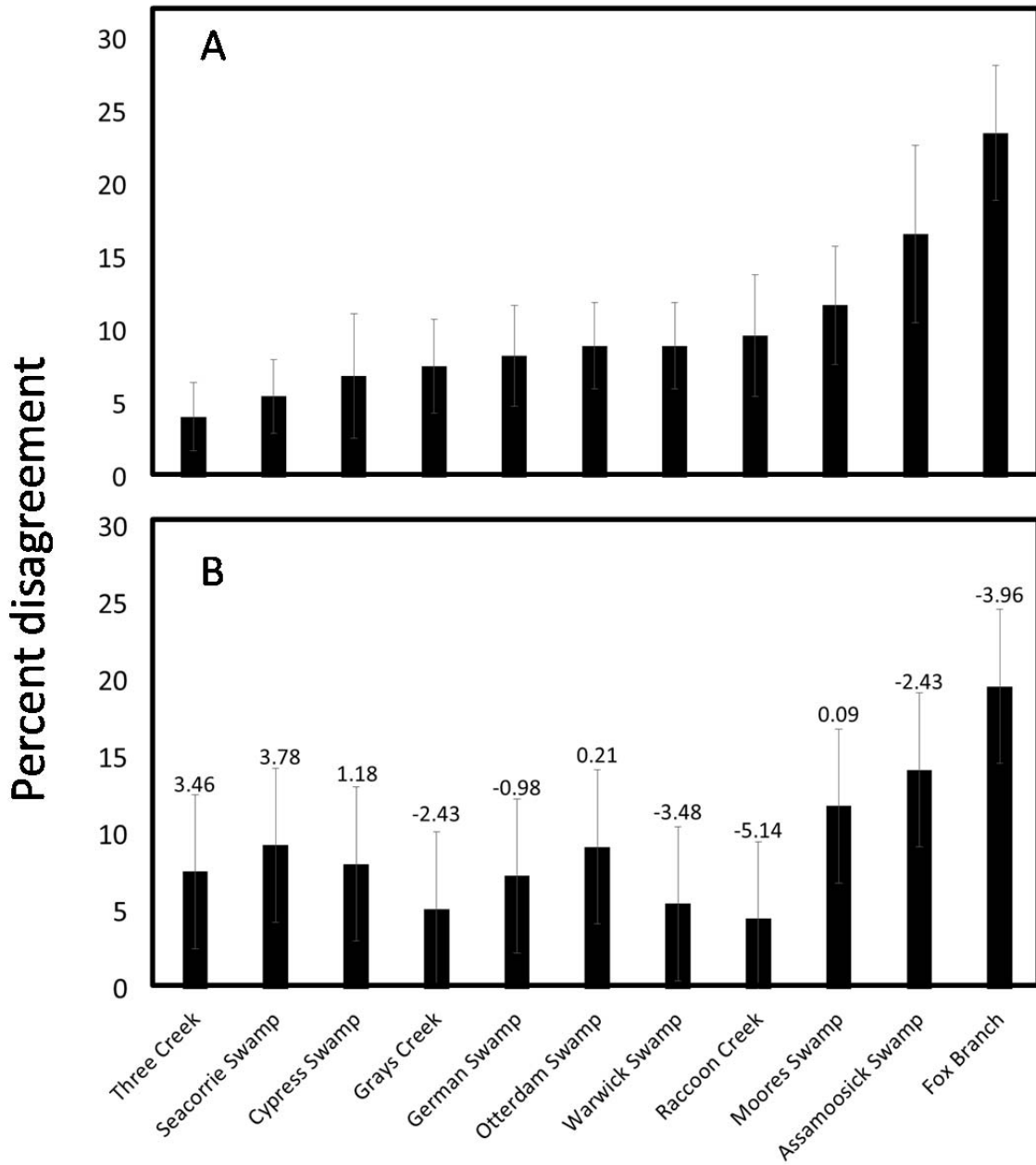|	Mean Disagreement (%)	Std. Error Disagreement (%)
Three Creek	4.2	2.3
Seacorrie Swamp	5.6	2.5
Cypress Swamp	6.9	4.3
Grays Creek	7.6	3.2
German Swamp	8.3	3.5
Otterdam Swamp	9.0	3.0
Warwick Swamp	9.0	3.0
Raccoon Creek	9.7	4.2
Moores Swamp	11.8	4.0
Assamoosick Swamp	16.7	6.1
Fox Branch	23.6	4.6

Figure 6: Mean percentage disagreement in the Blackwater Habitat Protocol metrics for the eleven test sites. Panel A: metric scores; Panel B: raw metric values. Error bars: +/- one standard error. Numbers in panel B denote changes in percent disagreement for each site between scores and raw values.

2.4 Conclusions Regarding the Current Blackwater Habitat Protocol

The observed classification performance of the current BHP and the relative agreement with the professional opinions of investigators indicate that the protocol generally supports the judgment of investigators regarding the correct system typology of the study sites. The preliminary guideline for Class VII designation of 16 is, at this point, relatively arbitrary, but it does serve as a benchmark for assessing the effectiveness of the BHP for site classification. Several sites that were originally designated by investigators as Class VII exhibited mean scores below 16, although the ranges in scores of all of these sites overlapped the 16-point benchmark. These small discrepancies between professional judgments are expected, given that the protocol is for rapid site evaluation that involves visual estimation, rather than discrete measurements of habitat characteristics (other than our recommendation that wetland widths be measured with a rangefinder). Moreover, stream and wetland systems do not exist within discrete classes, but instead they exist on a continuum from confined channel systems to open wetlands with no clearly-defined flow paths. Further study is needed to determine the relevance of the observed classification metrics for structuring benthic macroinvertebrate and fish communities. This information can then be used to determine the most important habitat characteristics for consideration in BCI construction.

The amount of disagreement among investigators is considered acceptable given the range of interpretations of each metric among the investigators at the beginning of the study. The evaluation exercise, discussions that followed it, and the resulting changes and clarifications to the BHP should increase consistency among evaluations at future sites. The interpretations of biologists on how to evaluate some metrics were, at first, quite variable (most especially flow scores), and this difference in interpretation was most probably the cause of much of the disagreement at sites that were evaluated early in the study. The changes to the BHP, which were motivated by discussions on how best to clarify each metric description, should serve to increase the overall accuracy of future evaluations. In its current form, the BHP represents a useful tool for site classification. The protocol is not intended to circumvent the professional judgment of experienced investigators because wetland stream systems are highly variable and exhibit a wide range of environmental conditions that have likely not been fully accounted for by the sites observed for this investigation. In addition, no classification metric is free from the potential confounding effects of anthropogenic impairment. It is incumbent upon the users of the protocol to determine whether conditions at the site indicate that the typology classification made using the BHP has the potential to be affected by anthropogenic factors (e.g., artificial flow diversions, removal of native riparian or in-stream vegetation). Finally, this protocol is intended for use only by experienced stream and wetland ecologists who are properly trained in its use. The variability among experienced investigators that was observed during the evaluation period highlights the importance of field-based training in the proper use of the BHP.

Task 3: Strategy for Validation of the Blackwater Habitat Protocol and Development of the Blackwater Condition Index

This section outlines a preliminary strategy for BCI development, given that the data requirements discussed in section one are satisfied. The development process includes validation and re-calibration of the BHP based on the following:

1) Habitat metrics that have been determined to be most relevant for shaping fish and macroinvertebrate assemblage structure among sites;

2) Classification of study sites into biologically relevant typology classes; and

3) Development of a multimetric index that most effectively discriminates between reference and stressed sites, as determined by a set of predetermined non-biological reference filters.

3.1 Blackwater Habitat Protocol Validation and Site Classification

Site classification will begin by determining patterns of variability among fish and macroinvertebrate assemblage structure that occur at sites that are relatively unimpaired. Final selection of these unimpaired reference sites will be determined by selecting those that do not violate any of the reference-filter criteria described in Table 1. The initial search for reference sites will be aided and informed by the large number of sites in the BCI database for which biologists have formulated initial hypotheses regarding impairment status (i.e., those initially classified as RBW7, R7 or RX). Several multivariate distance-based methods have been used extensively for classifying aquatic systems based on the assemblages present. The specific methods chosen will depend on the numerical structure of the biological datasets assembled. This method will likely include an agglomerative cluster analysis, such as used by Hawkins et al. (2000), as well as a distance-based, unconstrained ordination, such as non-metric multidimensional scaling as used by Burton and Gerritsen (2003) in the development of the Virginia Stream Condition Index.

Following classification of reference sites, statistical analyses will be performed to determine relationships between the BHP metrics and the biological assemblages at reference sites. The discrete 0-3 scores of the BHP function more like categories than on a continuous scale; therefore a categorical analysis, likely to include Multi-Response Permutation Procedure (MRPP), will be used to determine which variables provide the best site groupings, as indicated by the multivariate analyses on the biological assemblage data (see McCune and Grace 2002 for further explanation of this method). Linear discriminant analysis (LDA) will be used to determine which habitat metrics best discriminate between biological assemblage groupings among the reference sites. LDA is commonly applied to the problem of reference site classification (e.g., Poquet et al. 2009).

Stressed sites will also be designated based on the criteria listed in Table 1. Unlike reference sites, where all criteria for which data exist have to be satisfied for reference designations, stressed sites can be designated based on attainment of only one of the listed criteria. For

example, a reference system must exhibit specific conductance below 150 µS/cm and total nitrogen concentrations below 1.5 mg/L, but a system may be designated as stressed if either specific conductance exceeds 350 µS/cm or total nitrogen exceeds 3 mg/L.

Classification of stressed sites will not be based on biota because fish and macroinvertebrate communities are expected to vary in response to anthropogenic impairment. Instead, the typologies of stressed sites will be assigned based on the BHP variables determined to be the most biologically relevant for classifying reference sites. This technique assumes that the BHP metrics at stressed sites are not affected by the observed impairment. For validation of this assumption, it is important that investigators make careful note of potentially confounding anthropogenic factors and exclude sites that likely violate this assumption from the analysis.

3.2 Multimetric Index Development

A suite of candidate macroinvertebrate and fish metrics will be calculated for all sites that have been assigned statistically valid impairment and typology classes. The final list of metrics is to be determined, based on further review of the literature and on the observed assemblages in the final dataset. It will include metrics that reflect the following: richness and diversity, trophic status and functional feeding groups, ecological habit and habitat preference, pollution tolerance, and known specialization to blackwater conditions. The four fish metrics proposed in the 2013 report to AAC (number of blackwater species; catch per unit effort; number of opportunistic, generalist species; and Simpson's evenness index) will be tested and are expected to be especially effective in blackwater systems.

The discrimination efficiency of each candidate metric (i.e., the ability of the metric to discriminate between the pre-designated stressed and reference sites) will be evaluated using the method of Barbour et al. (1996), and metrics with high discrimination efficiency will be given priority for inclusion in the final BCI. The final suite of metrics to be included in the BCI will be selected primarily to maximize discrimination efficiency. Metric selection will also follow the guidelines of Schoolmaster et al. (2012) for minimizing metric redundancy as well as selecting the optimum number of metrics for maximizing the index's response to anthropogenic impairment while minimizing its response to the noise of stochastic environmental variability.

Task 4: Working with DEQ Personnel, Determine the Feasibility of Submitting a Proposal to EPA for Blackwater Condition Index Development

Based on recommendations from DEQ biologists, EPA Region 3 has indicated a preliminary agreement to allocate an additional $10,000 for the development of new biological assemblage data and water chemistry data to be used for this study. These funds, along with continued support from the AAC, are expected to fulfill the existing data needs for BCI development. The data-inventory computer program developed for this project should ensure that additional monetary and time expenditures for this project are extended in the most efficient manner possible.

Timeline and Recommendations for Index Development

Figure 7 presents an approximate timeline for future activities recommended for continued development of the BHP and BCI protocols. The initial data inventory should be complete by August 1, 2014. Based on this inventory, additional field activities will be planned in order to optimize the use of pre-existing data. Pending sufficient funding, additional data collection will begin in August 2014. Fish collections along with concurrent BHP evaluations and physicochemical water quality measurements will begin during this period. Late summer through fall is ideal for fish collection because water is low and warm, making capture easier and ensuring that fish are responsive to electrofishing. Fish collection will end in approximately mid-November when water temperatures drop below 5°C and fish become inactive. Land cover analysis will begin in August 2014 as well and will likely continue through mid-November. After the fish dataset is complete, analysis of fish data and construction of a preliminary fish BCI will begin; these activities will continue through late April 2015 or early May 2015. Macroinvertebrate collection and, if needed, concurrent physicochemical measurements and BHP surveys will begin in February 2015 and continue through April 30, 2015. Late winter to spring is ideal for macroinvertebrate collection because most taxa are large and can be most easily identified. After the macroinvertebrate dataset is complete, analysis of macroinvertebrate metrics and comparison of fish and macroinvertebrate models will begin. The prototype BCI model will likely be complete by fall 2015.

Figure 7: Timeline for recommended study activities. Bold dates in parentheses indicate the end of activity phases. **(Nov. 15, 2014)*** and **(April 30, 2015)*** indicate the end times for fish and macroinvertebrate surveys, respectively, which are highly variable and dependent on weather conditions. Red letters along the timeline indicate approximate due dates for progress reports. 1: due date for final draft of this report to DEQ (June 1, 2014); 2: due date for 2015 report to AAC (June 2015); 3: due date for report detailing use of EPA funding related to the project (Sept. 2015).

Literature Cited

Barbour, M.T., J. Gerritsen, G.E. Griffith, R. Frydenborg, E. McCarron, J.S. White, and M.L. Bastian. 1996. A framework for biological criteria for Florida streams using benthic macroinvertebrates. *Journal of the North American Benthological Society* **15(2)**: 185-211.

Burton J. and J. Gerritsen. 2003. *A Stream Condition Index for Virginia Non-Coastal Streams*. Tetra Tech, Inc., Owing Mills, Md. Report prepared for U.S. EPA Office of Science and Technology, Office of Water, Washington, D.C.; U.S. EPA Region 3 Environmental Services Division, Wheeling, W.V.; and Virginia Department of Environmental Quality, Richmond, Va.

Garman, G.C., L.A. Smock, A.L. Garey, and S.P. McIninch. 2013. *Development of Aquatic Life Use Assessment Protocols for Class VII Waters in Virginia: 2013 Report of the Academic Advisory Committee for Virginia Department of Environmental Qualit*y. SR53-2013. Virginia Water Resources Research Center, Blacksburg, Va. Available at http://vtechworks.lib.vt.edu/handle/10919/25871 (accessed September 22, 2014).

Garman, G.C., L.A. Smock, and A.L. Garey. 2012. *Development of Aquatic Life Use Assessment Protocols for Class VII Waters in Virginia: Report of the Academic Advisory Committee for Virginia Department of Environmental Qualit*y. SR52-2012. Virginia Water Resources Research Center, Blacksburg, Va. Available at http://vtechworks.lib.vt.edu/handle/10919/25871 (accessed September 22, 2014).

Hawkins, C.P., R.H. Norris, J.N. Hogue, and J.W. Feminella. 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. *Ecological Applications* **10**: 1456-1477. DOI: 10.1890/1051-0761(2000)010[1456:DAEOPM]2.0.CO;2.

McCune, B., J.B. Grace, and D.L. Urban. 2002. *Analysis of Ecological Communities*. MjM Software Design, Gleneden Beach, Ore.

Poquet, J.M., J. Alba-Tercedor, T. Puntí, M.d.M. Sánchez-Montoya, S. Robles, M. Álvarez, C. Zamora-Muñoz, C.E. Sáinz-Cantero, M.R. Vidal-Abarca, M.L. Suárez, M. Toro, A.M. Pujante, M. Rieradevall and N. Prat. 2009. The MEDiterranean Prediction And Classification System (MEDPACS): an implementation of the RIVPACS/AUSRIVAS predictive approach for assessing Mediterranean aquatic macroinvertebrate communities. *Hydrobiologia* **623**: 153-171. DOI: 10.1007/s10750-008-9655-y.

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at: http://www.R-project.org/.

Schoolmaster, Jr., D.R., J.B. Grace, and E.W. Schweiger. 2012. A general theory of multimetric indices and their properties. *Methods in Ecology and Evolution* **3**: 773-781. DOI: 10.1111/j.2041-210X.2012.00200.x.

Appendix A: Field Datasheet for the Blackwater Habitat Protocol (BHP)

Metric	Subcategories (metrics 1 and 2)	Raw values	Weights	Scoring
1) Channel Development		Percentage of channel	Weight	Percentage*weight
	One defined bed/bank flow path	_____	2	_____
	Multiple defined bed/bank flow paths	_____	1	_____
	No defined bed/bank flow paths	_____	0	_____
		Total weighted percentage:		_____
Comments:			Metric score:	_____
2) Flow Score		Percentage of channel	Weight	Percentage*weight
	No perceptible flow /flow extremely sluggish	_____	0	_____
	Slow flow (see reverse for description)	_____	1	_____
	Moderate, laminar flow (approx 0.1-0.2 m/s)	_____	2	_____
	Rapid, laminar flow (>0.2 m/s)	_____	3	_____
	Rapid, turbulent flow	_____	4	_____
		Total weighted percentage:		_____
Comments:			Metric score:	_____
3) Flood plain elevation		Value (cm): _____	Metric score:	_____
	Comments:			
4) Submerged and emergent vegetation	abs: 0, rare:1, common:2, abundant: 3		Metric score:	_____
	Comments:			
5) Benthic organic matter	Comments:	Percentage: _____	Metric Score:	_____
6) Forest type	(see criteria on reverse		Metric score:	_____
Dominant tree/vegetation cover species:		Comments:		
7) Wetland width	Percentage of 200-m area as wetland: _____		Metric score:	_____
Comments:				
8) Canopy	Percentage of overhead cover: _____		Metric score:	_____
Comments:				

Class 7 Status   __ Class 7 __Not Class 7

Stream type:

Comments:

Appendix B: Metric Interpretation and Scoring for the Blackwater Habitat Protocol (BHP)

Metric Explanations	Metric value calculation	Scoring criteria			
1) Channel Development Percentages of 100m reach comprised of bed/bank flow paths , multiple flow paths (i.e. braided-channel systems) and undefined flow paths (large standing-water areas).	Metrics 1 and 2: multiply perecentages for each subcategory by the subcategory weights to get weighted percentages, sum these values to get the total weighted percentage, and compare the total weighted percentage to the scoring criteria values listed to get the metric score	Total Weighted %	200	145-199	100-144	0-99
		Metric score	0	1	2	3
2) Flow Score Percentages of 100m reach comprised of each flow regime indicated on the field data sheet. Slow flow is clear evidence of slow, downstream flow visible from 3 meters away from wetted surface.		Total Weighted %	155 and greater	100-154	20-99	0-19
		Metric score	0	1	2	3
3) Flood plain elevation Average elevation of riparian floodplain above the wetted channel. Evaluate for a distance of 50m from either side of wetted area. Elevation may be highly variable. If so, estimate mean elevation and make a note regarding variability.	Compare the estimated value to the scoring criteria values	Elevation (cm)	40 and greater	25-39	10-24	0-9
		Metric score	0	1	2	3
4) Submerged and emergent vegetation Commonness of submerged and/or emergent vegetation within the wetted area	Score based on criteria indicated	Category	Absent	Rare	Common	Abundant
		Metric score	0	1	2	3
5) Benthic organic matter Percentage of the benthic area covered by large woody debris, coarse particulate organic matter, or fine particulate organic matter.		Percentage	0-49	50-84	85-95	96 and greater
		Metric score	0	1	2	3
6) Forest type Commonness of blackwater-indicator trees: *Nyssa aquatica* (water tupelo) and *Taxodium distichum* (bald cypress). Scores: Absent (0)- neither species present, Rare (1) other tree species are dominant, but either bald cypress or water tupelo are present, Common (2)- either bald cypress or water tupelo are present, and codominant with other tree species OR both species occur and are relatively common, but not the dominant species, Abundant (3)- water tupelo and/or bald cypress are the dominant tree species.	Score based on criteria indicated	Category	Absent	Rare	Common	Abundant
		Metric score	0	1	2	3
7) Wetland width Within 100 meters, laterally, from the center of the wetted area (200m total and including the wetted area), the percentage by area that is likely wetland (if variable, average over the 100m reach).	Compare the estimated percentage scoring criteria percentages	Percentage:	0-24	25-44	45-74	75 and greater
		Metric score	0	1	2	3
8) Canopy Percentage of the overhead view of the sky above the wetted area that, at full leaf-out, would be obstructed (requires estimation in late-fall to early spring).	Compare the estimated percentage scoring criteria percentages	Percent	0-19	20-49	50-69	70 and greater
		Metric score	0	1	2	3

Appendix C: Notes Regarding Changes to the Blackwater Habitat Protocol (BHP) Since 2013

Metric	Changes
1	Changed "channel" to "bed/bank flow paths".
2	Increased description on criteria page.
2	Changed "no flow" to "sluggish"; added approximate velocities.
3	Increased description on criteria page.
4	Added scoring criteria to field sheet.
6	Added note to see criteria page.
6	Added prompt for entry of actual species.
7	Changed "width" blank to read "Percentage of 200-m area as wetland".
8	Criteria page - Noted that canopy cover is above wetted area.
	Added section at bottom for Class 7 status, stream type, and general comments.
General	Added lines for comments throughout