SRNL-STI-2014-00050

DEVELOPMENT OF ECOLOGICAL REFERENCE MODELS AND AN ASSESSMENT FRAMEWORK FOR STREAMS ON THE ATLANTIC COASTAL PLAIN

SERDP Project RC-1694

October, 2014

Michael Paller
Savannah River National Laboratory, Aiken, SC, USA

Jack Feminella
Ely Kosnicki
Stephen Sefick
Miller Jarrell
Department of Biological Sciences, Auburn University, AL, USA

Tracey Tuberville
Dean Fletcher
Andrew Grosse
Bess Harris
Sean Sterrett
Savannah River Ecology Laboratory, Aiken, SC, USA

Blair Prusha
University of Georgia, Athens, GA, USA

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)* 03-09-2014	**2. REPORT TYPE** Final	**3. DATES COVERED** *(From - To)* Feb 2009 – Sep 2014

4. TITLE AND SUBTITLE Development of Ecological Reference Models and an Assessment Framework for Streams on the Atlantic Coastal Plain	**5a. CONTRACT NUMBER** **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER**
6. AUTHOR(S) Michael Paller, Jack Feminella, Ely Kosnicki, Stephen Sefick, Miller Jarrell, Tracey Tuberville, Dean Fletcher, Andrew Grosse, Bess Harris, Sean Sterrett, Blair Prusha	**5d. PROJECT NUMBER** **5e. TASK NUMBER** **5f. WORK UNIT NUMBER**
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Savannah River National Laboratory, Building 773-42A, Aiken, SC 29808 Dept. of Biological Sciences, 101 Life Sciences, Auburn University, Auburn, AL 36849-5407 Savannah River Ecology Laboratory, P.O. Drawer E, Aiken, SC 28902	**8. PERFORMING ORGANIZATION REPORT NUMBER** SRNL-STI-2014-00050
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Strategic Environmental Research and Development Program 4800 Mark Center Drive, Suite 17D08, Alexandria, VA 22350	**10. SPONSOR/MONITOR'S ACRONYM(S)** SERDP **11. SPONSOR/MONITOR'S REPORT NUMBER(S)** RC-1694

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT Military installations in the Sand Hills ecoregion of the Atlantic Coastal Plain protect unique ecosystems including blackwater streams. The Department of Defense is committed to the recovery of degraded ecosystems but lacks reference models that specify end-states representative of least disturbed conditions. This project developed ecological reference models and assessment frameworks for evaluating the ecological integrity of coastal plain streams. The models were based on biological, habitat, and watershed data collected from undeveloped areas within major DoD installations and other protected lands. Ecological reference models were developed for fish communities, macroinvertebrate communities, and stream hydrogeomorphic conditions. The models provided a basis for multimetric indices, predictive models, and other assessment frameworks that measured deviations from reference model objectives. Habitat data were analyzed to interpret biological responses and identify factors that could affect recovery success. Performance evaluations indicated that most methods performed well at highly disturbed sites and that the concomitant use of fish and macroinvertebrate methods resulted in greater ability to detect degradation. The assessment frameworks provide tools that can be used to monitor the ecological health of wadeable streams and assess their recovery to reference states characteristic of least disturbed conditions or best attainable conditions imposed by factors that limit recovery.

15. SUBJECT TERMS
Assemblage, Assessment, Bioassessment, Blackwater Stream, Coastal Plain, Fish, Habitat, Hydrogeomorphology, Macroinvertebrate, Multimetric Index, Predictive Model, Reference Model, Sand Hills, Southeast, Stream, Watershed

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Michael Paller
a. REPORT UU	**b. ABSTRACT** UU	**c. THIS PAGE** UU	UU	214	**19b. TELEPHONE NUMBER** *(include area code)* 803-725-5250

INSTRUCTIONS FOR COMPLETING SF 298

1. REPORT DATE. Full publication date, including day, month, if available. Must cite at least the year and be Year 2000 compliant, e.g. 30-06-1998; xx-06-1998-, xx-xx-1998.

2. REPORT TYPE. State the type of report, such as final, technical, interim, memorandum, master's thesis, progress, quarterly, research, special, group study, etc.

3. DATES COVERED. Indicate the time during which the work was performed and the report was written, e.g., Jun 1997 - Jun 1998; 1-10 Jun 1996; May - Nov 1998; Nov 1998.

4. TITLE. Enter title and subtitle with volume number and part number, if applicable. On classified documents, enter the title classification in parentheses.

Ba. CONTRACT NUMBER. Enter all contract numbers as they appear in the report, e.g. F33615-86-C-5169.

5b. GRANT NUMBER. Enter all grant numbers as they appear in the report, e.g. AFOSR-82-1234.

5c. PROGRAM ELEMENT NUMBER. Enter all program element numbers as they appear in the report, e.g. 61101A.

5d. PROJECT NUMBER. Enter all project nurnbers as they appear in the report, e.g. 1F665702D1257; ILIR.

5e. TASK NUMBER. Enter all task numbers as they appear in the report, e.g. 05; RF0330201; T4112.

5f. WORK UNIT NUMBER. Enter all work unit numbers as they appear in the report, e.g. 001; AFAPL30480105.

6. AUTHOR(S). Enter name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. The form of entry is the last name, first name, middle initial, and additional qualifiers separated by commas, e.g. Smith, Richard, J, Jr.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES). Self-explanatory.

8. PERFORMING ORGANIZATION REPORT NUMBER. Enter all unique alphanumeric report numbers assigned by the performing organization, e.g. BRL-1234; AFWL-TR-85-4017-Vol-21-PT-2.

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES). Enter the name and address of the organization(s) financially responsible for and monitoring the work.

10. SPONSOR/MONITOR'S ACRONYM(S). Enter, if available, e.g. BRL, ARDEC, NADC.

11. SPONSOR/MONITOR'S REPORT NUMBER(S). Enter report number as assigned by the sponsoring/ monitoring agency, if available, e.g. BRL-TR-829; -21 5.

12. DISTRIBUTION/AVAILABILITY STATEMENT. Use agency-mandated availability statements to indicate the public availability or distribution limitations of the report. If additional limitations/ restrictions or special markings are indicated, follow agency authorization procedures, e.g. RD/FRD, PROPIN, ITAR, etc. Include copyright information.

13. SUPPLEMENTARY NOTES. Enter information not included elsewhere such as: prepared in cooperation with; translation of; report supersedes; old edition number, etc.

14. ABSTRACT. A brief (approximately 200 words) factual summary of the most significant information.

15. SUBJECT TERMS. Key words or phrases identifying major concepts in the report.

16. SECURITY CLASSIFICATION. Enter security classification in accordance with security classification regulations, e.g. U, C, S, etc. If this form contains classified information, stamp classification level on the top and bottom of this page.

17. LIMITATION OF ABSTRACT. This block must be completed to assign a distribution limitation to the abstract. Enter UU (Unclassified Unlimited) or SAR (Same as Report). An entry in this block is necessary if the abstract is to be limited.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ACRONYMS

ANNA	assessment by nearest neighbor analysis
ANOVA	analysis of variance
BAC	best-attainable condition
BAM	best attainable model
BC	Bray Curtis
BI	biotic index
BOM	benthic organic matter
BPJ	best professional judgment
CCA	cannonical correspondence analysis
CQI	community quality index
CWD	coarse woody debris
DCA	detrended correspondence analyses
DEM	digital elevation model
DO	dissolved oxygen
DoD	Department of Defense
DOE	Department of Energy
DQO	data quality objectives
EPT	Ephemeroptera, Plecoptera and Trichoptera
GAMMI	Georgia Sand Hills macroinvertebrate multi-metric index
GDNR	Georgia Department of Natural Resources
GIS	geographic information system
GLM	generalized linear model
HGM	hydrogeomorphic
ISA	indicator species analysis
LDA	linear discriminant analysis
LDC	least disturbed condition
LDM	least disturbed model
LIDAR	light detection and ranging
LOOCV	leave one out cross-validation
LULC	land use/land cover
MDM	minimally disturbed model
MMI	multimetric index
NCBI	North Carolina biotic index
NCDENR	North Carolina Department of Environmental and Natural Resources
NMC	number misclassified
NMDS	nonmetric multidimensional scaling
O/E	observed/expected
OUT	operational taxonomic units
PAM	partitioning around mediods
PC	principal component
PCA	principal component analysis
PLS	partial least squares regression
RF	random forest
RIVPACS	river invertebrate prediction and classification System
SCDHEC	South Carolina Department of Health and Environmental Control
SD	standard deviation
SE	standard error
SERDP	Strategic Environmental Research and Development Program
SH	Sand Hills

SON	Statement of Need
SOP	standard operating procedures
SRNL	Savannah River National Laboratory
SRS	Savannah River Site
UGA	University of Georgia
USEPA	US Environmental Protection Agency
USFWS	US Fish and Wildlife Service
USGS	US Geological Survey
VIP	variable importance in projection
YOY	young of the year

LIST OF KEYWORDS

Assemblage
Assessment
Bioassessment
Blackwater Stream
Coastal Plain
Fish
Habitat
Hydrogeomorphology
Macroinvertebrate
Multimetric Index
Predictive Model
Reference Model
Sand Hills
Southeast
Stream
Watershed

ACKNOWLEDGEMENTS

This research was supported wholly by the U.S. Department of Defense, through the Strategic Environmental Research and Development Program (SERDP) under project RC-1694. Performers of this project would like to thank DoD natural resource managers in Fort Benning, Fort Bragg, and Fort Gordon, with special thanks to Hugh Westbury, Charles Bryan, and Robert Drumm. Appreciation is also extended to Colby Farrow for assisting with field sampling, Mary Freeman for helping with metric selection, and Bud Freeman for assisting with fish taxonomy.

1.0 ABSTRACT

Objective
There are several large military installations in the Sand Hills ecoregion of the Atlantic Coastal Plain that protect unique and valuable ecosystems including blackwater streams, a distinctive resource that supports a high diversity of vertebrates and invertebrates. Degradation of coastal plain streams can result in a significant loss of biodiversity and adversely affect ecological services such as nutrient cycling and natural mechanisms of water purification. The Department of Defense is committed to the recovery of degraded ecosystems but is constrained by a lack of reference models that specify end-states representative of least disturbed conditions.

This project had three objectives: 1) develop ecological reference models that quantify key fish and benthic macroinvertebrate variables representative of least disturbed conditions in wadeable Sand Hills streams; 2) assess relationships between the reference models and important habitat variables at watershed and stream reach habitat scales, 3) develop assessment frameworks based on the reference models that can be used to evaluate and communicate the status of coastal plain streams in relation to reference model objectives.

Technical Approach
The reference models were based on samples from stream reaches that represented least disturbed conditions, although some degraded sites were also sampled to characterize disturbance gradients. The sites were in Fort Benning GA, Fort Bragg NC, Fort Gordon SC, the Savannah River Site (a Department of Energy reservation in SC), state forests, a national wildlife refuge, and The Nature Conservancy lands. Digital elevation models and National Landcover Data were used to delineate watersheds, estimate disturbance levels, and calculate important watershed features. Field data were collected to characterize instream habitat including substrate, channel morphometry, snags and other structure, hydrology, and water chemistry. Fish assemblage data were collected by backpack electrofishing, and benthic macroinvertebrate assemblage data were collected using a multiple habitat sampling protocol.

Reference site screening criteria were developed to exclude disturbed sites that could bias recovery objectives. Three methods were combined to create a tiered system for definitive reference site selection that resulted in the identification of primary reference sites, which represented biological conditions of least disturbance useful for the development of reference models for full ecosystem recovery and secondary reference sites, which represented conditions suitable for the development of best attainable models useful where prevailing or legacy land uses prevent full recovery. The reference models for fish and macroinvertebrate assemblages served as benchmarks in assessment frameworks that measured the extent to which anthropogenic disturbance caused deviations from the least disturbed state. Assessment frameworks included multimetric indices that summarized scores from selected metrics (i.e., ecologically important variables that change predictably in response to human disturbance) and predictive models that related the observed taxonomic composition to the taxonomic composition expected under least disturbed conditions. The variability, sensitivity, accuracy, responsiveness, and statistical power of both types of frameworks were measured to evaluate model performance.

Results
Ecological reference models were developed for fish communities, macroinvertebrate communities, and stream hydrogeomorphic conditions. These models were the basis for three assessment frameworks based on fish assemblages, several assessment frameworks based on macroinvertebrate assemblages, and one assessment framework based on stream hydrogeomorphology. Habitat data were analyzed to provide background environmental information useful for interpreting biological responses and identifying factors that could affect recovery success. Data requirements for the assessment methods range from minimal to moderate and ease of implementation from simple to more complex. Performance evaluations indicated

that most methods performed well at highly disturbed sites but sometimes evaluated reference sites as disturbed. They also indicated that the concomitant use of fish and macroinvertebrate methods produced more accurate assessments with greater ability to detect environmental degradation and that bioassessment accuracy was increased by using more than one assessment method for each taxonomic group since the methods differed in relative ability to detect different types of community responses to degradation.

Benefits

The assessment frameworks provide an array of economical tools that can be used by DoD managers with varying resources and type of information to monitor the ecological health of wadeable streams and assess their recovery to reference conditions characteristic of the highest expectations for the ecoregion or, where necessary, best attainable conditions imposed by factors that limit full recovery. Because the frameworks represent different biotic communities as well as abiotic features, they provide a basis for comprehensive evaluations that include key stream organisms as well as the habitat conditions needed to support them. These tools can contribute to weight-of-evidence risk assessments and to monitoring programs that assess current conditions and changes associated with recovery efforts.

2.0 OBJECTIVE

This project (RC-1694) was conducted in response to SERDP SON Number: SISON-09-01, "Development of science-based recovery objectives for ecological systems in the southeastern United States." The intent of SISON-09-01 was to "develop the science to define and support recovery objectives that result in ecologically appropriate, mission supportive, and achievable end states and trajectories for southeastern United States ecological systems at multiple spatial and temporal scales." Research areas discussed under SISON-09-01 included the development of ecological reference models and the development of assessment frameworks. The objectives of RC- 1694 were 1) develop ecological reference models that specify the central tendency, variability, and other properties of key biotic variables representative of least disturbed conditions in Sand Hills streams; 2) assess relationships between the reference models and important habitat variables at watershed and stream/riparian habitat scales, and 3) develop assessment frameworks for evaluating and communicating the status of coastal plain streams in relation to reference model objectives. These objectives, including the focus on blackwater streams, a particularly important and threatened ecosystem type, are congruent with the requirements and intent of SISON-09-01.

3.0 BACKGROUND

Ecosystem degradation on Department of Defense (DoD) bases can jeopardize the ability to sustain military training and testing in addition to diminishing the nation's resources. The DoD is committed to the recovery of degraded ecosystems with the ultimate objective of maintaining sustainable water and soil dynamics and supporting appropriate densities and diversities of native plants and animals. Achievable recovery objectives that are ecologically appropriate and compatible with DoD mission requirements are needed for successful recovery. It is also necessary to have assessment frameworks for evaluating ecosystem status, measuring progress towards recovery objectives, and identifying useful recovery strategies.

The DoD has several large installations in the southeastern United States. Ecological systems in this region support high biodiversity, some of which is threatened or endangered. For example, nearly 40 percent of the fish species in North American are currently in jeopardy, with the southeastern United States containing especially high numbers of endangered species (Warren et al. 2000). Southeastern ecosystems are under pressure because of high population densities and extensive development that have degraded and fragmented the natural landscape. DoD installations in the Southeastern Plains (Level III) ecoregion protect some of the largest remaining relatively intact ecosystems in the area. These include longleaf pine (*Pinus palustris*) ecosystems, bottomland hardwoods and associated systems, and inland aquatic and wetland systems (HydroGeoLogic 2007). In the latter category are blackwater streams, a unique resource on the Southeastern Plains that has not been as thoroughly studied as other stream types. The Southeastern Region TER-S Workshop in 2007 recommended "the highest priority focus should be on freshwater systems, especially blackwater streams" (HydroGeoLogic 2007). Degradation of blackwater streams can cause biodiversity losses and adversely affect nutrient cycling, natural water purification, and other valuable ecological functions of floodplain ecosystems.

Ecological reference models are a prerequisite for the recovery of degraded ecosystems. Such models specify the central tendency, variability, and other properties of key biotic variables that are representative of relatively undisturbed conditions. They are needed to specify achievable recovery objectives that are ecologically appropriate and compatible with DoD mission requirements. In this study ecological reference models were developed for wadeable upper Southeastern Plains streams in NC, SC, and GA. Models were based largely on the characteristics of sample sites in reference stream reaches that represented least disturbed conditions, although a lesser number of moderately and highly disturbed sites were also sampled. Multiple sites were used to represent reference conditions because of sampling variability and natural variability related to geography, stream size, and other habitat features. Screening criteria were developed to exclude sites that were not minimally disturbed from the reference site pool.

Biological reference models were emphasized because biota integrate their environment and have intrinsic value representative of ecological health. Biotic communities incorporate habitat information at 3 hierarchical scales:

1. *Landscape*. This is the geographic area surrounding the watershed emphasizing connectivity to source pools and dispersal routes. Stream segments relatively isolated from source pools of colonizing aquatic organisms may be relatively impoverished compared with better connected segments and may recover more slowly from disturbance because of slower recolonization.
2. *Watershed*. Characteristics of the stream watershed or catchment (e.g., size, topographic gradient, amount of developed land) upland of the riparian zone affects instream habitat by affecting the hydrological regime and water quality.
3. *Stream/riparian habitat*. Stream channel and riparian habitat provide direct support for aquatic organisms, such that departure from natural conditions will have substantial effects on stream biotic structure and function.

3.1 Lotic Ecosystems on the Southeastern Coastal Plain

Several types of streams occur on the southeastern coastal plain including alluvial streams, seepage streams, spring-fed streams, and blackwater streams (Patrick 1996, Whitney et al. 2004, Figure 1). Seepage streams are typically small, nutrient-poor, and receive most of their water via groundwater seepage from their banks. Spring-fed streams are fed by ground water, which may be mineral rich if derived from limestone springs. Although important in some regions of the Southeast (e.g., central Florida), spring-fed streams are rare or absent throughout much of the southeastern Atlantic Coastal Plain. Alluvial streams occur in areas with relatively impervious soils and high surface runoff. They typically carry heavy sediment loads, have muddy substrates, and show wide fluctuations in flow. Alluvial streams offer diverse habitat for vertebrates and invertebrates and have a trophic base supported mainly by allochthonous detritus. However, blackwater streams are the most characteristic streams of the Southeastern Plains and among the most important and unique ecosystems on DoD holdings in the Region. These low-gradient, slow-flowing streams are fed by water seeping through sandy soils that underlie floodplains and swamps. The water is usually acidic, stained by decaying organic matter, and carries little sediment in undisturbed systems. The color is produced primarily by dissolved organic carbon (Sabater et al. 1993, Carlough 1994). Stream bottoms are typically sandy. Snags and other large woody material form debris dams that play an important role in detrital dynamics and provide habitat for invertebrates and fish (Benke et al. 1985).

Blackwater stream

Alluvial stream

Seepage stream

Figure 1. Stream types found within the upper coastal plains of North Carolina, South Carolina, and Georgia.

Strong connectivity with their floodplains mediates many of the ecological processes in blackwater streams (Meyer et al. 1997). Dense floodplain forests contribute substantial organic matter to the flood plain, which is subsequently processed by bacteria and invertebrates. Runoff and inundation of the flood plain by seasonal floods make fluvial fine particulate organic matter and dissolved organic carbon available to stream organisms and drive blackwater stream food chains, which are largely allochthonous rather than autochthonous, although the importance of instream primary productivity tends to increase with stream size. The importance of fine particulate organic matter is reflected in the composition of blackwater stream macroinvertebrate communities, which are usually dominated by collector-gatherer and collector-filterer trophic guilds, with lesser numbers of shredders, scrapers, and predators (Rader et al., 1994, Patrick 1996).

Blackwater streams support a diverse vertebrate and invertebrate fauna, the latter of which is poorly known. A one-year study in Upper Three Runs, a relatively undisturbed blackwater stream on the Savannah River Site (an 800 km^2 Department of Energy [DoE] reservation in SC) identified at least 551 species of aquatic insects, including 52 new species and two new genera (Morse et al. 1980 and 1983), which, at the time, was the highest species richness for any North American stream of comparable size.

Floyd et al. (1993) later identified over 650 species from this stream, including 93 species of caddisflies. Streams on the Savannah River Site also support over 60 species of fish, and streams on nearby Fort Gordon (DoD) support at least 44 species (Marcy et al. 2005). Headwater fish assemblages in blackwater streams are relatively species rich because of comparatively stable water flows resulting from steady groundwater recharge (Paller 1994). Southeastern Plains streams also support numerous reptiles and amphibians.

Coastal plains streams are ideal for the development of ecological reference models because they are intrinsically important, highly visible, and of strong public interest. In addition, the integrity of stream ecosystems is a surrogate for the integrity of other ecosystems because of the connectivity between streams and their watersheds. Stream flood plains filter contaminants, supply organic carbon, and provide habitat for critical life stages of aquatic organisms. Upland ecosystems within the watershed affect the quality and quantity of water flowing through the flood plain and into the stream. Streams are also influenced by regional and global factors that affect climate, movement of atmospheric pollution, and dispersal or extirpation of species.

3.2 Ecological Reference Models

Ecological reference models should represent the maximum level of process integrity, functionality, structural complexity, and biological diversity that can be achieved under prevailing conditions of land use and regional ecological integrity. However, the term “reference condition” has assumed several different but related meanings (Stoddard et al. 2006a). Ideally, an ecological reference model depicts the natural structure and function of the biota in the absence of disturbance by modern humans (with the possible exception of traditional uses by indigenous peoples). In practice, however, the model is usually a minimally disturbed model (MDM), which depicts the structure and function of the biota in the absence of significant disturbance and acknowledges that it is usually impossible to completely avoid the influence of human activities. A strict MDM for North American ecosystems would represent conditions that prevailed prior to European settlement with the exception of low (and ecologically insignificant) concentrations of pollutants deposited by atmospheric processes. The least disturbed model (LDM) represents the structure and function of the biota “under the best available physical, chemical, and biological habitat conditions given today’s state of the landscape.” The LDM varies among regions because of regional differences in historical and current land uses and can vary over time. In relatively pristine areas of North America that are largely unaffected by European settlement (e.g., parts of Alaska), the LDM and MDM are equivalent. In the southeastern coastal plain, which has been extensively modified by European settlement for hundreds of years, the difference between a strictly defined MDM and a LDM may be greater and difficult to quantify because of a lack of sites that have not experienced some degree of anthropogenic modification.

A remaining model that is sometimes useful for assessing ecological condition is the best attainable model (BAM). The BAM is equivalent to the LDM when the best possible management practices have been in place for an extended period (Stoddard et al. 2006a). Best attainable condition exists when and where the impact on biota of inevitable land use has been minimized to the greatest extent possible by best management practices. Except under special circumstances, the BAM should not be worse than the LDM (Figure 2). The BAM is most useful at sites that are disturbed, but can significantly improve with the application of practical management strategies.

It is important to distinguish among the preceding models because different models will encompass different distributions of reference sites and will result in different criteria for recovery goals. Application of a strictly defined MDM to Southeastern Plains ecosystems may result in recovery goals that are unachievable at many sites due to environmental modifications resulting from historical land uses. Insistence on a strictly defined MDM could also result in a dearth of reference sites needed to adequately define the range of natural and sampling variability associated with achievable reference conditions. Given the magnitude and ubiquity of historical environmental disturbances on the Southeastern Plains and in deference to the relatively strict definition of the MDM used by some researchers, we believe that the most appropriate ecological reference model for this ecoregion is the LDM. The LDMs used in our study are intended to represent conditions that prevail under the least amount of ambient human disturbance that can be found in the region; i.e., conditions that prevail at sites that represent the best ecological conditions.

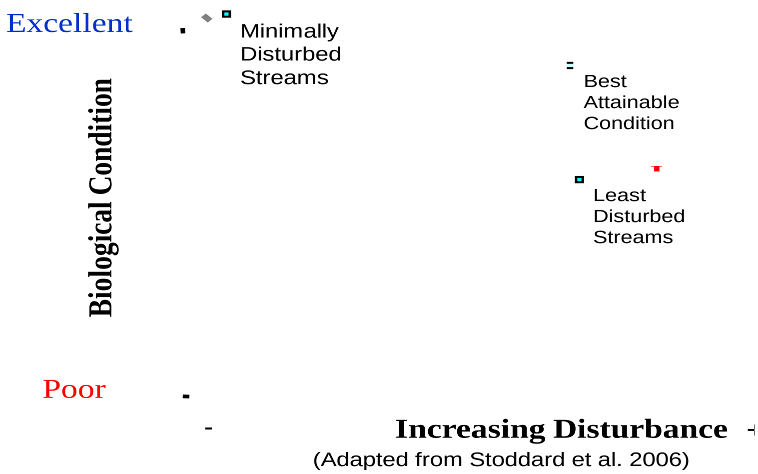


Figure 2. Theoretical position of different types of reference sites on the biological condition gradient.

Reference site selection is nearly always based on abiotic criteria to avoid preconceptions concerning biological structure and function at reference sites. The latter is important because ecological reference models for streams are usually defined on the basis of biotic structure and function (Karr et al. 1986, Barbour et al. 1999). Biota are exemplary for this purpose because they integrate all aspects of their chemical and physical environment, reflect both transient and chronic stressors, are generally more informative and easier to measure than chemical and physical indicators of environmental quality, and have intrinsic value to many people. Functional aspects of ecosystem quality, which are generally difficult to measure directly, can often be inferred from taxonomic and structural features of the biota; for example, macroinvertebrate functional groups can be used as indicators of carbon processing in streams and flood plains (Wallace and Webster 1996). Key to this process is the selection of appropriate indicators that are surrogates for important aspects of ecosystem structure and function.

The reference condition is usually described by data collected from sites selected according to a set of criteria that define what is least disturbed by human activities – a process referred to as the reference site approach (Hughes 1995, Bailey et al. 2004, Stoddard et al. 2006a). The selection of reference sites typically involves the application of abiotic screening criteria to identify sites that are minimally or least disturbed by human activities (Davis and Scott 2000). Historical reconstruction and predictive modeling can also be used if acceptable least disturbed sites are unavailable. The former method uses information from historical accounts, museum collections, or other sources to estimate undisturbed conditions or to adjust reference models developed from least disturbed site data (e.g., adding extirpated species or subtracting invasive species). Predictive modeling can also be used when reference site data and historical information are inadequate (EPA 2006). It may involve the extension of reference site results from similar or adjacent ecosystems or use of results from restoration/recovery studies. A less quantitative but sometimes used alternative to the preceding methods for developing reference models is thoroughly documented best professional judgment by experienced aquatic biologists. Best professional judgment can also be used to supplement other methods; e.g., to exclude sites (for sound ecological reasons) that have met other criteria.

3.3 Assessment Frameworks

After reference models are established, an assessment framework is needed to measure the difference between the objectives specified by the model and current ecosystem status. This framework should provide a means for conceptualizing the severity of degradation and the potential for recovery and specify metrics that reflect important ecosystem attributes. A metric is an ecological feature that shows predictable and quantifiable changes in response to increased human disturbance. Metrics should be sensitive to impacts occurring on different temporal and spatial scales and, where possible, act as surrogates for fundamental ecosystem properties or responses that are difficult to measure directly.

An assessment framework requires a methodology for summarizing and presenting metrics (e.g., taxonomic data) so that differences between disturbed states and reference model objectives are accurately measured. Some methods are listed below:
1) Simple univariate methods such as indices of species richness, diversity, and community similarity.
2) Robust multivariate methods summarizing multiple sources of evidence and graphical depictions of results. Methods such as cluster analysis, principal component analysis, and nonmetric multidimensional scaling represent an objective way to summarize and combine information from multiple variables and present the results in an interpretable graphical form (McCune et al. 2002).
3) Multimetric indices combining scores from metrics into a single number that indicates the difference between an assessment site and conditions represented by reference sites or a reference model (Karr and Chu 1999). Multimetric indices have been criticized for subjectivity and other shortcomings (Karr and Chu 1999, Simon and Lyons 1994) but can combine large amounts of information into a single number that is easy to understand.
4) Predictive models relating the observed taxonomic composition to the taxonomic composition expected in the absence of human disturbance. These methods, often termed observed/expected (O/E) models or probability models, are favored in Britain, Australia and in the United States.
5) Miscellaneous methods such as combining information on focal species ranked on the basis of indicator value (e.g., McCoy and Mushinsky 2002) .

Data requirements for assessment methods contrast somewhat. For example, multimetric indices ideally require information concerning biological responses along the entire stressor gradient for accurately evaluating metric ranges and scoring metric responses from best to worst. Most multivariate methods require similar information to accurately depict relationships among sites on stressor gradients. In contrast, predictive O/E models require information primarily on the physicochemical and biological characteristics of reference sites to estimate the probabilities of occurrence of taxa at reference sites and compare these probabilities to actual taxa occurrences at environmentally similar sites.

4.0 MATERIALS AND METHODS

Reference sites were in key military installation including Fort Benning GA, Fort Bragg NC, and Fort Gordon SC. Reference sites were also sampled on the Savannah River Site (SRS), a large and relatively undeveloped DOE reservation in SC that is also used for military training. A smaller number of additional reference sites were sampled in state forests, a national wildlife refuge, and The Nature Conservancy lands to include areas that experience patterns of current and legacy land use different from military lands. Reference sites in all locations were selected on the basis of screening criteria developed at the watershed, segment, and stream reach scales and implemented through field surveys and geographic information system (GIS) analyses. Sample areas were characterized at the watershed and stream segment spatial scales by using digital elevation models (DEMs) and National Landcover Data to delineate watersheds, estimate disturbance levels, drainage areas, drainage densities, lengths, shapes, and reliefs, land use/land covers, vegetation types, and other watershed features. Field data were collected to characterize habitat at the stream scale, including hydrology, and water chemistry. Data on changes in tree canopy coverage years were calculated at selected sites from recent and historical aerial photographs to investigate possible effects of historical disturbance. Fish assemblages were sampled by backpack electrofishing, and benthic macroinvertebrate assemblage were sampled using a multiple habitat sampling protocol.

Pertinent pre-existing biological data were compiled to provide ancillary information useful for model verification and other purposes. These data include the following:

1. Fort Benning
   a. Benthic macroinvertebrate data from several potential reference sites provided by consultants from Osage of Virginia;
   b. Benthic macroinvertebrate and habitat data collected as part of SI-1186 (Riparian Ecosystem Management at Military Installations: Determination of Impacts and Evaluation of Restoration and Enhancement Strategies);
   c. Fish assemblage data collected by Dr. William Birkhead, Columbus State University.
2) Fort Bragg
   a. Fish assemblage data collected by Chuck Bryan, Fort Bragg Endangered Species Branch;
3) Savannah River Site
   a. Fish and macroinvertebrate assemblage data collected by Michael Paller, Savannah River National Laboratory;
   b. In-stream habitat and watershed data collected from relatively undisturbed streams and streams suffering legacy impacts related to past agricultural practices collected by Dean Fletcher, Savannah (Savannah River Ecology Laboratory/University of Georgia [UGA]).

The reference models developed in this study served as the basis for assessment frameworks designed to measure the difference between the objectives specified by reference models and current ecosystem status. The frameworks provide a means for conceptualizing the severity of degradation and the potential for recovery, a methodology for summarizing and presenting data so that differences between disturbed states and reference model objectives are accurately measured, and a way of measuring progress towards the recovery objectives specified by the ecological reference models.

4.1 Study Area

The Southeastern Plains of the United States are a low-gradient region between the Piedmont, Atlantic Ocean, and Gulf of Mexico (Omernik 1997). Once dominated in places by longleaf pine (*Pinus palustris*) forests, the Southeastern Plains has undergone many anthropogenic modifications, reducing these communities to ~3% of their pre-settlement coverage (Landers et al. 1995). Legacy effects from historical agriculture contribute to present day impacts for many streams within this region (Loehle et al. 2009). More recently, however, this region has seen moderate amounts of succession, restoration, and improved management practices.

This study was restricted to a portion of the Southeastern Plain designated as the Sand Hills. The Sand Hills (Level IV) ecoregion (SH) represents the outer portion of the Southeastern Plains adjacent to the eastern part of the Piedmont along the fall line from west central GA to south central NC (Griffith et al. 2001). The SH spans ~20,600 km^2, consisting of some of the thickest and most extensive quartz sand deposits in the region, formed from the late Cretaceous to the Holocene (Markewich and Markewich 1994; Schmidt 2013). Predominant native vegetation in the SH consists of longleaf pine, turkey oak (*Quercus laevis*), and wire grass (*Aristida beyrichiana*) and was historically maintained by reoccurring low intensity fires.

Contemporary SH is characterized by a land use mosaic of agriculture, private lands, urban development, public natural areas, and military training facilities. Furthermore, the region has experienced extensive land use modifications since pre-settlement times (Loehle et al. 2009) and, therefore, the concept of pristine (Hughes 1995; Stoddard et al. 2006; Whittier et al. 2007) is virtually unattainable. For instance, some streams considered reference quality by government agencies occur within well-forested watersheds; however, these same watersheds were formerly clear-cut farms, and streams were impounded by low-head milldams. Finding reference sites in the SH is a challenge because of the difficulty in finding least disturbed conditions and because most land is privately owned or occupied by small municipalities. However, it is important to seek reference sites from within the ecoregion because models developed with reference sites outside of the ecoregion may produce inappropriate recovery objectives (Whittier et al. 2007).

Contacts were made with natural resource managers at Fort Bragg NC, Fort Benning GA, and Fort Stewart GA early in 2009 with the objectives of introducing our program to DoD managers at these installations, initiating work, selecting sample sites, and collecting existing data. Contacts were expanded to include DoD personnel at Fort Gordon GA, Nature Conservancy personnel managing easements contiguous with Fort Benning, and personnel managing Manchester State Forest SC, Sand Hills State Forest SC, Carolina Sandhills National Wildlife Refuge SC, and Sandhills Gamelands NC (Figure 3). Contacts were also made with Fort Jackson SC, although no sampling was conducted there. Inclusion of additional sites outside of the DoD/DoE system represented an effort to better define reference conditions by sampling a greater range of minimally impacted reference sites. Fort Stewart, located in the lower Sand Hills, was excluded from this project after site visits demonstrated that Fort Stewart streams were conspicuously lower in gradient and more swamp-like than streams in the other installations.

A total of 75 sample sites were included in the study distributed over 4 DoD installations (Fort Benning, Fort Bragg, Fort Gordon, and Camp McKall), one DoE installation (SRS, which is also used for military training) and 4 state, federal, and private (Nature Conservancy) natural areas (Table 1). Basic categories of data collected from each sample site are shown in Table 1. All data categories were represented at most sites; however, some were not because of logistical, budgetary, time, and labor constraints.

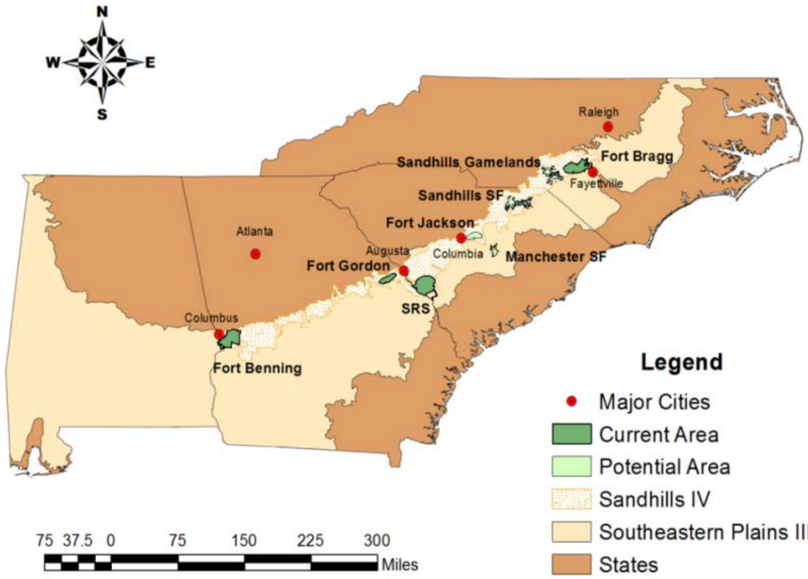


Figure 3. Map showing sampling locations in the upper Southeastern Plains (Sand Hills) ecoregion.

Table 1. Sample sites included in RC-1694.

Installation	Site name	Abbrev. 1	Abbrev. 2	Latitude	Longitude	Fish, SCDHEC & SRNL habitat protocols	Macroinvertbrates, geomorphology, water quality	N and P	Stream hydrology	GIS
Fort Benning	Bonham Creek Tributary (D12)	Bn-D12	d12	32.41187	-84.75690	X	X	X	X	X
Fort Benning	Bonham Creek Tributary (D13)	Bn-D13F	bc01	32.41546	-84.76090	X	X	X	X	X
Fort Benning	Bonham Creek Tributary (D13-C)	Bn-D13C	d13	32.41767	-84.76023	X	X	X	X	X
Fort Benning	Chatahoochee River Tributary	Bn-Chat		32.34904	-85.00444	X				X
Fort Benning	HCT-G2/5	Bn-HCT		32.33421	-84.70339	X				X
Fort Benning	Hollis Branch Mainstem(F4)	Bn-Holl	hbms	32.36395	-84.68455	X	X	X	X	X
Fort Benning	K13	Bn-K13	k013	32.49748	-84.70769	X	X	X	X	X
Fort Benning	King Mill Creek(K11E)	Bn-K11	kmms	32.51054	-84.64841	X	X	X	X	X
Fort Benning	Little Juniper Tributary (K10)	Bn-LJT	ljtb	32.51151	-84.63712	X	X	X	X	X
Fort Benning	Little Pine Knot (K20)	Bn-LPK	lpk_	32.39764	-84.67085	X	X	X	X	X
Fort Benning	Wolf Creek	Bn-wolf	wcms	32.41938	-84.83746	X	X	X	X	X
Fort Bragg	Beaver Creek trib	Br-bct	bvct	35.13342	-78.97667	X	X		X	X
Fort Bragg	Big Muddy Creek Main Stem	Br-bm	bmcm	35.01986	-79.51553	X	X	X	X	X
Fort Bragg	Cabin Creek	Br-cab	ccms	35.05333	-79.29673	X	X	X	X	X
Fort Bragg	Cypress Creek	Br-cyp	cypr	35.17974	-79.04723	X	X	X	X	X
Fort Bragg	Deep Creek	Br-deep	deep	35.14507	-79.15983	X	X	X	X	X
Fort Bragg	Field Branch Main Stem	Br-field	fbms	35.06282	-79.30483	X	X	X	X	X
Fort Bragg	Flat Creek Main Stem	Br-flat	fcms	35.17424	-79.18026	X	X	X	X	X
Fort Bragg	Gum Branch Main Stem	Br-gum	gbms	35.08958	-79.33523	X	X	X	X	X
Fort Bragg	Hector Creek Main Stem	Br-hect	hcms	35.18380	-79.09892	X	X	X	X	X
Fort Bragg	Horse Creek (HC)	Br-hrse		35.17208	-79.23305	X	X	X		X
Fort Bragg	Jennie Creek Main Stem	Br-jen	jcms	35.12032	79.33021	X	X	X	X	X
Fort Bragg	Juniper Creek Main Stem	Br-jun	jpms	35.07234	-79.25689	X	X	X	X	X
Fort Bragg	Little River Tributary	Br-lrt	lrtb	35.18748	-79.07469	X	X	X	X	X
Fort Bragg	Little Rockfish Main Stem North	Br-lrf	lrf1	35.17134	-79.08787	X	X	X	X	X
Fort Bragg	McPherson Creek	Br-mcp	mcph	35.14079	-79.04396	X	X	X	X	X
Fort Bragg	Rockfish Branch (RFB)	Br-rfb	rfb_	35.11509	79.32548	X	X	X	X	X
Fort Bragg	Tank Creek	Br-tank		35.14716	-79.02106	X		X		X
Fort Bragg	Upper Jennie Trib (UJT)	Br-ujt		35.13354	-79.34534	X		X		X
Fort Bragg	Upper Wolf Pit (UWP)	Br-uwp		35.11102	-79.35168	X		X		X
Fort Bragg	Wolf Pit Creek Main Stem	Br-wp	wpms	35.11320	79.33718	X	X	X	X	X
Fort Gordon	Bath Branch	Gr-bath	bbbb	33.35942	-82.15746	X	X	X	X	X
Fort Gordon	Boggy Gut Creek	Gr-Bog	bgut	33.34772	-82.29175	X	X	X	X	X
Fort Gordon	Headstall		head	33.34460	-82.34712		X	X	X	
Fort Gordon	McCoys Creek	Gr-Mcoy	mcoy	33.39939	-82.16021	X	X	X	X	X
Fort Gordon	South Prong	Gr-Prong	sspp	33.35942	-82.15746	X	X	X	X	X
Fort Gordon	Trib to Marcum Branch	Gr-Mrbtrb	mart	33.40980	-82.18657	X	X	X	X	X
Manchester State Forest	McCrays Creek	Mn-mcra	mccc	33.78431	-80.47599	X	X	X	X	X
Manchester State Forest	Tavern Creek	Mn-tav	tvms	33.75819	-80.52800	X	X	X	X	X
Sandhills Game Lands	Bones Fork Tributary	Sg-bone	bone	35.03531	-79.61385	X	X	X	X	X
Sandhills Game Lands	East Prong Juniper Creek		epjc	34.95324	-79.49592		X		X	X
Sandhills Game Lands	Joes Creek	Sg-joes	joec	34.88026	-79.62551	X	X	X	X	X
Sandhills Game Lands	Millstone Creek	Sg-mill	mist	35.06723	-79.66504	X	X	X	X	X
Sandhills Game Lands	Upper Beaverdam Creek		ubdc	34.85284	-79.57317		X		X	X
Sandhills Game Lands	West Prong Juniper Creek		wpjc	34.94988	-79.50353		X		X	X
Sandhills NWR	Big Black Creek Tributary	Sn-bbct	bbct	34.66050	-80.22655	X	X	X	X	X
Sandhills NWR	Hemp Creek	Sn-hemp	hemp	34.57139	-80.24704	X	X	X	X	X
Sandhills NWR	North Prong Swift Creek		npsc	34.52632	-80.29989		X		X	X
Sandhills NWR	Rogers Branch	sn-rogr	rgbc	34.60473	-80.20997	X	X	X	X	X
Sandhills State Forest	Little Cedar Creek	Sh-cedr	lcdr	34.51895	-79.99899	X	X	X	X	X
Sandhills State Forest	Mill Creek	Sh-mill	mlck	34.53822	-80.07601	X	X	X	X	X
Savannah River Site	Lower Three Runs (DS)	Sr-ltr		33.22353	-81.50894	X		X		X
Savannah River Site	McQueens Branch Headwater	Sr-Mqh	mqhw	33.29836	-81.62959	X	X	X	X	X
Savannah River Site	McQueens Branch Tributary (8)	Sr-mq8	mq08	33.30454	-81.62630	X	X	X	X	X
Savannah River Site	McQueens Branch Tributary (MQ10.1)	Sr-mq10		33.29810	-81.62629	X		X		X
Savannah River Site	Meyers Branch 6	Sr-mb6	mb06	33.17810	-81.56570	X	X	X	X	X
Savannah River Site	Meyers Branch 6.1	Sr-mb61	mb61	33.18051	-81.56335	X	X		X	X
Savannah River Site	Meyers Branch Headwaters	Sr-mbhw	mbhw	33.19357	-81.57880	X	X	X	X	X
Savannah River Site	Meyers Branch Main Stem	Sr-mbm	mbms	33.17613	-81.58174	X	X	X	X	X
Savannah River Site	Mill Creek Main Stem	Sr-mcm	mcms	33.30074	-81.58680	X	X	X	X	X
Savannah River Site	Mill Creek Tributary 5	Sr-mc5	mc05	33.31922	-81.58011	X	X	X	X	X
Savannah River Site	Mill Creek Tributary 6	Sr-mc6	mc06	33.31731	-81.59759	X	X	X	X	X
Savannah River Site	Mill Creek Tributary 6C	Sr-mc6c		33.31940	-81.59619	X				X
Savannah River Site	Mill_Creek Tributary 7	Sr-mc7	mc07	33.32440	-81.60101	X	X	X	X	X
Savannah River Site	Pen Branch Headwater	Sr-pbhw	pbms	33.23250	-81.62424	X	X	X	X	X
Savannah River Site	Pen Branch Main Stem	Sr-Pbm	pbm1	33.22576	-81.63570	X	X	X	X	X
Savannah River Site	Pen Branch Tributary (4)	Sr-pb4	pb04	33.23349	-81.63809	X	X	X	X	X
Savannah River Site	Tinker Creek Main	Sr-tcm	tink	33.36369	-81.55806	X	X	X	X	X
Savannah River Site	Tinker Creek Tributary (5)	Sr-tc5	tc05	33.37333	-81.54981	X	X	X	X	X
Savannah River Site	Tinker Creek Tributary (6)	Sr-tc6	tc06	33.36077	-81.55763	X	X	X	X	X
Savannah River Site	U10	Sr-u10		33.30036	-81.66618	X				X
The Nature Conservancy	Black Creek Tributary	Nc-bct	blkt	32.56942	-84.51484	X	X	X	X	X
The Nature Conservancy	Black Jack Creek	Nc-bjc	bctb	32.58038	-84.49602	X	X	X	X	X
The Nature Conservancy	Parkers Mill Creek Tributary	Nc-pmt	pmtb	32.45264	-84.57671	X	X	X	X	X
The Nature Conservancy	Pine Knot Creek	Nc-pkc		32.43953	-84.64734	X				X

4.2 Watershed Assessment

GIS data were used to assess environmental disturbance and compute several habitat descriptors at the watershed and landscape scales (Table 2). DEMs created from light detection and ranging (LIDAR) data were used as a basis for generating high-resolution (HUC-12) watershed units on which subsequent landscape-level calculations were based (SRS example, Figure 4). The resulting maps were used to compute a landscape disturbance index for drainages at each installation. The disturbance index was computed from 2006 National Land Cover Data, which included lands characterized by low-, medium-, and high-intensity development plus cultivated, pasture, and bare lands. The area of each land use category within each catchment was individually determined in ArcGIS, and the 6 disturbance categories were summed and converted to percentages to generate the disturbance index for each catchment (SRS example, Figure 5).

Watershed maps also served as the basis for a more detailed analysis of watersheds surrounding each sample site. Variables computed for drainages with biological sample sites included watershed area, perimeter, drainage density, length, shape, highest elevation, elevation of stream mouth; basin relief, basin relief ratio, entire stream gradient, drainage direction, sinuosity, % landcover types based on National Land Cover data, bifurcation ratio, cumulative stream length, stream length, stream order, stream magnitude, and length of mainstem tributaries (Table 2). These variables were also computed for 1000 m buffer zones upstream from each sampling point within selected subwatersheds. Additional GIS variables included the distance from each sample site to the point of confluence with a larger stream and to the nearest impoundment in the watershed.

We also assessed the potential effects of historical land use from aerial photographs taken before or soon after the Savannah River Site, Fort Benning, and Fort Gordon were established. Comparable data were unavailable for the other installations under study except for one site at Fort Bragg. Historical tree canopy coverages were calculated from select aerial photos (SRS: 1951 and 2010; Fort Benning: 1944, 1999, and 2009; Fort Bragg: 1951 and 2009; Fort Gordon: 1941, 1963, and 2006). Aerial photos were mosaicked to provide images of complete watersheds. Watersheds were clipped from each aerial photo using ArcGIS 9.3. Clips taken from color photos were converted to 1 band raster files. Colors were classified into 2 categories using a color ramp selected to provide the optimal contrast between forested and un-forested areas. Proportion of clip covered by colors corresponding to forested areas was calculated. Results from this methodology were found to be comparable to hand delineated coverages.

4.3 Instream Habitat Assessment

The preceding GIS analyses characterized habitat at the stream segment, watershed, and larger spatial scales. The instream habitat data, which were collected in the field, described local habitat features within the stream and adjacent riparian zone. Data were used to identify habitat features that influenced biological structure and to distinguish differences between disturbed and reference sites. Several types of instream habitat data were collected:

1. SCDHEC (South Carolina Department of Health and Environmental Control) protocol,
2. SRNL (Savannah River National Laboratory) protocol,
3. Geomorphological and organic material,
5. Hydrological,
6. Water quality

SCDHEC, SRNL, and Geomorphological variables are summarized in Tables 3, 4, and 5, respectively. More details are provided below.

Table 2. Watershed and landscape measures calculated for Sand Hills sample sites.

Variable	Definition
Catchment area (m2)	Total area of the catchment.
Catchment perimeter (m)	Perimeter of the catchment.
Cumulative stream length(m2)	Total length of all perennial stream segments in the catchment (Lt).
Catchment density	Total stream length divided by the catchment area or RD = Lt/A.
Catchment length (m)	Length of the catchment along a straight line most parallel to the mainstem (Cl).
Catchment shape	Rf = A/Lt^2.
Stream length (m)	Total length of the mainstem (L).
Highest point (m)	Highest point in the catchment.
Mouth elevation (m)	Elevation at the downstream portion of where sampling was conducted.
Catchment relief	h = high - mouth.
Catchment relief ratio	Rr = h/Cl.
Stream gradient	Slope of the stream Sc = (elevation at the top of Cl - mouth)/Cl.
Maloney Disturbance Index	Percentage of the catchment that is bare ground on >3% slope + non-paved roads.
Stream order (Strahler)	Strahler stream ordering method.
Stream magnitude (Shreve)	Shreve stream ordering method.
Catchment direction	Direction facing downstream along Cl.
Length of tributaries to main stem	Lb = Lt - L.
Sinuosity	Measured as L/Cl.
1km Catchment area(m2)	Catchment area within 1 km upstream of the sampling reach.
1km Cumulative stream length(m2)	Cumulative stream length within 1 km upstream of the sampling reach.
1km Catchment density	Catchment density within 1 km upstream of the sampling reach.
1km Catchment length (m)	Catchment length within 1 km upstream of the sampling reach.
1km Stream length	Stream length within 1 km upstream of the sampling reach.
1km Catchment relief	Catchment relief within 1 km upstream of the sampling reach.
1km Catchment relief ratio	Catchment relief ratio within 1 km upstream of the sampling reach.

Table 2. continued

Variable	Definition
1km Stream gradient	Stream gradient within 1 km upstream of the sampling reach.
1km Maloney Disturbance Index	Maloney Disturbance Index within 1 km upstream of the sampling reach.
1km Length of tributaries to main stem	Length of the tributaries to the main stream within 1 km upstream of the sampling reach.
1km Sinuosity	Sinuosity within 1 km upstream of the sampling reach.
Bare ground [a]	Percentage of the catchment that is bare ground.
Developed high [a]	Percentage of the catchment that is high intensity development.
Developed medium [a]	Percentage of the catchment that is medium intensity development.
Developed low [a]	Percentage of the catchment that is low intensity development
Developed open space [a]	Percentage of the catchment that is open space development.
Cultivated [a]	Percentage of the catchment that is cultivated land.
Pasture [a]	Percentage of the catchment that is in pastures or used for hay production.
Grassland [a]	Percentage of the catchment that is grassland.
Deciduous [a]	Percentage of the catchment that is deciduous forest.
Evergreen [a]	Percentage of the catchment that is evergreen forest.
Mixed forest [a]	Percentage of the catchment that is a mix of deciduous and evergreen forest.
Scrubland [a]	Percentage of the catchment that is shrubs and/or small trees.
Palustrine [a]	Palustrine Forested Wetland
Palustrine 1 [a]	Palustrine Scrub/Shrub Wetland
Palustrine 2 [a]	Palustrine Emergent Wetland (Persistent)
Water [a]	Percentage of the catchment that is open water.
Disturbance [a]	Percentage of the catchment that is developed, bare ground, and in agricultural use.
Impervious surfaces [a]	Calculated as developed high*0.90 + developed medium*0.65 + developed low*0.35 + developed open space * 0.10.
Agriculture [a]	Calculated as cultivated + pasture.
1km Bare ground [a]	Percentage of the catchment within 1 km upstream of the sampling reach that is bare ground.
1km Developed high [a]	Percentage of the catchment within 1 km upstream of the sampling reach that is high intensity development.

Table 2. continued

Variable	Definition
1km Developed medium [a]	Percentage of the catchment within 1 km upstream of the sampling reach that is medium intensity development.
1km Developed low [a]	Percentage of the catchment within 1 km upstream of the sampling reach that is low intensity development.
1km Developed open space [a]	Percentage of the catchment within 1 km upstream of the sampling reach that is open space development.
1km Cultivated [a]	Percentage of the catchment within 1 km upstream of the sampling reach that is cultivated land.
1km Pasture [a]	Percentage of the catchment within 1 km upstream of the sampling reach that is in pastures or used for hay production.
1km Grassland [a]	Percentage of the catchment within 1 km upstream of the sampling reach that is in grassland.
1km Deciduous [a]	Percentage of the catchment within 1 km upstream of the sampling reach that is deciduous forest.
1km Evergreen [a]	Percentage of the catchment within 1 km upstream of the sampling reach that is evergreen forest.
1km Mixed forest [a]	Percentage of the catchment within 1 km upstream of the sampling reach that is a mix of deciduous and evergreen forest.
1km Scrubland [a]	Percentage of the catchment within 1 km upstream of the sampling reach that is shrubs and/or small trees.
1km Palustrine [a]	Palustrine Forested Wetland
1km Palustrine 1 [a]	Palustrine Scrub/Shrub Wetland
1km Palustrine 2 [a]	Palustrine Emergent Wetland (Persistent)
1km Water [a]	Percentage of the catchment within 1 km upstream of the sampling reach that is open water.
1km Disturbance [a]	Percentage of the catchment within 1 km upstream of the sampling reach that is disturbance as described above.
1km Impervious surface [a]	Percentage of the catchment within 1 km upstream of the sampling reach that is impervious surface as described above.
1km Agriculture [a]	Percentage of the catchment within 1 km upstream of the sampling reach that is agriculture as described above.
Latitude	The latitude of the sample sites (decimal degrees).
Longitude	The longitude of the sample sites (decimal degrees).

Table 2 continued.

Variable	Definition
Distance to nearest upstream impoundment (m)	Distance of the sample site to the nearest upstream impoundment
Distance to nearest downstream impoundment (m)	Distance of the sample site to the nearest downstream impoundment
Impoundment area (m2)	Impoundment surface area
Paved area (including roads and other paved areas (m2)	Total pavement area in the catchment
Paved area density	Total paved area divided by total catchment area
Unpaved roads (including roads and other paved areas (m2)	Total unpaved road area in the catchment
Unpaved road density	Total unpaved area divided by total catchment area
Distance to nearest confluence with a larger stream (m)	Distance from the sample site to the confluence of a larger stream
Stream order (Strahler) of the larger stream	Stream order (Strahler) of the larger stream

a - NOAA’s Coastal Change Analysis Program (C-CAP): http://www.csc.noaa.gov/digitalcoast/data/ccapregional

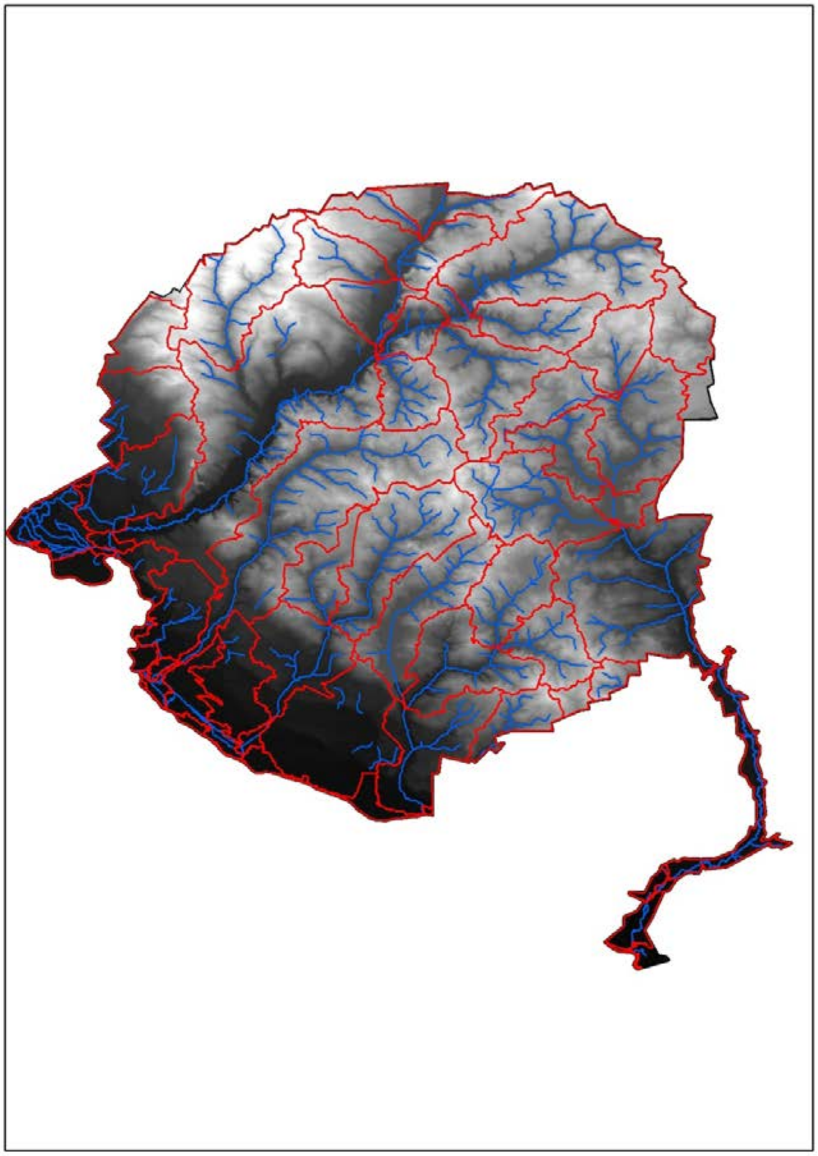


Figure 4. Digital Elevation Model for the Savannah River Site created from LIDAR data. Red borders depict watershed subunits.

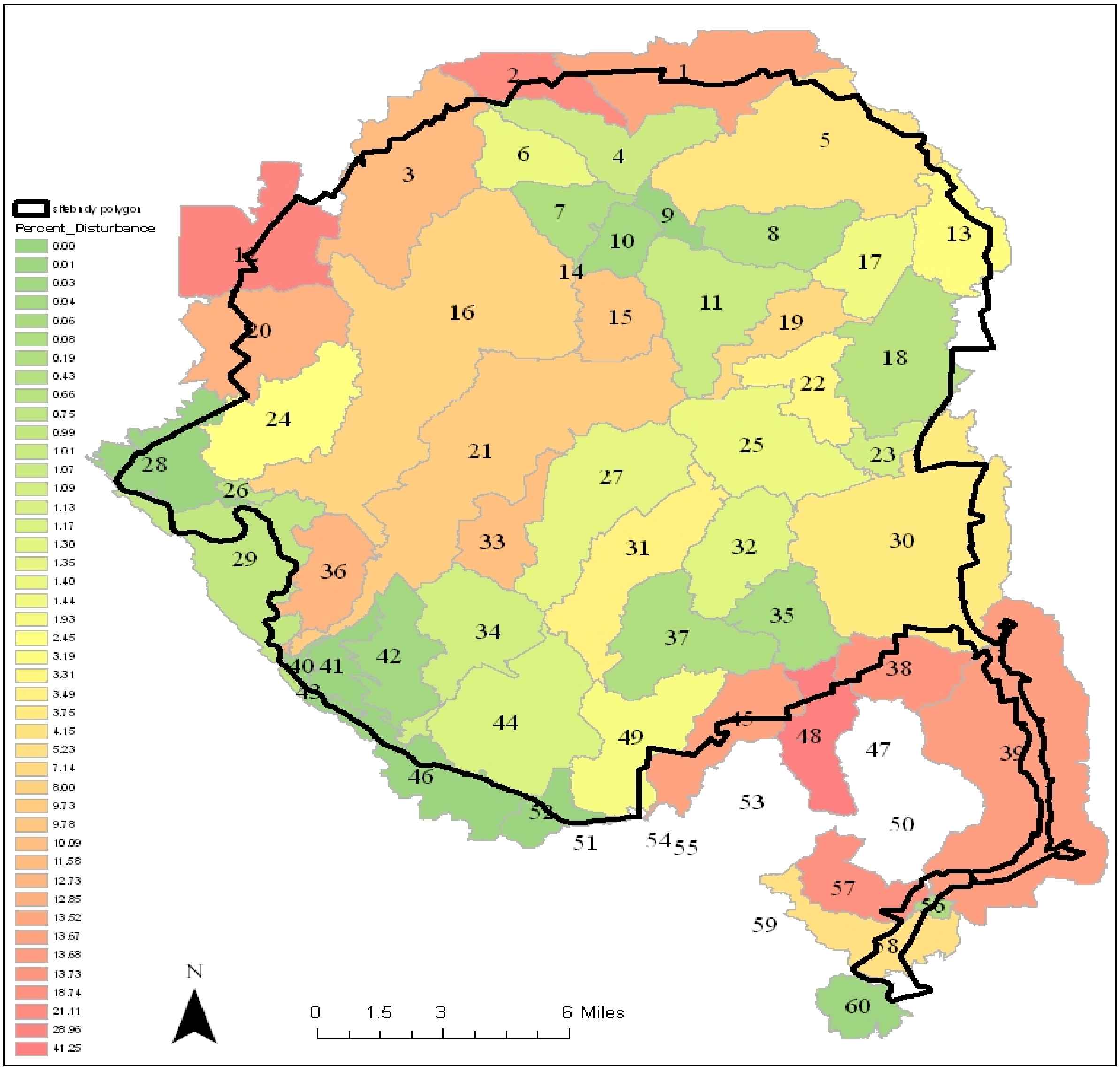


Figure 5. Landscape disturbance index for watershed subunits on the Savannah River Site. The disturbance index included lands characterized by low-, medium-, and high-intensity development plus cultivated, pasture, and bare lands.

4.3.1 *SCDHEC Habitat Protocol*

The SCDHEC habitat protocol was adopted from the South Carolina Department of Health and Environmental Control (SCDHEC 1998) biological sampling protocol for low-gradient streams. It was used at all sites where fish communities were assessed (Table 3). When conducting this protocol, it should be remembered that SH streams often lack well-defined pools, riffles, runs, and rocky substrates typically associated with good habitat in higher-gradient streams. To avoid inappropriately low scoring, inexperienced field crews should be "calibrated" by observing habitat at minimally disturbed SH reference sites before scoring test sites.

Field crews conducting fish sampling observed the habitat characteristics of the reach under study and collaboratively rated 10 habitat variables from a best of 20 to a worst of 0. Epifaunal substrate (snags, submerged logs, undercut banks and other stable colonizable habitat) received a high score at >50% and a low score at <10% coverage of the stream bottom. Pool substrates characterized by firm sand, gravel, roots, and vegetation scored near 20, whereas pools without vegetation or with large amounts of silt or hardpan clay scored less. Stream reaches with a large number and good mix of large, small, shallow and deep pools scored highly on the pool variability metric, whereas those with low numbers of predominantly small shallow pools scored poorly. Sediment deposition was estimated by the presence of fine particles and the formation of shifting bars and islands. A reach with no bars or islands and <20% of the substrate affected by sediment deposition ranked highly; the score diminished as the percentage of shifting sand increased. Scores for channel flow, a simple estimate of the amount of water present, ranged from high, where very little of the substrate was exposed, to low, where there was little water and standing pools predominated. Channel alteration, was assessed by the presence of cement banks, other artificial structures, and other evidence of stream channelization. Sites without these disturbances scored well on this metric. The last 4 metrics focused on the banks and riparian area. Channel sinuosity, a measure of the 'curviness' of a stream, rated near the top when the channel measured 3-4 times longer than if it was straight. Straight, channelized streams scored low on this measure. For the final 3 metrics, one score (from 0-10) was given to each bank (left and right) and then added together to give the final score for the assessment. Bank Stability estimated the level of bank failure and erosion on each side of the stream. A score of 20 was given when <5% of the bank was affected by erosion and a score near zero was given at sites where 60-100% of the bank had erosional scars. Vegetative Protection received a score near 20 when the banks displayed at least 90% coverage in naturally growing vegetation and a low score when coverage fell <50%. Finally, the riparian zone was assessed with the metric riparian vegetative zone width. Intact riparian areas of greater than 20 m received a score of 20. As anthropogenic impacts increased and the width decreased to <6 m, the score approached zero. Standard Operating Procedures (SOPs) for the SCDHEC protocol are given in Appendix 6.

4.3.2 *SRNL Habitat Protocol*

The SRNL habitat protocol measured 2 attributes directly and visually estimated another 17 (Table 4). Like the SCDHEC protocol, it was implemented in conjunction with fish sampling. Each fish sampling reach was divided into contiguous 10-15 m segments. A transect was established at the downstream end of each segment and each metric in the SRNL protocol was measured at or within a 1m band, 0.5m below and 0.5 m above each transect. Maximum depth and stream width at each transect were measured directly using a depth rod (nearest 0.05m) and a field tape measure (nearest 0.1m), respectively. The depth rod was also used to estimate left and right bank heights from the water level to the top of the bank. Left and right bank angles were visually estimated (nearest 5 degrees) as an average angle over the 1-m band associated with the transect. Anything >90 degrees indicated an undercut bank. Bank vegetative cover within an area up to 5 m on each side of the stream was visually estimated to the nearest 5%. A single qualitative score from 0 (none) to 3 (high)

Table 3. Variables measured under the SCDHEC instream habitat protocol.

Measure	Type	Definition - Score from 0 (Poor) to 20 (Excellent)
Epifaunal Substrate	Qualitative	Qualitative score for colonizable habitat within the total substrate area (>50% = Excellent, <10% = Poor).
Pool Substrates	Qualitative	Qualitative score for pool substrate (firm sand, gravel and roots and vegetation = Excellent, Hardpan clay = Poor)
Pool Variability	Qualitative	Qualitative score for variability in depth and size of pools (large number and good mix of large, small, shallow and deep pools = Excellent, Few small shallow pools = Poor)
Sediment Deposition	Qualitative	Qualitative score for presence of sediment (No bars or islands and less than 20% sedimentation = Excellent, 80% = Poor)
Channel Flow Status	Qualitative	Qualitative score for amount of water present (No substrate exposed = Excellent, Very little water and standing pools = Poor)
Channel Alteration	Qualitative	Qualitative score estimating in-stream anthropogenic disturbance (None = Excellent, Presence of cement banks, gabions, and stream channelization = Poor)
Channel Sinuosity	Qualitative	Qualitative score for the ‘curviness’ of a stream (Channel 3-4 times longer than if it was in a straight line = Excellent, Straight channel = Poor)
Bank Stability	Qualitative	Qualitative score for level of bank failure and erosion (<5% of the banks effected = Excellent, 60-100% effected = Poor)
Vegetative Protection	Qualitative	Qualitative score for percent bank cover in natural vegetation (>90% coverage = Excellent, <50% = Poor)
Riparian Vegetative Zone Width	Qualitative	Qualitative score for intact riparian areas (>18m = Excellent, <6m = Poor)

Table 4. Variables measured under the SRNL instream habitat protocol.

Measure	Type	Definition (Measured within a 1m band around transect)
Maximum Depth	Quantitative	Measured maximum water depth (nearest 0.05m)
Average Width	Quantitative	Measured wetted width (nearest 0.1m)
Right Bank Height	Quantitative	Measured distance from the water level to the top of the bank (nearest 0.05m)
Left Bank Height	Quantitative	Measured distance from the water level to the top of the bank (nearest 0.05m)
Right Bank Angle	Estimated	Visually estimated (nearest 5 degrees)
Left Bank Angle	Estimated	Visually estimated (nearest 5 degrees)
Bank Vegetative Cover	Estimated	Visually estimated percent vegetative cover within 5 meters on each side of the stream (nearest 5%)
Bank Erosion	Estimated	Visually estimated score from 0 (Excellent) to 3 (Poor) assessing erosion and bank instability
Channel Modifications	Estimated	Visually estimated score from 0 (Excellent) to 3 (Poor) assessing instream anthropogenic structures and channel modifications
Macrophytes	Estimated	Visually estimated areal coverage of the substrate by macrophytes (nearest 5%)
Overhanging Vegetation	Estimated	Visually estimated areal coverage of the substrate by overhanging vegetation (nearest 5%)
Root Mats	Estimated	Visually estimated areal coverage of the substrate by root mats (nearest 5%)
Large Woody Debris	Estimated	Visually estimated areal coverage of the substrate by large woody debris (diameter >3cm) (nearest 5%)
Small Woody Debris	Estimated	Visually estimated areal coverage of the substrate by small woody debris (diameter <3cm) (nearest 5%)
Mesohabitat Type	Estimated	Transect assigned to riffle, run, shallow pool or deep pool accounting for stream size
Undercut Banks	Estimated	Visually estimated percentage of the total wetted stream width under a bank (nearest 5%)
Riparian Zone Width	Estimated	Visually estimated (nearest 5m - 20m maximum)
Riparian Zone Vegetation	Estimated	Vegetation present within 20m from each bank categorized as hardwood (w), pine (p), shrubs (s) and herbaceous (h) and ranked from 1 (most dominant) to 4 (least dominant)

Table 5. Geomorphology and organic matter variables measured within the stream sampling reaches.

Variable	Symbol	Definition
Bankfull perimeter	BkfP	The distance along the cross-section of the stream channel at the level of bankfull capacity.
Bankfull area	BkfA	The area across the level of bankfull capacity.
Bankfull width	BkfW	The width across the level of bankfull capacity.
Bankfull depth	BkfD	The depth from the stream channel to the level of bankfull capacity measured as a mean or maximum.
Width to depth ratio	W/D	BkfW/BkfD
Wetted perimeter	WtP	The distance along the cross-section of the stream channel where it contacts water.
Wetted area	WtA	The area across the water line.
Wetted width	WtW	The width across the water level.
Wetted depth	WtD	The depth from the stream bed to the level of the water measured as a mean or maximum.
Canopy cover		The percentage of the overhead canopy that is covered by tree canopy.
Mean Surber depth [a]		The mean depth at the point where Surber (benthic) samples was taken.
Mean Surber velocity [a]		The mean velocity at the point where Surber samples was taken.
Bankfull area/ wetted area		Bankfull area divided by the wetter area.
Hydraulic radius	R	R = WtA/WtP
All CWD (coarse woody debris)		The proportion of stream channel surface area that is CWD > 0.025 m diameter.
Submerged CWD [a]		The proportion of the stream channel surface area that is CWD under water.
Buried CWD [a]		The proportion of the stream channel surface area that is CWD buried under ≤ 0.1 m sediment.
Dry CWD [a]		The proportion of the stream channel surface area that is CWD not under water.
Live CWD [a]		The proportion of the stream channel surface area that is living CWD.
Dead CWD [a]		The proportion of the stream channel surface area that is dead CWD.
BOM midstream [a b]		Benthic organic matter (BOM) taken from the stream margin.
BOM stream margin [a b]		Benthic organic matter (BOM) taken from the stream center.
BOM total [a b]		Average of BOM taken from the stream margin and center.
Clay cover		Percentage of the stream substrate that is visually estimated as clay.
Silt cover		Percentage of the stream substrate that is visually estimated as silt.

Table 5. concluded

Measure	Symbol	Definition
Sand cover		Percentage of the stream substrate that is visually estimated as sand.
Gravel cover		Percentage of the stream substrate that is visually estimated as gravel.
Rock cover		Percentage of the stream substrate that is visually estimated as rock.

a - see Appendix 3

b - see Appendix 4

was given at each transect for bank erosion by assessing the amount of erosional scaring, bank instability, and bank material entering the stream (see Figure 6 for example). The channel modification metric also received a score from 0 (none) to 3 (high) based on the amount of anthropogenic disturbance including instream modifications and channelization. The percent of areal coverage of the substrate by macrophytes, overhanging vegetation, root mats, coarse woody debris (diameter >3 cm) and small woody debris (diameter <3 cm) were visually estimated to the nearest 5% within the 1-m transect bands. The mesohabitat type was assigned to one of four categories; riffle, run, shallow pool or deep pool using best professional judgment accounting for the size of each stream. Undercut banks were estimated at each transect as the percentage of the total wetted stream width under a bank to the nearest 5%. The percent areal coverage of each of 5 approximate substrate size classes within each transect area (clay <0.005 mm, silt/muck 0.005-0.1 mm, sand 0.1-5 mm, gravel 5-30 mm and rocks >30 mm) was visually estimated to the nearest 5%. The riparian zone width and riparian zone vegetation type were also estimated. The width was visually estimated to the nearest 5 m. Vegetation type present within the first 20 m from each stream bank was categorized into 4 types; hardwood, pine, shrub and herbaceous. Each type was then ranked from 1 (most dominant) to 4 (least dominant). SOPs for the SRNL protocol are in Appendix 6.

Figure 6. Erosion scores in the Sand Hills streams ranged from 0 in the stream on the left to 3 (unprotected eroding banks) and 4 (collapsing banks) in the stream on the right.

4.3.3 *Geomorphology and Organic Matter Data Collection*

Hydraulic geometry relationships (Leopold and Maddock 1953) were evaluated to discriminate hydrologically disturbed sites from reference sites as a function of predicted natural channel configuration (Metcalf et al. 2009) (Table 5). Channel geomorphology was quantified by surveying 4 to 6 channel cross sections at each study reach. Transects were established in runs dividing the reach into approximately equidistant sections, with reach length fixed regardless of 4 or 6 transects, thus allowing comparison. Cross sections were established by staking rebar at determined top of bank height on either side of the channel perpendicular to the direction of flow. Top of bank height was determined as the point where water breaches the lowest of the 2 stream banks (Leopold 1994). A line level was used to establish the relative top of bank datum for each survey and top of bank depths and water depths were recorded every 20 cm along each transect.

Measures of instream habitat consisted of estimate of substrate size, amount of coarse woody debris (CWD, wood > 2.5 cm diameter) and benthic organic matter (BOM, organic matter material ≤1.6 cm diameter) at each transect (Wallace and Benke 1984). Substrate size and BOM were estimated from PVC cores (7.62cm diameter, 20.01 cm^2) inserted to a depth of 10cm near the center of the channel directly above each transect (substrate size) or at two locations (midstream and stream margin) along each transect. Samples were dried and combusted to remove BOM (below), and then dry sieved for representative particle sizes (i.e., *d5*, *d50*, *d95*, etc., phi scale -4-5, Lane 1947) and summarized as a reach-specific median. Geometric mean and standard deviation of particle size were calculated from reach medians after equations in Table 2.8 in Bunte and Apt (2001). For removed BOM, samples were oven-dried at 80°C for 24 to 48 h, weighed, and ashed in a muffle furnace at 550°C for 3 h. Samples were then cooled in a desiccator and reweighed; % BOM was determined as the difference between dry and ashed masses divided by total dry mass (Wallace and Grubaugh 1996). For CWD, length and width of each piece of CWD was measured and summed to obtain an areal estimate in square meters and divided by the wetted transect area to calculate percent coverage. BOM was processed in the laboratory according to standard methods. SOPs for geomorphology data collection are given in Appendix 3.

4.3.4 *Hydrology*

Stream level data were measured by instream pressure transducers from representative study sites at all installations (Table 1). Stream level data provide a measure of stream flashiness (as rate of decrease in the hydrograph from a storm event), which has been shown to correlate with watershed disturbance at Fort Benning (Maloney et al. 2005).

Stream stage was used as a measure of temporal variation in stream hydrology, estimated from Solinst Levelogger Junior pressure transducers (Model 3001, Solinst Canada Ltd., Canada), installed at the downstream end of each study reach. Leveloggers were adjusted for ambient atmospheric pressure using Solinst Barologger Gold pressure transducers or barometric pressure data obtained from local airport weather stations. Level and barometric pressure data were measured every 15 min for the duration of logger deployment at each stream reach. Temporary stilling wells were used to house leveloggers, constructed from schedule-40 PVC (3.81 cm ID) and perforated on the downstream side to allow water circulation. Stilling wells have been shown to produce stable water surface elevations in other SE streams (Schoonover et al. 2006). Continuous stream level was summarized using a suite of metrics reflecting variation in flow regime (McMahon et al. 2003, Helms et al. 2009).

Site-specific stream discharge (Q) was estimated in situ at cross sections perpendicular to the direction of flow using the velocity-area method (Gordon et al. 1993). Velocity was quantified with a Flo-Mate velocity meter (model 2000, Hach Co., CO) and depth was measured with a meter stick. Continuous Q profiles (hydrographs) were estimated from these data by building a stage-discharge relationship. Q estimates were collected at the sites over time and at contrasting flow conditions to construct rating curves between Q and stage-height data derived from pressure transducers. SOPs for the collection of hydrological (discharge) data are given in Appendix 4.

4.3.5 *Water Quality*

Specific conductance and pH were taken at the time of macroinvertebrate sampling with a YSI 56 MPS (Yellow Springs, OH, USA) and Orion 290A (Thermo Electron Corporation, Waltham, MA, USA). Total nitrogen (N) and total phosphorus (P) concentrations were calculated from water samples taken between 9 November and 8 December 2012. Nutrient water sample bottles (250 mL bottles 250 mL Nalgene®) were acid washed in 10% HCl acid overnight, triple rinsed and filled with bubble free deionized (DI) water. The DI water was emptied in the field, and the bottles were filled and emptied with stream water three times before collecting about 200 mL of flowing water with no sediments. Samples were immediately stored on ice and frozen within 24 h of sampling. Samples were transported and analyzed within 14 d of sampling for total N and P at the laboratory of Alan Wilson at the Department of Fisheries and Allied Aquacultures, Auburn University, 203 Swingle Hall, Auburn AL 36849.

Total carbon © was obtained by filling sulfuric acid prepped 40 mL vials with flowing stream water free of sediments. Samples were stored on ice and transported to Environmental Resource Analysts, Inc. for analysis within 7 d of sampling.

4.4 Biological Assessment

Fish assemblage composition was assessed by backpack electrofishing employing 2 passes and division of the site into segments to quantify species-reach length relationships. Benthic macroinvertebrate assemblage composition was assessed using a multiple habitat sampling approach. Biological sampling is discussed in detail below

4.4.1 *Fish Assemblages*

Electrofishing samples were collected from 70 samples sites during summer and fall in 2009, 2010, 2011, and 2012 (Table 1). A representative stream reach was selected for backpack electrofishing at each sample site. Reach lengths varied from 150 m for smaller streams to 285 m for larger streams based on species-area relationships and logistical feasibility; however, most reaches were ~ 200 m long. Each reach was divided into contiguous 10 to 15 m segments with longer segments in longer reaches (Figure 7). A fish storage bucket was placed at the top of each segment to hold fish collected from the segment. Block nets were not used because of the habitat disturbance associated with setting them and the difficulty of keeping them in place (especially when clogged with debris). Our field observations indicated little movement of most fish during sampling except for occasional larger fish that sometimes darted ahead or behind sampling personnel as they moved through stream segments. Fish were collected using one or two Smith-Root LR-24 Backpack Electrofishers. Electrofishers were calibrated for each stream below the sample reach prior to sampling using the Smith-Root 'Quick Setup' function, which automatically adjusted output to an acceptable level determined by conductivity at each site. In some cases further adjustments were necessary to optimize collection efficiency while minimizing fish mortality. Safety procedures were followed and protective equipment was worn to guard against electric shock.

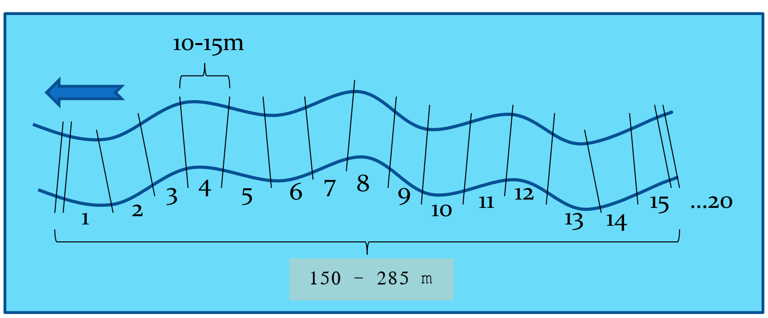


Figure 7. Electrofishing sampling reach divided into segments. Arrow depicts direction of flow.

A field team of 2-4 personnel was used for electrofishing, depending on stream size. Two passes were made through all segments at each site. At sites <3.5 m wide, two crew members sampled the stream, moving in an upstream direction, one in front with the backpack unit and a net and the second following behind with a net and a fish collection bucket. Sites measuring ~3.5 to 6 m wide required 2 backpack units in front with one netter following behind. Streams >6 m required two backpack units and two netters. This increase in effort with stream size was undertaken to maintain similar sampling efficiency in streams of different size. Care was taken to sample all available micro- and macro-habitats including riffles, pools runs, snags, logs, root-mats, undercut banks, etc. At the end of each segment, all fish captured within that segment were identified and enumerated, removed from the collection bucket and placed in the storage bucket at the top of the segment. Data collected at each segment included taxa present and the number of individuals of each taxon. Anomalies and disease were also noted. All fish, including young of the year (YOY), were counted. Fish not readily identified in the field were preserved in 10% formalin, identified, and verified at the Georgia Museum of Natural History (Athens, GA). After completing the first pass, the field crew recorded the shock time for the pass and returned to the first segment at the bottom (downstream end) of the reach for the second pass. All field-identified fish were returned to the stream from each storage bucket after the second pass. Total shock time (i.e., time that current was directed to the water) was commensurate with stream size and averaged 75 min (SD=40, range = 20-224) for the first pass and 56 min (SD=29, range = 12-163) for the second pass. Electrofishing samples were collected during the late spring, summer, and fall of 2009, 2010, and 2011. Electrofishing SOPs are provided in Appendix 2.

4.4.2 *Macroinvertebrate Assemblages*

Benthic macroinvertebrates were sampled from 64 wadeable streams of varying levels of disturbance during the summers of 2010-2012 (Table 1). Two multihabitat samples were taken with a 244 µm mesh D-frame dipnet over a reach of ~150 m. Each sample represented about 1 m^2 of combined depositional, CWD, root mat, and macrophyte habitats, respectively. Samples were preserved in the field with 95% ethanol and placed on ice during transportation to the laboratory. Samples were processed in the laboratory by emptying contents into stacked, graded sieves of 2 mm and 250 µm mesh and rinsing them thoroughly. Contents of the 2 mm sieve were transferred into a sorting tray and immersed in water. Large and rare organisms were coarse picked for 0.5 h (or more if the sample contained a large number of organisms). After the coarse pick was completed, the sample residue was re-washed through the sieves. Material retained in the 250-µm sieve was transferred to a 1000 mL beaker, which was filled with ~600 mL of salt solution (≥22 ppt NaCl) and agitated to separate organic matter from sand by elutriation. After elutriation, the remaining material was homogenized in a 1000 mL beaker of water, and two 25 mL aliquots were removed as a single subsample (5% of sample volume). Subsamples were microscopically sorted (4-10x) until 300 or more macroinvertebrates were removed. Oligochaetes were identified to the class level, whereas all other taxa were identified to lowest practical taxonomic level (usually genus or species group) with appropriate keys (Kowalyk 1985; Epler 1996; Epler 2001; Merritt et al. 2008; Thorp and Covich 2009). After resolving ambiguous taxa (Cuffney et al. 2007), counts of individuals were numerically combined to estimate densities for each sample reach. A complete description of the preceding procedures is given in Appendices 3 and 4.

5.0 RESULTS AND DISCUSSION

The Results and Discussion is divided into five sections:
1) Reference Site Selection. This section describes the selection of sampling sites and methods used to separate sites that represented the reference condition from the remaining sample sites (hereafter called test sites).
2) Fish Assemblages. This section characterizes the fish assemblages within the study area and describes the development of reference models and assessment frameworks for fish.
3) Macroinvertebrate Assemblages. This section characterizes the macroinvertebrate assemblages within the study area and describes the development of reference models and assessment frameworks for macroinvertebrates.
4) Stream Hydrogeomorphology. This section presents an assessment framework based on stream hydrogeomorphology for separating reference from non-reference sites.
5) Performance Assessment: This section presents an in-depth evaluation of the performance of the assessment frameworks developed during this study.

5.1 Reference Site Selection

Reference site screening criteria were needed to exclude disturbed sites that could negatively bias recovery objectives specified by reference models. Selection of reference sites was accomplished through application of screening criteria based on abiotic site characteristics. These included the potential effects of historical land use inferred from recent and historical aerial photographs taken at some of the sample sites. Such historical information can help determine the prevalence of legacy effects at sites considered to represent reference conditions.

Department of Defense (DoD) and Department of Energy (DOE) installations as well as public and non-profit lands represented the best locations for finding least disturbed habitats within the ecoregion because of the relative lack of disturbance in these or portions of these areas (Figure 8). All properties were well forested and many were managed with prescribed burning to promote longleaf pine ecosystems as recommended to support threatened and endangered species of the SH, particularly the red-cockaded woodpecker (*Picoides borealis*) (Jordan et al. 1997; USFWS 2003).

Preliminary criteria were initially developed for selecting probable reference sites and several disturbed sites for comparison with the reference sites. Refined criteria were later developed for the definitive selection of reference sites.

5.1.1 *Preliminary Reference Site Selection*

A comprehensive process was used for preliminary reference site selection on the SRS Site, where extensive data were available describing current and historical levels of disturbance. These criteria were based on a combination of existing information and extensive field surveys of SRS streams. Disturbance histories for proposed sites were established by field surveys and by viewing 1938, 1943, 1951, 1955 and more recent aerial photos, 1943 US Geological Survey (USGS) topo maps, and more recent maps. Reference sites were selected that had the least pre-SRS disturbance and no current SRS impacts. Sites were also selected that had known pre-SRS disturbance, but had been recovering for more than 50 years, and that had present-day impacts such as stormwater runoff. Criteria used to select reference sites included the following:
1) No un-natural present day runoff;

2) No evidence of excessive historical runoff; i.e. absence of large gullies along the valley walls and severe stream incision;
3) No active flow impediments (dams, road crossings, railroad crossings);
4) No abandoned flow impediments (based on field surveys, maps, and aerial photos) that appeared to have long-term hydrological impacts;
5) Present day, intact forest in the entire valley (floor and sides);
6) Connectivity to larger streams; and
7) Status of the watershed and drainage. Entire reference watersheds were preferred; however, both reference streams within the best available watersheds and reference streams in disturbed watersheds were included in the sampling plan.

The extensive preexisting data available for the SRS were unavailable for the other installations necessitating adoption of other preliminary screening criteria, as follows:
1) Minimal watershed disturbance as indicated by the % Disturbance Index. This catchment based metric was derived from the National Land Cover Database and consisted of the percentage of the catchment consisting of bare ground, developed land cover (low-, medium-, and high -intensity development), and agricultural land cover. This measure, which increased with disturbance, accounted for potential watershed-scale disturbances from impervious surfaces, sedimentation, and nutrient enrichment. Candidate reference sites for biological sampling were generally selected from watersheds with a Disturbance Index value <10%.
2) Lack of obvious disturbance based on visual examination during preliminary stream walkdowns. Specific criteria for candidate reference sites included the following:
   a) No evidence of un-natural present day or historical runoff;
   b) No active or abandoned flow impediments (dams) or point discharges; and
   c) Present day, intact forest in stream valley (floor and sides).

5.1.2 *Final Reference Site Selection*

Selection of reference sites for biological assessment has historically been based on best professional judgment (BPJ) (e.g. Stoddard et al. 2005). More recently, criteria driven selection methods based on chemical, physical, and geographic parameters are used to identify sites as reference quality (Waite et al. 2000; Whittier et al. 2007; Herlihy et al. 2008). We combined 3 independent methods to create a tiered system for selecting reference quality streams:
1) Method 1 used a disturbance gradient derived from a principal component analysis (PCA) of several measures of stress (described previously) including the SCDHEC habitat quality score, erosion score, channel modification score, bank height, paved road area (m^2), unpaved road area (m^2), number of road crossings, watershed disturbance (%), and percentage bare ground for 70 stream reaches (Table 6). PCA is an indirect ordination technique based on an eigenanalysis of a covariance or correlation matrix (in this case correlation, McCune et al. 2002). Rankings of candidate sites on the disturbance gradient were visually inspected for an inflection point that delineated reference from non-reference. PCA was conducted with PC-ORD (McCune and Mefford 1999).
2) Method 2 was based on the channel cross-sectional surveys of 62 stream reaches. Discrimination between reference and non-reference streams was based on the assumption that predictable relationships exist between channel morphology and watershed area and that streams deviating in

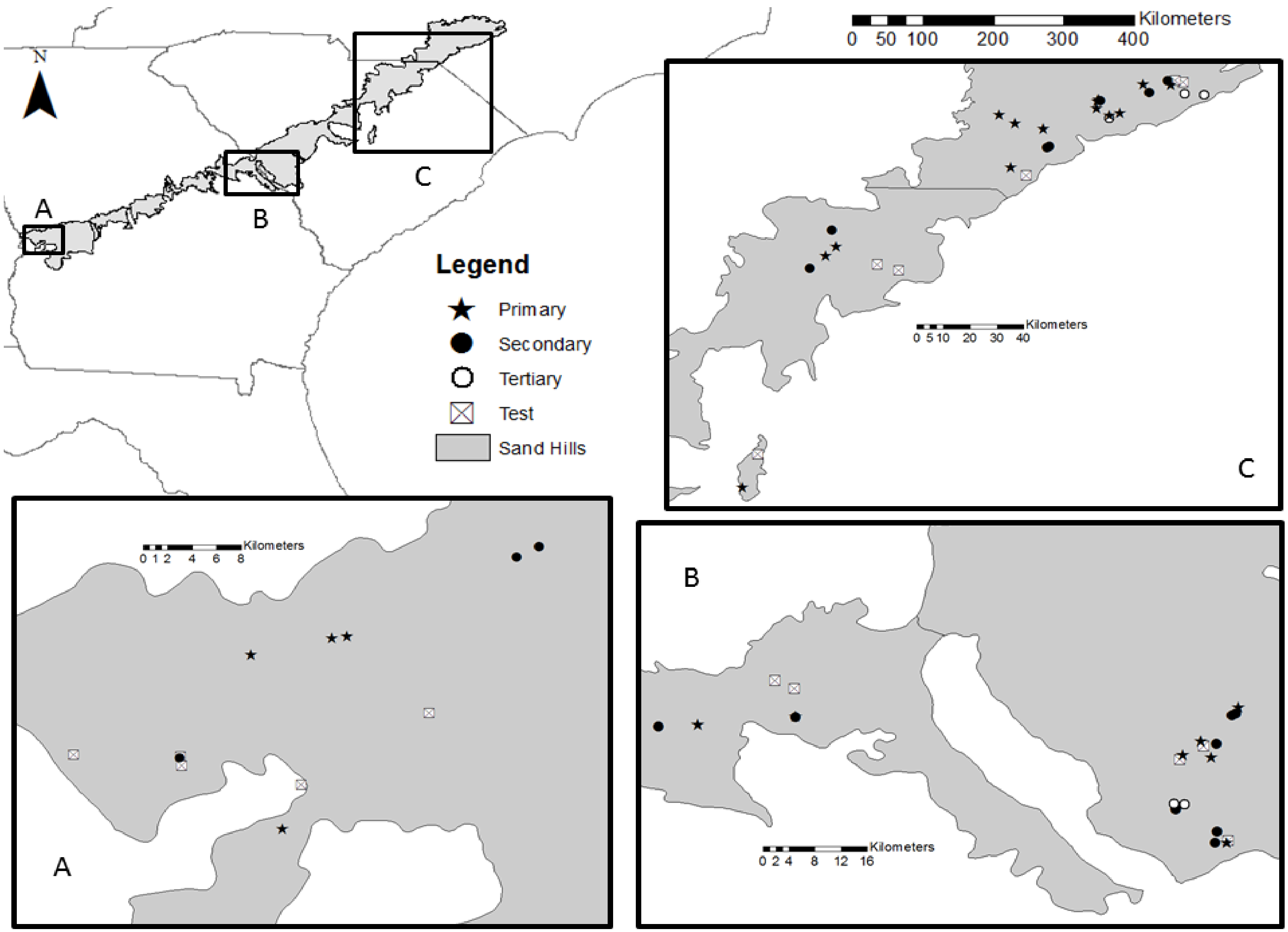


Figure 8. Sampling sites selected as reference or test within the Sand Hills Ecoregion. See text for further information.

morphology from the reference condition, given watershed area, represent a non-reference population. To quantify this deviation, residuals were generated from iteratively re-weighted least squares analysis of mean top of bank width, top of bank height, and top of bank area, separately on watershed area (Table 7). A multivariate distance (Cd_o) of channel morphology residual distance from a common centroid was calculated as a measure of variable interaction:

$$Cd_o = \sqrt[2]{(0-(A_r/3))^2 + (0-(T_r/3))^2 + (0-(W_r/3))^2} \quad (1)$$

where A_r was top of bank area residuals, T_r was top of bank mean height residuals, and W_r was top of bank width residuals. Model residuals and Cd_o were clustered used the partitioning around the mediods technique to delineate disturbed streams from non-disturbed (reference) sites. Robust regression was performed with the rlm function in the MASS package and cluster analysis with the cluster package for R software version 3.0.1 (R Development Core Team 2010).

3. Method 3 used 7 physicochemical stress measures as screening criteria (Table 8). Stream geomorphology date were used to compute a measure of incision termed the “modified bank height ratio (*BHRmod*)”:

$$BHRmod = \left|\left(\frac{T}{P}\right) - 1\right| \quad (2)$$

where *T* was the mean top of bank height measured in the field and *P* was the mean bankfull predicted from watershed area based on regional curves for the southeastern coastal plain hydrologic region (McCandless 2003). Cutoff points for stream water total N and P were from Herlihy et al. (2008); remaining variable cutoff points were determined by inspecting plotted ranks of each variable for an inflection point. A "two strike and out" rule was implemented so that a stream was considered reference if it passed 6 out of the 7 habitat criteria.

Table 6. Reference selection method 1 based on disturbance-related variables used in a principal component analysis (PCA) of sample sites. Standardized eigenvectors (i.e., correlation coefficients) and unstandardized eigenvectors of PCA axis 1 are also given. PCA axis-1 explained 41.9% of the variance and was the only significant axis. SCDHEC =South Carolina Department of Health and Environmental Control.

Variables used	Data derivation	PCA axis-1 standardized eigenvector	PCA axis-1 unstandardized eigenvector
SCDHEC habitat quality score	On-site evaluation	-0.79	-0.41
Erosion score	On-site evaluation	0.78	0.40
Channel modification score	On-site evaluation	0.70	0.36
Bank height (m)	On-site evaluation	0.69	0.36
Paved road area (m^2)	GIS	0.87	0.45
Unpaved road area (m^2)	GIS	-0.08	-0.04
Number of road crossings	GIS	-0.06	-0.03
Watershed disturbance (%)	GIS	0.88	0.45
Percentage bare ground (%)	GIS	-0.19	-0.10

Table 7. Reference selection method 2 based on regression analysis. Results and equation coefficients from robust regressions generating residuals that were used in cluster analysis to separate channel morphology reference from non-reference. Each variable was modeled with watershed area.

Variable	Coefficient	y-intercept	*F*	*p*
Top bank width	2.52	9.54e-08	$F = 43.14$	$p < 0.0001$
Top bank height	0.35	1.22e-06	$F = 25.21$	$p < 0.0001$
Top bank area	0.74	8.39e-08	$F = 75.69$	$p < 0.0001$

Table 8. Reference selection method 3based on 7 disturbance-related variables. Variables and criteria used to delineate reference from non-reference. Sites had to pass at least 6 criteria to be considered reference. [a] = from Herlihy, et al. 2008. BHR = modified bank height ratio.

Variable	Criteria
Stream water specific conductance (µS)	< 58
Watershed road density (%)	< 6.2
Stream water pH	< 6
Stream habitat quality scores	>158
BHR mod	< 0.46
Stream water Total N [a] (µg/L)	< 1000
Stream water Total P [a] (µg/L)	< 30

Some sites with clear signs of disturbance (e.g. presence of household trash, excessive sedimentation determined from previous studies) were designated as non-reference by BPJ, regardless of their outcome of each selection method. BPJ was used only to discount sites as reference, thereby acting as a final screening criterion to ensure that only the best sites were classified as least disturbed.

A site was characterized as a primary reference site if it passed all 3 screening methods, a secondary reference site if it failed 1 of 3 screening methods, and a tertiary reference site if it failed 2 of 3 screening methods. Sites that did not pass any of the criteria or were not considered reference due to BPJ were classified as test sites.

The generalized linear model (GLM) procedure in SAS® version 9.2 (SAS Institute Inc. ,2004) was used to conduct an unbalanced ANOVA to test for significant differences in key macroinvertebrate metrics across primary, secondary, and tertiary reference groups. The metrics included Ephemeroptera, Plecoptera, Trichoptera (EPT) richness, North Carolina biotic index (NCBI) (NCDENR 2006a), and the Georgia SH Macroinvertebrate Multi-Metric Index (GAMMI) (GDNR 2007). These metrics were chosen because they represent commonly used tools for assessing the biological integrity of macroinvertebrate assemblages in the region. Reference groups depicting "levels" or tiers of quality were delineated based on the test results.

The PCA used in Method 1 showed that most of the stress related variables were strongly correlated with PCA axis-1, which was interpreted as a disturbance gradient (Table 6). PCA axis 1 accounted for 41.9% of the variation among the sample sites and was the only significant PCA axis (randomization test p =

0.001). The ranking of candidate sites along PCA axis-1 identified 56 stream reaches as reference and 14 as non-reference based on a visually inspected inflection point (Figure 9).

Robust regression of Method 2 found all 3 variables to be significantly related to watershed area with the measure of mean top of bank area being the most significant (Table 7). The maximum average silhouette width indicated two clusters for partitioning around mediods; cluster analysis selected 51 streams as reference and 11 as non-reference.

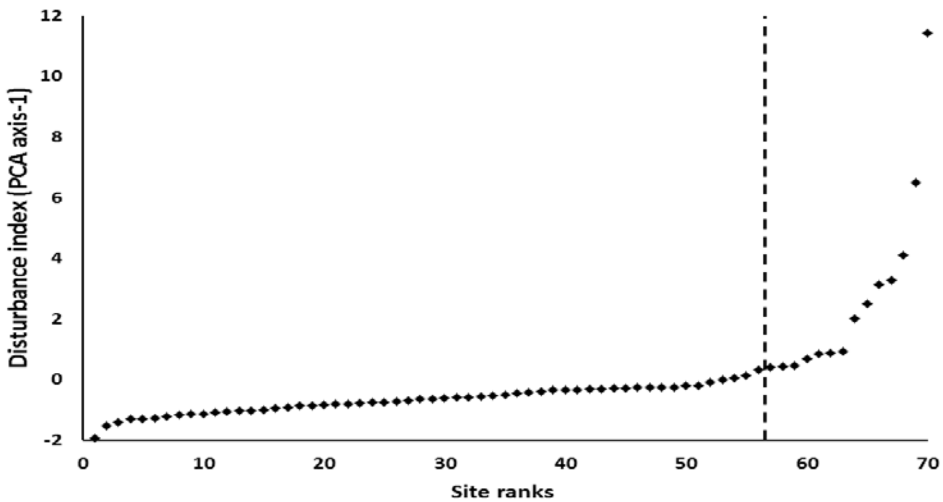


Figure 9. Plot of stream sites ranked along a disturbance gradient interpreted from PCA axis 1 scores. The vertical dotted line indicates the cut off between candidate reference (left side) and non-reference (right side) sites.

Method 3 selected 44 reference and 22 non-reference sites based on criteria developed from visual inspection of inflection points of the 7 physical-chemical stressors (Figure 10, Table 8). Eleven sites were considered non-reference due to BPJ even if they were selected as reference by any of the 3 methods. Primary reference sites were distributed throughout the northeastern and central regions of the SH, with only 4 sites represented in the southwestern region (Figure 8).

Of the 64 stream reaches included in this assessment, 57 were evaluated with all 3 reference site selection methods, and 7 were evaluated with a combination of 2 methods. Cross comparison of each stream reference designation resulted in 25 primary, 18 secondary, and 6 tertiary reference sites. Four sites were considered as non-reference due to failure to pass any of the reference condition screening

EPT richness was significantly different across reference groups ($df = 3$, $F = 8.34$, $p < 0.0001$); sites designated as primary and secondary were not significantly different from each other, but were significantly different from tertiary reference and test sites (Figure 11). The NCBI was significantly different among reference groups ($df = 3$, $F = 9.98$, $p < 0.0001$); primary and secondary reference sites were not significantly different from each other, but primary reference sites were significantly different from tertiary and test groups. The GAMMI ($df = 3$, $F = 2.19$, $p = 0.0987$) was highly variable

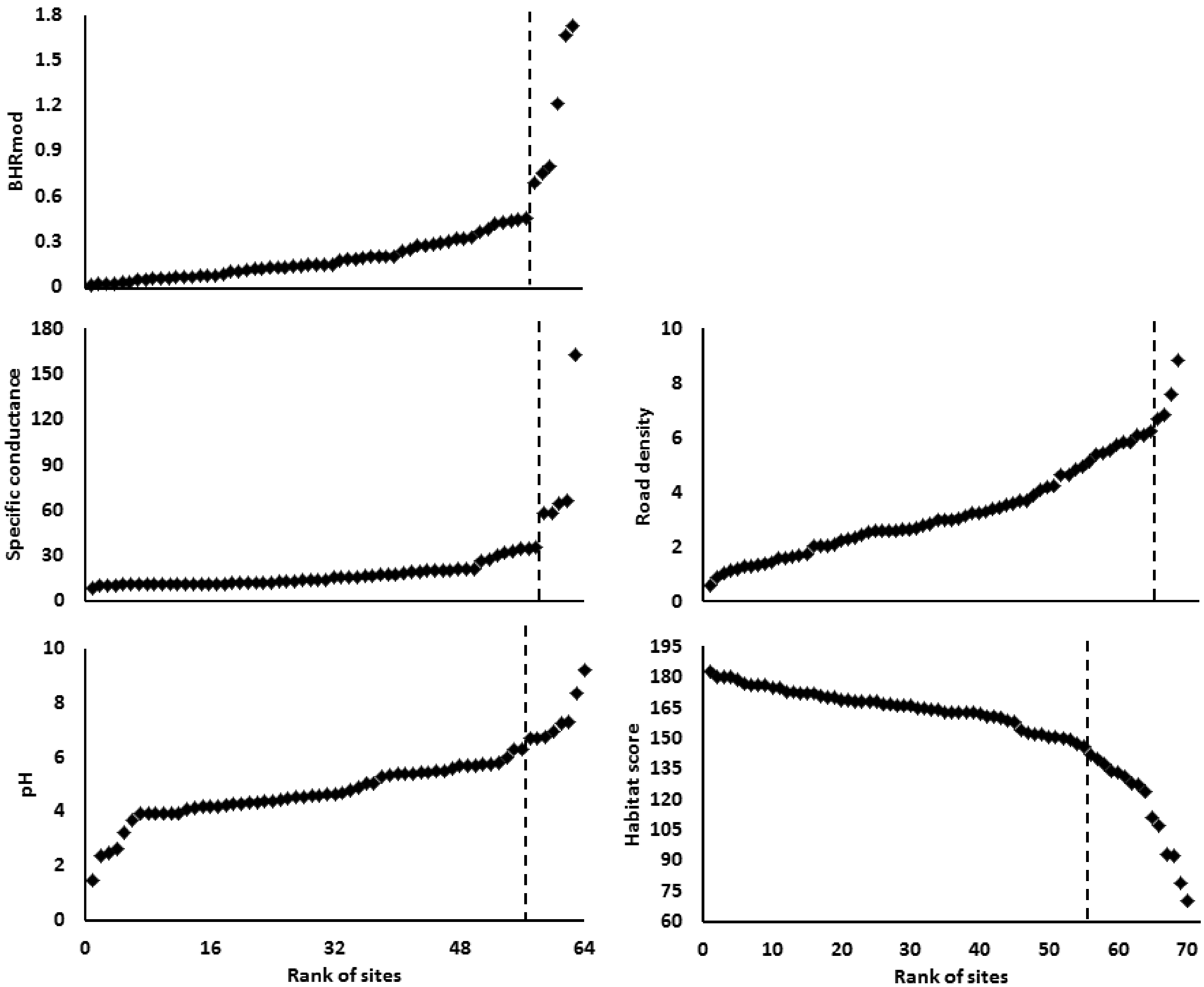


Figure 10. Reference criteria derived from visually inspected inflection points along disturbance gradients. The vertical dotted lines indicate the cut off between candidate reference (left side) and non-reference (right side) sites. Road density and habitat scores were plotting in association with 70 sites to establish reference criteria, but only 64 of these sites were evaluated with macroinvertebrate community measures.

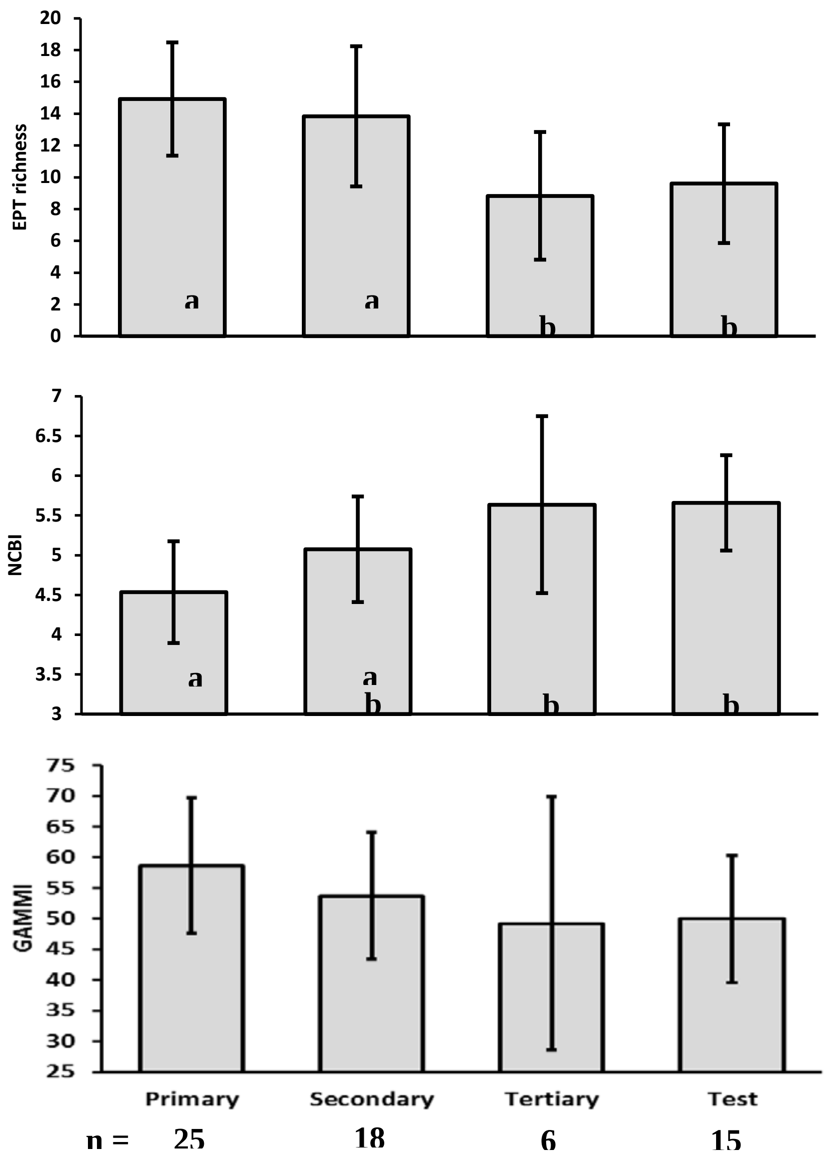


Figure 11. Means (+ 1SD) of sites determined as primary reference, secondary reference, tertiary reference, or test for EPT richness, North Carolina Biotic Index (NCBI), and the Georgia Macroinvertebrate Multimetric Index (GAMMI). Numbers below bars are sample size, and bars with the same letter for each metric indicate means were not significantly different.

and did not show any difference among reference groups and test sites (Figure 4). These results show that the method of reference site selection can have a significant impact on the biological expectations associated with the reference condition.

5.1.3 *Historical Impacts at Reference Sites*

Sites classified as "reference" in this study were from largely forested landscapes with little or no anthropogenic development within their watersheds. All were in large, protected government installations, primarily owned by the DoD and DOE, with relatively low population densities and no private ownership for decades (Table 9). Development for military and industrial functions was present, but the anthropogenic "footprint" was less than in the surrounding land as shown by examination of satellite photographs (Figure 12). Therefore, it is reasonable to conclude that the reference sites, which were in the least developed parts of the installations under study as well as in other protected land holdings, represented some of the best stream sites in the upper Southeastern Plains. However, ecosystems within the this ecoregion been extensively modified by agriculture, forestry, and other anthropogenic disturbances for well over 200 years, so it is likely that even these reference sites had a history of disturbance.

Table 9. Year that land holdings included in RC-1694 were removed from private ownership.

Location	Year established
Fort Bragg, NC	1918
Fort Benning, GA	1918
Manchester State Forest, SC	1939
Sandhills State Forest, SC	1939
Fort Gordon, GA	1917
Camp McKall, NC	1942
Sandhills Gamelands, NC	1942
Savannah River Site, SC	1950

To assess the disturbance history of our sites, we used historical and recent aerial photographs to estimate changes in tree canopy coverage at the reference sites. Historical photographs were available for 24 sites classified as reference and 10 sites classified as disturbed. Sites were located on the SRS (aerial photographs from 1951 and 2010), Fort Benning (1944, 1999, and 2009), and Fort Gordon (1941, 1963, and 2006). Historical photographs were available for one site on Fort Bragg (1951 and 2009).

Examination of historical photographs showed that canopy coverage at nearly all of the 34 sites with historical photographs was <50% and as low as ~20% during the 1940s to 1950s (the earliest years of record) (Figure 13). By 2000 to 2010 canopy coverage at all of the reference sites increased to >70% with many sites reaching 80 to 90%. Increases were observed at all sites on the SRS and most sites on Fort Gordon and Fort Benning. Canopy coverage also increased at most disturbed sites for which historic

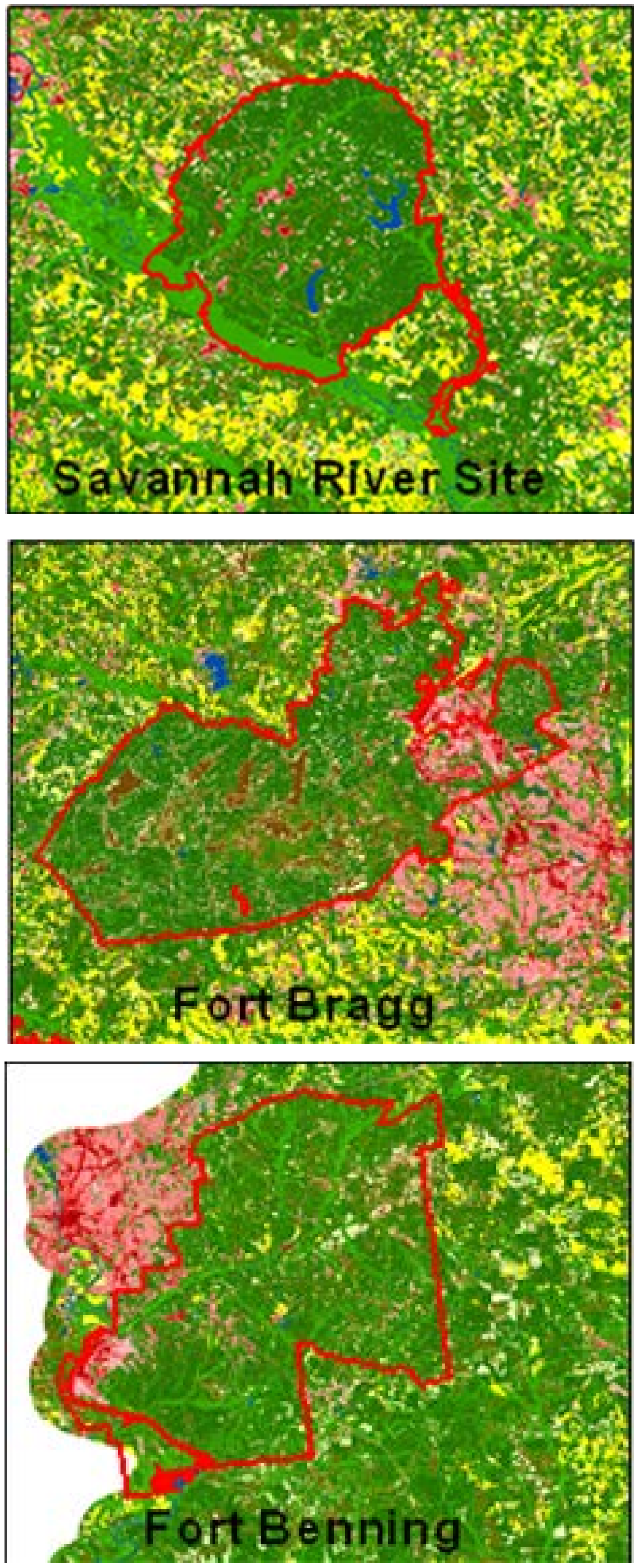


Figure 12. Satellite photographs of the Savannah River Site, Fort Bragg, and Fort Benning. Green areas represent forest cover.

photographs were available, although a few showed subsequent declines. This analysis showed that most study watersheds have been largely reforested during the last 60 to 70 y. It also showed the pervasiveness of historical disturbances within the upper Southeastern Plains, even at the best reference sites. It is possible that effects of this disturbance on aquatic communities are negligible following decades of secondary succession; however, this assertion is impossible to verify because of the unavailability of historically undisturbed sites that could serve as a benchmark. Therefore, following Stoddard et al. (2006), the reference sites in this study were categorized as “least disturbed” rather than “minimally disturbed.” As such, our sites represent the best of what is available within the ecoregion and are unlikely to improve further within a time span relevant for most management and recovery decisions.

5.1.4 *Reference Site Selection Summary*

Defining reference conditions on a basis of differing "levels of quality" allows for use of different benchmarks to meet different objectives. Use of secondary or even tertiary reference conditions to construct reference models may lower the benchmark; however, limitation to a strict set of core reference sites may fail to account for the range of natural variation that occurs across the SH. Utilizing some secondary reference sites may be warranted for regions that are underrepresented by the highest quality sites. Developing a reference model for the SH utilizing only primary reference sites, would mean that streams in the southwestern portion of the ecoregion would likely be evaluated by streams occurring outside of their range from completely different catchments. Although within the same ecoregion, streams in the southwest may differ because of natural factors or historical land uses so that recovery end-states differ compared with central or northeastern portions of the study area. Therefore, inclusion of secondary reference conditions would establish an assessment with more realistic expectations given the size and variable environmental conditions characterizing this region.

Adopting different sets of reference conditions also encourages the development of different models that may be appropriate for different objectives. For example, primary reference sites may be suitable for recovery programs oriented towards maximum ecological integrity. Tertiary reference conditions may be more appropriate when prevailing or historical land uses impose limitations that preclude this goal; i.e., when best attainable conditions constitute a more realistic benchmark (Van Sickle et al. 2006). Similarly, references of different overall quality may be useful with regard to developing models based on specific stressors that limit the ecological recovery of different taxonomic groups. Fish and macroinvertebrates responded differently to the reference site groupings, suggesting that reference site criteria may not be identical for these assemblages that differ in environmental sensitivities and requirements (Paller 2001).

Tiered reference model selection allows for construction of different models that target a broad range of questions. For instance, a strict model that evaluates sites with a high standard of quality would be based on models that only include primary reference sites. However, a more flexible model also could include secondary reference sites as a means of accounting for more regional variation, which in some cases may be more relevant for evaluating restoration success. DoD installations are managed for multiple land uses including, military training, logging, and endangered species protection. Military training can create considerable impacts to terrestrial and aquatic ecosystems (Quist et al. 2003; Maloney et al. 2005); however, military installations also represent some of the largest refuges for threatened and endangered species (Cohn 1996; Kaufman 2010). In line with the mission of military readiness and legal obligations such as the Endangered Species Act, land managers on military bases must balance these conflicting demands and therefore may need a broad set of tools for evaluating and assessing habitat conditions.

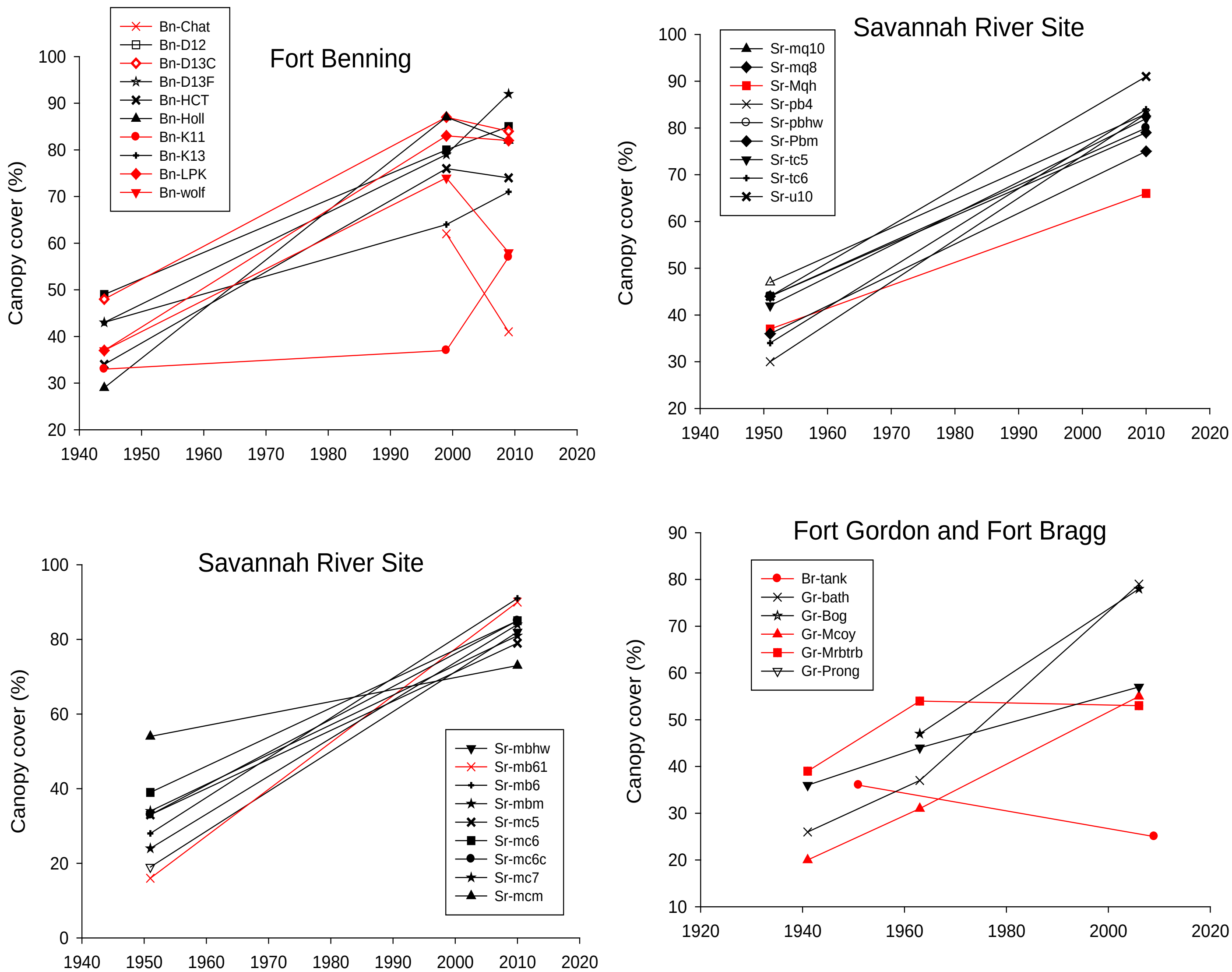


Figure 13. Changes in tree canopy cover calculated from selected aerial photographs (Savannah River Site: 1951 and 2010; Fort Benning: 1944, 1999, and 2009; Fort Bragg: 1951 and 2009; Fort Gordon: 1941, 1963, and 2006). Black = reference sites, Red = test sites. Sample sites described in Table 1.

We have retained the term “least disturbed” to define the quality of our reference conditions (Stoddard et al. 2006). Almost all of the areas of the SH in our study have undergone some type of anthropogenic alteration in recent history as shown by the analysis of historical aerial photographs; however, some areas have been under strict management and/or excluded from human disturbance for many years and may be approaching a state of minimal disturbance (Stoddard et al. 2006). Many of our primary reference sites occur within military installations that were formerly clear-cut farm lands and are currently managed for silviculture and fire control (bull-dosed fire breaks every 300 m) in addition to being subjected to training exercises. Such sites are influenced by a conglomerate of past and present human activities. However, they are less exposed to stressors than other areas in the SH, and are managed to maintain a high degree of naturalness, and are considered here as least disturbed, clearly representing the best reference conditions for the region.

Reference sites selected by differing methods and representing different degrees of rigor provide an effective means of designating the least disturbed sites within a region such as the SH that has undergone extensive historical land use, shifts in land use, and is still strongly influenced by human activities. Identifying reference sites as representing secondary or tertiary conditions allows for the construction of models that may serve different purposes and may be more robust in accounting for regional variation. In addition to developing such tiered models, we emphasize the recognition that best attainable conditions are not necessarily being attained in all cases. Other models that do not distinguish the quality of their references imply that the best conditions are being attained, whereas this may be misleading. In Sections 5.2.9 and 5.3.3 of this report, we specify the reference models (primary secondary or tertiary) used for different applications and the reasons for our choices.

5.2 Fish Assemblages

SH streams support diverse fish assemblages with many species in the families Cyprinidae, Ictaluridae, and Centrarchidae, although other taxonomic groups are also well represented (Marcy et al. 2005). Fish distributions within the SH is affected by geography: endemism among drainages is common, sometimes resulting in the need for drainage specific management and monitoring strategies (Swift et al. 1986, Warren et al. 2000). A site's position along the stream/river continuum is another important factor because it may influence habitat structure and trophic resources, thereby influencing species occurrence and abundance. A distinctive characteristic of SH streams is comparatively high species richness in headwater reaches, possibly because of relatively stable water levels from high groundwater recharge (Paller 1994). Localized differences in habitat due to stream geomorphology, hydrology, and the distribution of instream structure can also affect species occurrence and abundance on a smaller scale. Other important factors include connectivity with source pools of colonizing individuals such as larger streams and man-made lentic habitats such as ponds and reservoirs, the latter acting as a source of non-native species that can cause departures from the reference condition.

This section of the report describes fish assemblage data collected under RC-1694. It begins by describing the development of data quality objectives that ensured the adequacy of fish assemblage sampling methods. It proceeds by summarizing the biodiversity of fishes within the study area and identifying factors that contributed to the observed variability in assemblage distribution. From baselines established by the preceding analyses, reference models are identified that depict the expectations for least disturbed fish assemblages across the study area. Last, these reference models are used to build assessment frameworks that are sensitive to departures from the reference condition and that can be used to assess ecological health and progress towards recovery goals.

5.2.1 *Data Quality Objectives for Fish Sampling*

Data quality objectives (DQOs) were developed for fish assemblage sampling to ensure that data were of known quality, spatially and temporally comparable, representative of the parameters being measured, and suitable for addressing objectives of the project. Specific issues addressed for fish sampling included the reach length and number of electrofishing passes required to accurately represent species richness and composition.

Species richness, the number of species in a given area, is affected by the area sampled and by the probability of detecting each species within the sample area (Rosenzweig 1995, Cam et al. 2002, Rosenzweig et al. 2003). Probability of detection, in turn, is a function of other variables including sampling effort and efficiency. The division of sample sites into segments and the use of 2 electrofishing passes through the sampling area (i.e., two levels of effort) permitted us to assess the relationships among area sampled, level of effort, and species richness. This information helped us to identify protocols that produced adequate estimates of assemblage structure. Division of each sample site into smaller sample units (i.e., segments) also permitted use of an estimator to approximate true total species richness at each site. Estimated species richness is typically higher than observed species richness and more accurately approximates true species richness (McCune and Mefford 1999).

The preceding issues were investigated with data collected from the first 7 sites sampled in the study. All seven were located at the Savannah River Site. Fishes were collected by backpack electrofishing. Analyses were performed with PCORD (McCune and Mefford 1999), which includes methods for analyzing the relationships among area sampled, number of species, and species composition. The program used subsampling to determine the average number of species collected as a function of the size of the sample area (stream reach length in this case). It also calculates the average Sorenson (Bray-Curtis) distance based on species composition between subsamples of different size and the overall species composition as a method of determining the sample area needed to accurately represent species composition. PCORD includes several estimators for estimating total species richness from patterns of species accumulation in the sample segments. We selected the first-order jackknife estimator because it is generally accurate when the number of sample subunits is comparatively small.

The accuracy of the first-order jackknife estimator was verified by first computing estimated species richness individually for 2 sites, Sr-mc6 and Sr-mc6c (150 and 130 m, respectively) separated by ~200 m on the same stream and then combining the sites into a single 280-m aggregate site. Observed richness at each site was 7 at Sr-mc6 and 10 at Sr-mc6c, whereas estimated richness was 10 and 11, respectively. These estimates compared well with observed richness (11) for the combined sites. The jackknife estimator was then used for the previously described 7 sites to compute the estimated species number from simulated sample reaches of different length obtained by including different numbers of sample segments in the analysis.

A total of 1091 fish representing 21species was collected from the 7 sites. Richness per site ranged from 4 to 16, with fishes within the family Cyprinidae being numerically dominant. The second electrofishing pass through each sample site usually resulted in addition of at least one new species to for each stream, indicating that additional effort provided by the second pass contributed to the accurate estimation of species richness. These were often small, benthic species, such as madtoms or darters, that were either overlooked during the first pass or that might have been flushed from hiding places by the disturbance associated with the first pass.

Species-area (i.e., stream length) relationships were computed for each of the 7 sites (Figure 14). The results showed the expected asymptotic relationship between richness and sample reach length. They also showed the continued ascension of the species-area curves at the ends of the sample sites, indicating that observed richness underestimated true richness. The average difference between true richness for the stream reaches represented by the sample segments (estimated by the first-order jackknife estimator) and observed richness was a little more than 2: observed richness was 9.6 compared to an estimated true species richness of 12.1. Reducing the sample segment length to 100 m (by eliminating the last 3 to 5 sub-segments from each sample segment) resulted in a slight decline in observed richness to an average of 8.9, but estimated true richness remained about the same (average of 12.0). A further reduction of the sample segment length to 50 m resulted in a comparatively large reduction in both observed and estimated true richness (averages of 7.3 and 10.0, respectively) (Figure 15). Similarly, the relationship between segment length and Sorenson distance showed that segments of 100 m differed little (Sorenson distance <0.1) in species composition from the true species composition (as represented by the entire sample reach), whereas segments of 50 m sometimes misrepresented true species composition substantially (Figure 14). These results suggested that a sample segment length of 100 m (equivalent to about 60 to 70 stream widths in the 7 streams in this analysis) was sufficient to accurately estimate both true richness (with the first-order jackknife estimator) and species composition in small streams on the Savannah River Site when following our sampling protocols (stream site here being defined as a portion of a stream of comparatively uniform size and general similar habitat). However, to ensure accurate results, a minimum of 150 m was sampled at most sites, and 200-285 m was sampled in larger streams.

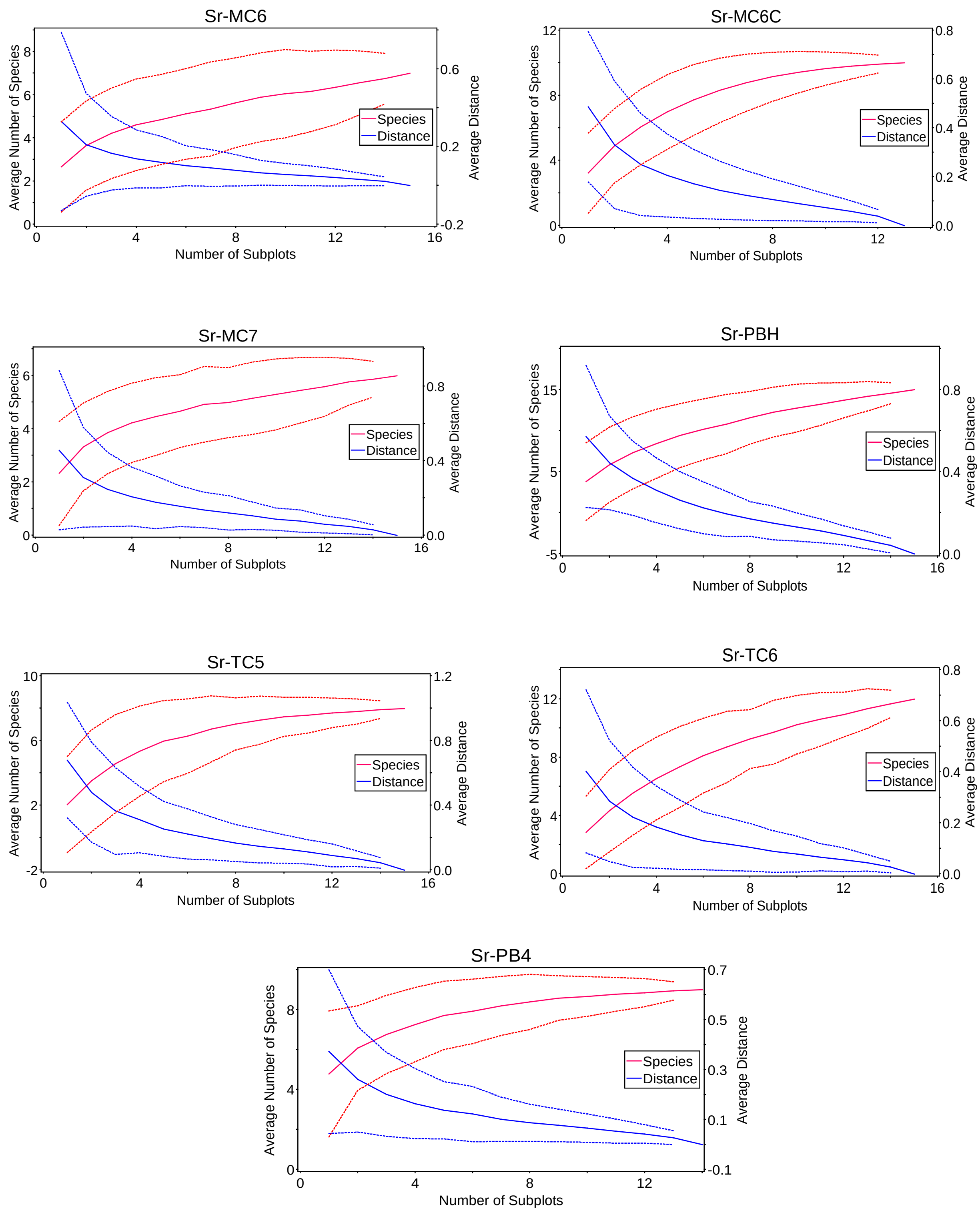


Figure 14. Relationships between sample segment length (m), number of fish species collected, and fish assemblage composition (as indicated by Sorenson distance) for 7 small streams on the Savannah River Site. Subplots represent 10-m stream sub-segments.

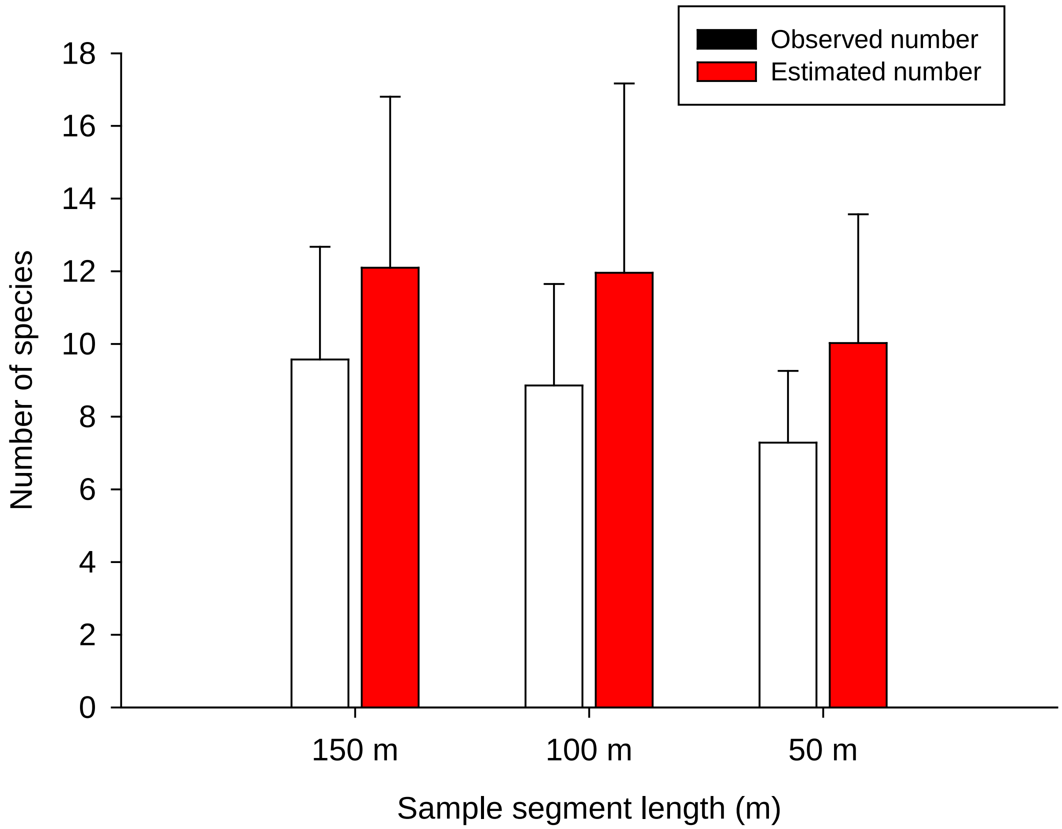


Figure 15. Mean (± 1SD) observed and estimated fish species richness for different electrofishing sample segment lengths. Estimated richness was computed with a first-order jackknife estimator.

5.2.2 *Fish Assemblage Structure*

A total of 55 fish species representing 15 families were collected from the study area (Table 10). The most speciose families in order were the Cyprinidae, Centrarchidae, Percidae, Ictaluridae, and Catostomidae. Identification of differences among installations is important because they can affect reference model development. A 2-way cluster analysis (based on Bray-Curtis similarities and using the group average linkage method) showed that several species including the yellow bullhead *Ameiurus natalis*, redfin pickerel *Esox americanus*, and pirate perch *Aphredoderus sayanus* were ubiquitous throughout the study area, whereas other species were more localized in distribution, which contributed to substantial geographic variability in assemblage structure (Figure 16). Fort Benning was distinguished by the broadstripe shiner *Pteronotropis euryzonus,* dixie chub *Semotilus thoreauianus,* and southern brook lamprey *Ichthyomyzon gagei*; Fort Bragg by the dusky shiner *Notropis cummingsae,* margined madtom *Noturus insignis,* and sandhills chub *Semotilus lumbee*; and the Savannah River Site by the yellowfin shiner *Notropis lutipinnis*, bluehead chub *Nocomis leptocephalus,* and creek chub *Semotilus atromaculatus*. There was also considerable variation among sample sites within installations. Prominent biotic differences among installations need to be incorporated into the structure of predictive reference models based on taxonomic structure. Multimetric reference models may be adjusted for such differences by developing installation-specific scoring criteria.

In describing ecological communities, it is often informative to supplement cluster analysis with ordination methods. Cluster analysis is a classification method that puts samples into groups. Ordination arranges samples along continuous environmental gradients and, in this respect, is a more ecologically meaningful approach. We used the ordination technique of non-metric multidimensional scaling (NMDS) to summarize patterns in fish assemblage structure across sample sites. NMDS is a relatively assumption free indirect ordination method based, in this case, on Bray-Curtis dissimilarity matrices (Clarke and Warwick 2001). Abundances were $\log_{10}(x+1)$ transformed to provide a balanced representation of common and rare taxa. The number of significant dimensions (axes) in the ordination was determined by a Monte Carlo procedure that compared the stress in the ordinations with the stress in randomized species arrangements (McCune and Mefford 1999). All ordinations were repeated with different random starting configurations to ensure a final solution with consistent and relatively low stress.

NMDS produced 3 significant axes ($p = 0.02$ for all) but only the first 2 displayed interpretable patterns. Sites were strongly segregated geographically, paralleling the patterns observed in the 2-way cluster analysis (Figure 17). Fort Bragg, Fort Benning, and the SRS produced well-defined groups. Nature Conservancy sites grouped with Fort Benning sites and Sandhills Gameland sites grouped with Fort Bragg sites, reflecting the geographic proximity of these preserves to the larger military installations. There was also substantial geographic segregation of species, with well-defined species assemblages associated with each installation, also paralleling the cluster analysis. In summary, both analyses indicated strong geographic patterns in fish taxonomic composition, but also showed considerable variation within locations.

Table 10. Fish collected from sample sites by backpack electrofishing.

Scientific Name	Common Name	Abbreviation	Family	Fort Bragg NC	Sandhills Game Lands NC	Savannah River Site SC	Manchester State Forest SC	Sandhills National Wildlife Refuge SC	Sandhills State Forest SC	Fort Gordon GA	Fort Benning GA	The Nature Conservancy GA	Totals
Anguilla rostrata	American Eel	eel	Anguillidae	14		24			15				53
Elassoma zonatum	Banded Pygmy Sunfish	bpsf	Elassomatidae						1	8			9
Percina nigrofasciata	Blackbanded Darter	bbdt	Percidae	2		12				8	68	14	104
Enneacanthus chaetodon	Blackbanded Sunfish	bbsf	Centrarchidae			2		1					3
Fundulus olivaceus	Blackspotted Topminnow	bstm	Fundulidae								4		4
Enneacanthus gloriosus	Blue Spotted Sunfish	bssf	Centrarchidae	12		3	4						19
Campostoma pauciradii	Bluefin Stoneroller	bfst	Cyprinidae								2		2
Lepomis macrochirus	Bluegill	bg	Centrarchidae	13		2				11	8		34
Nocomis leptcephalus	Bluehead Chub	bhc	Cyprinidae			322							322
Pteronotropis euryzonus	Broadstripe Shiner	bssh	Cyprinidae								612	417	1029
Labidesthes sicculus	Brook Silverside	bss	Atherinopsidae			1							1
Esox niger	Chain Pickerel	cp	Esocidae	5			1	1	3	6			16
Notropis petersoni	Coastal Shiner	csh	Cyprinidae	3									3
Semotilus atromaculatus	Creek Chub	cc	Cyprinidae			440	1	6					447
Erimyzon oblongus	Creek Chubsucker	ccs	Catostomidae	14	3	50			7		1		75
Semotilus thoreauianus	Dixie Chub	dc	Cyprinidae								355	172	527
Lepomis marginatus	Dollar Sunfish	dsf	Centrarchidae	26		51		4	4	28	2		115
Notropis cummingsae	Dusky Shiner	dsh	Cyprinidae	891	75	61							1027
Gambusia holbrooki	Eastern Mosquitofish	mf	Poeciliidae	1983		58							2041
Umbra pygmaea	Eastern Mudminnow	mm	Umbridae	28		4		5					37
Ameiurus platycephalus	Flat Bullhead	fbh	Ictaluridae							38			38
Centrarchus macropterus	Flier	fl	Centrarchidae						1		2		3
Notemigonus crysoleucas	Golden Shiner	gsh	Cyprinidae	1		2						2	5
Etheostoma parvipinne	Goldstripe Darter	gsdt	Percidae								46		46
Lepomis cyanellus	Green Sunfish	gsf	Centrarchidae								2		2
Etheostoma swaini	Gulf Darter	gdt	Percidae								6		6
Erimyzon sucetta	Lake Chubsucker	lcs	Catostomidae	36	9			3	1	2	1		52
Micropterus salmoides	Largemouth Bass	lmb	Centrarchidae	8	2	3				3			16
Fundulus lineolatus	Lined Topminnow	ltm	Fundulidae									1	1
Lepomis megalotis	Longear Sunfish	lesf	Centrarchidae								11		11
Pteronotropis stonei	Lowland Shiner	llsh	Cyprinidae			26	23		557	392			998
Noturus insignis	Margined Madtom	mmt	Ictaluridae	171	50	26	1		36				284
Acantharchus pomotis	Mud Sunfish	msf	Centrarchidae	38	8	9	22	12	2	8			99
Hypentilium nigricans	Northern Hogsucker	nhs	Catostomidae			5							5
Etheostoma mariae	Pinewoods Darter	pwdt	Percidae	1	87								88
Aphredoderus sayanus	Pirate Perch	pp	Aphredoderidae	286	57	348	98	54	15	49	123	33	1063
Lepomis gibbosus	Pumpkinseed	pks	Centrarchidae	19									19
Lepomis miniatus X L. punctatus integrade	Red-Black Spotted Sunfish Hybrid	rssf	Centrarchidae								7	10	17
Lepomis auritus	Redbreast Sunfish	rbsf	Centrarchidae	62		60					11		133
Esox americanus americanus	Redfin Pickerel	rfp	Esocidae	46	9	59	25	2	14	25	41	3	224
Semotilus lumbee	Sandhills Chub	shc	Cyprinidae	244	244								488
Etheostoma fricksium	Savannah Darter	sdt	Percidae			26				56			82
Etheostoma serrifer	Sawcheek Darter	swdt	Percidae	1	2		16						19
Ameiurus brunneus	Snail Bullhead	sbh	Ictaluridae			4							4
Ichthyomyzon gagei	Southern Brook Lamprey	lam	Petromyzontidae								492	360	852
Noturus leptacanthus	Speckled Madtom	smt	Ictaluridae			23				49	12	15	99
Minytrema melanops	Spotted Sucker	ss	Catostomidae			4							4
Lepomis punctatus	Spotted Sunfish	ssf	Centrarchidae			74				9	1		84
Chologaster cornuta	Swampfish	swf	Amblyopsidae				3						3
Noturus gyrinus	Tadpole Madtom	tmt	Ictaluridae	1		28			4	2			35
Etheostoma olmstedi	Tesselated Darter	tdt	Percidae	43		21			6				70
Lepomis gulosus	Warmouth	wm	Centrarchidae	52		2			1	11	16	1	83
Notropis texanus	Weed Shiner	wesh	Cyprinidae								19		19
Ameiurus natalis	Yellow Bullhead	ybh	Ictaluridae	160	11	23	28	22	42	39	39	8	372
Notropis lutipinnis	Yellowfin Shiner	ysh	Cyprinidae			1766							1766
				4160	557	3539	223	110	709	744	1881	1036	12959

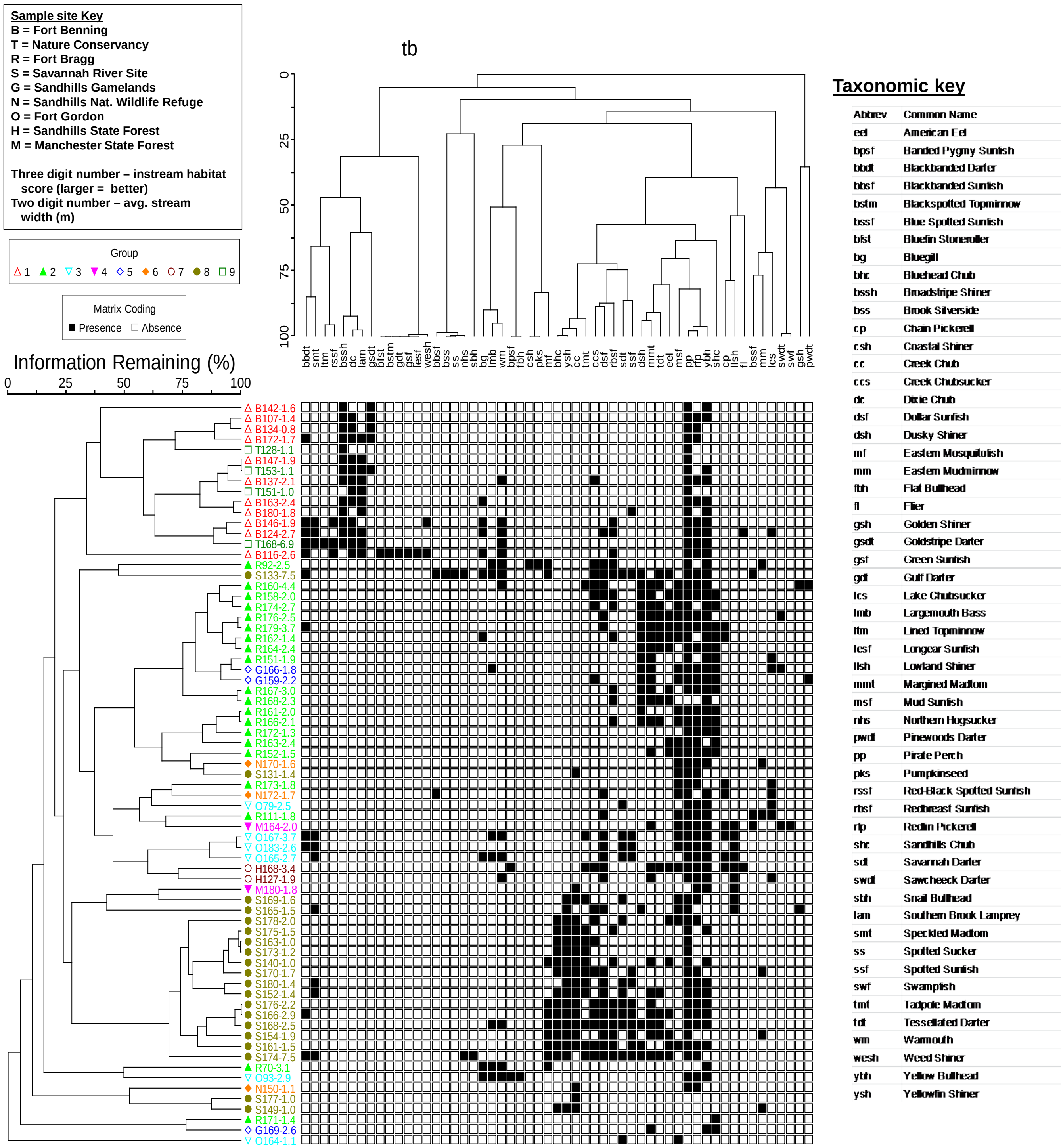


Figure 16. Two-way cluster analysis of fish assemblage data collected from Sand Hills sample sites. Site codes indicate location, disturbance, and stream size.

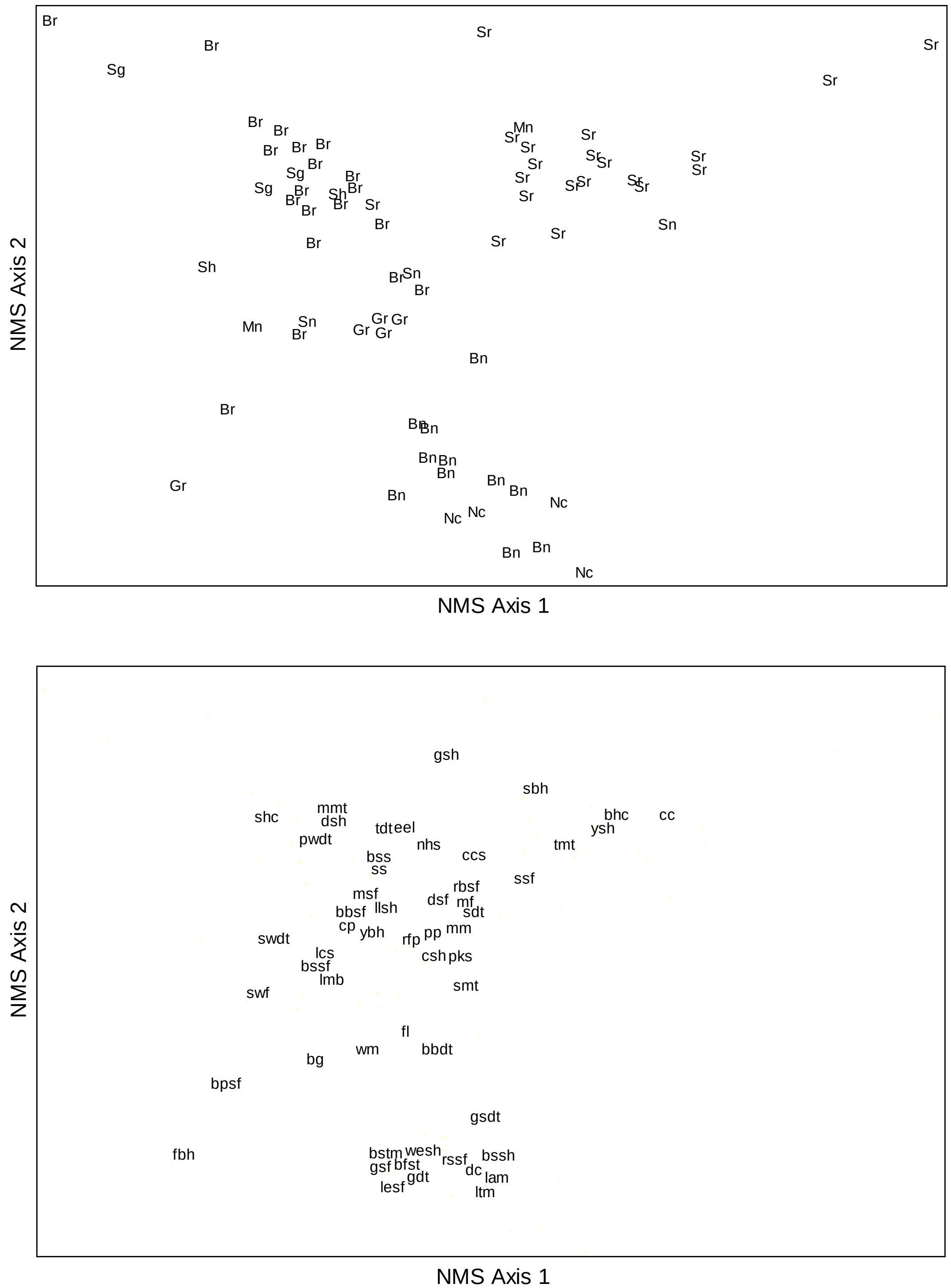


Figure 17. Nonmetric Multidimensional Scaling (NMDS) plots of samples sites (upper) and species (lower) based on fish assemblage taxonomic data collected from Sand Hills sample sites (Br = Fort Bragg, Bn = Fort Benning, Gr = Fort Gordon, Mn = Manchester State Forest, NC = Nature Conservancy, Sg = Sandhills Gamelands, and Sr = Savannah River Site). Juxtaposition of the plots indicates geographic occurrence of the species. Species abbreviations are presented in Table 10.

5.2.2.1 Habitat at Fish Sample Sites

The multi-scale habitat data were analyzed in conjunction with the fish assemblage data to better understand the factors potentially influencing fish assemblage structure within the study area. This was important for reference model development because reference models should be as general in scope as possible, but also specific enough to ensure that recovery objectives specified by the models are meaningful, appropriate, and achievable. To meet these goals, reference models must account for differences in habitat, zoogeography, and other putative casual factors. Small differences might justify the use of a single generic model, whereas large differences necessitate development of models that explicitly account for natural variation that exists among installations and/or inclusion of covariables that adjust for natural differences among and within installations. This section of the report summarizes watershed- and stream-scale habitat patterns within and among the installations included in this study.

As previously described, 2-way cluster analysis showed that there was considerable variability in fish assemblage structure among installations. Although geographic endemism was likely a contributing factor, there were also watershed and instream-scale habitat differences among and within installations that likely contributed to this variability. Most instream habitat variables were generally comparable among installations, but there were apparent differences such as higher streambed substrate diversity at the Savannah River Site (Table 11). Differences were also observed at the watershed-scale in terms of stream size (e.g., length, order, and magnitude – smaller at the SRS), geographic relief (e.g., mouth elevation, basin relief, and high point – higher at Fort Benning and lesser at the SRS), and degree of landscape disturbance (less at the SRS) (Table 12). Such factors, along with geographic endemism, may have contributed to differences among installations and among sites within installations.

Habitat differences among sample sites were investigated further with PCA, which was used to summarize and display patterns among sample sites based on co-variation in environmental data. By analyzing the correlation (rather than covariance) matrix, differently scaled environmental variables were normalized to avoid dominance of the analysis by variables with greater variance. The significance of each PCA axis was assessed by comparing the observed eigenvalues with eigenvalues generated by null models (McCune and Grace 2002). Because the total number of habitat variables was large in relation to the number of sample sites (which can lead to indeterminate results), they were divided into 3 groups for separate analysis: instream habitat variables (field), watershed morphometry (i.e., size and shape) (GIS), and watershed land cover (GIS).

Randomization tests showed that the first and second axes of the PCA of the watershed morphometry variables were significant (P=0.001 for both). These axes accounted for 36.2 and 21.9 %, respectively, of the variance in the watershed morphometry data. Variables with the strongest influence on PCA axis 1 were related to watershed and stream size, and variables with the strongest influence on axis 2 were related to watershed elevation and relief (Table 13). Examination of site position on axis 2 showed that Fort Benning catchments had the greatest relief, Savannah River Site catchments the least, and Fort Bragg and Fort Gordon catchments were intermediate. Examination of the position of the sites on axis 1 showed that most installations included a wide range of streams sizes, except for Fort Gordon, which lacked smaller stream sites (Figure 18).

Table 11. Average values of instream habitat variables for sample sites at Fort Benning, Fort Bragg, Fort Gordon, and the Savannah River Site (SRS). Data from smaller holdings with small sample sizes (e.g., Manchester State Forest) are not shown. Variables described in Table 1.

	Fort Benning		Fort Bragg		Fort Gordon		SRS	
Variable	**Avg**	Stdev	**Avg**	Stdev	**Avg**	Stdev	**Avg**	Stdev
Depth (m)	**0.25**	0.18	**0.31**	0.11	**0.34**	0.15	**0.18**	0.14
Width (m)	**2.11**	1.44	**2.34**	0.74	**2.92**	0.47	**2.23**	1.89
Bank height (m)	**0.70**	0.62	**0.70**	0.51	**0.65**	0.28	**0.75**	0.38
Bank angle (degrees)	**60.00**	15.03	**71.02**	19.00	**53.71**	11.75	**62.16**	11.95
Erosion (0-4)[a]	**0.68**	0.57	**0.47**	0.55	**0.89**	0.77	**0.52**	0.66
Channel modification (0-4)[a]	**0.05**	0.09	**0.17**	0.64	**0.01**	0.02	**0.00**	0.00
Bank vegetative cover (%)[b]	**0.63**	0.16	**0.72**	0.16	**0.74**	0.18	**0.58**	0.11
Large woody debris (%)[c]	**0.10**	0.04	**0.11**	0.04	**0.10**	0.04	**0.08**	0.03
Small woody debris (%)[c]	**0.08**	0.02	**0.09**	0.03	**0.05**	0.03	**0.13**	0.08
Overhanging vegetation (%)[c]	**0.08**	0.04	**0.14**	0.07	**0.11**	0.06	**0.12**	0.09
Undercut banks (%)[c]	**0.02**	0.01	**0.04**	0.03	**0.03**	0.03	**0.03**	0.02
Root masses (%)[c]	**0.06**	0.02	**0.09**	0.05	**0.10**	0.05	**0.05**	0.02
Macrophytes (%)[c]	**0.01**	0.02	**0.02**	0.04	**0.01**	0.01	**0.01**	0.04
Clay substrate (%)[c]	**0.02**	0.05	**0.02**	0.04	**0.00**	0.00	**0.00**	0.01
Silt/muck substrate (%)[c]	**0.31**	0.22	**0.20**	0.10	**0.17**	0.11	**0.34**	0.22
Sand substrate (%)[c]	**0.61**	0.26	**0.74**	0.11	**0.78**	0.17	**0.52**	0.26
Gravel substrate (%)[c]	**0.05**	0.10	**0.05**	0.05	**0.05**	0.08	**0.13**	0.13
Rock substrate (%)[c]	**0.00**	0.01	**0.00**	0.01	**0.00**	0.00	**0.01**	0.03
Riparian hardwoods (%)[d]	**3.65**	1.03	**3.73**	0.83	**3.80**	0.31	**3.97**	0.08
Riparian pines (%)[d]	**0.19**	0.43	**0.84**	1.05	**0.49**	0.68	**0.30**	0.32
Riparian shrubs (%)[d]	**2.38**	1.00	**2.03**	1.18	**2.84**	0.45	**1.18**	0.87
Riparian herbaceous (%)[d]	**0.94**	1.28	**0.91**	1.17	**0.00**	0.00	**2.18**	0.87
Proportion riffles	**0.12**	0.10	**0.08**	0.07	**0.00**	0.00	**0.17**	0.17
Proportion runs	**0.63**	0.17	**0.68**	0.14	**0.83**	0.15	**0.62**	0.21
Proportion deep pools	**0.08**	0.06	**0.11**	0.09	**0.09**	0.13	**0.11**	0.11
Proportion shallow pools	**0.17**	0.11	**0.13**	0.09	**0.08**	0.03	**0.10**	0.11

[a] 0 = absent, 4 = most severe
b Percent coverage of the stream bank by vegetation
c Percent coverage of the stream bottom area
d Percent coverage of the riparian zone out to 20 m on each side of the bank

Table 12. Average values of GIS (watershed scale) variables for sample sites at Fort Benning, Fort Bragg, Fort Gordon, and the Savannah River Site (SRS). Data from smaller holdings with small sample sizes (e.g., Manchester State Forest) are not shown. Variables described in Table 1.

	Fort Benning		Fort Bragg		Fort Gordon		SRS	
Variable	**Avg**	Stdev	**Avg**	Stdev	**Avg**	Stdev	**Avg**	Stdev
Drainage area (km2)	**9.29**	18.85	**9.19**	7.85	**9.36**	4.99	**7.35**	8.95
Drainage perimeter (km)	**10.02**	8.96	**12.68**	5.54	**12.77**	3.58	**10.76**	6.98
Cumulative stream length (km)	**12.15**	22.97	**13.55**	13.53	**9.25**	2.86	**4.90**	7.05
Stream order	**2.13**	0.83	**2.87**	0.81	**2.40**	0.55	**2.00**	0.79
Stream magnitude	**6.93**	12.49	**14.78**	16.69	**7.00**	1.22	**4.85**	4.61
Basin length (km)	**3.84**	3.01	**3.92**	1.62	**4.44**	1.53	**2.87**	1.41
Drainage shape	**0.20**	0.19	**0.16**	0.15	**0.11**	0.02	**0.84**	0.84
Stream length (m)	**3.33**	2.91	**3.80**	2.00	**4.39**	1.45	**2.03**	1.70
Watershed relief ratio	**0.06**	0.04	**0.03**	0.02	**0.04**	0.01	**0.03**	0.02
Stream gradient	**0.03**	0.02	**0.03**	0.02	**0.02**	0.00	**0.02**	0.01
Tributary length (m)	**14.23**	21.00	**5.78**	10.35	**2.12**	1.48	**3.37**	4.51
Sinuosity	**1.14**	0.07	**1.16**	0.13	**1.15**	0.11	**1.15**	0.08
High point (m)	**187.99**	33.54	**141.77**	19.09	**160.87**	6.09	**101.55**	9.19
Mouth elevation (m)	**108.42**	26.20	**78.44**	20.68	**86.95**	11.91	**54.75**	6.41
Basin relief (m)	**79.57**	22.28	**63.32**	21.06	**73.93**	12.02	**46.80**	7.80
High-intensity development (%)	**<0.00**	0.00	**0.01**	0.02	**0.05**	0.07	**0.01**	0.02
Medium-intensity development (%)	**0.00**	0.00	**0.02**	0.04	**0.05**	0.06	**0.01**	0.02
Low-intensity development (%)	**0.02**	0.01	**0.04**	0.09	**0.08**	0.08	**0.01**	0.02
Developed open space (%)	**0.01**	0.01	**0.02**	0.07	**0.06**	0.07	**0.01**	0.02
Cultivated (%)	**0.04**	0.08	**0.01**	0.02	**0.01**	0.02	**0.00**	0.00
Pasture (%)	**0.02**	0.04	**0.01**	0.03	**0.01**	0.01	**0.00**	0.01
Grassland (%)	**0.09**	0.06	**0.09**	0.05	**0.07**	0.05	**0.06**	0.05
Deciduous (%)	**0.14**	0.07	**<0.00**	0.01	**0.03**	0.01	**0.00**	0.01
Evergreen (%)	**0.23**	0.10	**0.53**	0.17	**0.24**	0.12	**0.65**	0.14
Mixed forest (%)	**0.08**	0.04	**0.02**	0.02	**0.02**	0.02	**0.02**	0.01
Scrubland (%)	**0.19**	0.11	**0.10**	0.08	**0.22**	0.14	**0.10**	0.06
Palustrine - all types (%)	**0.17**	0.07	**0.09**	0.05	**0.15**	0.03	**0.12**	0.06
Impervious surfaces (%)	**0.01**	0.01	**0.04**	0.06	**0.01**	0.01	**0.01**	0.02
Bare ground (%)	**0.01**	0.01	**0.05**	0.09	**0.01**	0.01	**0.01**	0.02
Paved roads (%)	**0.59**	0.65	**2.62**	5.81	**1.95**	1.76	**0.83**	1.32
Unpaved roads (%)	**2.56**	1.12	**3.14**	1.66	**3.68**	2.08	**1.67**	1.17
Number road crossings	**7.07**	10.46	**20.74**	23.46	**12.60**	5.27	**4.30**	7.16
Maloney disturbance index (%)	**2.57**	1.11	**3.13**	1.62	**3.68**	2.08	**1.68**	1.17
Disturbance index (%)	**0.08**	0.11	**0.14**	0.17	**0.21**	0.19	**0.03**	0.06

Table 13. Pearson correlation coefficients for axes 1 and 2 of a principal component analysis (PCA) of sample sites based on variables representing basin size, shape, and relief.

Variable	PCA axis 1	PCA axis 2
Drainage area	-0.92	0.04
Drainage perimeter	-0.71	0.15
Cumulative stream length	-0.93	-0.27
Basin length	-0.90	<0.01
Drainage shape	0.64	0.53
Stream length	-0.89	-0.23
Basin relief ratio	0.58	-0.52
Stream gradient	0.31	-0.45
Tributary length	-0.04	-0.21
Sinuosity	0.08	0.01
Drainage density	-0.08	-0.57
Basin high point	-0.22	-0.92
Stream mouth elevation	0.13	-0.79
Basin relief	-0.55	-0.64

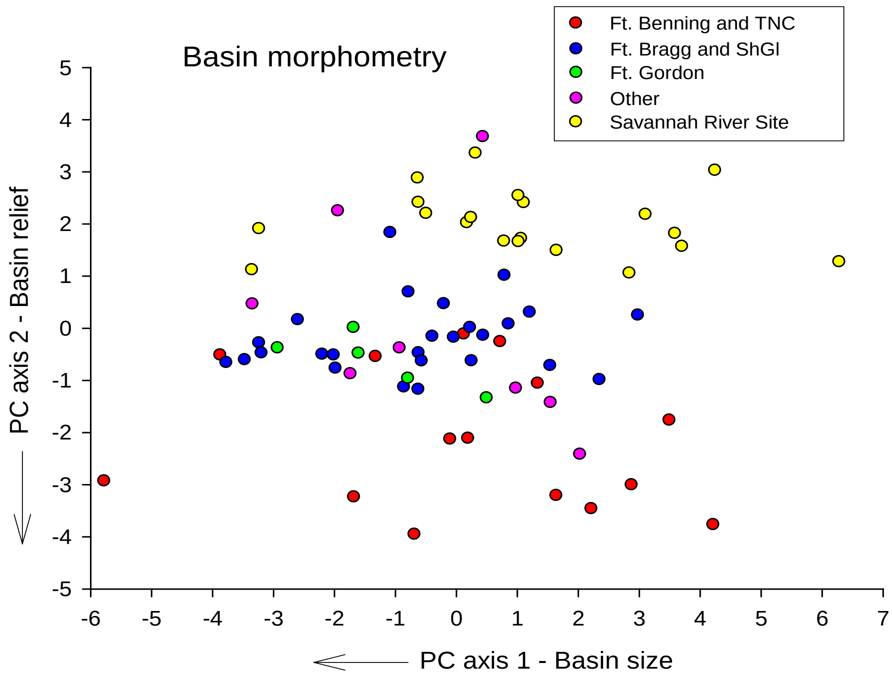


Figure 18. PCA of fish assemblage sample sites based on basin morphometry variables. Also shown are variables with the strongest effect on each axis plus their direction of increase.

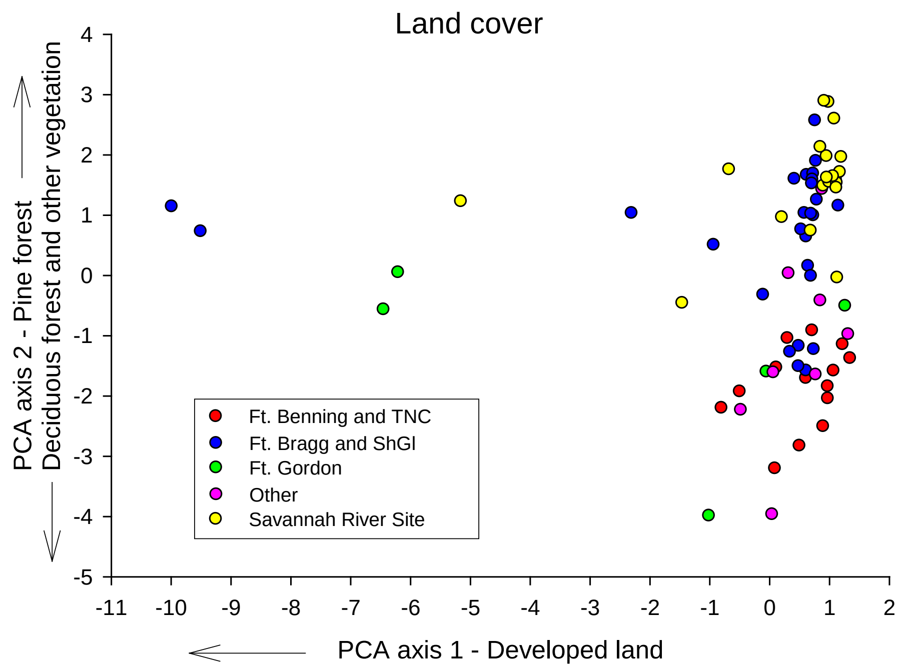
The first, second, and third axes of the PCA of the basin land cover variables were significant (p=0.001 for all), accounting for 26.5, 17.7, and 13.8%, respectively, of the variance in the watershed land cover data. The first PCA axis constituted a gradient that reflected the extent of anthropogenic development in the watershed (Table 14). Most sample sites clustered towards the low end of this gradient, reflecting the emphasis on sampling least disturbed sites and relative paucity of highly disturbed sites (Figure 19). The second axis represented a gradient of vegetation type, with pine forests on one end and deciduous trees, scrub/scrub, and other vegetation on the other. Examination of this gradient showed that, despite substantial overlap, SRS watersheds had the highest coverage by pine forests followed by Fort Bragg, and Fort Benning. The third axis was correlated with the number of unpaved roads in the watershed and the Maloney disturbance index (MDI, composite of unpaved roads and bare ground, see Table 20). Fort Bragg sites were generally characterized by higher MDI scores and greater coverage of unpaved roads (Figure 19). These results are related to the large number of unpaved firebreak roads that traverse Fort Bragg.

The third and final habitat related PCA was conducted on the instream habitat data. There were 3 significant p<0.001) PCA axes, which accounted for 18.3, 13.2, and 11.5%, respectively, of the variance in the data. The first axis largely represented a mesohabitat and substrate gradient extending from runs with sand bottoms to a greater frequency of riffles and deep pools with more varied substrates including gravel, silt, and small CWD (Table 15). SRS and Fort Bragg sites separated to some degree along this axis, with the former having more diverse substrates and mesohabitats and latter dominated by sandy runs (Figure 20). The second PCA axis largely represented a gradient of erosion, bank height (an indication of channel incision), and channel modification, with higher scores indicating greater prevalence of these features. Most sites scored low on this axis reflecting the emphasis on least disturbed sites in this study (Figure 20). The third axis was most strongly weighted by variables that reflected the occurrence of steep, undercut banks and clay substrates (Table 15). There was little difference among installations on this axis.

The preceding results indicate habitat differences among installations at both watershed and instream spatial scales. Fort Benning watersheds had the greatest relief and SRS watersheds the least. SRS watersheds generally had greater coverage by pine forests, while Fort Bragg and especially Fort Benning watersheds had greater coverage by deciduous trees and scrub/shrub vegetation communities. SRS streams also tended to have greater substrate and mesohabitat diversity than Fort Bragg streams, which were dominated to a greater degree by sandy-bottomed runs. Although there were prominent exceptions, most sites, regardless of installation, lacked extensive anthropogenic land development in their watersheds, extensive stream channel erosion, and artificial stream channel modifications. Despite these general patterns, there was high habitat variation within installations as indicated by the intra-site dispersion of sample sites in the PCAs. In the next section, habitat patterns observed among and between installations are explicitly linked with patterns of fish assemblage structure.

Table 14. Pearson correlation coefficients for axes 1, 2, and 3 of a principal component analysis (PCA) of sample sites based on land cover variables.

Land cover type (%)	PCA axis 1	PCA axis 2	PCA axis 3
Developed, high intensity	-0.89	0.08	0.08
Developed, medium intensity	-0.96	0.06	0.04
Developed, low intensity	-0.91	-0.12	-0.08
Developed, open space	-0.96	-0.05	0.05
Cultivated	-0.08	-0.64	-0.18
Pasture	0.00	-0.62	-0.06
Grassland/herbaceous	0.46	-0.17	-0.45
Deciduous trees	0.07	-0.65	0.46
Evergreen forest	0.40	0.80	-0.23
Mixed forest	0.32	-0.47	0.29
Shrub/scrub	0.09	-0.62	0.21
Palustrine	0.25	-0.32	0.53
Water	0.08	0.00	-0.39
Bare ground	0.08	0.17	-0.36
Paved roads	-0.87	0.09	-0.05
Unpaved roads	-0.07	-0.52	-0.72
Road crossings	-0.02	0.04	-0.49
Maloney Disturbance Index	-0.07	-0.52	-0.72



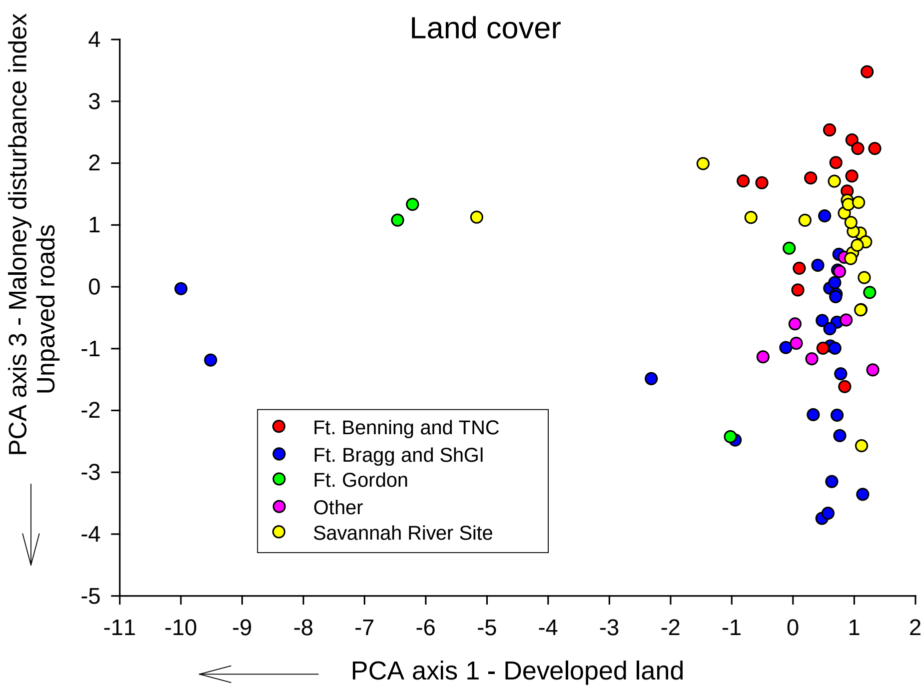
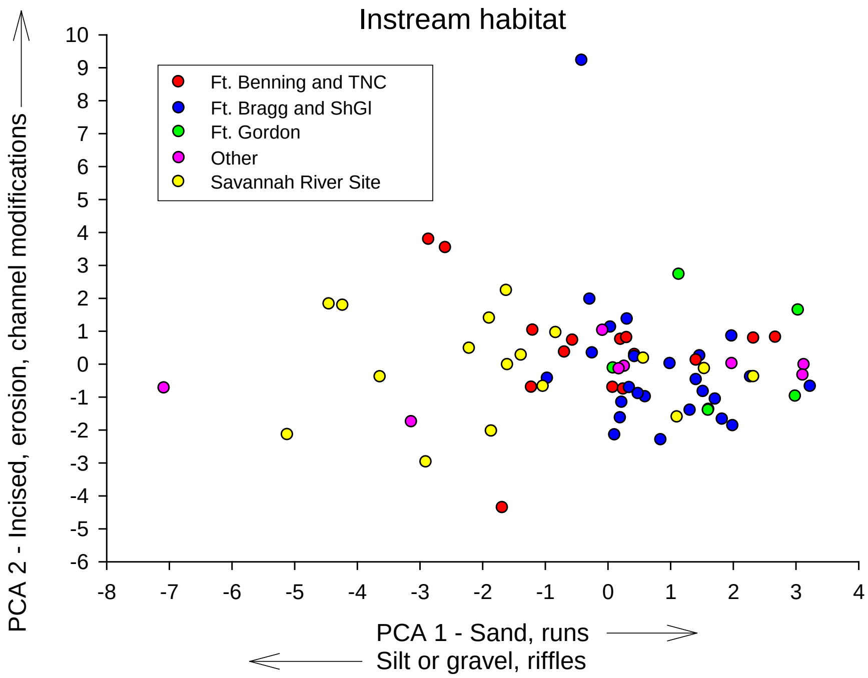


Figure 19. PCA of fish assemblage sample sites based on land cover variables (axes 1 and 2 on upper graph, axes 1 and 3 on lower graph). Also shown are variables with the strongest effect on each axis plus their direction of increase.

Table 15. Pearson correlation coefficients for axes 1, 2, and 3 of a principal component analysis (PCA) of sample sites based on variables representing instream habitat.

Variable	PCA axis 1	PCA axis 2	PCA axis 3
Sand substrate (percent)	0.84	0.02	0.04
Runs (average number)	0.67	0.02	0.02
Bank vegetation cover (percent)	0.45	-0.50	0.50
Riparian herbaceous shrubs (percent)	0.37	0.08	-0.08
Overhanging vegetation (percent)	0.34	-0.37	-0.08
Undercut banks (percent)	0.32	0.03	0.33
Macrophytes (percent)	0.32	-0.28	-0.23
Root mats (percent)	0.32	-0.08	0.01
Large woody debris (percent)	0.28	-0.27	-0.24
Average bank angle (degrees)	0.14	0.02	-0.76
Channel modifications (number)	0.04	0.61	-0.11
Riparian hardwoods (percent)	0.00	0.28	-0.11
Clay substrate (percent)	-0.01	0.58	-0.57
Riparian conifers (percent)	-0.03	0.10	-0.10
Shallow pools (average number)	-0.03	0.01	0.31
Erosion (visual score)	-0.17	0.73	-0.40
Average bank height (meters)	-0.28	0.81	-0.75
Deep pools (average number)	-0.48	-0.17	-0.44
Riparian herbaceous vegetation (percent)	-0.49	-0.33	-0.10
Rock substrate (percent)	-0.49	0.43	-0.01
Riffles (average number)	-0.56	0.09	-0.04
Small woody debris (percent)	-0.57	-0.39	0.33
Gravel substrate (%)	-0.64	0.37	-0.40
Silt/muck substrate (percent)	-0.67	-0.33	0.22



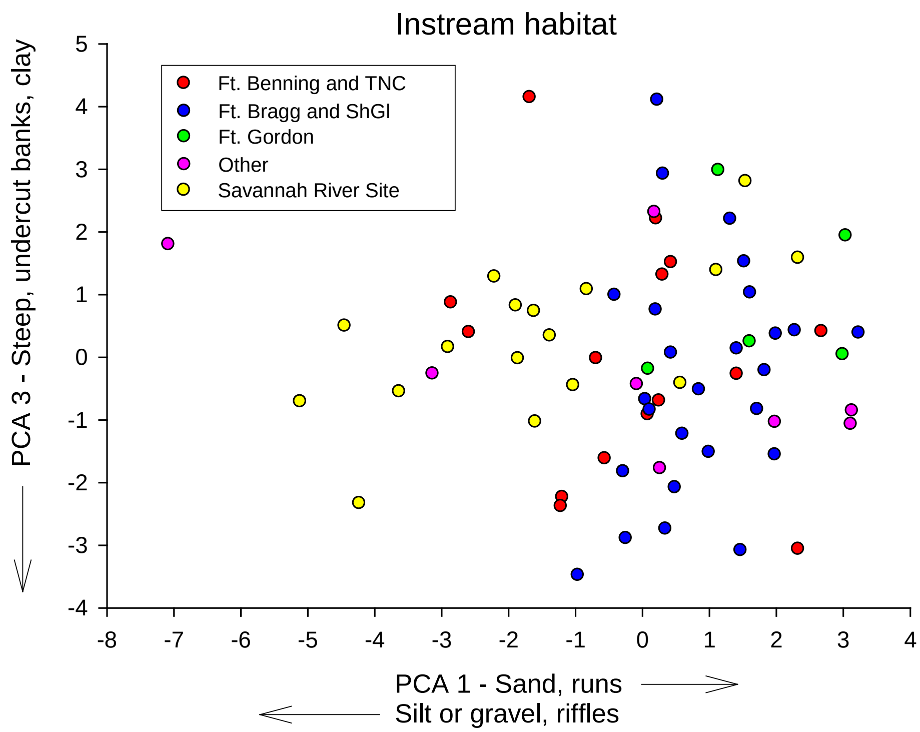


Figure 20. PCA of fish assemblage sample sites based on instream habitat variables (axes 1 and 2 on upper graph, axes 1 and 3 on lower graph). Also shown are variables with the strongest effect on each axis plus their direction of increase.

5.2.2.2 Relationship between Fish Assemblage Structure and Habitat

Canonical correspondence analysis (CCA), a direct ordination technique, was used to reveal relationships between fish assemblage structure and environmental factors. Direct ordination methods such as CCA identify patterns in biological data that are specifically associated with measured environmental variables. In contrast, indirect ordination methods (such as PCA and NMDS) identify patterns in biological data regardless of their source, thereby showing if environmental variables other than those under study are important. Variance partitioning methods were used in conjunction with CCA to assess the amount of unique and shared (i.e., joint) variance in fish assemblage structure that was associated with the environmental variables (Jongman et al. 1995, Lepš and Šmilauer 2007). Monte Carlo permutation tests were used to determine the significance of individual variance components.

Prior to CCA, detrended correspondence analyses (DCA) was used to estimate gradient lengths in the biological matrices following Lepš and Šmilauer (2007). The step was undertaken to verify that CCA, a unimodal ordination method, was an appropriate analytical method for the data under study. Unimodal ordination methods are appropriate when environmental gradients are relatively long and associated with substantial species turnover. The length of the first DCA gradient was 4.65. Gradient lengths <3.0 indicate that linear methods, such as redundancy analysis are appropriate, and gradient lengths >4.0 indicate that unimodal methods such as CCA are appropriate. The relatively long gradient associated with the fish assemblage data, which justified the use of CCA rather than linear methods, is not surprising given the geographic scope of the study. As with NMDS, the fish assemblage data were log (X+1) transformed prior to analysis.

The number of environmental variables measured in conjunction with the fish assemblage data was large relative to the number of sample sites in the CCA. Too many variables can create an excessively complex model in which overfitting results in the characterization of random error rather than underlying relationships. A related problem is multicollinearity in which highly correlated independent variables result in an inability to accurately identify key relationships between predictor and dependent variables. To overcome this problem, we used the PCA axes described in Section 5.2.2.1 as predictive environmental variables in the CCA. PCA axes have the advantage of combing the information from related variables into one summary variable, which can be assigned meaning based on the individual variables that weight the PCA axes most strongly (Tables 13, 14, 15). The environmental variables included in the CCA were as follows:

1) Longitude (Lon) – represented geographic variation. Latitude was not included with longitude because the 2 variables were strongly correlated (Pearson r = 0.95) and inclusion of both produced no increase in explanatory power.
2) Watershed relief (Relief) – represented by PCA axis 2 based on GIS variables representing watershed size, shape, and relief (Table 13).
3) Developed land (percentage of developed land in the watershed, Devlmnt) – represented by PCA axis 1 based on GIS variables representing watershed land cover (Table 14).
4) Watershed size (Size) – represented by PCA axis 1based on GIS variables representing watershed size (Table 13).
5) Instream habitat quality (Strhab) – based on the summary score produced by the SCDHEC instream habitat protocol, as described earlier.
6) Connectivity (Con); i.e., connection with larger streams – based on proximity of the sample site to a confluence with a larger stream that could serve as a source of colonists to the study site. Large streams, which typically support more species than small streams, can represent source pools for species to colonize nearby tributaries. Connectivity was represented as a categorical variable: a sample site located in a small stream (i.e., $\leq$ 1.5 m average width) was

considered to be connected with a larger stream (i.e., at least one order larger) if less than 250 m from the confluence with the larger stream. Sample sites located in larger streams (>1.5 m average width) were considered connected if located within 750 m of a larger stream. These distances were based on our field observations, literature concerning fish movements in small streams (Hill and Grossman 1987, Rodríguez 2002), and BPJ concerning the distances that stream fishes were likely to move into the study area.

7) Watershed vegetation type (Ldcov) – represented by PCA axis 2 based on GIS variables representing watershed land cover (Table 14).
8) Stream substrate and mesohabitat (Submeso) – represented by PCA axis 1 based on variables representing instream habitat (Table 15).
9) Undercut banks (Undct) – represented by PCA axis 3 based on variables representing instream habitat (Table 15).
10) Erosion and incision (Erosbnk) – represented by PCA axis 2 based on variables representing instream habitat (Table 15)
11) MDI and unpaved roads (MDI) – represented by PCA axis 3 based on variables representing instream habitat (Table 14).

The CCA showed that 39.6% of the variance in the fish assemblage data was explained by the environmental variables (p=0.002). The sample sites showed distinct geographic groupings with substantial variation within groups paralleling the results of the NMDS, described earlier (Figure 21). Correspondence between both the direct CCA and the indirect NMDS, suggests that the measured variables were among the key factors affecting distribution of taxa among sites. The CCA also showed clearly demarcated species assemblages associated with sampling location (Figure 21). This pattern, too, paralleled the results shown by the NMDS and 2-way CCA.

An automated stepwise forward-selection procedure was used to rank the environmental variables in order of their influence on the CCA results (Table 16). Marginal effects, which rank the environmental variables in the order of the amount of variance that they explain individually, showed that geography (represented by longitude), watershed relief, vegetation cover type, instream habitat quality, and the amount of developed land in the watershed had the most influence on fish assemblage structure (Table 16). However, marginal effects do not account for the covariance that individual variables share with other environmental variables. Conditional effects, on the other hand, show the unique effect of each environmental variable independent of shared variance. Eight of the 11 independent variables had significant (p<0.05) conditional effects in the following order of importance: longitude (i.e., geography), watershed relief, developed land, watershed size, instream habitat quality, connection with a larger stream, vegetation cover type, and stream substrate and mesohabitat type. These analyses showed that several environmental factors influenced assemblage structure apart from the influence of geographic variables. These included factors at the watershed scale (e.g., size and vegetation cover type) and factors at the instream habitat scale (substrate and mesohabitat). They also included factors that were largely natural (watershed size and connectivity with large streams) and factors that mainly reflected anthropogenic disturbance (% of developed land in the watershed and instream habitat quality).

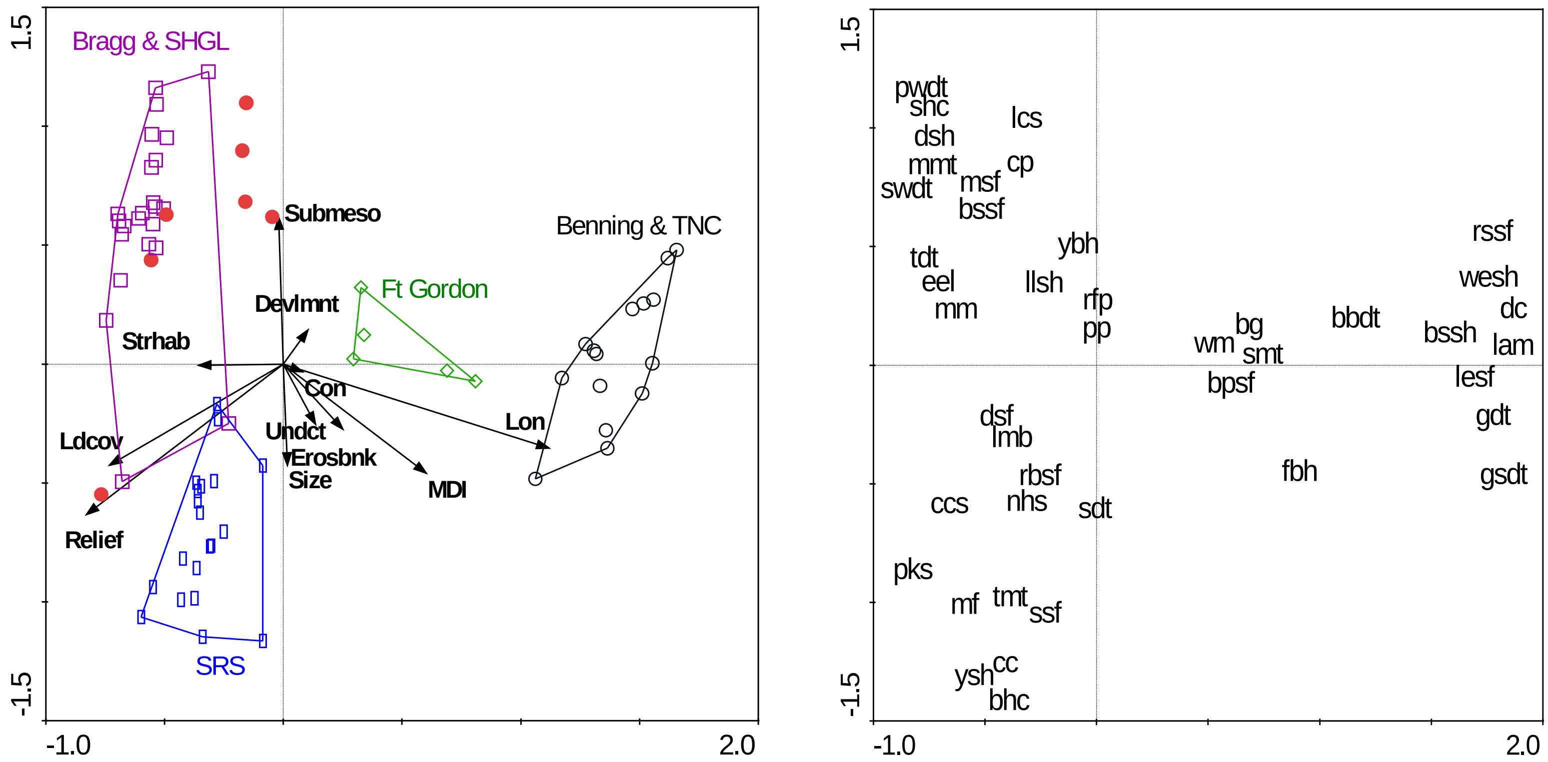


Figure 21. Canonical correspondence analysis (CCA) of samples sites (left) and species (right) based on fish assemblage taxonomic data collected from study sites. Juxtaposition of the plots indicates geographic occurrence of the species. The sample site biplot (left) depicts the relationship between metrics and environmental variables. Environmental variables are represented by arrows that show direction of increase and strength of correlation. Environmental variable abbreviations can be found in the text. Species abbreviations are presented in Table 10.

Table 16. Permutation test results for environmental variables included in a canonical correspondence analysis (CCA) of fish assemblage structure in southeastern coastal plain streams. See text for a description of the variables.

		Independent effect	Conditional effect (additional to other variables)		
Variable	Abbrev.	Lambda	Lambda	P	F
Longitude	Lon	0.55	0.55	0.002	8.64
Basin relief	Relief	0.44	0.35	0.002	6.10
Developed land	Devlmnt	0.25	0.26	0.002	4.59
Basin size	Size	0.19	0.16	0.002	2.99
Instream habitat quality	Strhab	0.26	0.09	0.014	1.83
Connection with larger streams	Con	0.13	0.14	0.018	2.53
Vegetation cover type	Ldcov	0.32	0.10	0.024	1.99
Stream substrate and mesohabitat	Submeso	0.22	0.09	0.042	1.65
Undercut banks	Undct	0.13	0.07	0.086	1.43
Erosion and incision	Erosbnk	0.20	0.07	0.094	1.49
MDI and unpaved roads	MDI	0.24	0.04	0.890	0.63

As a check on the accuracy of the preceding results, another CCA was conducted using independent variables that were not transformed nor summarized by PCA as described above. This analysis made use of individual habitat variables including the GIS variables representing watershed and landscape characteristics (Table 2), the field-collected variables representing instream habitat (Table 3), longitude representing geographic differences among sites, and the previously described connectivity variable representing proximity to a larger stream. The total number of independent variables was large relative to the number of sample sites (which could result in overfitting as previously described), so variables were divided into 3 groups: watershed size and shape, land cover, and instream habitat. Each group was analyzed in a preliminary CCA intended to select the best variables within the group, defined as those with significant ($p<0.05$) conditional effects identified by a forward-selection procedure. Next, a comprehensive CCA was conducted using a combination of the best variables in each category, plus longitude and connectivity. Following the comprehensive CCA, individual environmental variables or groups of variables were assessed separately by specifying the other environmental variables as covariables, which identified the unique effect of the variable (or variable group) independent of variance shared with other variables.

Twenty-one environmental variables were included in the comprehensive CCA, accounting for 53.5% of the variance in fish assemblage structure (p=0.002 for all canonical axes) (Table 17). Individual variables with significant ($p<0.05$) conditional effects included longitude, watershed high point, paved road coverage, instream habitat quality (SCDHEC instream habitat score), watershed drainage area, connectivity, coniferous forest cover, mixed forest cover, and percentage of developed land (disturbance index). The effects of geography (i.e., longitude), after fitting the GIS and instream habitat variables as covariables, accounted for 4.1% of the variance in assemblage structure (p=0.002 for all canonical axes).

The effects of the watershed-level variables (i.e., GIS variables including basin size, shape, and landcover variables), after fitting longitude, and the instream habitat variables as covariables, accounted for 18.6% of the variance in assemblage structure (p=0.012 for all canonical axes). The effects of instream and riparian habitat variables (including connectivity), after fitting the GIS variables and longitude as covariables, accounted for 11.7% of the variance in fish assemblage structure (p=0.026 for all canonical axes). Together the unique effects of the preceding sources of variation accounted for 34.4% of the variance in assemblage structure. The remaining 19.1% of the variance accounted for by the model was shared and unable to be uniquely attributed to a particular source or, perhaps, related to unmeasured factors correlated with measured variables. Like the CCA based on PCA summary variables, these analyses show that, although geography had an important effect on fish assemblage structure, several environmental and habitat factors were also influential.

Table 17. Permutation test results (conditional effects only) for individual environmental variables included in a cannonical correspondence analysis (CCA) of fish assemblage structure in Southeastern Plains streams. See Tables 2 and 3 for a description of the variables.

Variable	Lambda A	P	F
Longitude	0.55	0.002	8.64
Basin high point	0.34	0.002	5.83
Instream habitat quality	0.18	0.002	3.23
Basin drainage area	0.15	0.002	3.05
Coniferous forest cover	0.1	0.010	1.95
Connectivity	0.12	0.012	2.32
Mixed forest cover	0.09	0.016	1.81
Paved road coverage	0.27	0.018	4.87
Developed land	0.09	0.034	1.73
Bank erosion	0.08	0.056	1.66
Basin drainage perimeter	0.08	0.060	1.65
Stream depth	0.07	0.066	1.61
Stream width	0.07	0.100	1.48
Macrophyte cover	0.06	0.198	1.28
Open water	0.06	0.254	1.17
Deciduous forest cover	0.06	0.260	1.15
Grassland	0.05	0.328	1.09
Shallow pools	0.05	0.334	1.14
Riparian shrub cover	0.04	0.444	0.99
Drainage density	0.04	0.568	0.89
Root mats	0.04	0.792	0.7

5.2.2.3 Relationship between Fish Species Richness and Habitat

Fish species richness was measured at 72 sites over the study area. Two sites, both very small disturbed sites on the SRS, yielded no fish and were excluded from further analysis. Observed species richness ranged from 1 to 20 at the remaining 70 sites. Estimated species richness (first-order jackknife) ranged from 1 to 27.4 (Table 18). The difference between observed and estimated richness averaged 1.8 but tended to be greater in large than smaller streams (Pearson r = 0.52 between stream width and difference, P<0.001), suggesting that samples from large streams underestimated richness to a greater degree than smaller streams. One factor that likely contributed to this pattern was a decline in sampling efficiency with stream size due to smaller electrical field size in relation to the area sampled and associated difficulties in seeing and netting fish. However, ecological factors may have also played a role. Reach lengths adequate to include all habitat types in small streams (and perhaps represent them several times over) may have been inadequate in larger streams where habitats (e.g., pools and runs) occurred on a larger scale. Our efforts to sample longer reaches (up to 285 m) in larger streams may have been insufficient to fully counteract this trend.

Estimated species richness was a better measure than observed richness of true richness, especially, in larger streams, where observed richness was likely a negatively biased estimate of true richness. Therefore, we used the former as the dependent variable in general linear models constructed to identify important determinants of richness within the study area. The independent variables included measures of stream size (average width, average depth, average cross-sectional area, and order), measures of location (latitude and a categorical variable that represented each sampling location [e.g., Fort Bragg, SRS, Manchester State Forest, etc.]), watershed characteristics (size and relief), PCA-derived summary variables described in previous section, a categorical connectivity variable representing connection with larger streams (described in previous section), and a variable representing instream habitat quality (SCDHEC instream habitat score, Section 4.3.1). The instream habitat quality variable was also represented by squaring the value, thereby allowing for representation of a curvilinear relationship between instream habitat quality and richness. We considered a curvilinear relationship a possibility because of the “intermediate disturbance hypothesis,” which holds that richness is maximized when ecological disturbance is neither too rare nor frequent (Connell 1978). We relied on theoretical considerations rather than automated stepwise procedures to select variables for the final model.

After examining a number of potential general linear models, we selected a definitive model with the following independent variables that were significantly (p<0.05) related to estimated richness: average stream width, watershed size, sampling location, connectivity with a larger stream, instream habitat quality, and instream habitat quality squared (Table 19). The coefficient of determination (R^2) for the model was 0.76. Interaction terms were not significant. The least square mean number of species (LSM mean, i.e.; mean computed at the mean values of the continuous covariables) was highest in SRS streams followed by 2 streams in the Sandhills National Wildlife Refuge (Figure 22), with little difference among other sites.

The general linear model indicated that richness was generally comparable among sites except for SRS where it was higher. It also showed that richness was affected by stream and watershed size. Further analysis showed that this relationship was linear for all locations except SRS, where the rate of increase in species richness decreased with increasing stream size. This relationship was more accurately described by an exponential decay model than a linear model (R^2 = 0.81 compared with 0.59) (Figure 23). A similar relationship might have also been observed at the other sites if more large streams had been sampled. The general linear model further showed that the relationship between stream size and richness was affected by proximity to a larger stream that could serve as a source of immigrating species.

Table 18. Observed and estimated (using first-order jackknife estimator) fish species richness at Sand Hills sites (see Table 1 for full site names). Sample reaches were subdivided into segments making it possible to compute species-accumulation curves. Sites defined in Table 1.

Site	Reach length (m)	Number segments	Stream width (m)	Observed richness	Estimated richness	Difference
Bn-chat	170	17	2.6	16	16.9	0.9
Bn-d12	180	18	1.7	4	5.9	1.9
Bn-d13f	150	15	0.9	5	5.9	0.9
Bn-d13c	150	15	1.5	6	6.9	0.9
Bn-hct	190	19	1.9	12	12.9	0.9
Bn-holl	160	16	1.9	4	4.0	0.0
Bn-k11	190	19	2.4	7	8.9	1.9
Bn-k13	170	17	1.8	7	8.9	1.9
Bn-ljt	150	15	1.8	5	6.9	1.9
Bn-lpk	170	17	2.1	8	10.8	2.8
Bn-wolf	200	20	2.7	14	15.9	1.9
Br-bct	200	20	2.6	11	12.0	1.0
Br-bm	210	14	4.4	13	18.6	5.6
Br-cab	200	20	2.0	6	7.9	1.9
Br-cyp	200	20	2.1	11	13.9	2.9
Br-deep	200	20	2.0	6	6.9	0.9
Br-field	200	20	2.4	5	5.0	0.0
Br-flat	190	19	3.0	8	9.9	1.9
Br-gum	190	19	1.8	5	5.0	0.0
Br-hect	200	20	2.3	6	7.9	1.9
Br-hrse	140	14	1.4	4	4.9	0.9
Br-jen	190	19	2.8	9	10.9	1.9
Br-jun	200	20	2.5	10	11.9	1.9
Br-lrf	200	20	1.5	11	12.9	1.9
Br-lrt	180	18	1.5	7	10.8	3.8
Br-mcp	200	20	1.9	8	8.0	0.0
Br-rfb	210	14	3.8	12	15.7	3.7
Br-tank	210	14	3.2	5	5.0	0.0
Br-ujt	180	18	2.4	8	8.9	0.9
Br-uwp	200	21	1.4	1	1.0	0.0
Br-wp	200	20	2.2	8	11.8	3.8
Gr-bath	240	16	3.7	14	15.9	1.9
Gr-bog	170	17	2.7	10	10.9	0.9
Gr-mcoy	210	14	3.0	8	11.7	3.7
Gr-mrbtrb	190	19	2.6	5	5.0	0.0
Gr-prong	240	16	2.7	12	15.8	3.8
Mn-mcra	170	17	2.0	10	11.8	1.8
Mn-tav	170	17	1.9	4	5.9	1.9
Nc-bct	190	19	1.2	6	6.0	0.0
Nc-bjc	150	15	1.1	3	3.0	0.0
Nc-pkc	285	19	6.9	11	14.8	3.8
Nc-pmt	150	15	1.2	2	2.0	0.0
Sg-bone	190	19	1.8	10	10.0	0.0

Table 18. concluded.

Site	Reach length (m)	Number segments	Stream width (m)	Observed richness	Estimated richness	Difference
Sg-joes	180	18	2.2	9	9.0	0.0
Sg-millst	200	20	2.7	3	3.0	0.0
Sh-cedr	240	16	3.4	14	15.9	1.9
Sh-mill	170	17	2.0	8	11.8	3.8
Sn-bbct	160	16	1.2	3	3.9	0.9
Sn-hemp	170	17	1.7	7	9.8	2.8
Sn-rogr	170	17	1.7	5	5.9	0.9
Sr-ltr	195	13	7.6	21	25.6	4.6
Sr-mb6	150	15	2.0	9	11.8	2.8
Sr-mb61	150	15	1.5	4	4.0	0.0
Sr-mbhw	160	16	2.3	13	13.9	0.9
Sr-mbm	200	20	2.9	16	19.8	3.8
Sr-mc5	150	15	1.9	12	16.7	4.7
Sr-mc6	150	15	1.5	7	9.8	2.8
Sr-mc6c	130	13	1.5	10	10.9	0.9
Sr-mc7	150	15	1.1	6	7.9	1.9
Sr-mcm	150	15	1.5	**6**	**9.9**	3.9
Sr-mq10	90	10	1.0	1	1.0	0.0
Sr-mq8	150	15	1.2	5	5.0	0.0
Sr-mqh	150	15	1.1	10	10.9	0.9
Sr-pb4	140	14	1.7	9	9.9	0.9
Sr-pbhw	150	15	1.6	15	20.6	5.6
Sr-Pbm	200	20	2.6	18	20.9	2.9
Sr-tc5	150	15	1.7	8	8.9	0.9
Sr-tc6	150	15	1.5	12	16.7	4.7
Sr-tcm	210	14	7.5	20	27.4	7.4
Sr-u10	150	15	0.9	4	4.9	0.9

The latter factor resulted in an average richness increase of 2 compared with sites more distant from larger streams (Figure 24). Last, the model showed that richness was related to instream habitat quality as represented by the modified SCDHEC instream habit assessment protocol. The latter relationship was not linear but curvilinear, as indicated by a significant (p<0.03) squared instream habitat term. This curvilinear relationship was depicted by plotting residuals from a model with estimated species richness as the dependent variable and stream width, basin size, location, and connectivity as the independent variables against SCDHEC instream habitat quality scores (Figure 25). The resulting relationship was characterized by a significant (p=0.03) curvilinear trend showing that richness was low when instream habitat quality was low, peaked at intermediate to moderately high levels of instream habitat quality, and decreased at very high levels of instream habitat quality.

Table 19. Factors influencing fish species richness in the study area as shown by a general linear model incorporating continuous (stream width, watershed size, and instream habitat) and categorical (geographic location and connection to larger streams) variables*.

Source	Type III Sum Squares	Df	Mean Squares	*F*	*p*
Stream width	251.6	1	251.6	27.48	<0.001
Watershed size	218.7	1	218.7	23.93	<0.001
Geographic location	281.8	8	35.2	3.84	0.001
Connection to larger stream	80.2	1	80.2	8.79	0.005
Instream habitat	72.7	1	72.7	7.85	0.007
Instream habitat squared	71.0	1	71.0	7.81	0.007
Error	503.6	55	9.2		$R^2 = 0.76$

*Species richness values were estimates of true species richness derived from a first-order jackknife estimator.

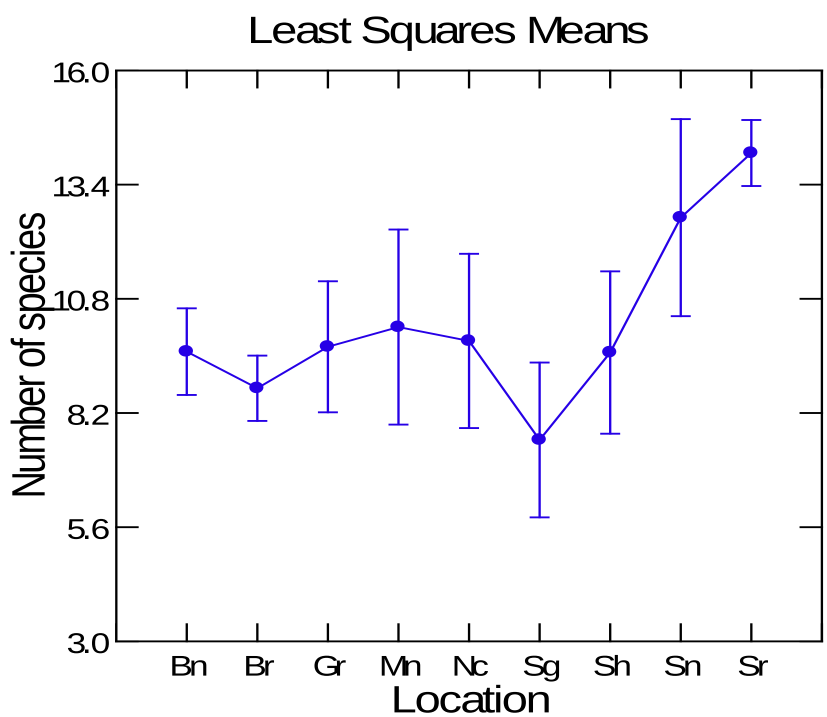


Figure 22. Differences in estimated mean (±1 SE) number of species (calculated using the first-order jackknife estimator) among installations (Bn = Fort Benning, Br = Fort Bragg, Gr = Fort Gordon, Mn = Manchester State Forest, Nc = The Nature Conservancy, Sg = Sandhills Gamelands, Sh = Sand Hills State Forest, Sn = Carolina Sandhills National Wildlife Refuge and Sr = Savannah River Site) after accounting for the effects of differences in stream size among installations.

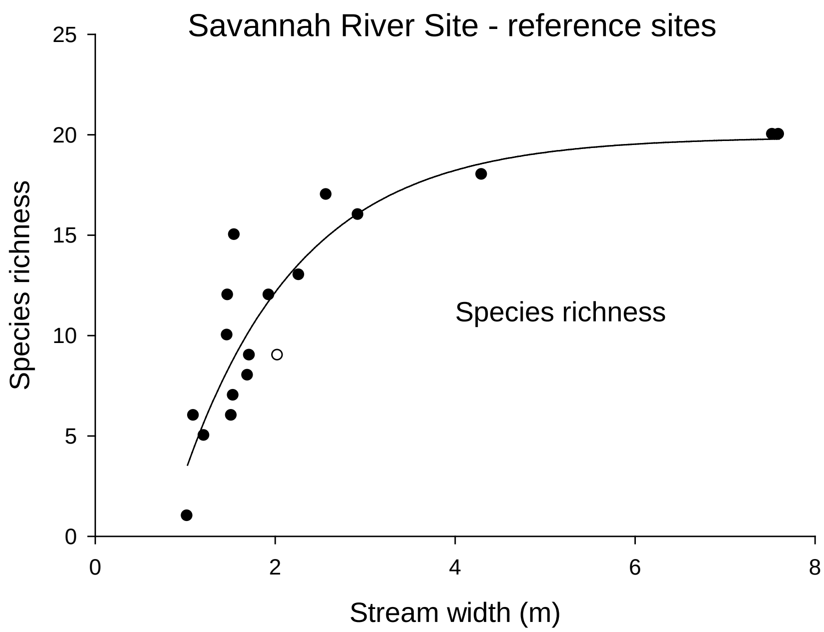


Figure 23. Relationship between species richness and stream width for Savannah River Site study streams.

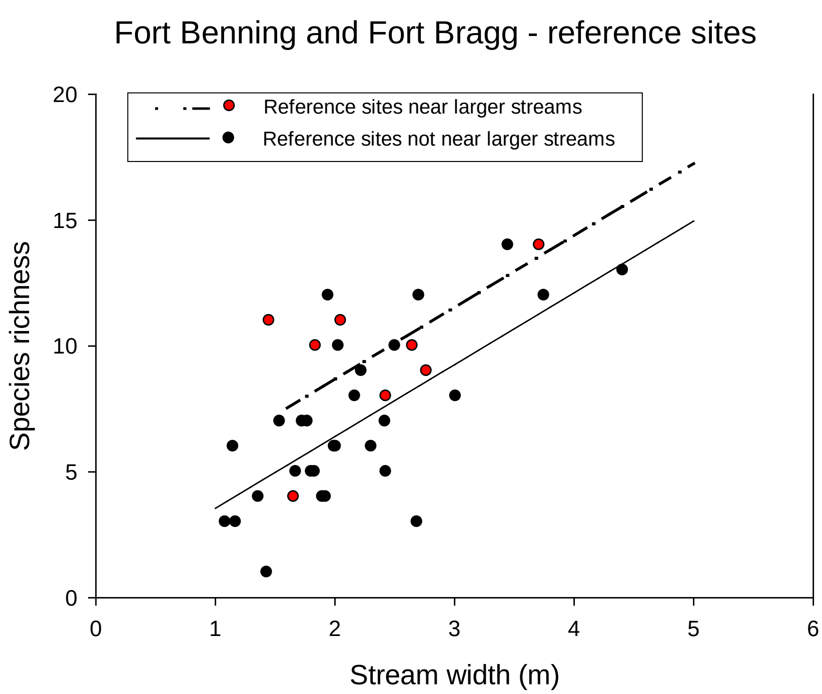


Figure 24. Relationship between species richness and stream width for Fort Benning and Fort Bragg streams.

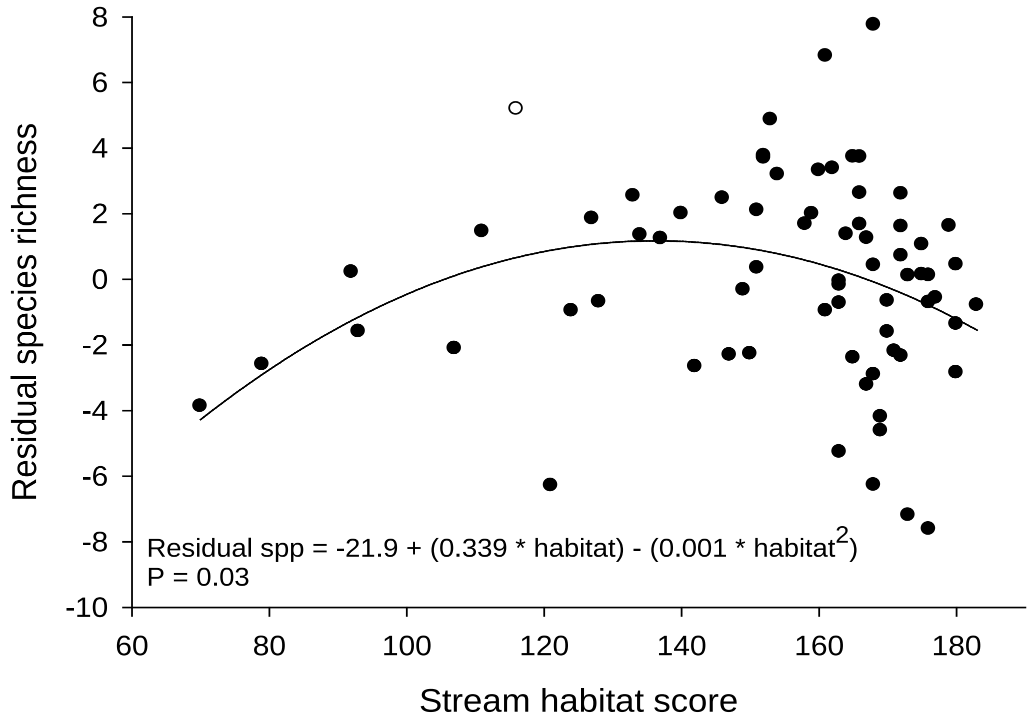


Figure 25. Relationship between fish species richness residuals at al sample sites and instream habitat quality as represented by the SCDHEC instream habitat assessment protocol. The residuals represent species richness after the statistical removal of variance associated with stream width, basin size, geographic location, and connectivity to larger streams.

5.2.2.4 Summary of Factors Affecting Fish Assemblage Structure

The importance of geography in determining stream fish assemblage composition within the study area was not surprising given the relatively high level of endemism that is characteristic of fish faunas in the southeastern United States. Southeastern rivers were not subject to Pleistocene glaciation resulting in a relatively long history of environmental stability (Swift et al. 1986, Warren et al. 2000). This factor, combined with the geographic isolation of many southeastern river systems and the relatively high habitat diversity within the region, has contributed to development of regionally distinct faunas (Swift et al. 1986, Warren et al. 2000). Differences in fish assemblage composition among watersheds indicate that a single reference model is inadequate for fish taxonomic structure across the study area and that fish-based assessment frameworks need to be responsive to geographic variations in assemblage composition.

Like geography, stream size and related measures (i.e., stream order and watershed size) are known to affect fish richness and assemblage composition, both within Southeastern Plain ecosystems and elsewhere (Sheldon 1968, Paller 1994). Its importance in this study was reflected in the significance of basin size in the CCA models of fish assemblage composition and in the GLM of fish richness. Fish composition changes with stream size through addition and replacement of species, although the relative importance of these 2 processes may vary. Downstream increases in richness tend to be large in streams and rivers where there is a strong longitudinal gradient of decreasing environmental variability (Horwitz 1978), and species replacements, rather than additions, are more common in streams with large longitudinal changes in thermal regime, habitat, or geomorphology (Rahel and Hubert 1991). Previous research indicated that headwater richness was relatively high in SRS streams compared with streams in other regions due to mild climate (e.g., lack of drought) and relatively weak longitudinal elevation gradients (Paller 1994). It also showed that species replacements were prominent across a gradient of first- through fourth-order streams.

Increases in richness with stream size observed in this study was largely the result of species additions rather than the loss of species characteristic of small streams, and their replacement by different species in larger streams. Most species found in the smallest streams under study were also found in all but the largest streams, albeit often in smaller numbers (Table 20). Examples include the sandhill chub (shc in Table 20) in Fort Bragg streams; the dixie chub (dc) in Fort Benning streams; the creek chub (cc), bluehead chub (bhc), and yellowfin shiner (ysh) in Savannah River Site streams, and the pirate perch (pp) in streams throughout the study area. This pattern is probably the result of sampling a comparatively narrow range of stream sizes (mostly between 1 and 4 m average width); inclusion of more large streams would have likely resulted in a greater frequency of species replacements. Prominent increases in species number with stream size need to be reflected in the development of stream reference models and assessment frameworks, although this problem is somewhat simpler when the increase is due primarily to additions rather than to the combination of addition and replacement.

In addition to stream size, fish assemblage composition can be significantly affected by the position of the site within the larger spatial scale of the entire stream network since this factor can influence resource availability (e.g., habitats, food), likelihood of immigration, and risk of extinction. Research has shown that proximity to a confluence with a larger stream can increase richness, particularly for taxa such as catostomids, cyprinids, and darters due to immigration from the larger stream (Osborne and Wiley 1992, Osborne et al. 1992). Immigration may be particularly strong when the connection is to a relatively species rich stream of significantly larger size and may extend up to 20 km from the confluence in mid-size to larger streams (Hitt and Angermeier 2011). A result is that isolated headwater stream sites often have fewer species than stream sites of the same size that are connected to larger streams, the latter being sometimes referred to as adventitious streams (Thornbrugh and Gido 2010). Processes of immigration

Table 20. Changes in fish species composition with stream size at Fort Benning, Fort Bragg, and the Savannah River Site (see Table 1 for site abbreviations and Table 10 for species abbreviations).

Fort Bragg

Site	Width (m)	shc	pp	mm	lcs	mf	lmb	msf	ccs	ybh	rbsf	dsf	rfp	bg	mmt	dsh	cp	wm	tdt	eel
Br-hrse	1.4	77	27							6			1							
Br-uwp	1.4	42																		
Br-lrf	1.5	10	5					1		9	3			3	39	154	1		2	3
Br-lrt	1.5	1	17					1		2			1		3					1
Br-gum	1.8		5		6			5					2				2			
Br-mcp	1.9		69	21	30			7		3		6	5							
Br-cab	2.0	19	25		1					1					7	70				
Br-deep	2.0	14	7					2		8			1			4				
Br-cyp	2.1	3	15					6	1	4	6	4	3		1	172				1
Br-wp	2.2	17	8					2		4	1		1		6	6			1	
Br-hect	2.3									6	1				26	15			1	2
Br-ujt	2.4	18	8							5			5		37	61			17	1
Br-field	2.4	36	10					4					6							2
Br-jun	2.5		15					5		15		2	6		19	232			11	1
Br-bct	2.6		37	7		1944	1		8	55	49	10						3		
Br-jen	2.8	3	15					2	1	9	1				3	40			4	
Br-flat	3.0	3	7							7	1		1		19	26				2
Br-tank	3.2					39	7			3				10				48		
Br-rfb	3.8	1	5					1		19		2	6		7	21	2		6	1
Br-bm	4.4		11					2	3	4		2	8		4	90		1	1	

Fort Benning

Site	Width (m)	gsdt	dc	pp	lam	wesh	bssh	ybh	rbsf	rfp	bg	wm	smt	bbdt	rssf
Bn-D13F	0.9	22	19	7			40			1					
Bn-D13C	1.5	18	1	15			325	8		4					
Bn-D12	1.7	5		20			1	1							
Bn-K13	1.8	1	3	7	5		47			5				1	
Bn-LJT	1.8			3	86		35	1							
Bn-Holl	1.9		107	12	115		40								
Bn-HCT	1.9		17	8		4	31	7	5	3	1	4	9	10	5
Bn-LPK	2.1		93	18	37		1	11		6		1			
Bn-K11	2.4		1	14	165		49	3		17	1				
Bn-Chat	2.6		24	6	1	15		3	4	2	4	2		39	2
Bn-wolf	2.7		90	13	83		43	5	2	3	2	9	3	18	

Savannah River Site

Site	Width (m)	cc	bhc	ysh	tmt	mm	pp	ssf	mf	sdt	lmb	msf	ccs	ybh	rbsf	dsf	rfp	llsh	mmt	nhs	dsh	wm	tdt	eel	smt	bbdt
Sr-u10	0.9	58	4	3		1																				
Sr-mq10	1.0	120																								
Sr-Mqh	1.1	57	44	146	7		8		4					6	5				3					1		
Sr-mc7	1.1	37	9	28	1		6						1													
Sr-mq8	1.2	36	35	122	4		11																			
Sr-mc6c	1.5	5		71	1		26	3		2				2		5	4								3	
Sr-tc6	1.5	9	7	129	2		5			5				1		1	1		2				1		5	
Sr-mb61	1.5	2					22					3					2									
Sr-mcm	1.5			3			4					1	1			2		6			5				1	
Sr-mc6	1.5	9	7	163	2		21							1	1											
Sr-pbhw	1.6	28	2	23	2		18	15	13			1	1	3	2	1	1				6			1		
Sr-tc5	1.7	5		6	2		10			1		3					2	20								
Sr-pb4	1.7	50	50	158	4	1	67	12					4			1	3									
Sr-mc5	1.9	4	13	51		2	20	1	16						10	7			1				1	1		
Sr-mb6	2.0	1	6	53			19			1		1	2				2							3		
Sr-mbhw	2.3	9	46	150			18	2	2	4			9	5	6	3	7		10							
Sr-Pbm	2.6	2	48	296	2		50	28	20	9	1		14	3	17	22	3		1		3	1	8			
Sr-mbm	2.9	8	44	233			14	10	1	1			7	1	8	3	22		5				5	2		1
Sr-tcm	7.5		7	131	1		6	1	1	2			1		2	1	2		4	1	2		3	6	14	3
Sr-ltr	7.6						23	2	1	1	2		10	1	9	5	10			4	45	1	3	10		8

and extinction likely contribute to this phenomenon although environmental factors such as the presence of more habitats and shallower elevation gradients may also contribute to higher richness in adventitious streams. The position of a stream site within a stream network may have important consequences for bioassessment frameworks because it can influence community metrics independently of environmental quality, which is the real target of such programs.

In this study, influence of a site's position in a stream network was reflected by the significance of the stream connectivity variable in CCAs of assemblage structure and species GLM. A site was considered connected if it was near the confluence with a larger stream (i.e., at least one order larger) that could serve as a source of immigrants to the site. Our definition of "near" was different for relatively small ($\leq$1.5 m average width) and relatively large (>1.5 m average width) streams and considerably less (250 and 750 m, respectively) than in some studies where proximity-related effects were observed several km from the confluence. The distances we chose were based on literature concerning fish movements in small streams (which indicates that many stream fishes are relatively sedentary and characterized by movements of several 100 m or less, Hill and Grossman 1987, Rodríguez 2002), and on field observations. Small streams often had shallow riffles and runs, log/brush jambs, and plunge pools that could serve as barriers to upstream movement of fish from higher-order streams, especially movement of large fish exposed to predators in shallow water. Such barriers were less common in larger streams, which possessed habitats that were generally similar to habitats in the higher-order streams to which they were joined and less likely to present obstacles to fish movement. Our study did not provide detailed information on effects of stream network position on fish assemblage structure, and it is likely that such effects extended farther than cut-off distances used in our analyses, although not as far as observed in studies on larger streams (Hitt and Angermeier 2011). The effects of immigration from a larger stream would be expected to weaken with distance from the confluence and have progressively smaller effects on bioassessment results.

The SCDHEC instream habitat quality score was significant in both CCAs of fish assemblage structure and the species richness general linear model. This variable was designed as an indicator of disturbance (lower scores indicate greater habitat degradation) and was related to other measures of disturbance including percentage of developed land in the watershed (Pearson r = -0.69), surface area of paved roads in the watershed (r = -0.58), and estimated bank erosion (r = -0.63). Given that instream habitat quality is a general indicator of disturbance, the previously described curvilinear relationship between quality and richness indicates that richness peaked at intermediate levels of disturbance and declined at sites where disturbance was lowest. The lowest levels of disturbance would be expected at the highest quality reference site; i.e., those sites least influenced by anthropogenic activities. This pattern of increased species number at sites of slight to moderate disturbance might be a manifestation of biotic homogenization – increased similarity of biotas over time caused by replacement of native species with nonindigenous species – which is a major threat to biodiversity and often the first sign of ecosystem degradation (Rahel 2000, Scott and Helfman 2001). Homogenization can result in an initial increase in number of species followed by a decrease as degradation worsens and endemic species are extirpated and replaced by a uniform assemblage of nonindigenous species (Figure 26).

Homogenization of fish assemblages in the southeastern United States typically involves replacement of endemic species characteristic of small, undisturbed, highland streams with generalist species characteristic of lowland areas (Scott and Helfman 2001). Endemic highland species are relatively specialized and typically associated with hard (e.g., rocky) bottoms that are used for feeding and spawning. Many are darters, but several shiners and other species also fall into this category (Scott and Helfman 2001). The generalists are typically widespread native taxa that are able to utilize a variety of foods including insects, zooplankton, detritus, and plant material and able to spawn in a range of habitats. Generalist species are well represented in the Centrarchidae, Cyprinidae, Ictaluridae, and other families (Scott and Helfman 2001) and are characteristic of downstream reaches at lower elevations. Streams in the southern Appalachian highlands may be particularly susceptible to faunal homogenization because they support comparatively large numbers of endemic taxa. Southeastern Plains streams support fewer endemic taxa (Warren et al. 2000); however, headwater fish assemblages in coastal plains streams are comparatively species-rich and distinctive in composition compared with assemblages farther downstream (Paller 1994). These characteristics make them susceptible to homogenization, although this process may be more subtle than in regions that support more endemic species.

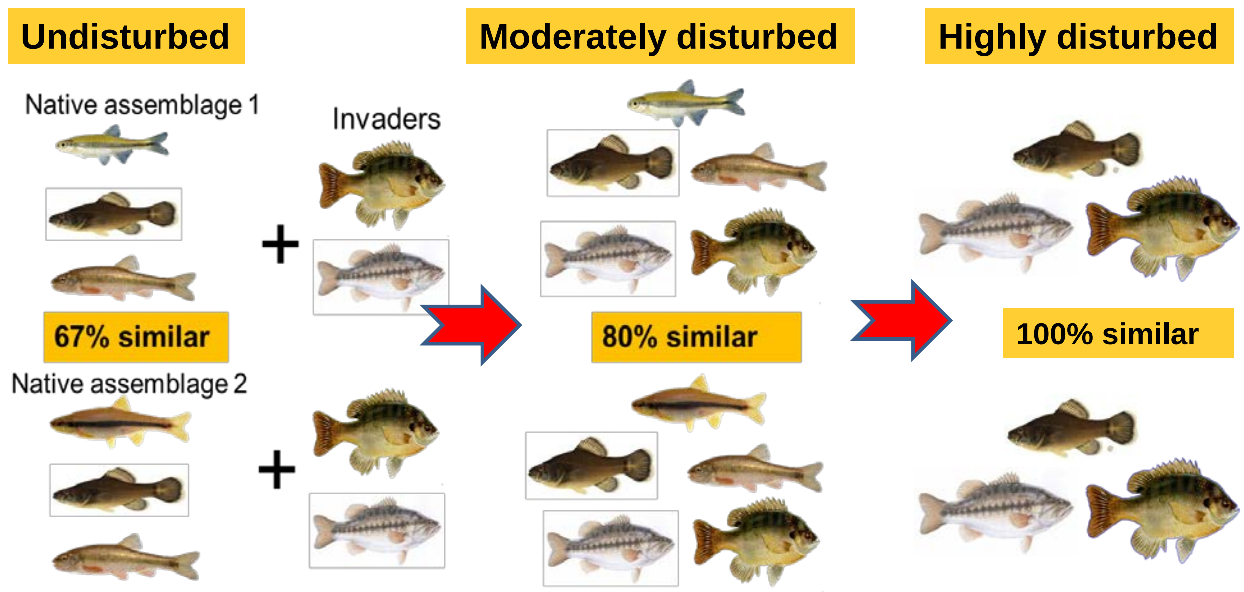


Figure 26. Biotic homogenization: a process by which species invasions and extinctions increase the genetic, taxonomic, or functional similarity of two or more locations over time.

In faunal homogenization, generalist species that invade headwater reaches as disturbance increases are usually native species characteristic of downstream reaches. This pattern makes it difficult to detect early stages of homogenization in which the assemblage consists of a mix of endemic upstream species plus a small number of downstream generalists. This difficulty is exacerbated by the fact that downstream generalists may occasionally move upstream as a result of natural disturbance, such as high water levels, rather than anthropogenic habitat degradation. However, sites reaching an advanced state of anthropogenic degradation are more easily distinguishable because expected endemic species are scarce or lacking and generalists are numerically dominant. Therefore, comparing assemblage composition at these sites with undisturbed sites may help to identify species that are typical of disturbance and likely to contribute to homogenization.

Assemblage composition was compared among sites that were separated into 2 groups along a disturbance gradient defined by PCA ordination of the sample sites on several measures of environmental disturbance, as explained earlier. Group 1 consisted of 14 sites characterized by the highest levels of disturbance (see Section 5.2.3 for site list), and group 2 consisted of remaining less-disturbed sites. Indicator species analysis (ISA) was used to identify fish species that were most characteristic of each disturbance group. This methodology assigns indicator values to species found in particular groups of sites (in this case disturbance groups) based on differences in relative abundance and frequency of occurrence between groups. Indicator values for each species are tested for significance using a Monte Carlo procedure (McCune and Grace 2002). The ISA identified 5 species that were more abundant and occurred more frequently at disturbed sites including the bluegill, largemouth bass, warmouth, eastern mosquitofish, and lake chubsucker. These species, although native to the study area, are more common in large streams or lentic habitats than in small, undisturbed (i.e., reference) streams and their appearance in the latter may represent faunal homogenization. Largemouth bass and bluegill, in particular, are often stocked in man-made impoundments from which they may invade contiguous stream reaches. The recognition of species likely to be associated with faunal homogenization is needed for the design of assessment frameworks that are sensitive to the early stages of degradation resulting from environmental disturbance.

5.2.3 *Reference Models for Fish*

A reference model is a description of the biota that is naturally present in the absence of significant human disturbance or alteration. Reference models are defined on the basis of biota because aquatic organisms integrate their chemical and physical environment, and therefore represent the summation of ambient abiotic conditions. Organisms are also intrinsically relevant due to the ecological services they provide and their symbolic importance as indicators of ecosystem health. The reference model serves as a criterion or benchmark that defines thresholds for biological impairment. References models are generally multivariate because of the complexity of ecological communities; they may consist of the matrix of species found under undisturbed conditions or of metrics that represent key aspects of ecological structure and function. Variables included in reference models are usually represented by a range of values to encompass variability inherent in natural systems and associated with sampling error. However, if the range of natural variability represented by a reference model is too great in magnitude; e.g., if the model seeks to encompass too large a geographic area, it may lose specificity and, hence, sensitivity to disturbance-related changes in biotic composition may be reduced.

The most common approach for defining a reference model or models is the reference site approach in which the model is defined by measuring biological variables at a set of minimally disturbed or least-disturbed sites (Bailey et al. 2004). Sites are selected by application of appropriate reference site screening criteria to abiotic data. As described is Section 5.1, reference site screening can be a tiered approach that incorporates progressively greater rigor by the successive application of more screening criteria. More rigorous screening guards against inclusion of sites that may inappropriately lower reference site expectations but increases difficulty of finding enough reference sites to adequately represent natural variation. Only one method was used to identify reference sites for establishment of fish assemblage reference models: Method 1 (Section 5.1) in which sites were ranked on a disturbance gradient developed by PCA of abiotic measures of disturbance. The application of multiple screening methods to the fish assemblage data was unproductive because it resulted in exclusion of sites that were similar in fish assemblage composition to sites remaining in the reference site pool.

In this study, 54 sites met the reference site screening criteria associated with Method 1 and, therefore, qualified as least disturbed sites: 20 from Fort Bragg and the nearby Sandhills Gamelands, 16 from the SRS, 9 from Fort Benning and nearby Nature Conservancy holdings, 3 from Fort Gordon, and 6 from Manchester State Forest and the Sand Hills State Forest (Table 21).

5.3.2.1 Reference Site Fish Assemblages

There are naturally occurring environmental factors or gradients within the study area that affected fish assemblage composition at reference sites – chief among them being geographic location and stream/watershed size. These need to be considered when defining reference site expectations so that natural variation is not confounded with effects of disturbance. Geographic location significantly affected taxonomic composition, relative abundance, and species richness resulting in different reference site expectations for each installation. Many species were installation-specific such as the yellowfin shiner and bluehead chub at SRS, broadstripe shiner and dixie chub at Fort Benning, and sandhills chub at Fort Bragg. Species richness, an important metric in many assessment frameworks, was higher at the SRS than at the other installations and was also strongly influenced by stream size throughout the study area. Taxonomic composition and richness at reference sites in each of the major installations included in this study constituted the basis of the fish assemblage reference models (Figures 27 – 29). These data will serve in the Section 5.2.9 as benchmarks for establishing assessment frameworks.

Reference model variability (i.e., variability in fish assemblage structure among reference sites) resulted from numerous sources that have been extensively discussed. Relationships between these variables and fish assemblage metrics are used in Section 5.2.9 to adjust for natural sources of variation that could be confounded with disturbance related effects or that could reduce the sensitivity of assessment methods by increasing background "noise." However, not all variability could be assigned to specific causes. Some unexplained variability may have been related to sampling error, such as the failure to collect some of the rare species at typical sampling scales (Lohr and Fausch, 1997). Temporal change in fish assemblage structure was another potential source of variability that was not explicitly quantified. Although stream fish assemblages exhibit significant persistence (constancy of species composition) and stability (constancy of relative abundance), they also show considerable unpredictable temporal variation (Matthews et al. 1988, Meffe and Berra 1988, Schlosser 1990). Temporal variation can be related to sampling methodology; e.g., sampling regimes synchronized with seasonal cycles of abundance may show less variability than sampling regimes that are not (Taylor et al. 1996). However, substantive ecological factors may also play a role. Hydrological fluctuations have been associated with changes in fish assemblage structure over time (Horwitz 1978, Matthews 1986), and fluctuations in temperature and dissolved oxygen may contribute to the relatively high temporal variability characteristic of fish assemblages in headwater streams (Schlosser 1990).

Ten randomly selected sites on the SRS and Fort Bragg were resampled to quantify temporal variability in fish assemblage structure. All were reference sites with the exception of Sr-mqhw. Sites were initially sampled in 2009 or 2010 and resampled with the same methods in 2012 (Table 22). Most sites exhibited little change in richness and abundance between samples. An exception was Sr-pbhw, which showed marked reductions in both variables (Table 22). Differences in assemblage structure between samples were summarized with NMDS, which connected original and repeat samples with "successional vectors" in ordination space – the length of vectors being proportional to the difference in assemblage structure over time. NMDS produced 3 significant ($P<0.05$) axes. Most repeat samples were connected by short vectors indicating little change in assemblage structure over time (Figure 30). Exceptions included Sr-pbhw and Br-gum. Br-gum did not show large changes in species richness and abundance like Sr-pbhw but did experience large changes in species composition between samples. This resulted in the long successional vector for this site. We hypothesize that large changes in Sr-pbhw occurred because this small stream was largely dewatered by low rainfall between samples as indicated by observations made between the 2 samples. Br-gum, the other site with substantial changes between samples, was also a small stream. In summary, the repeat samples suggested relatively little temporal variability with the exception of some sample sites in small streams. These results concur with the findings of others that headwater fish assemblages are more variable than downstream assemblages because of greater environmental fluctuations (Horwitz 1978). They also suggest that assessment frameworks that are based on fish assemblage structure will show more variable results in headwater streams than in downstream reaches.

Table 21. Reference (R) and test (T) sites for fish sampling.

Installation	Site name	SRNL abbrev	DN-LAT	DN-LONG	GIS
Fort Benning	Chatahoochee River Tributary	Bn-Chat	32.34904	85.00444	T
Fort Benning	Bonham Creek Tributary (D12)	Bn-D12	32.41187	84.75690	R
Fort Benning	Bonham Creek Tributary (D13)	Bn-D13F	32.41546	84.76090	R
Fort Benning	Bonham Creek Tributary (D13-C)	Bn-D13C	32.41767	84.76023	T
Fort Benning	HCT-G2/5	Bn-HCT	32.33421	84.70339	R
Fort Benning	Hollis Branch Mainstem(F4)	Bn-Holl	32.36395	84.68455	R
Fort Benning	King Mill Creek(K11E)	Bn-K11	32.51054	84.64841	R
Fort Benning	K13	Bn-K13	32.49748	84.70769	R
Fort Benning	Little Juniper Tributary (K10)	Bn-LJT	32.51151	84.63712	R
Fort Benning	Little Pine Knot (K20)	Bn-LPK	32.39764	84.67085	T
Fort Benning	Wolf Creek	Bn-wolf	32.41938	-84.83746	T
Fort Bragg	Beaver Creek trib	Br-bct	35.13342	-78.97667	T
Fort Bragg	Big Muddy Creek Main Stem	Br-bm	35.01986	79.51553	R
Fort Bragg	Cabin Creek	Br-cab	35.05333	-79.29673	T
Fort Bragg	Cypress Creek	Br-cyp	35.17974	-79.04723	R
Fort Bragg	Deep Creek	Br-deep	35.14507	-79.15983	R
Fort Bragg	Field Branch Main Stem	Br-field	35.06282	79.30483	R
Fort Bragg	Flat Creek Main Stem	Br-flat	35.17424	79.18026	R
Fort Bragg	Gum Branch Main Stem	Br-gum	35.08958	79.33523	R
Fort Bragg	Hector Creek Main Stem	Br-hect	35.18380	79.09892	R
Fort Bragg	Horse Creek (HC)	Br-hrse	35.17208	79.23305	R
Fort Bragg	Jennie Creek Main Stem	Br-jen	35.12032	79.33021	R
Fort Bragg	Juniper Creek Main Stem	Br-jun	35.07234	79.25689	R
Fort Bragg	Little Rockfish Main Stem North	Br-lrf	35.17134	79.08787	R
Fort Bragg	Little River Tributary	Br-lrt	35.18748	79.07469	R
Fort Bragg	McPherson Creek	Br-mcp	35.14079	-79.04396	T
Fort Bragg	Rockfish Branch (RFB)	Br-rfb	35.11509	79.32548	R
Fort Bragg	Tank Creek	Br-tank	35.14716	-79.02106	T
Fort Bragg	Upper Jennie Trib (UJT)	Br-ujt	35.13354	79.34534	R
Fort Bragg	Upper Wolf Pit (UWP)	Br-uwp	35.11102	79.35168	R
Fort Bragg	Wolf Pit Creek Main Stem	Br-wp	35.1132	79.33718	R
Fort Gordon	Bath Branch	Gr-bath	33.35942	-82.15746	R
Fort Gordon	Boggy Gut Creek	Gr-Bog	33.34772	82.29175	R
Fort Gordon	McCoys Creek	Gr-Mcoy	33.39939	-82.16021	T
Fort Gordon	Trib to Marcum Branch	Gr-Mrbtrb	33.40980	-82.18657	T
Fort Gordon	South Prong	Gr-Prong	33.35942	-82.15746	R
Manchester State Forest	McCrays Creek	Mn-mcra	33.78431	-80.47599	R
Manchester State Forest	Tavern Creek	Mn-tav	33.75819	-80.52800	R
The Nature Conservancy	Black Jack Creek	Nc-bjc	32.58038	-84.49602	R
The Nature Conservancy	Black Creek Tributary	Nc-bct	32.56942	-84.51484	R
The Nature Conservancy	Pine Knot Creek	Nc-pkc	32.43953	-84.64734	R
The Nature Conservancy	Parkers Mill Creek Tributary	Nc-pmt	32.45264	-84.57671	R
Sandhills Game Lands	Bones Fork Tributary	Sg-bone	35.03531	-79.61385	R
Sandhills Game Lands	Joes Creek	Sg-joes	34.88026	-79.62551	R
Sandhills Game Lands	Millstone Creek	Sg-mill	35.06723	-79.66504	T
Sandhills State Forest	Little Cedar Creek	Sh-cedr	34.51895	-79.99899	R
Sandhills State Forest	Mill Creek	Sh-mill	34.53822	-80.07601	R
Sandhills NWR	Big Black Creek Tributary	Sn-bbct	34.66050	-80.22655	T
Sandhills NWR	Hemp Creek	Sn-hemp	34.57139	-80.24704	R
Sandhills NWR	Rogers Branch	sn-rogr	34.60473	-80.20997	R
Savannah River Site	Lower Three Runs (DS)	Sr-ltr	33.22353	81.50894	R
Savannah River Site	Meyers Branch 6	Sr-mb6	33.1781	81.56570	R
Savannah River Site	Meyers Branch 6.1	Sr-mb61	33.18051	81.56335	T
Savannah River Site	Meyers Branch Headwaters	Sr-mbhw	33.19357	81.57880	R
Savannah River Site	Meyers Branch Main Stem	Sr-mbm	33.17613	81.58174	R
Savannah River Site	Mill Creek Tributary 5	Sr-mc5	33.31922	81.58011	R
Savannah River Site	Mill Creek Tributary 6	Sr-mc6	33.31731	81.59759	R
Savannah River Site	Mill Creek Tributary 6C	Sr-mc6c	33.31940	81.59619	R
Savannah River Site	Mill_Creek Tributary 7	Sr-mc7	33.32440	81.60101	R
Savannah River Site	Mill Creek Main Stem	Sr-mcm	33.30074	81.58680	R
Savannah River Site	McQueens Branch Tributary (MQ10.1)	Sr-mq10	33.29810	81.62629	R
Savannah River Site	McQueens Branch Tributary (8)	Sr-mq8	33.30454	81.62630	R
Savannah River Site	McQueens Branch Headwater	Sr-Mqh	33.29836	81.62959	T
Savannah River Site	Pen Branch Tributary (4)	Sr-pb4	33.23349	81.63809	R
Savannah River Site	Pen Branch Headwater	Sr-pbhw	33.23250	81.62424	R
Savannah River Site	Pen Branch Main Stem	Sr-Pbm	33.22576	81.63570	R
Savannah River Site	Tinker Creek Tributary (5)	Sr-tc5	33.37333	81.54981	R
Savannah River Site	Tinker Creek Tributary (6)	Sr-tc6	33.36077	81.55763	R
Savannah River Site	Tinker Creek Main	Sr-tcm	33.36369	81.55806	R
Savannah River Site	U10	Sr-u10	33.30036	81.66618	R

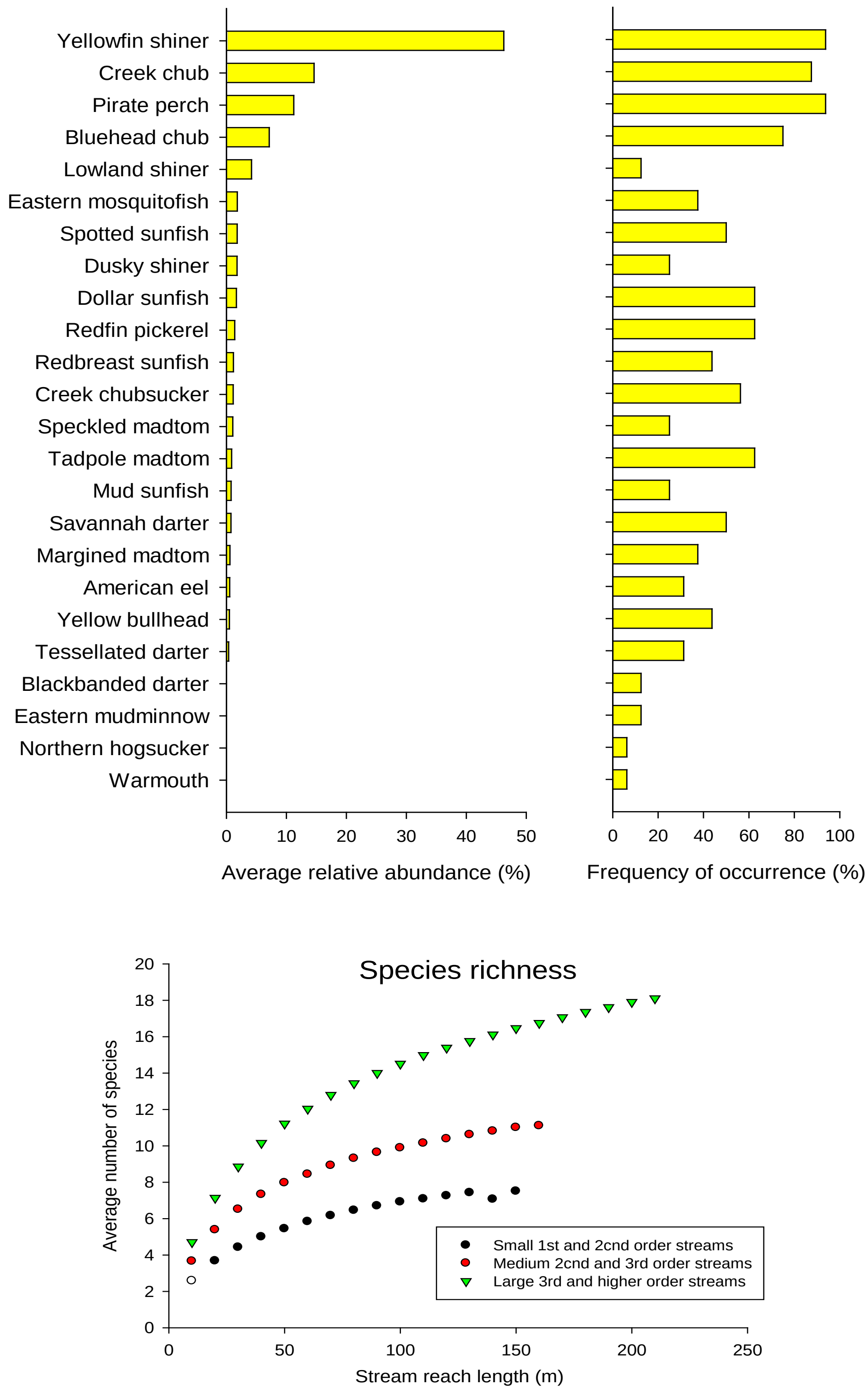


Figure 27. Fish species composition and species richness for reference sites at the Savannah River Site.

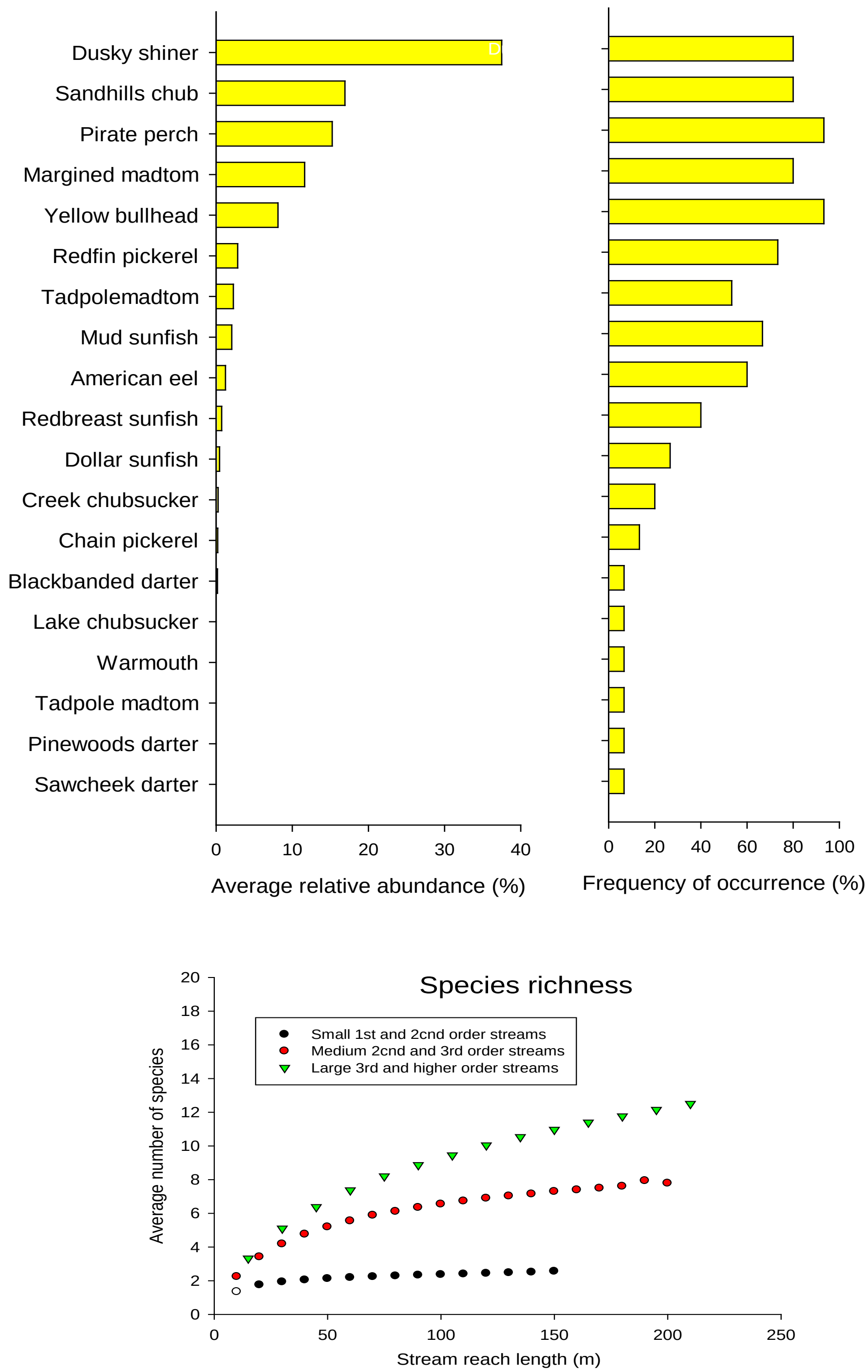


Figure 28. Fish species composition and species richness for reference sites at Fort Bragg.

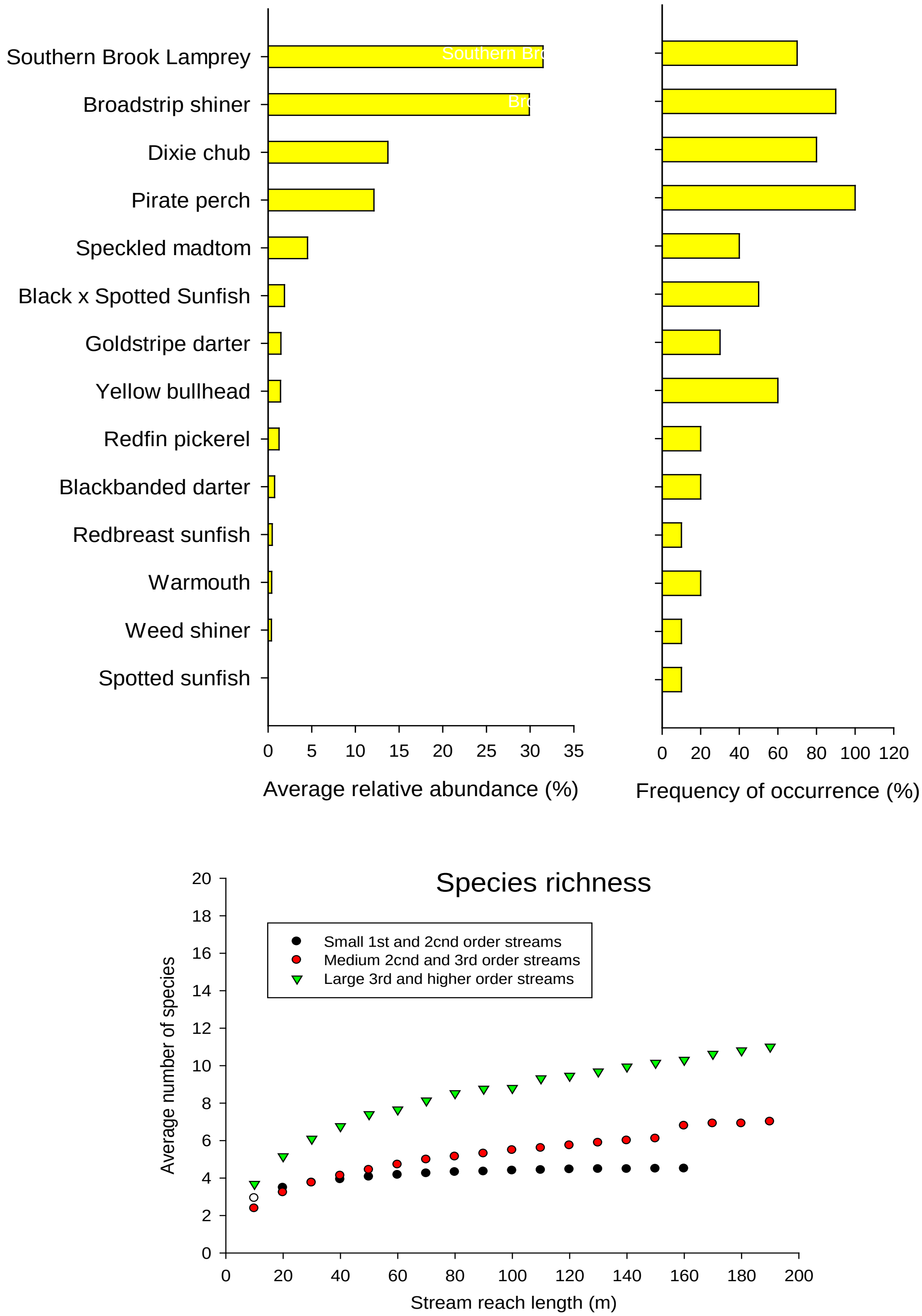


Figure 29. Fish species composition and species richness for reference sites at Fort Benning.

5.2.3.2 Environmental Factors Associated with Reference Conditions

Although ecological reference models consist of biological data, it is important to recognize key environmental characteristics that can aid in the recognition of relatively undisturbed conditions and assist in their discrimination from more disturbed sites. Environmental characteristics associated with the reference condition can be described at landscape, watershed, and stream segment spatial scales. These scales constitute a nested hierarchy, with each level having a strong influence on the levels subsumed within it. Environmental factors that were typically associated with reference sites and/or that influenced biota at reference sites are listed below. Support for cut-off points (referenced below) that delimit reference conditions will be provided in Section 5.2.9.3.

1) Landscape condition - affects connectivity to population sources of colonizing fishes
   Key variables:
   A. Proximity to (connection with) larger streams – a natural factor that tends to increase species richness due to proximity to sources of colonizing species. This factor does not determine whether a site represents the reference condition but, along with stream size, influences the number of species likely to occur at a reference site. Stream reaches located near (within 250 m for streams <1.5 m wide and 750 m for streams $\geq$1.5 m wide) a confluence with a larger (at least one order greater) stream had an average of 2 more species than isolated reaches.
   B. Proximity to artificial impoundments – an anthropogenic factor that influences the likelihood of encountering invasive lentic species and that may perturb stream hydrology. Artificial impoundments increase the likelihood of invasion by lentic species (e.g., bluegill and largemouth bass) that are not normally found in the undisturbed reaches of small streams (first through fourth order). Sites in streams with artificial impoundments are unlikely to represent reference conditions.
2. Watershed condition – affects stream/riparian habitat and stream hydrology
   Key variables:
   A. Percent coverage of anthropogenically developed lands. Watersheds with >20% developed land (as indicated by the previously defined % Disturbance Index) are unlikely to support reference sites.
   B. Percent forest cover (evergreen and/or deciduous). Watersheds with <50% forest cover are unlikely to support reference sites regardless of the status of the rest of the watershed.
   C. Watershed size – a natural factor that affects species richness. Larger watersheds support larger streams that contain greater richness. Within the mostly first- through fourth-order streams under study, this increase was primarily due to species additions rather than replacements. Watershed size does not determine whether a site represents the reference condition but rather strongly influences the number of species likely to occur at reference sites.
3. Stream/riparian habitat condition determines the habitat template for aquatic biota, which constitutes the reference model core.
   Key variables:
   A. Instream habitat quality as measured by the SCDHEC instream habitat protocol "calibrated" for SH streams. Guidance for scoring the variables included in the protocol is provided in Section 4.3. Sites with scores <140 are unlikely to represent reference conditions.
   B. Stream bank/channel erosion. Guidance for assessing this variable is included in Section 4.3. Sites with average erosion scores >1.0 are unlikely to represent reference conditions.
   C. Bank height and bank angle. High, steep banks are indicative of channel incision that can result from discharge fluctuations (i.e., flashiness) that are often associated with watershed degradation. Steep stream banks in $\geq$1 m in height within the study area were often associated with impaired fish assemblages.

D. Forested riparian zone. Lack of forestation in the riparian zone is indicative of environmental degradation and incompatible with the reference condition. An intact riparian forest should be present throughout the watershed upstream of potential reference sites.
E. Channel modifications. Stream reaches with any but minor channel modifications (e.g., a small bridge on an unpaved road that does not obstruct flow) are unlikely to support reference fish assemblages.

Table 22. Comparisons between repeated fish assemblage samples collected from 10 sample sites*.

	Br-field		Br-gum		Br-jen		Br-ujt		Br-uwp		Sr-mb6		Sr-mbhw		Sr-mc6		Sr-Mqh		Sr-pbhw	
Species	2009	2012	2010	2012	2009	2012	2010	2012	2010	2012	2009	2012	2009	2012	2009	2012	2009	2012	2009	2012
Blackbanded sunfish	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Bluegill	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Bluehead chub	0	0	0	0	0	0	0	0	0	0	6	0	46	57	7	18	50	45	2	0
Creek chub	0	0	0	0	0	0	0	0	0	0	1	17	9	18	9	34	57	33	28	4
Creek chubsucker	0	0	0	0	1	0	0	0	0	0	2	0	9	5	0	0	0	1	1	1
Chain pickerel	0	0	2	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0
Dollar sunfish	0	0	0	0	0	0	0	0	0	0	0	3	3	5	0	0	0	0	1	0
Dusky shiner	0	0	0	0	40	5	61	27	0	0	0	0	0	0	0	0	0	0	6	0
American eel	2	1	0	0	0	1	1	0	0	0	3	0	0	0	0	1	1	0	1	0
Lake chubsucker	0	0	6	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	1
Eastern mosquitofish	0	0	0	0	0	0	0	0	0	0	0	0	2	8	0	0	4	0	13	1
Margined madtom	0	0	0	0	3	2	37	28	0	1	0	0	10	5	0	5	3	1	0	0
Mud sunfish	4	0	5	2	2	2	0	4	0	0	1	2	0	0	0	0	0	0	1	0
Pirate perch	10	14	5	31	15	14	8	11	0	0	19	22	18	28	21	0	8	3	18	1
Redbreast sunfish	0	0	0	0	1	0	0	0	0	0	0	0	6	0	1	0	5	1	2	0
Redfin pickerel	6	2	2	1	0	4	5	25	0	0	2	4	7	10	0	0	0	0	1	4
Savannah darter	0	0	0	0	0	0	0	0	0	0	1	3	4	2	0	0	0	0	0	0
Sandhills chub	36	36	0	18	3	1	18	14	42	58	0	0	0	0	0	0	0	0	0	0
Speckled madtom	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
Spotted sunfish	0	0	0	0	0	0	0	0	0	0	0	1	2	16	0	0	0	0	15	0
Tesselated darter	0	0	0	0	4	7	17	9	0	0	0	0	0	0	0	0	0	0	0	0
Tadpole madtom	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2	3	7	4	2	0
Yellow bullhead	0	0	0	1	9	3	5	6	0	0	0	1	5	1	1	0	6	3	3	0
Yellowfin shiner	0	0	0	0	0	0	0	0	0	0	53	57	150	241	163	146	146	56	23	0
Number species	5	4	5	5	10	11	9	10	1	2	9	11	13	13	8	7	10	9	15	6
Number fish	58	53	20	53	79	41	153	126	42	59	88	112	271	404	205	208	287	147	117	12

* See Table 1 for a description of the sample sites.

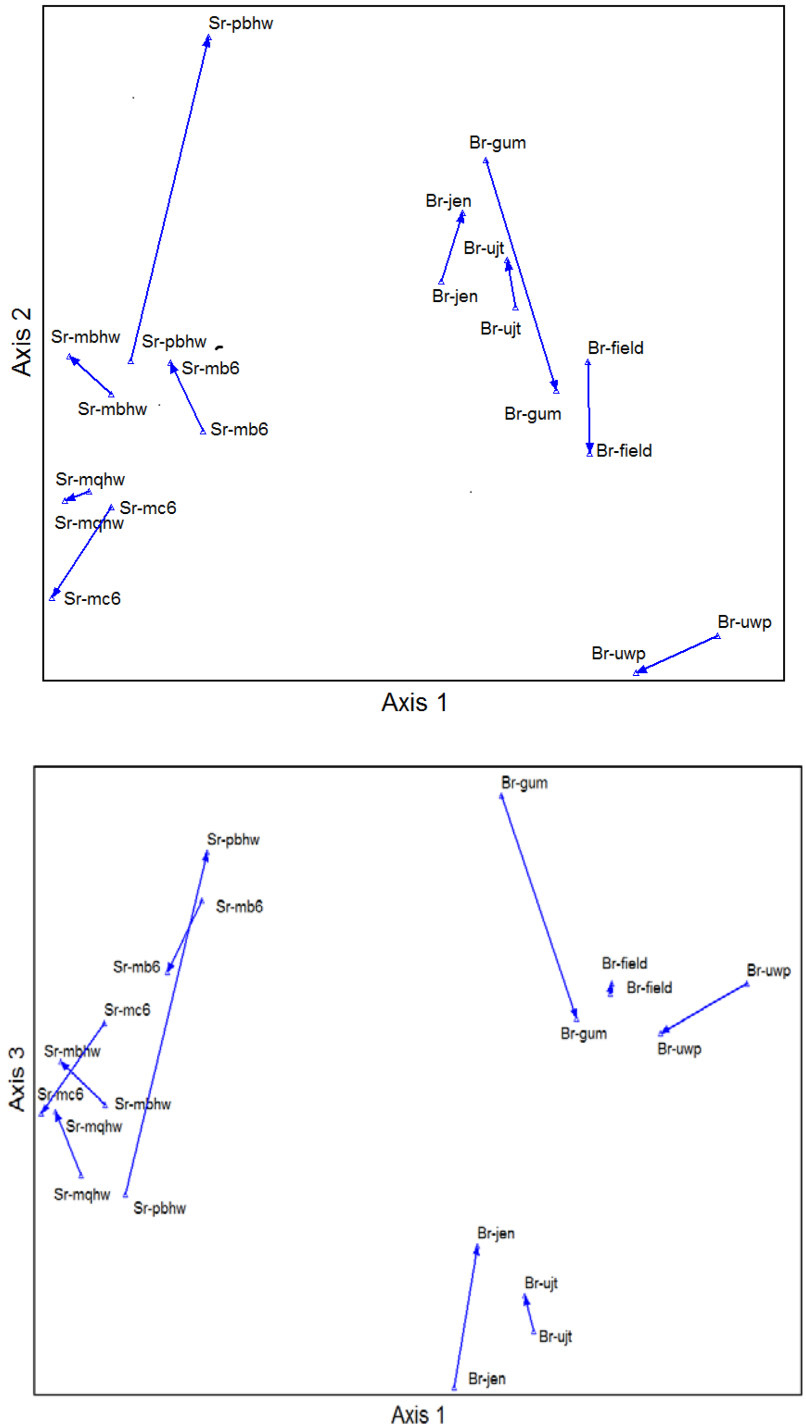


Figure 30. Nonmetric multidimensional scaling (NMDS) of repeated fish samples from 10 sample sites (axes 1 and 2 on upper graph, axes 1 and 3 on lower graph). Repeated samples from the same sites (2009 or 2010 versus 2012) are connected by successional vectors.

5.2.4 *Assessment Frameworks for Fish*

In this study, 3 assessment frameworks were developed for use with fish assemblage data: an ANNA (Assessment by Nearest Neighbor Analysis) O/E model, a SH multimetric index (MMI), and an original model termed the Community Quality Index (CQI).

5.2.4.1 Assessment by Nearest Neighbor Analysis (ANNA)

<u>Model description</u> – ANNA models fall within the category of assessment frameworks known as “predictive models” or “observed/expected (O/E)” models. These types of models were initially developed for macroinvertebrates in the UK (i.e., RIVPACS, Wright 1995), adopted in Australia (AUSRIV, Davies 2000) and Canada (BEAST), and are now used in the US (Hawkins et al. 2000). Predictive models statistically predict the taxonomic composition of the fauna that would be expected at a test site if it was in reference site condition. This prediction is based on key environmental characteristics of the test site including geography, habitat, and other important natural factors that influence assemblage structure. The expected fauna is then compared with fauna actually observed at the test site, and O/E is calculated from observed and expected assemblages. Knowledge of the reference site fauna is typically acquired by surveying taxonomic composition at reference sites that have met suitable screening criteria. Test sites are matched with appropriate reference sites based on similarities in naturally occurring geographic and environmental factors that affect assemblage composition. Unlike multimetric indices, predictive models deal with taxonomic composition (generally presence/absence) rather than metrics.

Predictive models provide a probability-based, relatively objective method for biological assessment that takes account of natural environmental factors that affect taxonomic composition under reference conditions. However, they have been criticized for their initial biological classification step, which assumes that reference biotic assemblages occur in discrete classes rather than varying continuously along environmental gradients, as is usually the case. To counter this objection, Linke et al. (2005) devised a methodology known as Assessment by Nearest Neighbor Analysis (ANNA), which avoids the initial step of classifying reference sites into groups and instead predicts expected taxonomic composition from individual reference sites whose contribution is weighted by their similarity to the test site on key environmental variables. This aspect of ANNA imparts a theoretical advantage that may be important when the geographic scope of the model is restricted to a single ecoregion resulting in overlap of communities within the study area. In such circumstances, which obtain to some degree in this study, the classification algorithms associated with conventional predictive models are particularly likely to produce results that are indeterminate and somewhat subjective.

We used an ANNA predictive model for the aforementioned reasons. Being relatively novel, ANNA models are seldom used with fish assemblage data and have not been developed for fish assemblages in the Southeast. ANNA models are similar to other predictive models but have some important differences, especially in the initial calculations.

<u>Model development</u> – Development of predictive models generally involves several steps (Wright 1995, Hawkins et al. 2000). Details are provided below:

STEP 1. In step 1 of ANNA, the biological data collected from reference sites is ordinated using (NMDS), a robust indirect ordination technique that “maps” sites in ordination space so that similar sites are in close proximity. The data used by NMDS is a matrix of Bray-Curtis coefficients derived from the species by sample site matrix. The Bray-Curtis coefficients represent distances among sample sites based on similarities in taxonomic composition. NMDS maximizes rank-order correlations between distance measures/distances in ordination space. Points are adjusted to minimize "stress", which is a measure of

the mismatch between the two kinds of distance. Monte Carlo permutation procedures can be used to assess the significance of the ordination axes.

To perform step 1, we ordinated the fish assemblage data from the reference sample sites using NMDS. Two sites, Br-uwp and Sr-mq10 (Table 1), were excluded from the reference site pool because they were in very small, likely intermittent streams that supported only 1 species each (the sandhills chub in "uwp" and the creek chub in "mq10." NMDS produced 3 significant (P<0.05) axes, but only the first two were retained for analysis as the third was not significantly correlated with environmental data.

STEP 2. Step 2 of ANNA identifies the natural environmental variables that influence species composition at the reference sites. These include factors such as geography, elevation, watershed size, etc. Step 2 is typically accomplished by stepwise multiple regression of each ordination axis with environmental variables as predictors and ordination axis scores as the dependent variable. Environmental variables that are significant predictors of the ordination scores are used in step 3 as are the regression equations that describe relationships between these predictors and ordination scores.

To accomplish step 2, we performed stepwise multiple regression on each of the 3 NMDS ordination axes using 19 variables that represented geographic location, habitat features at watershed- and stream-reach scales, and water quality (N and P concentrations). Significant models were produced for axis 1 and axis 2 but not axis 3. Models for axis 1 and 2 were further refined by substituting environmental variables that are difficult to measure (e.g., watershed size) with correlated environmental variables that are comparatively easy to measure (e.g., stream width). This step was done, as a practical matter, to enhance the utility of the ANNA model as an assessment tool. Models were then retested with these predictors. The final regression models for axis 1 and 2 included 5 and 4 predictors, respectively, and had R^2 values of 0.83 and 0.80, respectively (Tables 23 and 24).

STEP 3. Step 3 of ANNA computes the distances (in NMDS space) from each reference site to the test site (Linke et al. 2005). This process is initiated by calculating "q," which is the predicted value for each reference site (i) on each NMDS axis (k). Intercepts are excluded from this equation because they would be subtracted from each other in the subsequent step:

$$q_{ik} = \sum_{m=1}^{M} a_{km} X_{im} \qquad (3)$$

where a_{km} is the multiple regression coefficient for environmental variable m on NMDS axis k and X_m is the value for environmental m for site i. Computation of q_{ik} for NMDS axis 1 and 2 are shown for the reference sites (Tables 25 and 26). Values of q_{ik} for the test sites are computed similarly.

Values of q are used in the distance equation, which computes the modified Euclidean distance (d) between 2 sites i and j for the 2 NMDS axes:

$$d_{ij} = \sqrt{\sum_{k=1}^{2} (q_{ik} - q_{jk})^2} \qquad (4)$$

STEP 4. The probabilities of occurrence for each taxon, weighted by proximity to environmentally similar reference sites by d_{ij} values, are calculated in step 4, following the formula of Linke et al. (2005):

$$p = \frac{\sum_{i=1}^{n} x_i \frac{1}{\sqrt{d_i}}}{\sum_{i=1}^{n} \frac{1}{\sqrt{d_i}}} \quad (5)$$

where *p* is the probability of occurrence of a particular taxon, *n* is the number of reference sites, $x_i = 0$ for absence of the taxon at reference site *i* and 1 for presence of the taxon at reference site *i*, and d_i is the distance to reference site *i*. The denominator in equation 5 is the sum of the square root of the reciprocal distances of the test site from each reference site. The numerator is similar except that each reciprocal distance value is multiplied by a 1 or 0. The square root of the reciprocal distance rather than the reciprocal distance itself is used to avoid overemphasizing very close sites (Linke et. al. 2005).

STEP 5. The last step in ANNA is the computation of the O/E ratio: E being the number of taxa expected at the test site and O being the number of these taxa that are actually observed at the test site. Taxa with a probability of collection near 1.0 should be present if the test site is minimally disturbed. About 3 in 4 and about 1 in 2 taxa should be collected at probabilities of 0.75 and 0.50, respectively. "E" in the O/E ratio is computed by summing the collection probabilities of all taxa expected above a designated probability level such as 0.50 or 0.75. The observed number of taxa is computed by counting the number of expected taxa that are actually observed at the test site. O/E ratios near one indicate that a test site meets reference site expectations. Wright (1995) provides a clear example of the method of O/E calculation.

It is important to determine the number of nearest neighbor reference sites to be included in an ANNA model. Although ANNA weights closer (i.e., environmentally similar) reference sites more strongly, the influence of distant sites can decrease assessment accuracy. The optimal number of nearest neighbor sites for predicting assemblage composition depends on strength of ecological gradients, amount of unexplained variation in the model, the model's geographic scope, number and spatial density of reference sites, and other factors (Linke et al. 2005). The number of nearest neighbor sites for optimal predictive accuracy is model specific and best defined empirically by determining the relative accuracy of models with different numbers of sites. This can be done by plotting observed versus expected (i.e., predicted) reference site values for models with different numbers of sites: plots with a high R^2, slope near 1, and intercept near 0 are indicative of an accurate model that generates average O/E site values near 1 and is unbiased. Expected values for each reference site are obtained by excluding the site from the reference site pool and computing the expected value from the remaining sites. This methodology can also be used to test the relative accuracy of models that use different probabilities of collection.

The preceding method was used to determine the number of nearest neighbor reference sites and the probability level that produced the most accurate model with the fish assemblage data. Models based on a 0.50 probability level of collection were comparatively biased and inaccurate as indicated by slopes that clearly departed from one and y-intercepts that departed from 0 (Figure 31). Models based on a 0.25

Table 23. Regression model (R^2 = 0.83) of environmental factors that influenced the first axis of a nonmetric multidimensional scaling (NMDS) of fish assemblages at reference sites.

Effect	Coefficient	Standard error	Standardized coefficient	*t*	*p*
Constant	-13.77	1.448	0.000	-9.508	0.000
Stream width (m)	0.092	0.044	0.158	2.109	0.004
Stream order (Strahler)	0.133	0.054	0.191	2.490	0.016
Latitude (degrees format)	0.410	0.044	0.691	9.390	0.000
Highest point in basin (m)	-0.003	0.001	-0.204	-3.076	0.004
Stream gradient*	-6.523	2.666	-0.176	-2.447	0.018

* (highest point in basin [m] – elevation at stream mouth [m])/basin length (m)

Table 24. Regression model (R^2 = 0.80) of environmental factors that influenced the second axis of a nonmetric multidimensional scaling (NMDS) of fish assemblages at reference sites.

Effect	Coefficient	Standard error	Standardized coefficient	*t*	*p*
Constant	129.084	9.814	0.000	13.152	0.000
Latitude (degrees format)	-1.725	0.131	-2.806	-13.121	0.000
Longitude (degrees format)	-0.867	0.068	-2.636	-12.722	0.000
Stream order (Strahler)	-0.161	0.08	-0.223	-2.020	0.049
Stream magnitude (Shreve)	0.017	0.006	0.332	3.022	0.004

Table 25. Computation of qik for NMDS axis 2. Coefficients are derived from the regression model shown in Table 24.

Site	Latitude coefficient	Latitude	$a_{km}X_{km}$	Longitude coefficient	Longitude	$a_{km}X_{km}$	Order coefficient	Stream order	$a_{km}X_{km}$	Magnitude coefficient	Stream magnitude	$a_{km}X_{km}$	$q_{ik=2}$
Bn-D12	-1.73	32.4119	-55.910	-0.87	84.7569	-73.484	-0.16	2	-0.322	0.017	4	0.068	-129.649
Bn-D13F	-1.73	32.4155	-55.917	-0.87	84.7609	-73.488	-0.16	2	-0.322	0.017	2	0.034	-129.692
Bn-HCT	-1.73	32.3342	-55.777	-0.87	84.7034	-73.438	-0.16	2	-0.322	0.017	7	0.119	-129.417
Bn-Holl	-1.73	32.3640	-55.828	-0.87	84.6846	-73.422	-0.16	2	-0.322	0.017	2	0.034	-129.537
Bn-K11	-1.73	32.5105	-56.081	-0.87	84.6484	-73.390	-0.16	2	-0.322	0.017	4	0.068	-129.725
Bn-K13	-1.73	32.4975	-56.058	-0.87	84.7077	-73.442	-0.16	2	-0.322	0.017	5	0.085	-129.737
Bn-LJT	-1.73	32.5115	-56.082	-0.87	84.6371	-73.380	-0.16	3	-0.483	0.017	6	0.102	-129.844
Br-bm	-1.73	35.0199	-60.409	-0.87	79.5155	-68.940	-0.16	5	-0.805	0.017	62	1.054	-129.100
Br-cyp	-1.73	35.1797	-60.685	-0.87	79.0472	-68.534	-0.16	3	-0.483	0.017	18	0.306	-129.396
Br-deep	-1.73	35.1451	-60.625	-0.87	79.1598	-68.632	-0.16	2	-0.322	0.017	3	0.051	-129.528
Br-field	-1.73	35.0628	-60.483	-0.87	79.3048	-68.757	-0.16	3	-0.483	0.017	6	0.102	-129.622
Br-flat	-1.73	35.1742	-60.676	-0.87	79.1803	-68.649	-0.16	3	-0.483	0.017	20	0.34	-129.468
Br-gum	-1.73	35.0896	-60.530	-0.87	79.3352	-68.784	-0.16	2	-0.322	0.017	3	0.051	-129.584
Br-hect	-1.73	35.1838	-60.692	-0.87	79.0989	-68.579	-0.16	4	-0.644	0.017	39	0.663	-129.252
Br-hrse	-1.73	35.1721	-60.672	-0.87	79.2331	-68.695	-0.16	2	-0.322	0.017	2	0.034	-129.655
Br-jen	-1.73	35.1203	-60.583	-0.87	79.3302	-68.779	-0.16	3	-0.483	0.017	12	0.204	-129.641
Br-jun	-1.73	35.0723	-60.500	-0.87	79.2569	-68.716	-0.16	3	-0.483	0.017	26	0.442	-129.257
Br-lrf	-1.73	35.1713	-60.671	-0.87	79.0879	-68.569	-0.16	3	-0.483	0.017	12	0.204	-129.519
Br-lrt	-1.73	35.1875	-60.698	-0.87	79.0747	-68.558	-0.16	4	-0.644	0.017	27	0.459	-129.441
Br-rfb	-1.73	35.1151	-60.574	-0.87	79.3255	-68.775	-0.16	4	-0.644	0.017	54	0.918	-129.075
Br-ujt	-1.73	35.1335	-60.605	-0.87	79.3453	-68.792	-0.16	2	-0.322	0.017	6	0.102	-129.618
Br-wp	-1.73	35.1132	-60.570	-0.87	79.3372	-68.785	-0.16	3	-0.483	0.017	6	0.102	-129.737
Gr-bath	-1.73	33.3581	-57.543	-0.87	82.1580	-71.231	-0.16	2	-0.322	0.017	8	0.136	-128.960
Gr-Bog	-1.73	33.3477	-57.525	-0.87	82.2918	-71.347	-0.16	2	-0.322	0.017	5	0.085	-129.109
Gr-Prong	-1.73	33.3594	-57.545	-0.87	82.1575	-71.231	-0.16	3	-0.483	0.017	7	0.119	-129.140
Mn-mcra	-1.73	33.7843	-58.278	-0.87	80.4760	-69.773	-0.16	2	-0.322	0.017	3	0.051	-128.322
Mn-tav	-1.73	33.7582	-58.233	-0.87	80.5280	-69.818	-0.16	1	-0.161	0.017	1	0.017	-128.195
Nc-bct	-1.73	32.5694	-56.182	-0.87	84.5148	-73.274	-0.16	3	-0.483	0.017	4	0.068	-129.872
Nc-bjc	-1.73	32.5804	-56.201	-0.87	84.4960	-73.258	-0.16	1	-0.161	0.017	1	0.017	-129.603
Nc-pmt	-1.73	32.4526	-55.981	-0.87	84.5767	-73.328	-0.16	1	-0.161	0.017	1	0.017	-129.453
Sg-bone	-1.73	35.0353	-60.436	-0.87	79.6139	-69.025	-0.16	2	-0.322	0.017	4	0.068	-129.715
Sg-joes	-1.73	34.8803	-60.168	-0.87	79.6255	-69.035	-0.16	2	-0.322	0.017	3	0.051	-129.475
Sg-mlst	-1.73	35.0672	-60.491	-0.87	79.6650	-69.070	-0.16	3	-0.483	0.017	7	0.119	-129.925
Sh-cedr	-1.73	34.5190	-59.545	-0.87	79.9990	-69.359	-0.16	3	-0.483	0.017	15	0.255	-129.132
Sn-hemp	-1.73	34.5714	-59.636	-0.87	80.2470	-69.574	-0.16	2	-0.322	0.017	2	0.034	-129.498
sn-rogr	-1.73	34.6047	-59.693	-0.87	80.2100	-69.542	-0.16	2	-0.322	0.017	2	0.034	-129.523
Sr-mb6	-1.73	33.1781	-57.232	-0.87	81.5657	-70.717	-0.16	2	-0.322	0.017	4	0.068	-128.204
Sr-mbhw	-1.73	33.1936	-57.259	-0.87	81.5788	-70.729	-0.16	2	-0.322	0.017	4	0.068	-128.242
Sr-mbm	-1.73	33.1761	-57.229	-0.87	81.5817	-70.731	-0.16	4	-0.644	0.017	20	0.34	-128.264
Sr-mc5	-1.73	33.3192	-57.476	-0.87	81.5801	-70.730	-0.16	1	-0.161	0.017	1	0.017	-128.350
Sr-mc6	-1.73	33.3173	-57.472	-0.87	81.5976	-70.745	-0.16	2	-0.322	0.017	3	0.051	-128.488
Sr-mc6c	-1.73	33.3194	-57.476	-0.87	81.5962	-70.744	-0.16	1	-0.161	0.017	1	0.017	-128.364
Sr-mc7	-1.73	33.3244	-57.485	-0.87	81.6010	-70.748	-0.16	1	-0.161	0.017	1	0.017	-128.377
Sr-mcm	-1.73	33.3007	-57.444	-0.87	81.5868	-70.736	-0.16	3	-0.483	0.017	5	0.085	-128.578
Sr-mq8	-1.73	33.3045	-57.450	-0.87	81.6263	-70.770	-0.16	2	-0.322	0.017	5	0.085	-128.457
Sr-pb4	-1.73	33.2335	-57.328	-0.87	81.6381	-70.780	-0.16	2	-0.322	0.017	3	0.051	-128.379
Sr-pbhw	-1.73	33.2325	-57.326	-0.87	81.6242	-70.768	-0.16	2	-0.322	0.017	3	0.051	-128.365
Sr-Pbm	-1.73	33.2258	-57.314	-0.87	81.6357	-70.778	-0.16	3	-0.483	0.017	9	0.153	-128.423
Sr-tc5	-1.73	33.3733	-57.569	-0.87	81.5498	-70.704	-0.16	2	-0.322	0.017	3	0.051	-128.544
Sr-tc6	-1.73	33.3608	-57.547	-0.87	81.5576	-70.710	-0.16	2	-0.322	0.017	6	0.102	-128.478
Sr-tcm	-1.73	33.3637	-57.552	-0.87	81.5581	-70.711	-0.16	3	-0.483	0.017	13	0.221	-128.525
Sr-u10	-1.73	33.3004	-57.443	-0.87	81.6662	-70.805	-0.16	1	-0.161	0.017	1	0.017	-128.392

Table 26. Computation of qik for NMDS axis 1. Coefficients are derived from the regression model shown in Table 23.

Site	Width (m) coefficien	Width	$a_{km}X_{km}$	Order coefficien	Order	$a_{km}X_{km}$	Latitude coefficient	Latitude	$a_{km}X_{km}$	Gradient coefficien	Gradient	$a_{km}X_{km}$	High point coefficient	High point	$a_{km}X_{km}$	$q_{ik=2}$
Bn-D12	0.09	1.66	0.152	0.13	2	0.266	0.411	32.4119	13.321	-6.52	0.024	-0.155	-0.003	158	-0.474	13.111
Bn-D13F	0.09	0.88	0.081	0.13	2	0.266	0.411	32.4155	13.323	-6.52	0.048	-0.316	-0.003	145	-0.435	12.919
Bn-HCT	0.09	1.95	0.179	0.13	2	0.266	0.411	32.3342	13.289	-6.52	0.022	-0.146	-0.003	225	-0.675	12.914
Bn-Holl	0.09	1.92	0.177	0.13	2	0.266	0.411	32.3640	13.302	-6.52	0.043	-0.281	-0.003	222	-0.666	12.797
Bn-K11	0.09	2.42	0.223	0.13	2	0.266	0.411	32.5105	13.362	-6.52	0.027	-0.173	-0.003	191	-0.573	13.104
Bn-K13	0.09	1.77	0.163	0.13	2	0.266	0.411	32.4975	13.356	-6.52	0.031	-0.205	-0.003	180	-0.540	13.041
Bn-LJT	0.09	1.80	0.166	0.13	3	0.399	0.411	32.5115	13.362	-6.52	0.003	-0.022	-0.003	185	-0.554	13.351
Br-bm	0.09	4.41	0.406	0.13	5	0.665	0.411	35.0199	14.393	-6.52	0.019	-0.125	-0.003	144	-0.431	14.908
Br-cyp	0.09	2.05	0.189	0.13	3	0.399	0.411	35.1797	14.459	-6.52	0.019	-0.124	-0.003	138	-0.415	14.507
Br-deep	0.09	2.01	0.185	0.13	2	0.266	0.411	35.1451	14.445	-6.52	0.028	-0.183	-0.003	162	-0.487	14.225
Br-field	0.09	2.43	0.223	0.13	3	0.399	0.411	35.0628	14.411	-6.52	0.047	-0.304	-0.003	158	-0.473	14.256
Br-flat	0.09	3.01	0.277	0.13	3	0.399	0.411	35.1742	14.457	-6.52	0.031	-0.200	-0.003	158	-0.474	14.458
Br-gum	0.09	1.83	0.169	0.13	2	0.266	0.411	35.0896	14.422	-6.52	0.086	-0.562	-0.003	142	-0.425	13.870
Br-hect	0.09	2.31	0.212	0.13	4	0.532	0.411	35.1838	14.461	-6.52	0.032	-0.206	-0.003	109	-0.328	14.671
Br-hrse	0.09	1.36	0.125	0.13	2	0.266	0.411	35.1721	14.456	-6.52	0.023	-0.152	-0.003	126	-0.379	14.316
Br-jen	0.09	2.77	0.255	0.13	3	0.399	0.411	35.1203	14.434	-6.52	0.029	-0.192	-0.003	160	-0.481	14.415
Br-jun	0.09	2.50	0.230	0.13	3	0.399	0.411	35.0723	14.415	-6.52	0.030	-0.195	-0.003	155	-0.466	14.382
Br-lrf	0.09	1.45	0.134	0.13	3	0.399	0.411	35.1713	14.455	-6.52	0.063	-0.409	-0.003	134	-0.402	14.177
Br-lrt	0.09	1.54	0.142	0.13	4	0.532	0.411	35.1875	14.462	-6.52	0.051	-0.330	-0.003	106	-0.317	14.488
Br-rfb	0.09	3.75	0.345	0.13	4	0.532	0.411	35.1151	14.432	-6.52	0.015	-0.095	-0.003	158	-0.473	14.741
Br-ujt	0.09	2.43	0.223	0.13	2	0.266	0.411	35.1335	14.440	-6.52	0.021	-0.135	-0.003	159	-0.477	14.317
Br-wp	0.09	2.17	0.200	0.13	3	0.399	0.411	35.1132	14.432	-6.52	0.059	-0.384	-0.003	161	-0.482	14.165
Gr-bath	0.09	3.71	0.341	0.13	2	0.266	0.411	33.3581	13.710	-6.52	0.017	-0.112	-0.003	151	-0.452	13.753
Gr-Bog	0.09	2.65	0.244	0.13	2	0.266	0.411	33.3477	13.706	-6.52	0.017	-0.111	-0.003	162	-0.486	13.619
Gr-Prong	0.09	2.70	0.249	0.13	3	0.399	0.411	33.3594	13.711	-6.52	0.014	-0.089	-0.003	167	-0.502	13.767
Mn-mcra	0.09	2.03	0.187	0.13	2	0.266	0.411	33.7843	13.885	-6.52	0.010	-0.063	-0.003	95	-0.286	13.990
Mn-tav	0.09	1.90	0.175	0.13	1	0.133	0.411	33.7582	13.875	-6.52	0.008	-0.052	-0.003	70	-0.209	13.921
Nc-bct	0.09	1.15	0.106	0.13	3	0.399	0.411	32.5694	13.386	-6.52	0.034	-0.220	-0.003	201	-0.604	13.067
Nc-bjc	0.09	1.08	0.100	0.13	1	0.133	0.411	32.5804	13.391	-6.52	0.041	-0.271	-0.003	204	-0.611	12.742
Nc-pmt	0.09	1.16	0.107	0.13	1	0.133	0.411	32.4526	13.338	-6.52	0.040	-0.261	-0.003	213	-0.639	12.678
Sg-bone	0.09	1.84	0.169	0.13	2	0.266	0.411	35.0353	14.400	-6.52	0.015	-0.098	-0.003	133	-0.398	14.339
Sg-joes	0.09	2.22	0.205	0.13	2	0.266	0.411	34.8803	14.336	-6.52	0.012	-0.081	-0.003	133	-0.400	14.325
Sg-mlst	0.09	2.69	0.247	0.13	3	0.399	0.411	35.0672	14.413	-6.52	0.004	-0.025	-0.003	156	-0.469	14.564
Sh-cedr	0.09	3.45	0.317	0.13	3	0.399	0.411	34.5190	14.187	-6.52	0.016	-0.102	-0.003	137	-0.410	14.392
Sn-hemp	0.09	1.73	0.159	0.13	2	0.266	0.411	34.5714	14.209	-6.52	0.036	-0.232	-0.003	161	-0.482	13.920
sn-rogr	0.09	1.68	0.154	0.13	2	0.266	0.411	34.6047	14.223	-6.52	0.042	-0.274	-0.003	174	-0.523	13.845
Sr-mb6	0.09	2.03	0.187	0.13	2	0.266	0.411	33.1781	13.636	-6.52	0.013	-0.088	-0.003	103	-0.309	13.692
Sr-mbhw	0.09	2.27	0.209	0.13	2	0.266	0.411	33.1936	13.643	-6.52	0.007	-0.045	-0.003	98	-0.294	13.778
Sr-mbm	0.09	2.93	0.269	0.13	4	0.532	0.411	33.1761	13.635	-6.52	0.009	-0.059	-0.003	103	-0.309	14.068
Sr-mc5	0.09	1.93	0.178	0.13	1	0.133	0.411	33.3192	13.694	-6.52	0.011	-0.073	-0.003	96	-0.288	13.645
Sr-mc6	0.09	1.54	0.142	0.13	2	0.266	0.411	33.3173	13.693	-6.52	0.026	-0.169	-0.003	97	-0.291	13.641
Sr-mc6c	0.09	1.47	0.135	0.13	1	0.133	0.411	33.3194	13.694	-6.52	0.022	-0.143	-0.003	97	-0.291	13.529
Sr-mc7	0.09	1.10	0.101	0.13	1	0.133	0.411	33.3244	13.696	-6.52	0.026	-0.173	-0.003	99	-0.297	13.461
Sr-mcm	0.09	1.52	0.140	0.13	3	0.399	0.411	33.3007	13.687	-6.52	0.010	-0.067	-0.003	104	-0.312	13.846
Sr-mq8	0.09	1.21	0.112	0.13	2	0.266	0.411	33.3045	13.688	-6.52	0.024	-0.157	-0.003	105	-0.315	13.594
Sr-pb4	0.09	1.72	0.158	0.13	2	0.266	0.411	33.2335	13.659	-6.52	0.024	-0.155	-0.003	93	-0.279	13.649
Sr-pbhw	0.09	1.55	0.143	0.13	2	0.266	0.411	33.2325	13.659	-6.52	0.016	-0.107	-0.003	110	-0.330	13.630
Sr-Pbm	0.09	2.57	0.237	0.13	3	0.399	0.411	33.2258	13.656	-6.52	0.009	-0.060	-0.003	110	-0.330	13.901
Sr-tc5	0.09	1.70	0.156	0.13	2	0.266	0.411	33.3733	13.716	-6.52	0.018	-0.114	-0.003	118	-0.354	13.670
Sr-tc6	0.09	1.48	0.136	0.13	2	0.266	0.411	33.3608	13.711	-6.52	0.013	-0.084	-0.003	104	-0.312	13.717
Sr-tcm	0.09	7.53	0.693	0.13	3	0.399	0.411	33.3637	13.712	-6.52	0.011	-0.070	-0.003	122	-0.366	14.369
Sr-u10	0.09	1.04	0.096	0.13	1	0.133	0.411	33.3004	13.686	-6.52	0.050	-0.324	-0.003	85	-0.255	13.337

probability level generally exhibited greater accuracy and less bias. Those with 10, 12, and 15 nearest neighbors had slopes near 1 and y-intercepts near 0 (Figure 32). R^2 values for these models ranged from 0.30 to 0.36, with the highest R^2 belonging to the model with 12 nearest neighbors. All 3 models were significant, and the R^2 values for the models were within the range observed by Linke et al. (2005) for their models. Additionally, all R^2 values were above the cut-off level (0.22) used by Linke et al (2005) to identify potentially useful models. None of the models were excessively affected by influential data points as indicated by Cook's distance values under 1.0 for all data points in all regressions. The 0.25 model with 12 nearest neighbors was used in subsequent testing of the ANNA methodology.

Model results – The ANNA model produced an average O/E value of 0.99 (SD=0.30, 0.38 – 1.55) for the reference sites. The lower tenth percentile of the distribution of reference site O/Es (i.e., 0.63) was selected as the cut-off O/E value for separating reference from impaired sites. This value was selected because reference sites encompass a range of variability that is affected by various biological factors as well as sampling issues and errors in reference site selection. Errors in the selection of reference sites can result from disturbances that are undetected during the process of reference site screening. The accidental inclusion of disturbed sites in the reference site pool or the inclusion of reference sites that scored poorly because of sampling problems can lower reference site standards, which results in the failure to accurately identify disturbed sites. The lowest scoring portion of the reference site distribution can be excluded from the reference model to minimize this problem. The lower 10% of the reference site distribution is a commonly selected cutoff point (e.g., Wright 1995, Simpson and Norris 2000Linke 2005).

Calculation of O/E values for all of the sample sites (disturbed and reference) showed that 8% of the sites with good instream habitat (i.e., SCDHEC habitat assessment scores $\geq$160) had O/E scores below <0.63, 19% of the sites with fair habitat (i.e., habitat assessment scores $\geq$120 and <160) had O/E scores <0.63, and 71% of the sites with poor habitat had O/E scores <0.63 (Figure 33). The ANNA model results for the 14 sites that were classified as disturbed (using the PCA gradient approach, Section 5.1.2) was particularly important because none of these sites were used to develop the ANNA model, thereby avoiding circularity. Success of the model would be indicated by O/E values <0.63 for most or all sites that were classified as disturbed. Of these 14 sites, 7 (50%) had O/E ratios <0.63 indicating biotic impairment. The remaining 7 sites had scores approaching or exceeding 1 (0.94-1.52). This latter group was more similar to the sites classified as reference, as indicated by position on the PCA disturbance gradient, than were the 7 sites that received scores <0.63 (Figure 9). Sites within this group can be considered slightly or moderately rather than highly disturbed.

Further examination of the 14 disturbed sites included in the assessment showed that those with either low species richness (after considering stream size) or aberrant species composition received low O/E scores; examples included Br-tank and Gr-mcoy. However, slightly or moderately disturbed sites, especially those with relatively high richness, had O/E scores within the reference site range. A plausible explanation is that degradation at the latter group of sites was as yet insufficient to eliminate the native species expected at reference sites. As shown previously, moderately disturbed sites were often characterized by normal or even elevated species richness as a consequence of biotic homogenization. We hypothesize that the ANNA method may not be sensitive to early stages of fish assemblage degradation that are associated with biotic homogenization because these stages are characterized by the addition of

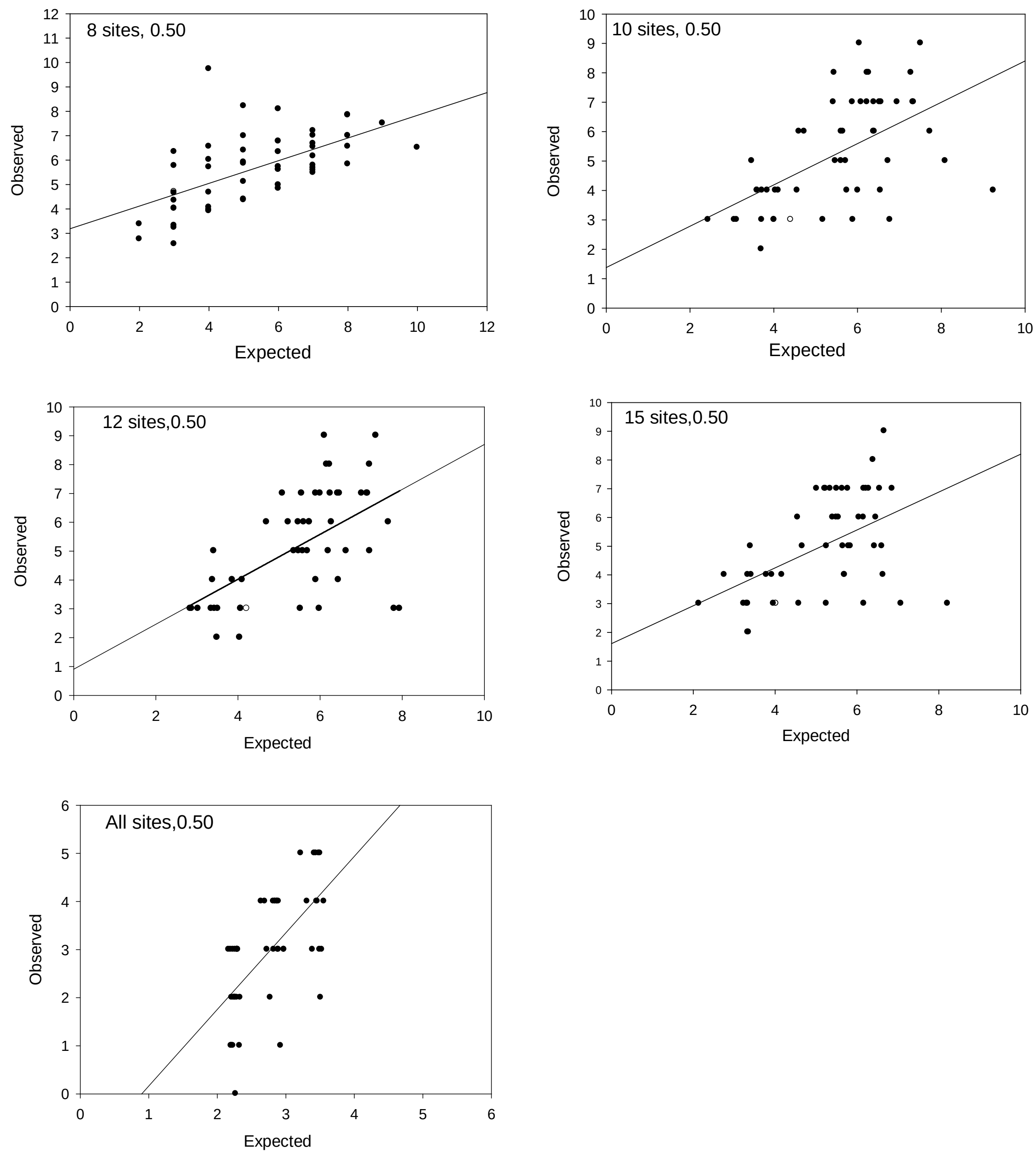


Figure 31. Observed versus expected number of taxa excluding fish taxa with an expected probability of occurrence >= 0.50 using the Assessment by Nearest Neighbor Analysis (ANNA) predictive modeling approach.

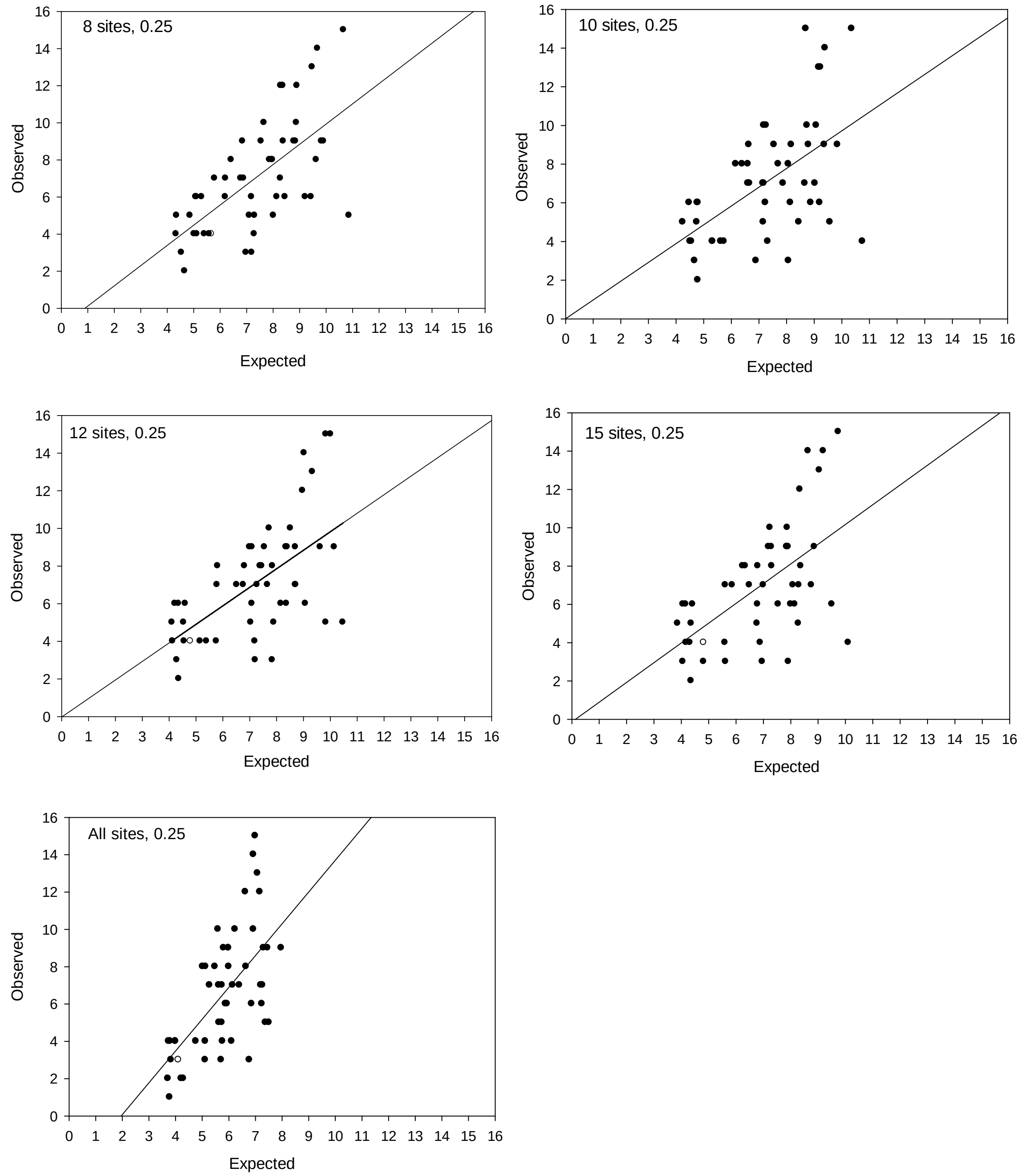


Figure 32. Observed versus expected number of fish taxa excluding taxa with an expected probability of occurrence >= 0.25 using the Assessment by Nearest Neighbor Analysis (ANNA) predictive modeling approach.

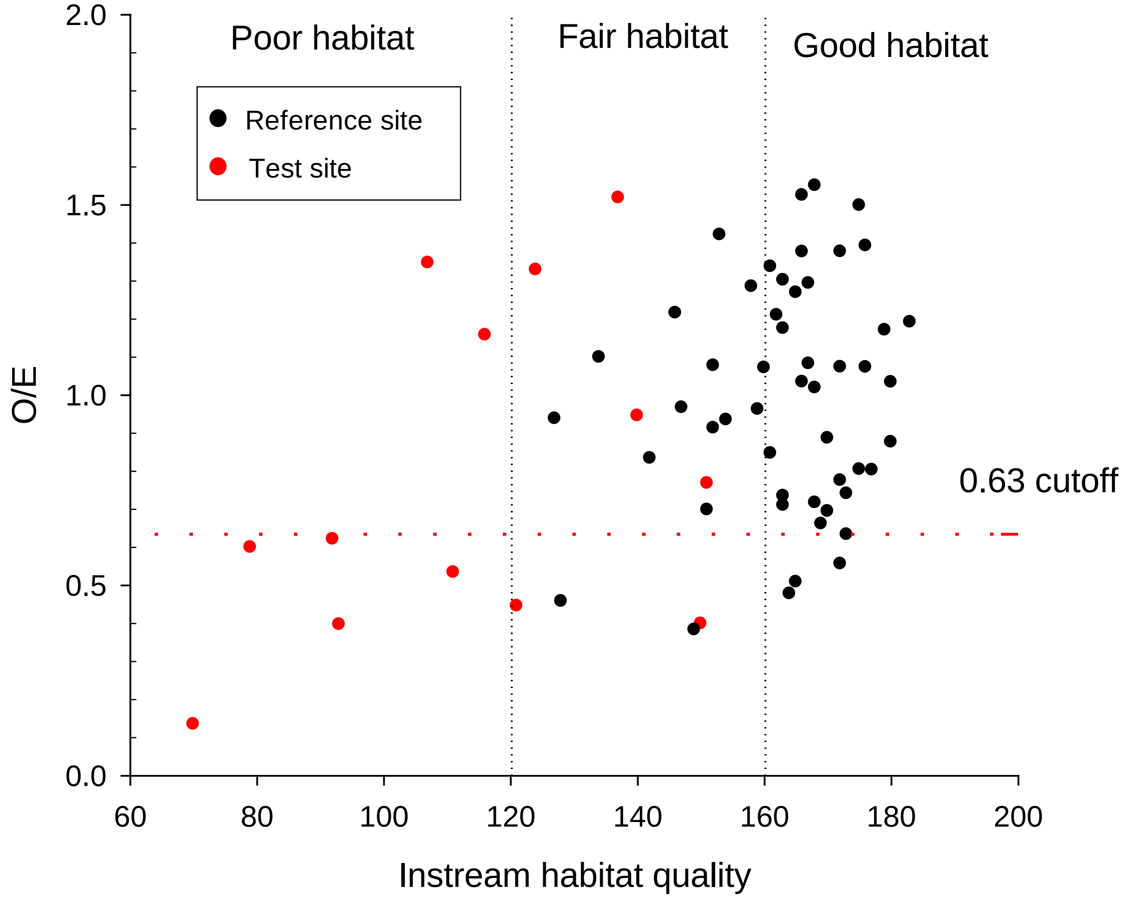


Figure 33. Observed/Expected (O/E) values generated with the Assessment by Nearest Neighbor Analysis (ANNA) method versus instream habitat quality determined by the SCDHEC instream habitat protocol. Data are shown for reference and test (i.e., disturbed) sites.

generalist species rather than the loss of reference site species. However, the ANNA method is highly sensitive to later stages of degradation as species characteristic of reference conditions decline in number.

5.2.9.2 Fish Multimetric index (MMI)

Model description – Karr (1981) and Karr et al (1986) proposed a method of using reference information in the Index of Biotic Integrity (IBI), which is a multimetric index (MMI). Metrics are community, population, and organism level measures that respond in a predictable fashion to disturbance. MMIs combine scores derived from several metrics into a single number that indicates the difference between an assessment site and conditions represented by reference sites or a reference model (Karr and Chu 1999). MMIs for fish have been implemented by many state agencies including the North Carolina Biological Assessment Unit (NCDENR 2006) and the Georgia Department of Natural Resources Stream Survey Team (Georgia Department of Natural Resources, GDNR, 2005). Both states have identified fish assemblage metrics that may be appropriate for the SH Ecoregion. The selection of metrics requires consideration of a number of issues including circularity, subjectivity, variability, colinearity, ecological relevance, and ability to discern between reference and impacted sites.

Model development – North Carolina uses a field procedure similar to the procedure we used, but Georgia uses a different methodology. In addition, both states sample streams that are larger than some of the streams we sampled. Finally, young of year (YOY) were included by us but not by North Carolina and Georgia. Therefore, it was important to choose only metrics appropriate for small to moderate-size streams in the SH Ecoregion. It was also important to identify naturally occurring factors such as stream size that could affect metric scoring criteria. Twenty-six coastal plain metrics used by state agencies were chosen as a starting point in creating a MMI specific to the SH Ecoregion (Table 27). Metrics were calculated following the Standard Operating Procedures (SOP's) of both North Carolina and Georgia. However, some fish were unassigned for metric calculation purposes because this study included fish community data from South Carolina as well as Georgia and North Carolina. Assignments for these fishes were established using various references (Table 28). Final fish assignments are shown in Table 29.

In addition to SC and GA state agency metrics, we also evaluated three other metrics for inclusion in the SH MMI. These included number of "invader" species defined as fishes likely to originate from artificial impoundments (e.g., bluegill, largemouth bass, and black crappie) as well as true alien species, percent abundance of invader species, and number of native species. The latter variable was equal to total species richness minus the number of invader species. These variables were included to account for the effects of biotic homogenization.

The following steps were used to identify and score metrics for potential inclusion in the SH MMI:

STEP 1: GLM procedures were used to identify metrics that were significantly related to disturbance at the sample sites in this study. Two GLM procedures using data from all sites were conducted on each metric. The dependent variable in both was the metric under study. Independent variables included location (e.g., DoD installation), stream width, and a categorical variable indicating whether a site was disturbed or reference in the first GLM and location, stream width, and the instream habitat assessment score in the second GLM. Stream width and location were included in both GLMs to account for variance associated with the two most important natural sources of variation in the study area (i.e., geography and stream size). Including these variables accounted for extraneous variance that would otherwise be subsumed into the error term and identified significant natural sources of variation for incorporation into the metric scoring criteria. The use of a categorical disturbance variable in one model and a continuous instream habitat assessment score in the other provided two ways to identify metrics that were potentially related to

Table 27. North Carolina and Georgia fish metrics considered for multimetric index development.

State	Metric Name	Used in RC-1694
NC	# Species Suckers	Yes
NC	# Species Darters	Yes
NC	# Intolerant Species	Yes
NC	# Species Sunfish	Yes
NC	# Species	Yes
NC	% Individuals Piscivores	Yes
NC	% Tolerant Individuals	Yes
NC	% Individuals Insectivores	Yes
GA-NC	% Individuals Omni-Herb Species	Yes
GA-NC	Abundance (200m)	Yes
GA	# Benthic Invertivore Species	Yes
GA	# Sensitive Species	Yes
GA	# Native Sunfish Species	Yes
GA	# Native Insectivorous Cyprinid Species	Yes
GA	# Native Species	Yes
GA	% Individuals *Lepomis*	Yes
GA	% Individuals Benthic Fluvial Specialist Species	Yes
GA	% Individuals Insectivorous Cyprinids	Yes
GA	# Native Round Bodied Sucker Species	Yes
GA	Evenness	Yes
NC	% Individuals Diseased Fish	Yes
NC	% Species with multiple Age Groups	No
GA	# Native Centrarchid Species	No – For large streams only
GA	# Intolerant Species	Yes
GA	% Individuals Piscivores	No – For large streams only
GA	% Individuals External Anomalies	Yes

Table 28. References used for categorical fish assignments (in order of use).

Assignment	Ref 1	Ref 2	Ref 3	Ref 4	Ref 5
Tolerance Level	NC SOP[a]	GA SOP[b]	EPA SOP[c]	Marcinek BPJ[d]	Freeman BPJ[e]
Feeding Group	NC SOP	GA SOP	Fishes of SC[f]	TX Website[g]	
Sensitivity	NC SOP	GA SOP*	EPA SOP	Marcinek BPJ	Freeman BPJ
Benthic Fluvial Specialist	GA SOP	Prusha BPJ			
Native Chattahoochee River	GA SOP	Warren	FOG[h]		
Native Savannah River	GA SOP	Warren	FOG		
Native Lumber River	Warren[i]	Prusha BPJ			

[a]NC SOP = NCDENR 2006
[b]GA SOP = GDNR 2005
[c]EPA SOP = Barbour et al 1999
[d]BPJ = Best Professional Judgment. P. Marcinek (GA DNR 2010), M. C. Freeman (USGS/UGA 2010),
[e]B. Prusha (UGA2010).
[f]Fishes of SC = Rhodes et al 2009
[g]TX Website = Hassan-Williams et al 2010
[h]FOG = Straight et al 2010
[i]Warren = Warren et. al. 2000.

Table 29. Final fish assignments for metric calculations.

Common Name	Scientific Name	Tolerance	*Lep omis*	Feeding Group	Dart er	Catost omid	Cypr inid	Round Bodied Sucker	GA Sensitive	Benthic Fluvial Specialist	Benthic Insectivore	GA SF Specie	NC SF Specie	Native (Chatt)	Native (Sav)	Native (Lum)
American Eel	*Anguilla rostrata*	Intermediate	no	Piscivore	no	no	no	no	no	no	no	no	no	yes	yes	yes
Blackbanded Darter	*Percina nigrofasciata*	Intolerant	no	Insectivore	yes	no	no	no	no	yes	yes	no	no	yes	yes	no
Blackbanded Sunfish	*Enneacanthus chaetodon*	Intermediate	no	Insectivore	no	no	no	no	no	no	no	yes	yes	no	yes	yes
Blackspotted Topminnow	*Fundulus olivaceus*	Intermediate	no	Insectivore	no	no	no	no	no	no	no	no	no	yes	no	no
Blue Spotted Sunfish	*Enneacanthus gloriosus*	Intermediate	no	Insectivore	no	no	no	no	no	no	no	yes	yes	no	yes	yes
Bluefin Stoneroller	*Campostoma pauciradii*	Tolerant	no	Herbivore	no	no	yes	no	no	no	no	no	no	yes	no	no
Bluegill	*Lepomis macochirus*	Intermediate	yes	Insectivore	no	no	no	no	no	no	no	yes	yes	yes	yes	yes
Bluehead Chub	*Nocomis leptcephalus*	Intermediate	no	Omnivore	no	no	yes	no	no	no	no	no	no	yes	yes	yes
Broadstripe Shiner	*Pteonotropis euryzonus*	Intermediate	no	Insectivore	no	no	yes	no	no	no	no	no	no	yes	no	no
Brook Silverside	*Labidesthes sicculus*	Intermediate	no	Insectivore	no	no	no	no	no	no	no	no	no	yes	yes	no
Chain Pickerell	*Esox niger*	Intermediate	no	Piscivore	no	no	no	no	no	no	no	no	no	yes	yes	yes
Creek Chub	*Semotilus atromaculatus*	Tolerant	no	Insectivore	no	no	yes	no	no	no	no	no	no	no	yes	yes
Creek Chubsucker	*Erimyzon oblongus*	Intermediate	no	Omnivore	no	yes	no	yes	no	yes	no	no	no	yes	yes	yes
Dixie Chub	*Semotilus thoreauianus*	Tolerant	no	Omnivore	no	no	yes	no	no	no	no	no	no	yes	no	no
Dollar Sunfish	*Lepomis marginatus*	Intermediate	yes	Insectivore	no	no	no	no	no	no	no	yes	yes	yes	yes	yes
Dusky Shiner	*Notropis cummingsae*	Intermediate	no	Insectivore	no	no	yes	no	no	no	no	no	no	yes	yes	yes
Eastern Mosquitofish	*Gambusia holbrooki*	Tolerant	no	Insectivore	no	no	no	no	no	no	no	no	no	yes	yes	yes
Eastern Mudminnow	*Umbra pygmaea*	Intermediate	no	Insectivore	no	no	no	no	no	no	no	no	no	no	yes	yes
Golden Shiner	*Notemigonus crysoleucas*	Tolerant	no	Omnivore	no	no	yes	no	no	no	no	no	no	yes	yes	yes
Goldstripe Darter	*Etheostoma parvipinne*	Intolerant	no	Insectivore	yes	no	no	no	no	yes	yes	no	no	yes	no	no
Green Sunfish	*Lepomis cyanellus*	Tolerant	yes	Insectivore	no	no	no	no	no	no	no	no	yes	no	no	no
Gulf Darter	*Etheostoma swaini*	Intolerant	no	Insectivore	yes	no	no	no	yes	yes	yes	no	no	yes	no	no
Lake Chubsucker	*Erimyzon sucetta*	Intermediate	no	Insectivore	no	yes	no	yes	yes	yes	no	no	no	yes	yes	yes
Largemouth Bass	*Micropterus salmoides*	Intermediate	no	Piscivore	no	no	no	no	no	no	no	no	no	yes	yes	yes
Longear Sunfish	*Lepomis megalotis*	Intolerant	yes	Insectivore	no	no	no	no	no	no	no	no	yes	no	no	no
Lowland Shiner	*Pteronotropis stonei*	Intermediate	no	Insectivore	no	no	yes	no	no	no	no	no	no	no	yes	no
Margined Madtom	*Noturus insignis*	Intermediate	no	Insectivore	no	no	no	no	yes	yes	yes	no	no	no	yes	yes
Mud Sunfish	*Acantharchus pomotis*	Intermediate	no	Insectivore	no	no	no	no	no	no	no	yes	yes	no	yes	yes
Northern Hogsucker	*Hypentilium nigricans*	Intermediate	no	Insectivore	no	yes	no	yes	no	yes	no	no	no	no	yes	yes
Pinewoods Darter	*Etheostoma mariae*	Intolerant	no	Insectivore	no	no	no	no	yes	yes	yes	no	no	no	no	yes
Pirate Perch	*Aphredoderus sayanus*	Intermediate	no	Insectivore	no	no	no	no	no	no	no	no	no	yes	yes	yes
Red-Black Spotted Sunfish Hybrid	*Lepomis miniatus X L. punctatus integrade*	Intermediate	yes	Insectivore	no	no	no	no	no	no	no	no	yes	yes	no	no
Redbreast Sunfish	*Lepomis auritus*	Tolerant	yes	Insectivore	no	no	no	no	no	no	no	yes	yes	yes	yes	yes
Redfin Pickerell	*Esox americanus americanus*	Intermediate	no	Piscivore	no	no	no	no	no	no	no	no	no	yes	yes	yes
Sandhills Chub	*Semotilus lumbee*	Intolerant	no	Insectivore	no	no	yes	no	yes	no	no	no	no	no	no	yes
Savannah Darter	*Etheostoma fricksium*	Intolerant	no	Insectivore	yes	no	no	no	no	yes	yes	no	no	no	yes	no
Sawcheeck Darter	*Etheostoma serrifer*	Intolerant	no	Insectivore	yes	no	no	no	no	yes	yes	no	no	no	yes	yes
Snail Bullhead	*Ameiurus brunneus*	Intermediate	no	Insectivore	no	no	no	no	no	no	no	no	no	yes	yes	yes
Southern Brook Lamprey	*Ichthyomyzon gagei*	Intolerant	no	Herbivore	no	no	no	no	no	no	no	no	no	yes	no	no
Speckled Madtom	*Noturus leptacanthus*	Intermediate	no	Insectivore	no	no	no	no	yes	yes	yes	no	no	yes	yes	no
Spotted Sucker	*Minytrema melanops*	Intermediate	no	Insectivore	no	yes	no	yes	no	yes	no	no	no	yes	yes	yes
Spotted Sunfish	*Lepomis punctatus*	Intermediate	yes	Insectivore	no	no	no	no	no	no	no	yes	yes	yes	yes	yes
Tadpole Madtom	*Noturus gyrinus*	Intermediate	no	Insectivore	no	no	no	no	yes	yes	no	no	no	yes	yes	yes
Tessellated Darter	*Etheostoma olmstedi*	Intermediate	no	Insectivore	yes	no	no	no	yes	yes	yes	no	no	no	yes	yes
Warmouth	*Lepomis gulosus*	Intermediate	yes	Insectivore	no	no	no	no	no	no	no	yes	yes	yes	yes	yes
Weed Shiner	*Notropis texanus*	Intolerant	no	Insectivore	no	no	yes	no	no	no	no	no	no	yes	no	no
Yellow Bullhead	*Ameiurus natalis*	Tolerant	no	Omnivore	no	no	no	no	no	no	no	no	no	yes	yes	yes
Yellowfin Shiner	*Notropis lutipinnis*	Intermediate	no	Insectivore	no	no	yes	no	no	no	no	no	no	no	yes	no

disturbance; hence, possibilities for inclusion in the MMI. Because the goal was to identify potential metrics for further consideration, we were more concerned with missing a potential metric (i.e., type two error) than with mistakenly identifying a metric with significant effects (i.e., type one error); therefore, no attempt was made to control for accumulating type one error associated with multiple tests. Metrics selected for potential inclusion in the MMI were those that exhibited a significant difference between disturbed and reference sites and/or that showed a significant relationship with the instream habitat quality assessment score. Significance was defined as $P \leq 0.10$ rather than $P \leq 0.05$ reflecting an emphasis on type two rather than type one error.

STEP 2: Assess the redundancy among the metrics selected in Step 1. Pearson correlations (r) were calculated among potential metrics to identify redundant (highly correlated) metric pairs. Only one member of highly correlated (i.e., $r > 0.8$) pairs was included in the MMI.

STEP 3: Develop scoring criteria for the metrics. Individual metrics in MMIs are usually normalized by assigning each a score of one (worst), three (mid-level), or five (best) based on metric scoring criteria. The final MMI value is computed by summing the scores for the individual metrics. In this study, metric scores were assigned so that about 10% or less of the reference sites received a score of one. Scoring criteria for five and three were assigned by approximately evenly dividing the remaining sites.

The GLM procedures identified 11 metrics that were significantly related to disturbance including number of cyprinid species, number of darter species, number of benthic fluvial specialist species, number of benthic insectivorous species, number of tolerant species, number of invader species, number of fish, percent abundance of tolerant fish, percent abundance of cyprinids, percent abundance of darters, percent abundance of invaders, and percent abundance of sunfishes. Number of benthic insectivorous species was strongly correlated with number of benthic fluvial specialist species and number of darter species, so was excluded. Number of fish differed marginally ($P = 0.09$) between disturbed and undisturbed sites but was highly variable so it, too, was excluded. Three metrics that were not significantly related were included in the MMI: total number of native species, percent with disease or anomalies, and percent hybrids. These metrics, which are often used in MMIs, were included because of their potential utility in quantifying disturbance at highly disturbed sites. Diseased fish and fish with anomalies were uncommon in our samples as were hybrids, with the exception of hybrids of *L. punctatus* x *L. miniatus* (spotted sunfish x redspotted sunfish).

Percent hybrids usually increases with environmental degradation. However, some hybrids occur naturally in high quality streams. This was the case with spotted sunfish x redspotted sunfish hybrid. These two species were formerly considered subspecies, and they naturally hybridize in some rivers in Georgia. We did not include this hybrid in a metric that is expected to increase with disturbance because it occurs naturally in high quality streams,. The spotted sunfish x redspotted sunfish hybrid occasionally occurred in the Fort Benning area and did not appear to be associated with disturbed conditions.

The final MMI included 12 metrics distributed among five metric categories including species richness, trophic guild composition, species composition, indicator species, and fish condition (Table 30). Most species richness metrics were adjusted for stream width to account for the positive relationship between these two variables as shown by the GLM procedures and other analyses (Section 5.2.2). Number of cyprinid species was adjusted for location to accommodate the relatively high number of cyprinid species on the SRS. Percent composition metrics were unaffected by stream size with the exception of percent sunfish, which was adjusted to account for the frequent absence of sunfish in small least disturbed streams. Percent sunfish also differed from the other percent composition metrics in that higher rather than lower percentages were associated with degradation, a pattern also observed in other studies of coastal plain streams (Paller et al. 1996 and 2000). Scoring criteria for most metrics were derived as described previously; however, criteria for percent hybrids and percent disease were taken from other studies (Paller et al. 1996). The highest possible MMI score was 60, with departures from this value indicating progressive degradation. The average MMI score for the reference sites was 47.6 (34-58). The cut-off MMI score to determine an impaired site was 42, which was about equal to the lower tenth percentile of the distribution of reference site MMI scores.

Table 30. Metrics and scoring criteria used in Sand Hills multimetric index for streams under 5 m average width. Individual metrics are assigned scores of one, three, or five.

	Scoring criteria		
Metrics	5	3	1
Species richness			
Number darter species			
Stream width >= 2m	>=2	1	0
Stream width < 2 m	>=1	0	
Number cyprinid species			
SRS streams	>=4	2-3	0-1
All streams except SRS	>=3	1-2	0
Number native species			
Stream width >= 2m	>=12	7-11	<=6
Stream width < 2 m	>=9	4-8	<=3
Trophic composition			
Number benthic fluvial specialist species			
Stream width >= 2m	>=4	2-3	0-1
Stream width < 2 m	>=2	1	0
Species composition			
Percent darters	>2	>0-2	0
Percent cyprinids	>10	>1-10	0
Percent sunfish			
Stream width >= 2m	>0 and <=15	>15-25	>25
Stream width < 2 m	<=15	>15-25	>25
Indicator species			
Percent tolerant fish	0-5	>5-20	>20
Percent invaders	0	>0-5	>5
Number invader species			
Stream width >= 2m	0	1-2	>=3
Stream width < 2 m	0	1	>=2
Fish condition			
Percent disease or anomalies	<2	2-5	>5
Percent hybrids	0	>0-2	>2

Model results – Nine of the 14 sites identified as disturbed by the PCA method had MMI scores under 42, and nine of the 56 sites classified as reference had an MMI score under 42. MMI values for all sites were compared with SCDHEC instream habitat assessment scores (Figure 34). Twenty percent of the sites with good habitat (i.e., habitat assessment scores ≥160) had MMI scores below 42, 22% of the sites with fair habitat (i.e., habitat assessment scores ≥120 and <160) had MMI scores below 42, and 71% of the sites with poor habitat had MMI scores below 42.

The MMI was tested with independent fish assemblage data collected from the SRS, Fort Benning, and Fort Bragg. Five sites were sampled at the SRS by one of the authors (M. Paller) and four were sampled at Fort Benning by W. Birkhead (Columbus State University). More extensive test data were provided for Fort Bragg by Charles Bryan, a fisheries biologist with the Fort Bragg Endangered Species Branch. This data set included 55 fish assemblage samples collected from 32 sites between 2002 and 2008 with methods comparable to those used in this study. Bryan also qualitatively rated the habitat at each site as “poor,” “fair,” “good,” or “excellent.” Box plots showed that the MMI averaged about 48 at the “excellent” sites, and that few sites in this category had MMI scores below the cutoff of 42. In contrast, the average MMI score at “poor” sites was 35, and most MMI scores were below the cutoff (Figure 35). “Fair” and “good” sites were intermediate with MMI scores averaging 41 and 38, respectively. These comparisons suggest that the MMI scores corresponded reasonably well with expectations of biotic integrity based on assessments of habitat and and extent of anthropogenic disturbance.

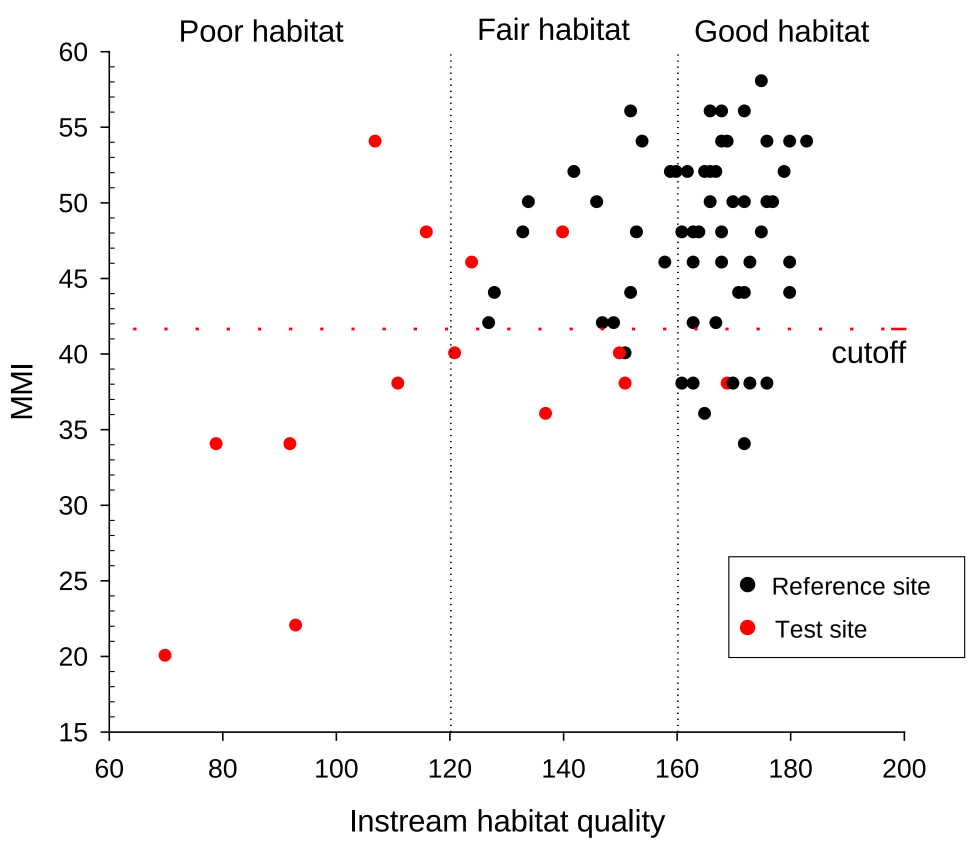


Figure 34. MMI values for reference sites and test (i.e., disturbed) sites versus instream habitat quality determined by the SCDHEC instream habitat protocol.

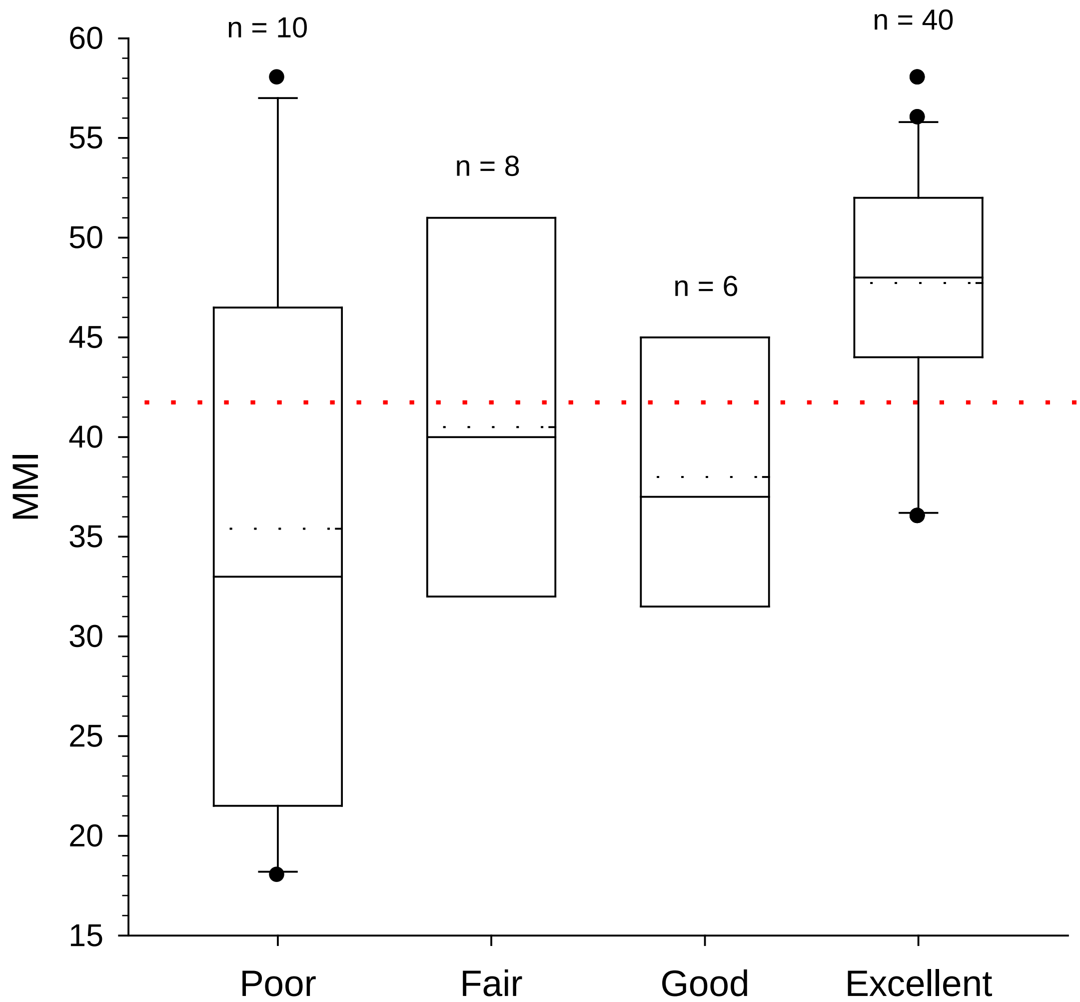


Figure 35. MMI values for stream sites with poor, fair, good, and excellent habitat quality as determined by expert professional judgment. Box boundaries indicate 25th and 75th percentiles, lines within the box indicate means (dashed) and medians (solid), and error bars indicate 90th and 10th percentiles. The latter are shown only when n≥9. Data from C. Bryan, Fort Bragg Endangered Species Branch.

5.2.4.3 Community Quality Index (CQI)

Model description – Current stream monitoring and assessment protocols may be insensitive to biotic homogenization because they ignore the endemism that characterizes fish assemblages in minimally disturbed headwater stream reaches. These assemblages lose their integrity when habitat degradation encourages invasion by generalist fish species characteristic of downstream reaches or man-made ponds and reservoirs. The invasions result in an initial increase in number of species followed by a decrease as degradation worsens and endemic species are extirpated and replaced by generalist species. Assessment protocols are needed that emphasize the early stages of this process rather than the reduction in number of species and related changes characteristic of later stages of degradation.

Most assessment methods for streams do not explicitly address the problem of biotic homogenization and are largely insensitive to the early stage of this phenomenon because, rather than involving the loss of species, it involves their addition. Furthermore, the added species are not necessarily non-native species (e.g., carp *cyprinus carpio*) but native generalist species that immigrate from farther downstream or that emigrate from impoundments within the watershed. Another complicating factor is that downstream generalists may occasionally move upstream as a result of natural disturbances, such as high water levels, rather than anthropogenic habitat degradation. The Community Quality Index (CQI), an original assessment framework that is sensitive to biotic homogenization as well as more advanced stages of habitat degradation, was developed to address this issue. As with the other assessment frameworks described herein, this methodology was developed for small streams (1-5 m average width) within the SH study area.

Model development – The CQI is based on species composition data and species richness/stream width curves from the reference sites. It can be computed for a test site as follows:

STEP 1. Determine the number of native species expected at the site assuming it is least disturbed (i.e., in reference condition) based on its geographic location (specific DoD or DOE installation under study in this case), stream size, and connectivity with (proximity to) a larger stream. This can be visually estimated from species number/stream size curves (or computed from the regression equations for these curves) derived from reference site data (Figure 36)

STEP 2. Create a master list of the species that occur within the study area (specific DoD or DOE installation under study in this case) and divide them into the following categories:

- Core species: species commonly found in minimally disturbed (i.e., reference) streams
- Secondary species: species that may be present in low numbers in minimally disturbed streams
- Itinerants: species that are atypical of minimally disturbed streams and may have entered the stream from contiguous habitats such as wetlands or beaver ponds
- Invaders: species that are nonnative or characteristic of man-made habitats such as farm ponds and reservoirs

Lists of species within each of the preceding categories were derived from the reference site data for each installation (Table 31). Core species were species that were found at 50% of more of the reference sites from an installation and had an average relative abundance of 5% across all installation reference sites. An exception was made for the lowland shiner at the SRS, which was considered a core species even though it occurred at only two out of 16 reference sites. This species was a significant part of the assemblage at these sites (26 -41% relative abundance) where it appeared to replace the yellowfin shiner in low gradient, sandy reaches without significant gravel (which is needed by the latter species for successful reproduction, Marcy et al. 2005). Secondary species were the remaining species that occurred at the reference sites with lower frequencies and lower average relative abundance, usually under 2.0%. Invaders were species

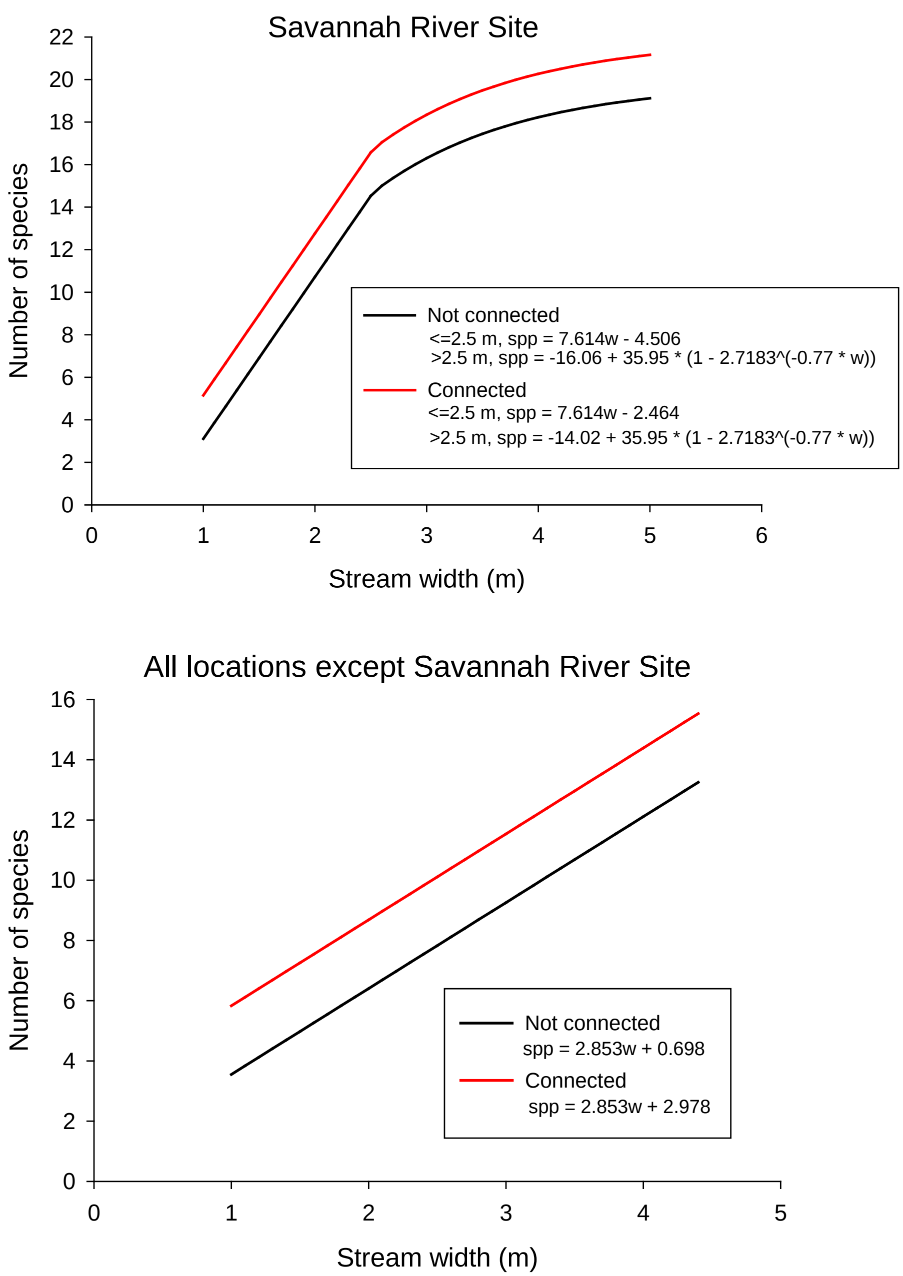


Figure 36. Relationships between species number and stream width at sample sites in the study area. Connected sample sites were located near the confluence with a larger (i.e., at least one order greater) stream. A sample site in a small stream ($\leq$ 1.5 m average width) was considered to be connected with a larger stream if less than 250 m from the confluence with the larger stream. Sample sites in larger streams (>1.5 m average width) were considered connected if within 750 m of a larger stream.

Table 31. Species occurring within the study area divided into ecological categories (C=core, S=secondary, I=itinerant, V=invader, see text for explanation).

SRS and Fort Gordon		Fort Bragg		Fort Benning	
Species	Category	Species	Category	Species	Category
Bluehead chub	C	Dusky shiner	C	Southern brook lamprey	C
Creek chub	C	Sandhills chub	C	Broadstripe shiner	C
Lowland shiner	C	Pirate perch	C	Dixie chub	C
Pirate perch	C	Margined madtom	C	Pirate perch	C
Yellowfin shiner	C	Yellow bullhead	C	Goldstripe darter	S
Blackbanded darter	S	Redfin pickerel	S	Redfin pickerel	S
Creek chubsucker	S	Tesselated darter	S	Blackbanded darter	S
Dollar sunfish	S	Mud sunfish	S	Yellow bullhead	S
Dusky shiner	S	American eel	S	Speckled madtom	S
American eel	S	Redbreast sunfish	S	Red-Black spotted sunfish hybrid	S
Eastern mosquitofish	S	Dollar sunfish	S	Redbreast sunfish	S
Eastern mudminnow	S	Creek chubsucker	S	Warmouth	S
Margined madtom	S	Chain pickerel	S	Weed shiner	S
Mud sunfish	S	Blackbanded darter	S	Spotted sunfish	S
Northern hogsucker	S	Bluegill	S	Creek chubsucker	S
Redbreast sunfish	S	Lake chubsucker	S	Dollar sunfish	S
Redfin pickerel	S	Pinewoods darter	S	Longear sunfish	I
Savannah darter	S	Tadpole madtom	S	Gulf darter	I
Speckled madtom	S	Warmouth	S	Lake chubsucker	I
Spotted sunfish	S	Sawcheek darter	S	Bluegill	V
Tesselated darter	S	Eastern mosquitofish	S		
Tadpole madtom	S	Pumpkinseed	I		
Warmouth	S	Blue spotted sunfish	I		
Yellow bullhead	S	Eastern mudminnow	I		
Flat bullhead	S	Largemouth bass	V		
Blackbanded sunfish	I	Bluegill	V		
Blue Spotted sunfish	I				
Brook silverside	I				
Spotted sucker	I				
Largemouth bass	V				
Bluegill	V				

such as bluegill, largemouth bass, black crappie, and true alien species (e.g., carp *cyprinus carpio*) that are not normally found in small, least disturbed streams unless an impoundment is present in the watershed or the habitat is highly disturbed. Itinerants were species that are atypical of reference streams although they may be found in wetland habitats or larger streams that are contiguous with small streams. Discriminating itinerant species from secondary species is based on reference site data, knowledge of the autecology of stream fishes, professional judgment, and local experience. In this respect, the CQI provides a framework for the application of expert knowledge, which imparts a degree of subjectivity but also adaptability and flexibility.

STEP 3. Compare the master list for the study area with the list of species at the sample site and score as follows:

- Core species = 1
- Secondary species = 1if <15% relative abundance, 0 if $\geq$15% and $\leq$25% relative abundance, and -1 if > 25% relative abundance
- Itinerant species = 0 if <15% relative abundance and -1 if $\geq$15% relative abundance
- Invaders = -1 if present, regardless of abundance

Core species are given a score of one regardless of abundance because they are found in most reference streams and are often very abundant within these streams. Secondary species, which are often found in low numbers at reference sites, are given a score of one if present in expected relative abundances (<15% based on reference site data). However, the score is reduced for higher relative abundances, which suggest that habitat at the test site is incongruent with reference site conditions. This could occur because the test site is anthropogenically disturbed or because the test site is atypical of the stream sites used to develop the reference site criteria (e.g., it is intermittent or affected by beaver impoundments). Itinerant species, which are not typically found at reference sites but cannot be clearly linked with anthropogenic disturbance, are given a score of zero if found in low numbers as might occasionally be expected; e.g., due to occasional upstream immigration from downstream reaches. However, they are given a negative score if present in higher numbers that likely reflect biotic homogenization. Invaders, being strongly indicative of anthropogenic disturbance, are always assigned a score of negative one regardless of abundance.

STEP 4. Sum the scores for all species and divide the sum by the expected number of native species determined in STEP 1 to calculate the CQI. Summing the scores gives credit for species expected under reference conditions, subtracts for species indicative of anthropogenic disturbance, and places the remaining species in a neutral category. Division of the sum by the number of species expected under reference conditions normalizes the results to a scale of one. The cut-off CQI value to determine an impaired site was 0.75, which was equal to the lower tenth percentile of the distribution of reference site CQI scores. Sample CQI computations for a stream site from the SRS are shown for illustration (Table 32).

<u>Model results</u> – CQI values were calculated for the sample sites at Fort Bragg, Fort Benning, Fort Gordon, and the SRS and compared with SCDHEC instream habitat assessment scores (Figure 37). Only 7% of the sites with good habitat (i.e., habitat assessment scores $\geq$160) had a CQI score below 0.75. In contrast, 32% of the sites with fair habitat (i.e., habitat assessment scores $\geq$120 and <160) had CQI scores below 0.75, and 86% of the sites with poor habitat had CQI scores below 0.75.

The CQI method was tested further in two ways. First, paralleling the method used with the ANNA model, CQI scores were tested for sites sampled in this study that were classified as disturbed (using the PCA gradient approach). None of these sites were used in the development of the CQI, thereby avoiding the circularity of testing the method with data used to develop it. Success of the model would be

Table 32. Community Quality Index (CQI) computations.

Site: SRS - Mill Creek tributary (mc6)				
Stream width (m)		1.5		
Distance to larger stream (m)		650		
SRS species list	SRS species category	Site species list	Site species percent abundance	Site species tally
Bluehead chub	C	Bluehead chub	3.4	1
Creek chub	C	Creek chub	4.4	1
Lowland shiner	C	Redbreast sunfish	0.5	1
Pirate perch	C	Pirate perch	10.3	1
Yellowfin shiner	C	Tadpole madtom	1.0	1
Blackbanded darter	S	Yellow bullhead	0.5	1
Creek chubsucker	S	Yellowfin shiner	79.9	1
Dollar sunfish	S			
Dusky shiner	S			
American eel	S			
Flat bullhead	S			
Eastern mosquitofish	S			
Eastern mudminnow	S			
Margined madtom	S			
Mud sunfish	S			
Northern hogsucker	S			
Redbreast sunfish	S			
Redfin pickerel	S			
Savannah darter	S			
Snail bullhead	S			
Speckled madtom	S			
Spotted sunfish	S	SUM		7
Tesselated darter	S			
Tadpole madtom	S	Expected number of native species		
Warmouth	S	in connected stream reach		
Yellow bullhead	S	Spp = 7.614w - 2.464		
Blackbanded sunfish	I			
Brook silverside	I	Expected number of native species		
Bluespotted sunfish	I	in unconnected stream reach		
Golden shiner	I	Spp = 7.614w - 4.506		7.2
Spotted sucker	I			
Bluegill	V	Community Quality Index		
Largemouth bass	V	(Sum/Expected)		0.97
C (core species): Score "1" regardless of abundance				
S (secondary species): Score "1" if under 15%, 0 if >= 15% and <=25%, and -1 if >25%				
I (itinerants): Score 0 if <15%, "-1" if >=15%				
V (invaders): Score "-1" regardless of abundance				
Connected if width <=1.5 m and <250 m from larger stream or if width >1.5m and <750 m from larger stream				

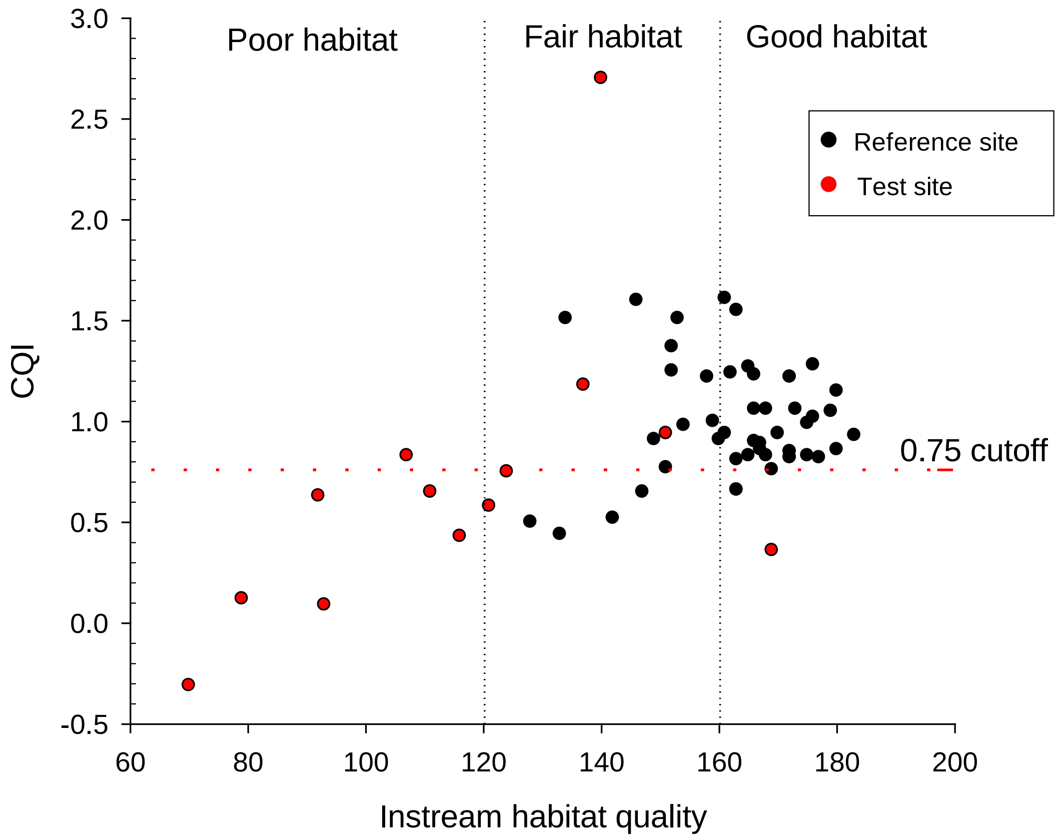


Figure 37. Community Quality Index (CQI) values based on fish assemblage structure plotted against instream habitat quality determined by the SCDHEC instream habitat protocol.

indicated by CQI values <0.75 for most or all of the disturbed sites. Of the 13 disturbed sites included in this analysis, 8 (62%)had O/E ratios <0.75 indicating biotic impairment. One disturbed site in the Carolina Sandhills National Wildlife Refuge (sn-bbct) was excluded from this analysis because there was insufficient fish assemblage data from this geographic area to compute the CQI.

The CQI was also tested with independent fish assemblage data collected from the SRS, Fort Benning, and Fort Bragg using methods similar to those in this study. Five sites were sampled at the SRS, and 4 were sampled at Fort Benning. A more extensive test data set was provided for Fort Bragg by Charles Bryan, a fisheries biologist with the Fort Bragg Endangered Species Branch. This data set included 55 fish assemblage samples collected from 32 sites between 2002 and 2008 with methods comparable to those of this study. Bryan also qualitatively rated the habitat at each site as “poor,” “fair,” “good,” or “excellent.” A box plot showed that CQI scores progressively increased as habitat quality increased from “poor” to “excellent” (Figure 38) and that ~75% of the CQI scores from sites rates as “excellent” exceeded the 0.75 cut-off between impaired and non-impaired sites. These comparisons indicated that the

CQI scores corresponded well with expectations of biotic integrity based on assessments of habitat and accurately reflected the effects of anthropogenic disturbance on fish assemblage structure.

CQI scores corresponded with expectations of biotic integrity across a gradient of disturbance, so they were compared with various disturbance variables to identify those with the greatest influence on fish assemblages. Pearson correlations indicated 5 disturbance variables with highly significant effects on the CQI: erosion score (r = -0.53, P<0.001), average bank height (r = -0.53, p=0.002), watershed disturbance index (r = -0.58, p<0.001), instream habitat quality score (r = 0.59, p<0.001), and watershed coverage by evergreen and/or deciduous forest (r = 0.53, p<0.001). Cook's distance values for all data points in all regressions were under 1.0 indicating an absence of influential data points. Paved road density and extent of channel modification were also significantly related to CQI scores, but the former was highly correlated with the easier to calculate watershed disturbance index and the latter was strongly influenced by a single data point (BR-tank). Examination of plots of these disturbance variables versus CQI scores can provide insights into levels of disturbance likely to be associated with impaired fish assemblages (Figure 39). For example, instream habitat quality scores <140 were typically associated with CQI scores indicative of biotic degradation (i.e., <0.75).

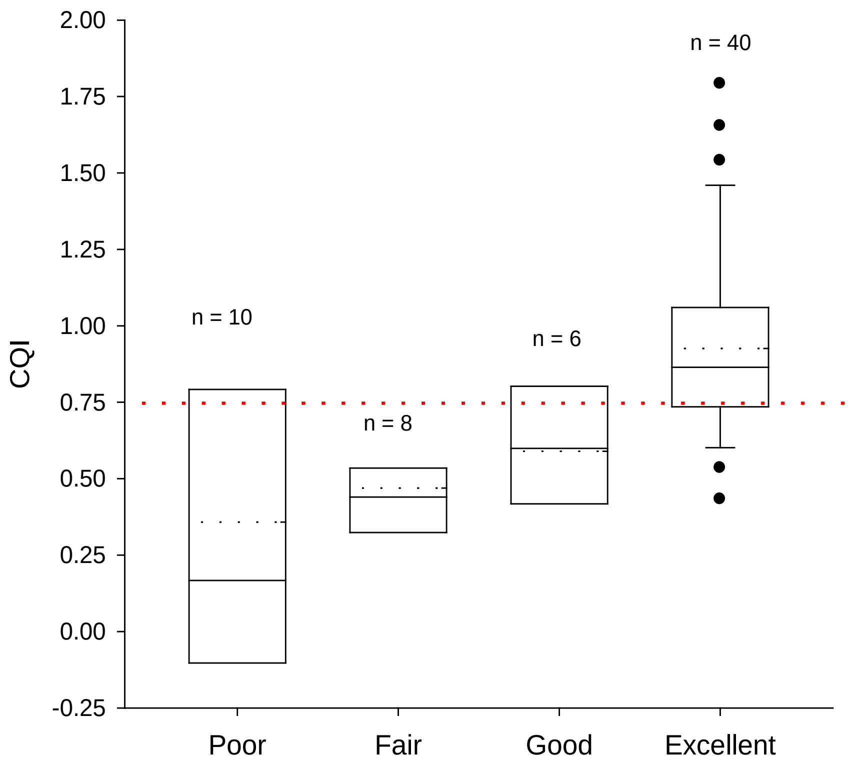


Figure 38. Community Quality Index (CQI) values for stream sites with poor, fair, good, and excellent habitat quality as determined by expert professional judgment. Box boundaries indicate 25th and 75th percentiles, lines within the box indicate means (dashed) and medians (solid), and error bars indicate 90th and 10th percentiles. The latter are shown only when n≥9. Data from C. Bryan, Fort Bragg Endangered Species Branch.

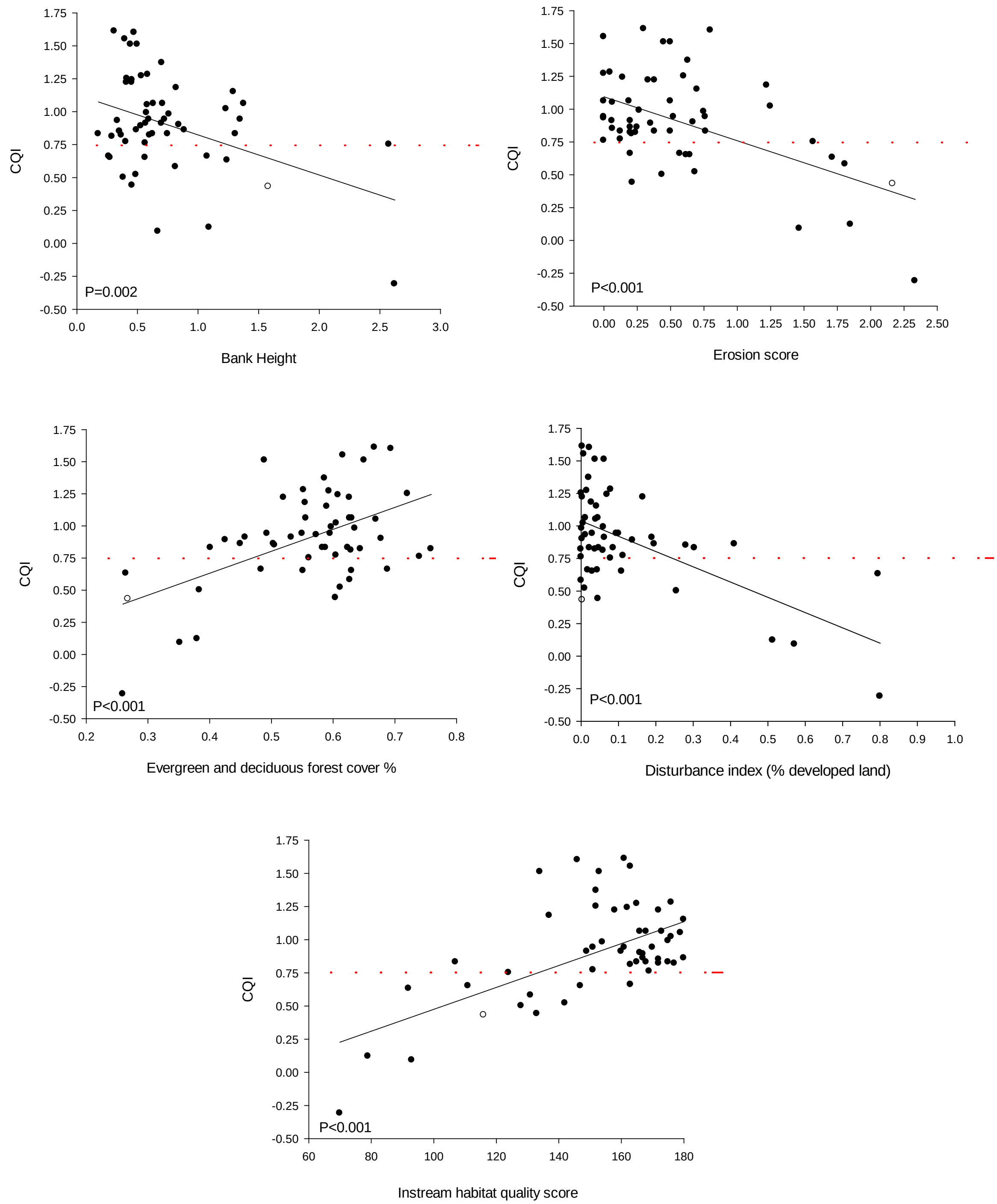


Figure 39. Relationships between Community Quality Index (CQI) values and key environmental variables.

5.2.4.4 Summary for Fish Models

The 3 fish assessment frameworks were designed for small (<5 m average width) streams in the SH Ecoregion sampled with 2-pass backpack electrofishing over a reach length of ~200 m. All 3 methods were highly sensitive to advanced stages of degradation as represented by sites such as Br-tank and Gr-mcoy. The ANNA model was the most objective and mathematically rigorous of the 3 fish assessment frameworks but was less sensitive to moderate levels of biotic degradation. The CQI was more sensitive but was more subjective and required expert judgment to classify species into ecological categories. This step, which requires an understanding of species-specific ecology, is not necessarily problematic because it necessitates careful consideration of the data. The MMI incorporated some of the strengths and weaknesses of both methods. Although not as objective as ANNA, MMIs are well accepted in the US where they have been evaluated extensively (Fore et al. 1996). The incorporation of metrics designed to reflect biotic homogenization (i.e., percent invaders, number of invader species, and number of native species) increased the MMI's sensitivity to the early stages of biotic degradation. The performance of the three assessment frameworks is analyzed in detail in Section 5.5.

It will often be useful to combine the information from biotic assessment frameworks with information concerning key environmental features at landscape, watershed, and instream scales (Section 5.2.3.2), especially when attempting to assess the feasibility of recovery to the reference condition. Recovery at each level in this hierarchy is feasible if levels above it meet thresholds for minimal levels of disturbance. For example, if key instream habitat variables meet values for minimal disturbance, recovery of the biota to reference expectations is feasible assuming access to source pools for locally extirpated species and the absence of established populations of invasive species. Conversely, recovery of the biota to the reference condition would be difficult or impossible if instream habitat conditions fell below reference site thresholds. Moving up the hierarchy, recovery of degraded instream habitat is influenced by conditions in the watershed. Correction of watershed-level problems would create the potential for the recovery of instream habitat, although such changes could take considerable time (multiple generations) unless expedited by habitat recovery programs. Generally, limiting levels in the hierarchy of habitat scales can be examined for key environmental features with low values, and these can be targeted for recovery actions that will result in maximum benefits.

In some cases, the reference condition will not be attainable because of environmental disturbances that are largely irreversible over a pragmatic time-scale. For example, the occurrence of "invader species" that originate from an upstream impoundment represent a departure from the reference condition that is difficult to correct. In such cases, the reference model defined as the least disturbed condition is unachievable and should be replaced by a best-attainable model (BAM). Best-attainable conditions exist when and where the impact on biota of inevitable land use has been minimized to the greatest extent possible by best management practices (Stoddard et al. 2006a). BAMs are likely to vary with the type of impact that is causing the biota to depart from reference expectations. For example, a BAM for a site near an impoundment should account for the inevitable occurrence of lentic invader species. Other cases may be less straightforward. Poor biotic assessment scores in a watershed with extensive anthropogenic development necessitate careful scrutiny for management improvements that could facilitate biotic recovery. In such cases, a BAM based on current biotic conditions, may represent a benchmark from which to gage future changes. Such site- and time-specific BAMs should document current conditions, indicate how current conditions differ from the reference condition, and identify barriers to recovery. Specific protocols for developing BAMs for macroinvertebrate assessment frameworks are provided in Section 5.3.2 of this report.

5.3 Macroinvertebrate Assemblages

We sampled benthic macroinvertebrates throughout the SH Ecoregion of the Southeastern Plains (Table 1). Sample sites were well distributed across each of the states within the study area with 26, 27, and 21 sampling events taken in NC, SC, and GA, respectively (Table 33).

5.3.1 *Macroinvertebrate Assemblage Structure*

The Class Insecta was the most abundant group of macroinvertebrates encountered in our collections (Table 34). A total of 268 unique macroinvertebrate taxa were identified from the sampling sites (Table 35). Many taxa were ubiquitous in their distribution; however, some taxa were associated with specific installations, regions, or stream sizes (Table 35). Streams with the highest species richness were generally located in the center of our study area; the 5 sites with the highest taxa richness samples were from the SRS (tc06=102, mb06=99, tink=97, mcms=95; Table 1) and Fort Gordon (bbbb=95; Table 1Figure 40). Macroinvertebrates from the orders Ephemeroptera, Plecoptera, and Trichoptera (EPT) are considered sensitive taxa (Lenat, 1993; Barbour et al., 1999). In general, EPT richness was low and variables across

Table 33. Total number of macroinvertebrate sampling events from each site.

	Fort Gordon (GA)	Fort Benning (GA)	The Nature Conservancy (GA)	Fort Bragg (NC)	Sandhills Gamelands (NC)	Manchester State Forest (SC)	Sandhills State Forest (SC)	Sandhills National Wildlife Refuge (SC)	Savannah River Site (SC)
Sampling events	6	12	3	20	6	2	2	4	19

Table 34. Macroinvertebrate classes collected from localities in the Sand Hills Ecoregion. Total sample abundances reflect all sampling events within each location.

Class	Fort Gordon (GA)	Fort Benning (GA)	The Nature Conservancy (GA)	Fort Bragg (NC)	Sandhills Gamelands (NC)	Manchester State Forest (SC)	Sand Hills State Forest (SC)	Sandhills National Wildlife Refuge (SC)	Savannah River Site (SC)
Arachnida	3008	4235	804	6878	2317	1031	143	1657	13449
Bivalvia	18	133		399	10	20	88		1337
Branchiopoda	347	4253		461	3753				3547
Clitellata	1157	2339	464	5677	2209	498	806	890	3036
Gastropoda	80	101			247	60			509
Insecta	62632	107675	32282	150377	51222	15480	15197	82180	264981
Malacostraca	20	79	1	689	83	20	7	90	321
Maxillopoda	3699	9844	3882	7426	2806	2890	310	11057	30338
Ostracoda	203	993		1681	460	50	1710		2851
Turbellaria									20

Table 35. Taxonomic orders, families, and operational taxonomic units (OTUs). Sample abundances reflect estimates from all specimens identified from that site throughout the study period.

Order	Family	OTUs	Fort Bragg (NC)	Sandhills Gamelands (NC)	Manchester State Forest (SC)	Sand Hills State Forest (SC)	Sandhills National Wildlife Refuge (SC)	Savannah River Site (SC)	Fort Gordon (GA)	Fort Benning (GA)	The Nature Conservancy (GA)
Trombidiformes		Acari	6473	2317	1031	143	1617	13369	3008	4227	797
		Acari A	405				40	80		8	7
Veneroida	Corbiculidae	*Corbicula fluminea*						17			
	Pisidiidae		399	10	20	88		1320	18	133	
Diplostraca		Cladocera	461	3753				3547	347	4253	
Oligochaeta			5677	2209	498	806	890	3036	1157	2339	464
Architaenioglossa	Viviparidae	*Campeloma*						87			
Basommatophora	Ancylidae	*Ferrissia*		247	60			335	80	100	
	Planorbidae	*Gyraulus*						80		1	
		Menetus								1	
Neotaenioglossa	Pleuroceridae	*Elimia*						7			
Coleoptera	Dryopidae	*Helichus basalis*						2			
		Helichus fastigiatus						3			
		Helichus lithophilus						1			
	Dytiscidae	*Neoporus*								22	
		unidentified	23					147	1	24	
	Elmidae	*Ancyronyx variegatus*	428	380	31	30		387	43		

Table 35. continued

Order	Family	OTUs	Fort Bragg (NC)	Sandhills Gamelands (NC)	Manchester State Forest (SC)	Sandhills State Forest (SC)	Sandhills National Wildlife Refuge (SC)	Savannah River Site (SC)	Fort Gordon (GA)	Fort Benning (GA)	The Nature Conservancy (GA)
		Dubiraphia	11		121	22		322			
		Gonielmis dietrichi	963	742	20	80	169	81	287		21
		Macronychus glabratus	3					345			
		Microcylloepus pusillus	392	60				230			
		Oulimnius		210	20			1067			
		Stenelmis	7159	1909	147	1537	875	2180	1726	1127	446
	Gyrinidae	*Dineutus*	14	1	8	4		6	15	16	
		Gyrinus	4					1			
	Hydraenidae	*Hydraena*	25	31							
	Hydrophilidae			1				4		1	
	Psephenidae	*Ectopria*						51		12	
	Ptilodactylidae	*Anchytarsus bicolor*	320	5			1	88	10	21	8
Diptera		Cyclorrhapha	72	10			30	130	40	74	
		Orthorrhaphous	12					21	20	10	
	Ceratopogonidae	*Atrichopogon*	20	30					10	21	
		Bezzia/Palpomyia Complex	1602	577	140	117	505	2852	455	1169	193
		Ceratopogon	75	77			11	247	21	159	
		Culicoides	124	90			10	360	20	399	7
		Forcipomyia	13							20	
		Probezzia	222	44	10	7	153	376	87	457	33
		Sphaeromias					1	84			
		Ceratopogoninae	830	127	160	80	229	2061	784	303	288

Table 35. continued

Order	Family	OTUs	Fort Bragg (NC)	Sandhills Gamelands (NC)	Manchester State Forest (SC)	Sandhills State Forest (SC)	Sandhills National Wildlife Refuge (SC)	Savannah River Site (SC)	Fort Gordon (GA)	Fort Benning (GA)	The Nature Conservancy (GA)
		Ceratopogoninae A	18	30	40		30	30	57	30	20
	Chaoboridae							10			
	Chironomidae	*Ablabesmyia annulata*								1	
		Ablabesmyia hauberi	194	109		20			3	52	
		Ablabesmyia mallochi	1538	431	187	399	43	990	820	1059	142
		Ablabesmyia monilis/rhamphe group	673	147	133	238	1	185	315	474	40
		Alotanypus aris	3							1	
		Antillocladius								5	
		Apsectrotanypus johnsoni	503	374	252	20	1189	872	203	1706	60
		Brillia		4				573		1	
		Brundiniella eumorpha						2			
		Cantopelopia gesta								1	
		Chaetocladius ligni	1								
		Chironomus	104	172				278		20	
		Cladotanytarsus D						117			
		Cladotanytarsus daviesi	1717	30		133		362	197	1451	1963
		Cladotanytarsus F						20			
		Cladotanytarsus I	1958	604	70		260	419	40	55	
		Clinotanypus	87	73	2	3	24	189	2	107	
		Constempellina						20			40
		Corynoneura	3641	713	380	379	1847	28962	1579	2863	1019
		Cricotopus/Orthocladius complex	51					111		44	
		Cryptochironomus	358	184	31	70	51	280	52	143	41
		Cryptotendipes							4	99	

Table 35. continued

Order	Family	OTUs	Fort Bragg (NC)	Sandhills Gamelands (NC)	Manchester State Forest (SC)	Sandhills State Forest (SC)	Sandhills National Wildlife Refuge (SC)	Savannah River Site (SC)	Fort Gordon (GA)	Fort Benning (GA)	The Nature Conservancy (GA)
		Demicryptochironomus	50	1	1			58	1	7	1
		Dicrotendipes	33							7	
		Djalmabatista pulchra variant	158	64				3	100	193	
		Gillotia alboviridus	1								
		Guttipelopia guttipennis		16							
		Gymnometriocnemus	31								
		Harnischia complex unidentified	20					55	7		
		Heterotrissocladius marcidus	623	2	32	7	1164	695	109	197	176
		Hyporhygma quadripunctatus								1	
		Kloosia dorsenna	133					64	4	7	
		Krenosmittia	123	10		1	240	350	11	111	
		Labrundinia becki								20	
		Labrundinia becki/virescens	141	181	10	60		22	97	74	
		Labrundinia pilosella	680	15	181	453	351	716	122	184	20
		Larsia	8							2	
		Limnophyes	16					200	20	130	
		Lopescladius	353					53	100	44	
		Mesosmittia								20	
		Microchironomus	10								
		Micropsectra						97			
		Microtendipes pedellus group	317	337	22	76	694	5503	181	4024	20
		Microtendipes rydalensis group	1929	91	20	71	103	958	389		
		Monopelopia		30				40			
		Nanocladius balticus group	10	17		13		61	10		
		Nanocladius cf. *crassicornus/rectinervis*	693	100	70	173	20	555	60	80	7

Table 35. continued

Order	Family	OTUs	Fort Bragg (NC)	Sandhills Gamelands (NC)	Manchester State Forest (SC)	Sandhills State Forest (SC)	Sandhills National Wildlife Refuge (SC)	Savannah River Site (SC)	Fort Gordon (GA)	Fort Benning (GA)	The Nature Conservancy (GA)
		Nanocladius D						10		17	
		Nanocladius spiniplenus						4			
		Natarsia	16		2			503		8	
		Neozavrelia						20			
		Nilotanypus americanus	1359	1298	300	7	878	703	70	188	60
		Nilotanypus fimbriatus						2873			
		Nilothauma	77	51	10		260	149	190	478	20
		Orthocladius annectens	96			13	133	669	20	20	1
		Orthocladius lignicola	191	97	40		37	1189	160	35	40
		Pagastiella	50				10	115	20	40	
		Parachaetocladius abnobaeus	1812	21	3	20	544	238	151	262	68
		Parachironomus	85								
		Paracladopelma unidentified						40	10		9
		Paracladopelma doris	45					7			
		Paracladopelma undine	125	193	110		40	197		57	
		Parakiefferiella	277		10		10	409	80	320	
		Paralauterborniella nigrohalterale	437	90		27		749	80	35	7
		Paramerina	342	308	101	60	338	2026	262	493	167
		Parametriocnemus	2669	6044	1	220	434	27774	751	1916	347
		Paraphaenocladius	443					227	13	146	
		Parasmittia carinata						4			
		Paratanytarsus	10					10	7		
		Paratendipes albimanus	20	10				947	7	52	
		Paratendipes basidens							7		
		Paratendipes subaequalis	24		10			35	53	64	
		Pentaneura inconspicua	443		20	695	40	1	761		

Table 35. continued

Order	Family	OTUs	Fort Bragg (NC)	Sandhills Gamelands (NC)	Manchester State Forest (SC)	Sandhills State Forest (SC)	Sandhills National Wildlife Refuge (SC)	Savannah River Site (SC)	Fort Gordon (GA)	Fort Benning (GA)	The Nature Conservancy (GA)
		Phaenopsectra	57					233			
		Polypedilum aviceps/flavum grp	881	1073	54	406	935	7453	819	1290	230
		Polypedilum braseniae						20			
		Polypedilum fallax group	64	10				360	56	22	20
		Polypedilum halterale group	1384	257	80	67	77	893	240	1068	9
		Polypedilum illinoense group	1276	105	2	129	471	3116	779	848	173
		Polypedilum laetum						120		7	
		Polypedilum scalaenum group	2561	195	261	330	581	494	403	524	172
		Polypedilum tritum	9	2	51			442	27	74	160
		Potthastia longimana	111	20			30	40	10	10	
		Procladius	1	10	70	1		287	22	75	
		Psectrocladius psilopterus group	82	44				42	7	903	
		Pseudochironomus	151					12	30		
		Pseudosmittia	7		20				7	47	
		Psilometriocnemus triannulatus								7	
		Rheocricotopus robacki	2668	763		337	34	4351	997	360	29
		Rheocricotopus tuberculatus	2410	576	131	70	3611	1994	593	1426	584
		Rheosmittia arcuata	1185	933	150		360	4228	35	463	107
		Rheotanytarsus	14300	3274	300	678	864	15975	2112	1404	499
		Robackia demeijerei	261				10	33	118		
		Saetheria hirta							54		
		Saetheria sp. 1						69	7		
		Saetheria tylus	197					157	22	10	
		Stelechomyia perpulchra	73			20		29		3	
		Stempellina A	197				6000	269	267	157	13

Table 35. continued

Order	Family	OTUs	Fort Bragg (NC)	Sandhills Gamelands (NC)	Manchester State Forest (SC)	Sandhills State Forest (SC)	Sandhills National Wildlife Refuge (SC)	Savannah River Site (SC)	Fort Gordon (GA)	Fort Benning (GA)	The Nature Conservancy (GA)
		Stempellina C	1676	580		10	247	2214	263	273	28
		Stempellinella A	3142	2127	300	347	2911	6367	1371	4509	780
		Stempellinella cf. *leptocelloides*	1734	384	152	360	17568	2292	1700	3478	3264
		Stenochironomus	535	114	12	1	57	451	261	344	87
		Stictochironomus	205					457		48	58
		Synorthocladius						339	30		
		Tanypus		2	10			20			
		Tanytarsus	8218	4707	610	643	9207	23663	7096	12179	5732
		Telopelopia okoboji	20								
		Thienemanniella lobapodema	655	141	40	107	720	3221	413	221	100
		Thienemanniella taurocapita	7					81	20		
		Thienemanniella xena	479	187	190	70	90	5523	67	404	25
		Thienemannimyia group	6302	3742	783	414	2859	6091	2056	3783	935
		Tribelos	881	446	49	145	85	838	138	5641	14
		Tvetenia bavarica group	766	212	82		204	10697	48	1382	27
		Tvetenia vitracies	101					21			
		Unniella multivirga	886	30	120	20	220	2014	263	930	30
		Xenochironomus xenolabis	11								
		Xylotopus par	194	21	13	2		92	1	116	2
		Zalutschia							7		
		Zavrelimyia	1794	397	211	60	3645	4833	444	1995	1672
	Culicidae		10	7				65	7		20
	Dixidae	*Dixa*						1039			
		Dixella						131		53	

Table 35. continued

Order	Family	OTUs	Fort Bragg (NC)	Sandhills Gamelands (NC)	Manchester State Forest (SC)	Sandhills State Forest (SC)	Sandhills National Wildlife Refuge (SC)	Savannah River Site (SC)	Fort Gordon (GA)	Fort Benning (GA)	The Nature Conservancy (GA)
	Dolichopodidae									1	
	Empididae	*Chelifera*	40					20			
		Clinocera		20				100	60	31	
		Hemerodromia	301	158	30		127	374	481	227	141
		Neoplasta	99	30			307	565	33	103	40
		Roederiodes						40		10	
	Mycetophilidae							31		30	
	Psychodidae	*Pericoma*						140	27		
		Psychoda							10		
	Ptychopteridae	*Bittacomorpha*	20					1			
		Ptychoptera						4			
	Sciomyzidae		10								
	Simuliidae		2964	405	34	250	463	2907	372	910	426
	Tabanidae	*Chrysops*	102	11	10	1	20	20	21	50	1
		Tabanus						1		1	
	Tipulidae	*Dicranota*	30	2			32	49		48	27
		Gonomyia						1	1		
		Hexatoma	193	12	28	9	59	243	50	111	17
		Limnophila	1				3			31	

Table 35. continued

Order	Family	OTUs	Fort Bragg (NC)	Sandhills Gamelands (NC)	Manchester State Forest (SC)	Sandhills State Forest (SC)	Sandhills National Wildlife Refuge (SC)	Savannah River Site (SC)	Fort Gordon (GA)	Fort Benning (GA)	The Nature Conservancy (GA)
		Limonia						1			
		Ormosia						1			1
		Pilaria	4	7	3		2	106	1	21	
		Pseudolimnophila	1		4			47	12	15	
		Tipula	9	1			4	55	57	20	4
Ephemeroptera	Baetidae	*Acerpenna*	20	789				749	300	217	40
		Baetis complex	196	245				790			
		Procloeon		10				35			
		Pseudocloeon	1219	10	30	20		1072	71	71	27
	Caenidae	*Caenis*	11	89				2327	27	10	
	Ephemerellidae	*Attenella attenuata*	1								
		Eurylophella	3040	451	1190	212	1958	10801	1788	1091	220
		Serratella						36			
	Ephemeridae	*Hexagenia*	7							11	15
	Heptageniidae	*Maccaffertium*	3840	957	196	513	1020	1272	156	365	71
		Stenacron						35		21	
	Isonychiidae	*Isonychia*						2			
	Leptophlebiidae		1029	1275	1570	454	90	1341	388	4477	1300
	Tricorythidae	*Tricorythodes*						1845			

Table 35. continued

Order	Family	OTUs	Fort Bragg (NC)	Sandhills Gamelands (NC)	Manchester State Forest (SC)	Sandhills State Forest (SC)	Sandhills National Wildlife Refuge (SC)	Savannah River Site (SC)	Fort Gordon (GA)	Fort Benning (GA)	The Nature Conservancy (GA)
Lepidoptera	Crambidae	*Parapoynx*	3	1				1		5	
	Noctuidae			1				1			
Megaloptera	Corydalidae	*Corydalus cornutus*	3					1			
		Nigronia serricornis	257	39	3	52	240	404	248	176	56
	Sialidae	*Sialis*	108	21	12	21	60	145	1	139	20
Neuroptera	Sisyridae	*Sisyra*	20	12							
Odonata	Aeshnidae	*Boyeria vinosa*	123	11	6	8	13	67	14	15	5
	Calopterygidae	*Calopteryx*	413	191	21	107	23	44	415	95	
		Hetaerina	102		51			20		6	
	Coenagrionidae	*Argia*	42	3	62	9		113		67	20
		Enallagma	20			3					
		Ischnura	42		1						
		unidentified	54	10				27	17	16	
	Cordulegastridae	*Cordulegaster*	61	19	3		18	97	66	2	6

Table 35. continued

Order	Family	OTUs	Fort Bragg (NC)	Sandhills Gamelands (NC)	Manchester State Forest (SC)	Sandhills State Forest (SC)	Sandhills National Wildlife Refuge (SC)	Savannah River Site (SC)	Fort Gordon (GA)	Fort Benning (GA)	The Nature Conservancy (GA)
	Corduliidae	*Helocordulia*		5					10		
		Neurocordulia	290	20	47	59	15	15	22	42	6
	Gomphidae	*Dromogomphus*	37	13	4	36		24		7	
		Gomphus	145	119	33		37	910	125	98	5
		Hagenius brevistylus	24	6	4		3	2	4		
		Ophiogomphus						75			
		Progomphus	61	4	21	2	3	19	9	83	
		Stylurus	1					6			
	Libellulidae		22					1	20	16	
	Macromiidae	*Macromia*	19	3		2		1			1
Plecoptera	Leuctridae	*Leuctra*	27537	4216	4382	1198	10021	20762	21586	28072	5710
	Peltoperlidae	*Tallaperla*						181	1	495	324
	Perlidae	*Acroneuria*	129	28	1		11	180	63	1	1
		Attaneuria ruralis						1			
		Eccoptura xanthenes	61		2		2	274		49	7
		Neoperla	1								
		Perlesta	98	45	2		53	238	4	106	2
		Perlinella					100	44			
	Pteronarcyidae	*Pteronarcys dorsata*						1			

Table 35. continued

Order	Family	OTUs	Fort Bragg (NC)	Sandhills Gamelands (NC)	Manchester State Forest (SC)	Sandhills State Forest (SC)	Sandhills National Wildlife Refuge (SC)	Savannah River Site (SC)	Fort Gordon (GA)	Fort Benning (GA)	The Nature Conservancy (GA)
Trichoptera	Brachycentridae	*Brachycentrus chelatus*	160			30			162		1
		Brachycentrus nigrosoma	2				10	137			
		unidentified						8			
	Calamoceratidae	*Anisocentropus pyraloides*	447	286	242	15	643	619	790	131	1795
		Heteroplectron americanum	920	40	113	1	154	1761	289	707	220
	Dipseudopsidae	*Phylocentropus*	265				1	7	8		
	Hydropsychidae	*Cheumatopsyche*	266	46	1	1665		657	71		
		Diplectrona modesta	5087	1604	448		2960	2493	760	2457	1052
		Hydropsyche	974	847	12	459		299	288		
		Macrostemum	46			1					
	Hydroptilidae	*Hydroptila*	899	861	120	40	160	91	350	850	49
		Mayatrichia	90			13			20		
		Neotrichia	60					104	20		
		Orthotrichia	60								
		Oxyethira	5	30						40	
	Lepidostomatidae	*Lepidostoma*	731	654	91		901	519	52	2	
	Leptoceridae	*Ceraclea*		1							
		Oecetis	498	254	151	405	382	275	624	131	160
		Triaenodes	2584	656	63	218	268	1993	484	209	94

Table 35. concluded.

Order	Family	OTUs	Fort Bragg (NC)	Sandhills Gamelands (NC)	Manchester State Forest (SC)	Sandhills State Forest (SC)	Sandhills National Wildlife Refuge (SC)	Savannah River Site (SC)	Fort Gordon (GA)	Fort Benning (GA)	The Nature Conservancy (GA)
	Limnephilidae	*Pycnopsyche*	2	2				6			
	Molannidae	*Molanna*	80	53	12		2	479	320	37	80
	Odontoceridae	*Psilotreta*		2				81	53	1	220
	Philopotamidae	*Chimarra*	1401	11	21	15	319	1475	74	1	1
	Polycentropodidae	*Neureclipsis*	79			41			47		
		Nyctiophylax		71	10			106		1	
		Polycentropus	240	1		78	62	115	1	65	10
	Psychomyiidae	*Lype*	85	40	59	3	40	492		103	47
	Rhyacophilidae	*Rhyacophila*	5	25	1		5	88	10	3	
Amphipoda	Crangonyctidae	*Crangonyx*	689	83	20	7	90	321	20	79	1
Decapoda	Palaemonidae	*Palaemonetes*						79			
Isopoda	Asellidae		221	18			20	23		20	
Cyclopoida			6299	1763	1430	170	8873	13404	1818	8737	3410
Harpacticoida			837	1042	1460	140	2163	16934	1881	1108	472
Ostracoda			1681	460	50	1710		2851	203	993	
Tricladida	Planariidae							20			

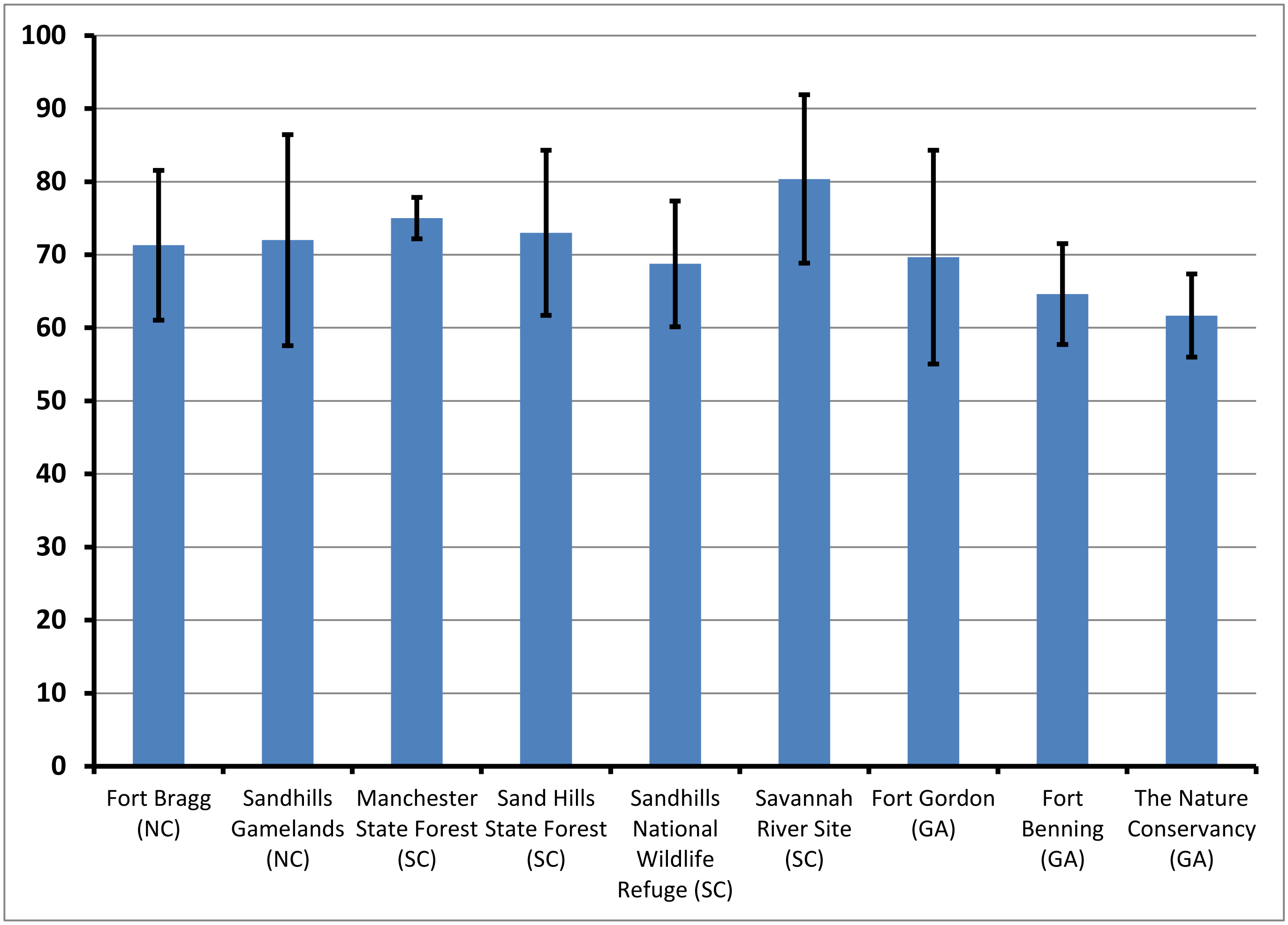


Figure 40. Mean (+1 SD) taxa richness of macroinvertebrate samples taken from sites over the study area. Number of samples per site is indicated in Table 33.

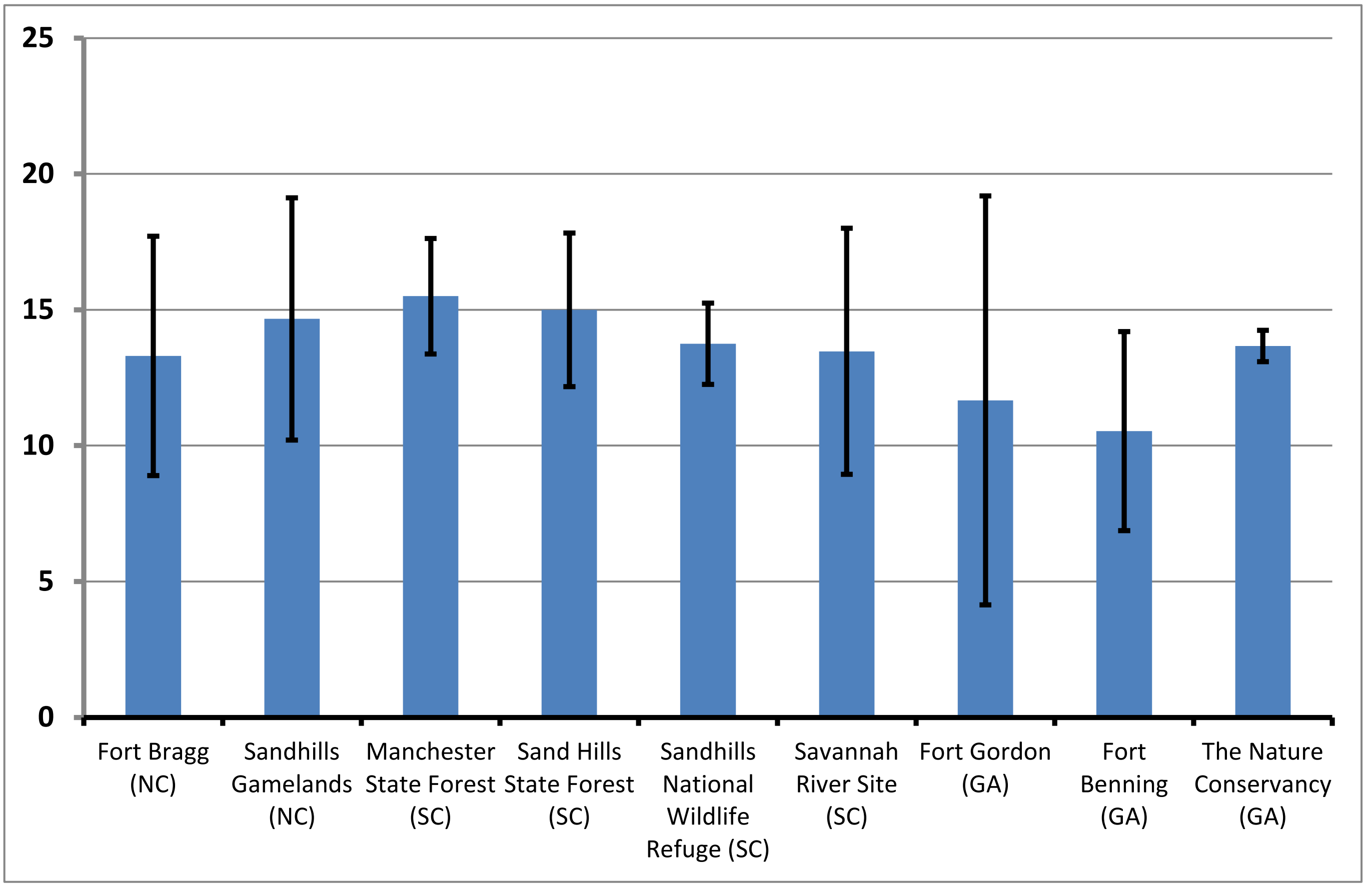


Figure 41. Mean (+1 SD) Ephemeroptera, Plecoptera, Trichoptera (EPT) richness from macroinvertebrate samples taken from sites over the study area. Number of samples per site is indicated in Table 33.

the study area (Figure 41). These low values reflect the inherently low numbers of mayflies and stoneflies in the study area with only modest numbers of caddisfly species. The sandy substrates of Southeastern Plains streams are less than ideal habitats for EPT taxa. This indicated that we should consider alternative community measures in developing reference models.

Twenty-nine taxa were only found at the SRS. Most of these taxa occurred once or few time; however, *Palaemonetes* sp., *Tricorythodes* sp., *Ophiogomphus* sp., *Campeloma* sp., *Dixa* sp., *Cladotanytarsus* sp. "D", and *Nilotanypus fimbriatus*, were found throughout the installation and at several sample sites. *Palaemonetes* sp. is the freshwater grass shrimp. As their name suggests, they are usually associated with aquatic vegetation but are evidently omnivorous (Hobbs and Lodge, 2010). The mayfly genus *Tricorythodes* sp. is facultative and is associated with depositional areas where silt collects (Beaty, 2011a). Our collections indicate that we may have 2 species of *Ophiogomphus* sp. (*O. carolus* and *O. mainensis*). Little documentation is available about these species, but these records are within their known distributions. The snail genus *Campeloma* sp. is ovoviviparous and frequently parthenogenic (Dillon et al., 2006). The dixid fly genus *Dixa* sp. is similar in habit to mosquitos, floating at the surface of the water feeding on microorganisms and detritus. It is a stream taxon typically associated with clean water (Merrit et al. 2008). The chironomid morpho-type *Cladotanytarsus* sp. "D" (Epler, 2001) is know from a few localities in NC and is evidently widespread throughout FL. Our records from the SRS represent the first records of this morpho-type from SC (Scott Castleberry and John Epler, personal communications). *Nilotanypus fimbriatus*, was found only at SRS sites, whereas *Nilotanypus americanus* was found throughout the study area. Both taxa are considered intolerant of pollution (Epler, 2001). Our finding of high biodiversity at SRS is not surprising as it has been documented as an aquatic macroinvertebrate species-rich location (Morse et al., 1980,1983).

Saetheria hirta was only found from Fort Gordon sites "mart" and "bbbb" (Table 1). Epler (2001) noted this species is distributed throughout NC and SC, but it is unknown from GA and our specimens may represent the first records of this species for the state.

The taxa *Tallaperla* sp. was distributed from the central to southern regions of our study range, found only at the SRS, Fort Gordon, Nature Conservancy, and Fort Benning. Nymphs are generally semivoltine, feeding on leaves in clean waters. Our records confirm previous studies that show this taxon has not been recorded in the northern extent of the SH (Beaty, 2011b). Similarly, the beetle *Ectopria* sp. was only found at a few sites on the SRS and Fort Benning. Larvae are often associated with clean water and are adapted to persist in swift-flowing water. Our collections indicate that they are present at 2 installations, but it is likely they occur throughout the SH, although infrequently (Beaty, 2011c)

The chironomid *Parachironomus* was only found in urban streams at Fort Bragg. This is a specious genus with members occupying a range of habitat quality conditions (Epler, 2001). Even though we were not able to identify the species of this taxon, it is likely that more urban samples within the SH will reveal more species within this genus. The net-spinning caddisfly *Macrostemum* sp. was only found in watersheds >24 km^2. These larvae are typically found in larger systems where they collect fine particulate organic matter by filtering with a complex silken retreat (Wiggins, 1996). The case-making caddisflies *Bracycentrus nigrosoma* and *B. chelatus* attach their cases to substrates facing upstream and use their legs to filter detritus and other material from flowing water. Although members of the same genus, these species are easily distinguishable as filtering setae of the ventral portion of the mid- and hind femora of *B. chelatus* consist of evenly spaced larger setae separated by fan-like arrays of finer setae, whereas setae of *B. nigrosoma* is even across the ventrum of the femora (Flint,,1984). These species never co-occurred within the same sample from our collections, and it is noted that *B. chelatus* is considered to be extremely intolerant of organic pollution (Barbour et al., 1999).

Correspondence analysis was performed to ordinate samples representing each state across macroinvertebrate space. Macroinvertebrate density data were selected by stepwise analysis, associating taxa with each state at $\alpha = 0.05$. This analysis indicated that Georgia and South Carolina showed considerable overlap, while both Georgia and South Carolina tended to differ from North Carolina along axis-1. These findings indicate that there are regional differences throughout our study range (Figure 42), and that predictive approaches may be a suitable framework for developing reference models within our study region. It should be noted that our analyses implies that some species occur within specific regions or installations. Unlike fish, which may be restricted to dispersal within a basin, macroinvertebrates have excellent dispersal abilities, especially insects with flight capabilities in the adult phase. Therefore, it is likely that the distributions we observed are not based on range limitations, but instead reflect different management styles or past land uses that have acted as filters in conjunction with biotic processes. For instance, competition during a recovery period from small scale agriculture may results in different assemblages. Furthermore, some of these differences may reflect overall conditions across the study region; hypothetically speaking, the best sites in Georgia may not meet the level of quality found in North Carolina. Therefore, several frameworks should be considered for reference modeling.

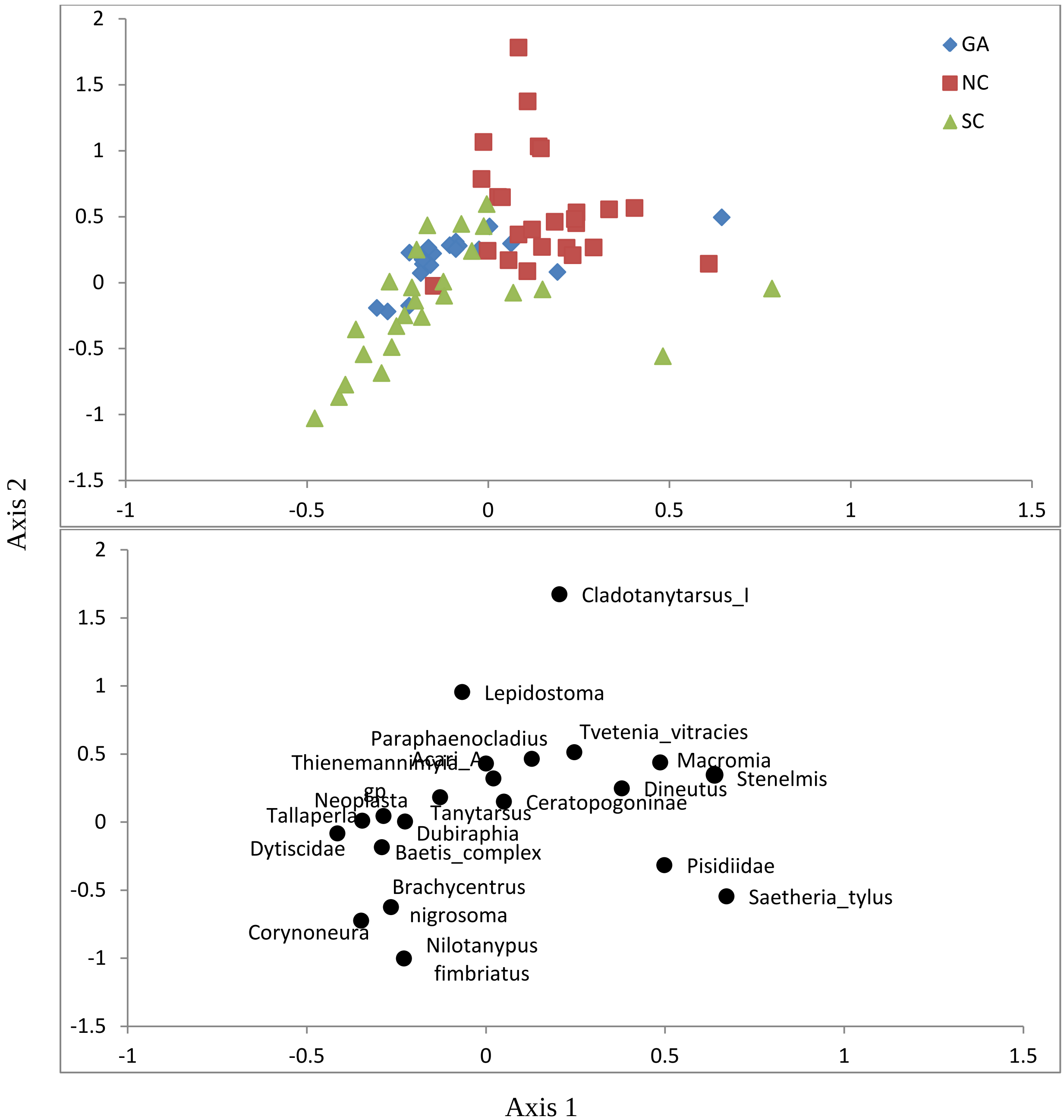


Figure 42. Correspondence analysis of study sites (top) representing North Carolina (NC), South Carolina (SC), and Georgia (GE) and benthic macroinvertebrate taxonomic structure (bottom). Axis-1 & 2 account for 21.3% and 16.5% of the variation, respectively. Two sites from SC, 1 from NC, and *Cheumatopsyche* are omitted.

5.3.2 *Macroinvertebrate Reference Models and Assessment Frameworks: Approach*

The objective of this part of the project was to develop biological modeling frameworks for evaluating habitat conditions based on benthic macroinvertebrate assemblages from reference-condition streams in the SH Ecoregion of GA, SC, and NC. Specifically, we developed reference models that could be used by managers with contrasting resources. We acknowledge that some managers may only have the capabilities to take samples and send them off for identification, whereas others will be able to take multiple field measurements while acquiring and processing GIS and climate data. However, stream quality assessment will be important to all managers; therefore, our report describes a full suite of sampling and assessment approaches that should accommodate most needs.

Below we describe a 3-level approach for an assessment framework. First, we developed a multi-metric model with benthic macroinvertebrate data that could be used for simple stream biomonitoring and assessment. A second set of assessment models was developed for managers with the ability to obtain additional habitat information, allowing for a predictive O/E model approach. Third, we developed predictive O/E models that use an expanded set of predictor variables including land use as well as climate data. Additionally, we built each of these model frameworks with sites representing primary reference conditions and sites representing primary + secondary reference conditions (see reference site selection, Section 5.1). Our purpose in using a tiered reference approach for model construction was twofold. First, we wanted to ensure that we had an assessment tool that used the best conditions within the region to establish the highest benchmarks for evaluation. Second, we wanted to describe a secondary (lower-level) reference condition. Habitats represented by secondary sites are relatively undisturbed and may not differ in community composition from primary reference sites, and, in some instances, can be viewed as sustainable. We acknowledge that inclusion of secondary reference conditions will effectively lower the bar for some assessments; however, it may be impractical to apply stringent evaluations to streams that currently include several land uses or streams undergoing restoration/recovery from impact. Under such conditions a best-attainable model (BAM) might be preferred, which will be better represented by a model based on secondary + primary reference conditions than a model based on the primary reference condition only or a reference model that incorporates varying reference conditions without acknowledging them.

5.3.3 *Model Construction*

Macroinvertebrates, instream habitat, and landscape variables were measured as described in Section 4.0. In addition to these variables, climate data were obtained from the National Oceanic and Atmospheric Administration (http://www.ncdc.noaa.gov/cdo-web/datasets). Data from seven weather stations with daily air temperature highs and precipitation were associated with sample sites within the closest proximity. Missing data were supplemented by the next closest weather station or modeled from several weather stations in the approximate area. Daily precipitation was converted to cm/day, and measures of cumulative precipitation were calculated for 2 weeks, and 1, 3, 6, and 12 months before the sampling date. Cumulative degree-days were calculated from daily highs greater than 0^{o} C with the addition of a measure of cumulative degree days from the Winter Solstice (December 21 or 22) of the previous calendar year to the day of sampling. Sub-catchments were also delineated to build a stressor gradient for model evaluations (below).

We sampled 64 streams across the SH Ecoregion representing a range of landscape conditions (Table 36). Sites were assigned to the reference grades primary, secondary, tertiary, or test, based on criteria described in Section 5.1. Repeat sampling in a subset of sites was done for validation and evaluation of model performance and precision. The sampling date used as a training sample for building models was chosen randomly if a reference site was sampled more than once. About 10% of the reference sites were randomly selected from the total reference site pool, not including repeat samples, and used as samples for validating models. Validation samples were not obtained for tertiary reference sites since they were not used in model construction. Furthermore, tertiary reference sites were not used as test sites and were only used in analyses that assessed overall model performance (e.g., gradient response described below).

5.3.2.1 Predictive Models (RIVPACS)

Modeling framework – We used a River Invertebrate Prediction and Classification System (RIVPACS) type model (Wright et al., 1984) for a predictive O/E (Observed/Expected) model reference framework. We used primary reference sites and primary + secondary reference sites to build two separate sets of models. Within each reference grade we also constructed predictive models from two sets of predictors. First, we used the entire set of instream, GIS, and climate data as predictive variables ("expanded predictor set"; Table 37). We also understand that obtaining a full set of predictors may be difficult for some managers and consultants; thus, as a second approach, we used a smaller set of predictor variables that were easily obtained in the field and from basic maps ("simple predictor set"; Table 37). In all, we constructed the following four sets of predictive models: 1) primary reference with simple predictor set; 2) primary reference with expanded predictor set; 3) secondary reference with simple predictor set; and 4) secondary reference with expanded predictor set.

Theoretical considerations – We designed predictive models that take into account local variations within a single ecoregion (i.e., the SH) because different locations have undergone different land uses and are currently experience differing management practices. Therefore, it is appropriate to specify reference conditions representative of the variability within the ecoregion. Regional distributions of species occur within the study area (as previously shown), and RIPACAS-type modeling is useful in comparing a stream site being assessed to stream groups that are similar to it. If comparison with a single best obtainable reference condition (i.e., the entire SH) is of interest, a multi-metric index (MMI) may be more appropriate.

Large-scale RIVPACS-type studies incorporate many ecoregions and often use precipitation and temperature variables based on return frequencies or annual averages (Wright, 2000; Hawkins et al., 2010; Tsang et al., 2011). In contrast, our study is based on a single ecoregion and, therefore, long-term measures are expected to be similar across the region; however, short term temporal measures from within the sample year or the preceding year may provide information on local effects that account for assemblage differences (Kosnicki and Sites, 2011). Classically, RIVPACS-type models do not include land coverage as a potential predictor because it is often influenced by human activities (Clarke et al., 2003). However, our expanded predictive models included land cover variables as potential predictors because our study sites had varying types and levels of land management. Therefore, we constructed models that could account for the influence of management style on biotic integrity. For instance, the SH is largely recognized as a long leaf pine dominated ecosystem; however, some regions have substantially more deciduous tree coverage than others, which may be a result of prescribed burning frequency.

Classification and prediction – The first step in developing a predictive model is classification of reference sites into distinct groups based on macroinvertebrate assemblage composition. Rare and ubiquitous taxa can create noise that makes it difficult to identify separate assemblage types (Hawkins et

Table 36. Macroinvertebrate stream site locations, sample dates, reference grade (see text), and catchment size.

Site Property	Stream Name	Stream abbreviation	Sample Dates	Reference Designation	Drainage Area (km2)	Northing	Westing
Carolina Gamelands	Joes Creek	sgamjoec	2011-06-21	Primary	4.64	34.88225	-79.62581
Carolina Gamelands	Millstone Creek	sgammist	2011-07-08	Primary	8.12	35.06794	-79.66463
Carolina Gamelands	Bones Fork trib	sgambone	2011-07-08	Primary Validation	3.06	35.03550	-79.61193
Carolina Gamelands	East Prong Juniper Creek	sgamepjc	2011-08-13	Secondary	3.73	34.95324	-79.49592
Carolina Gamelands	West Prong Juniper Creek	sgamwpjc	2011-08-13	Secondary	3.46	34.94988	-79.50353
Carolina Gamelands	Upper Beaverdam Creek	sgamubdc	2011-06-21	Test	9.28	34.85284	-79.57317
Fort Benning	Hollis_Branch_Mainstem(F4)	ftbnhbms	2010-05-13 & 2011-08-01	Primary	1.80	32.36384	-84.68472
Fort Benning	K13	ftbnk013	2010-05-15 & 2012-05-26	Primary	3.04	32.49753	-84.70758
Fort Benning	King_Mills_Creek(K11E)	ftbnkmms	2011-08-01 & 2010-05-12	Primary	2.84	32.51034	-84.64820
Fort Benning	Little_Juniper_Trib(K10)	ftbnljtb	2010-05-15	Primary Validation	8.17	32.51159	-84.63724
Fort Benning	Bonham_Creek_Trib(D13)_up	ftbnbc01	2010-07-12	Secondary	0.61	32.41799	-84.76011
Fort Benning	BCT-D12	ftbnd012	2010-08-04 & 2012-05-27	Test	2.24	32.41195	-84.75909
Fort Benning	BCT-D13-C	ftbnd013	2010-08-14	Test	0.76	32.41900	-84.76000
Fort Benning	LPK K20	ftbnlpk_	2011-06-09	Test	6.30	32.39764	-84.67085
Fort Benning	Wolf Creek	ftbnwcms	2011-07-31	Test	24.46	32.42058	-84.83861
Fort Bragg	Big_Muddy_Creek_Main_Stem	ftbrbmcm	2010-06-18 & 2011-08-11	Primary	29.39	35.01880	-79.51692
Fort Bragg	Field_Branch_Main_Stem	ftbrfbms	2010-06-19 & 2012-08-31	Primary	10.25	35.06520	-79.29351
Fort Bragg	Flat_Creek_Main_Stem	ftbrfcms	2010-06-21	Primary	18.48	35.17428	-79.18029
Fort Bragg	Gum_Branch_Main_Stem	ftbrgbms	2010-06-19	Primary	3.39	35.08975	-79.33671
Fort Bragg	Jennie_Creek_Main_Stem	ftbrjcms	2010-06-20	Primary	11.78	35.12029	-79.33044
Fort Bragg	Juniper_Creek_Main_Stem	ftbrjpms	2010-06-21 & 2011-08-12 & 2012-08-31	Primary	24.10	35.07248	-79.25893
Fort Bragg	Wolf_Pit_Creek_Main_Stem	ftbrwpms	2010-06-20	Primary	7.17	35.11413	-79.33589
Fort Bragg	Little_Rockfish_Main_Stem_North	ftbrlrf1	2010-06-18	Primary Validation	5.58	35.17136	-79.08782
Fort Bragg	Deep Creek	ftbrdeep	2011-08-12	Secondary	5.31	35.14429	-79.16144
Fort Bragg	Hector_Creek_Main_Stem	ftbrhcms	2010-06-22	Secondary	11.85	35.18379	-79.09895
Fort Bragg	Rockfish Branch (RFB)	ftbrrfb_	2012-09-01	Secondary	25.96	35.11509	-79.32548
Fort Bragg	Beaver Creek trib	ftbrbvct	2012-09-01	Tertiary	3.33	35.13448	-78.97676
Fort Bragg	Cabin Creek	ftbrccms	2012-08-31	Tertiary	6.37	35.05382	-79.29671
Fort Bragg	McPherson Creek	ftbrmcph	2012-09-01	Tertiary	4.36	35.13944	-79.04402
Fort Bragg	Cypress Creek	ftbrcypr	2012-09-01	Test	8.53	35.17956	-79.04770
Fort Bragg	Little_River_Tributary	ftbrlrtb	2010-06-22	Test	4.94	35.18655	-79.07469
Fort Gordon	Boggy Gut Creek	fgorbgut	2011-06-13	Primary	6.80	33.34804	-82.29226

Table 36. continued.

Site Property	Stream Name	Stream abbreviation	Sample Dates	Reference Designation	Drainage Area (km2)	Northing	Westing
Fort Gordon	South Prong	fgorsspp	2011-08-06	Primary	16.74	33.35998	-82.15797
Fort Gordon	Bath Branch	fgorbbbb	2011-06-12	Secondary	11.71	33.35817	-82.15793
Fort Gordon	Headstall	fgorhead	2012-06-01	Secondary	0.94	33.34460	-82.34712
Fort Gordon	McCoys Creek	fgormcoy	2011-06-12	Test	7.69	33.39975	-82.16031
Fort Gordon	Trib to Marcum Branch	fgormart	2011-06-13	Test	3.88	33.41086	-82.18649
Manchester State Forest	Tavern Creek	manftvms	2011-07-22	Primary	7.24	33.75840	-80.52736
Manchester State Forest	McCrays Creek	manfmccc	2011-08-10	Test	12.88	33.87429	-80.47611
Sand Hills State Forest	Little Cedar Creek	shsflcdr	2011-06-18	Test	25.55	34.51913	-80.00169
Sand Hills State Forest	Mill Creek	shsfmick	2011-06-18	Test	9.15	34.53971	-80.07420
Sandhills National Wildlife Refuge	Hemp Creek	snwrhemp	2011-06-19	Primary	4.44	34.57151	-80.24763
Sandhills National Wildlife Refuge	Rogers Branch	snwrrgbc	2011-06-17	Primary	2.50	34.60537	-80.21088
Sandhills National Wildlife Refuge	Big Black Creek trib	snwrbbct	2011-06-17	Secondary	3.11	34.66010	-80.22661
Sandhills National Wildlife Refuge	North Prong Swift Creek	snwrnpsc	2011-06-19	Secondary	11.70	34.52632	-80.29989
Savannah River Site	McQueen's_Branch_8	srs_mq08	2010-05-24	Primary	3.50	33.30464	-81.62626
Savannah River Site	Meyer's_Branch_6	srs_mb06	2010-05-22	Primary	5.23	33.17828	-81.56560
Savannah River Site	Mill_Creek_7	srs_mc07	2010-05-23	Primary	1.52	33.32440	-81.60101
Savannah River Site	Tinker_Creek_5	srs_tc05	2010-05-21	Primary	3.08	33.37333	-81.54981
Savannah River Site	Mill_Creek_Main_Stem	srs_mcms	2010-05-25	Primary Validation	7.25	33.30074	-81.58680
Savannah River Site	MC5	srs_mc05	2011-06-14	Secondary	4.02	33.31922	-81.58011
Savannah River Site	Meyer's_Branch_Headwaters	srs_mbhw	2010-05-22	Secondary	10.40	33.19373	-81.57880
Savannah River Site	Pen_Branch_Main_Stem	srs_pbm1	2010-07-27	Secondary	14.30	33.22576	-81.63570
Savannah River Site	Tinker_Creek_6_trib	srs_tc06	2010-05-21 2011-08-03	Secondary	3.40	33.36059	-81.55813
Savannah River Site	Meyer's_Branch_Main_Stem	srs_mbms	2010-07-28	Secondary Validation	28.78	33.17761	-81.58163
Savannah River Site	TCM	srs_tink	2011-08-03	Secondary Validation	31.90	33.36322	-81.55294
Savannah River Site	Mill_Creek_6	srs_mc06	2010-05-23 & 2012-06-02	Tertiary	1.32	33.31719	-81.59757
Savannah River Site	Pen_Branch_4	srs_pb04	2010-07-27	Tertiary	1.38	33.23349	-81.63809
Savannah River Site	Pen_Branch_Headwater	srs_pbms	2010-05-25 & 2012-06-02	Tertiary	4.96	33.23263	-81.62399
Savannah River Site	McQueen's_Branch_Headwater	srs_mqhw	2010-05-24	Test	2.45	33.29734	-81.63051
Savannah River Site	Meyer's_Branch_6.1	srs_mb61	2010-07-28	Test	4.76	33.18051	-81.56335
The Natrue Conservancy	Black Creek trib	tnc_bctb	2011-06-07	Secondary	2.33	32.57205	-84.51212
The Natrue Conservancy	Black Jack Creek	tnc_blkt	2011-06-07	Secondary	1.24	32.57991	-84.49581
The Natrue Conservancy	Parkers Mill Creek trib	tnc_pmbt	2011-06-08	Test	2.36	32.45246	-84.57668

Table 37. Predictor variables used for Random Forest (*RF*) model construction. Simple predictors are indicated.

Predictor code	Simple	Definition
percentDECIDUOUS_		Percentage of the catchment as deciduous tree cover
percentEVERGREEN_		Percentage of the catchment as evergreen tree cover
percentMIXED_FORE		Percentage of the catchment as mixed tree cover
percentSCRUB_SHRU		Percentage of the catchment as shrubs cover
basin_relief_ratio		Highest point in the catchment minus elevation of sample point divided by basin length along mainstem
cumulative_stream_length		(Lt) Total length of all perennial stream segments in catchment
drainage_area__m2_		(A) Total area of catchment (m2)
drainage_density		Lt/A
drainage_perimeter__m_		Perimeter of catchment
drainage_shape		A/Lt^2
highest_point__m_		Highest point in catchment
stream_length		(L) Total length of stream mainstem
length_of_tribs_of_the_mainstem		Lt - L
Q		Discharge (m3/s) measured at time of sample
mon3_precip		Cumulative precipitation (cm) 3 mo. prior to sampling
yr1_precip		Cumulative precipitation (cm) 1 year prior to sampling
wk2_temp		Cumulative daily maximum air temperature (^{o}C) 2 wk. prior to sample date
mon6_precip		Cumulative precipitation (cm) 3 mo. prior to sampling
mon6_temp		Cumulative daily maximum air temperature (^{o}C) 6 mo. prior to sampling
mon3_temp		Cumulative daily maximum air temperature (^{o}C) 3 mo. prior to sampling

Table 37. continued.

Predictor code	Simple	Definition
ddcum		Cumulative daily maximum air temperature (^oC) from the Winter Solstice to sampling date
mon1_temp		Cumulative daily maximum air temperature (^oC) 1 mo. prior to sampling
Distance_nearest_Pond_Lake_down_		Distance (m) to the nearest pond or lake upstream from sampling location
sandHUC_A	yes	Sample site occurs within Hydrologic Unit Code Subregion Apalachicola (1 or 0)
sandHUC_OS	yes	Sample site occurs within Hydrologic Unit Code Subregion Ogeechee-Savannah (1 or 0)
HUC	yes	Hydrologic Unit Code 4 (Subregion)
wetTave	yes	Average (n=4) maximum wetted depth
lat	yes	Latitude (decimal degrees)
long	yes	Longitude (decimal degrees)
mouth_elevation__m_	yes	Elevation at the lower portion of sample site
liveave	yes	Average (n=4) area of stream channel covered by living coarse particulate organic matter >2.5 cm in diameter
fieldtemp	yes	Water temperature (^oC) measured at time of sampling
wetWave	yes	Average (n=4) maximum wetted width
wetWDave	yes	Average (n=4) maximum wetted width/depth
deadave	yes	Average (n=4) area of stream channel covered by dead coarse particulate organic matter >2.5 cm in diameter
dryave	yes	Average (n=4) area of dry stream channel covered by coarse particulate organic matter >2.5 cm in diameter
submergave	yes	Average (n=4) area of wetted stream channel covered by coarse particulate organic matter >2.5 cm in diameter
large	yes	Sample catchment area >20 km^2 (1 or 0)

al., 2000); therefore, we removed taxa occurring at > 90% or < 10% of the sample sites for cluster analysis. Jaccard's dissimilarity was used with the "flexible" method in the cluster package for R. Nonmetric multi-dimensional scaling (NMDS) was used with the same data sets. Training sites were plotted over NMDS dimensions 1 and 2 and color coded based on their groupings from the cluster analysis.

Several predictive analyses can be used in O/E models to identify environmental variables associated with the biological groups identified as described above. These include discriminate analysis, logistic regression, nearest neighbor analysis, and others. We chose random forests (*RF*), which has recently been used successfully in ecology (Cutler, et al. 2007) and has been noted for its superior ability to map species distributions without overfitting (Prasad, et al. 2006). *RF* is developed by producing a large number of regression trees that are grown individually from a randomized subset of predictors. Prediction of class membership is based on the "average" of all trees.

We delineated a simple set of predictors based on the construction of preliminary models and beta testing of *RF* capabilities (Table 37). For the expanded predictor set, we screened 74 predictor variables by building five consecutive *RF* models and identifying the top 30 most frequently occurring predictors from these runs. Then, we constructed three additional models, and the variables appearing most important based on mean decrease accuracy in two out of three models were selected and cross referenced with the 30 most frequently occurring predictors. The final set of variables was considered the expanded predictor set. This procedure was performed separately for the primary and primary + secondary classes of reference data sets. We developed *RF*s for each of the predictive model sets with the one simple and two expanded predictor sets just described. Our models were built from 500 bootstrap samples, using in-bag predictions for cross-validating accuracies and error rates for new observations.

<u>Scoring methods</u> – Probabilities of group membership and frequency of occurrence for each taxon of each reference class were multiplied to find the capture probability for each taxon at an assessment site. The summation of all capture probabilities was the expected (*E*) species richness. Observed taxa expected to occur are summed for the observed species richness (*O*). The ratio of *O/E* is then calculated with a value close to one indicating that the assessment site has an assemblage representative of reference quality and a value less than one indicating an assemblage deviating from reference quality.

An issue with classic *O/E* scoring methods is that test scores may be inflated when many rare taxa with low capture probabilities are observed (Hawkins et al., 2000). A way to compensate for such inflation is to count only taxa with a capture probability ≥ 0.5 (OE_{50} model, Simpson and Norris, 2000). This method has been shown to be more sensitive (Hawkins et al., 2000; Van Sickle et al., 2007), but the exclusion of infrequently occurring taxa is highly contentious because of lost information (Cao et al. 1998; Clarke and Murphy, 2006), and it is acknowledged that inclusion of such taxa may warrant revisiting (Van Sickle, 2008). Van Sickle (2008) demonstrated the effectiveness of Bray-Curtis dissimilarities (*BC*) in comparing *O* and *E* assemblages in RIVPACS-type models:

$$BC = \frac{\sum |O_i - E_i|}{\sum (O_i + E_i)} \tag{6}$$

The advantage to using this type of a measure is that inclusion of rare taxa may be beneficial in showing response to disturbance gradients. *BC* also can be calculated by limiting taxa to only those that have a capture probability ≥ 0.5 (BC_{50}), although these models were shown to be weaker than *BC* models that included the full set of taxa in responding to stressor gradients.

We used a novel method of scoring *O/E* models ("capped", OE_{CAP}) in which inflation of *O/*E scores was controlled while including all taxa in the model:

$$CAP_j = \sum_{i=1}^{n} O_i \, P_i \tag{7}$$

$$OE_{CAP} = 100 \times \frac{CAP_j}{E_j} \tag{8}$$

where *O* is the presence or absence (1 or 0) and *P* is the capture probability of taxon *i*, and *E* is the expected taxa richness for site *j*. The maximum score of any observed taxon (O_i) is capped by the capture probability (P_i) of that taxon occurring at that site, and OE_{CAP} cannot exceed 100. This equation can be thought of as a similarity index that evaluates predictive performance of RIVPACS-type models and can be applied to any *O/*E model that estimates expected species richness based on capture probabilities.

We built RIVPACS-type *RF* models with modified versions of *R* scripts provided by Van Sickle at the Environmental Protection Agency website:
http://www.epa.gov/wed/pages/models/rivpacs/rivpacs.htm

We developed R scripts for calculating predictive model scores for any macroinvertebrate dataset using our taxonomic standards (Appendix MM1). A database for assigning correct Operational Taxonomic Units (OTUs) and scripts for calculating *OE*, OE_{50}, *BC*, BC_{50}, and OE_{CAP}, and instructions for using these products are available for download at the following website:
http://www.auburn.edu/~ezk0004/sandhills/ **Under Construction to be delivered 1 April 2014**

5.3.2.2 Macroinvertebrate Multimetric Index (MMI)

Multi-metric benthic macroinvertebrate models have been designed and implemented by state agencies for many years as part of basic water quality monitoring (NCDENR, 2006; GDNR, 2007). These assessment tools tend to be calibrated to local conditions that may be overlooked in large national assessments (e.g. USEPA 2006). State *MMI*s are usually restricted to data derived from within the state. We suggest that construction of an *MMI* that includes the same sampling protocol across state boundaries will have added value in accounting for variation specific to the SH Ecoregion.

Metric screening and evaluation – We calculated 86 macroinvertebrate metrics representing measures of richness, diversity, composition, sensitivity to pollution, functional feeding group composition, and habit use (Table 38). Metric values were screened across all sample sites for range limitations (< 4, for non-diversity measures) and excessive zero values (threshold 66% of samples with 0 values). Metrics from reference sites were initially inspected for their variability over catchment area, sample site elevation, and wetted width to adjust metric scores for natural gradients (Fausch et al., 1984). However, this process was cumbersome, and many metrics showed no response over these natural gradients, similar to other studies (Stoddard et al., 2008). Thus, instead of adjusting metrics for variability over natural gradients, we screened all metrics for variability (e.g. Barbour et al. 1999) where CV > 0.5 was considered to be excessive and these measures were excluded from *MMI* modeling (Table 38).

For sites sampled repeatedly we evaluated signal to noise (S/N) effects, with the signal representing the variance of sites sampled in 2010 and the noise the pooled covariance from the same sites with repeated

Table 38. Benthic macroinvertebrate metrics considered for development of a Macroinvertebrate Multi-Metric Index (*MMI*). Evaluation of candidate metrics for signal noise ratio (S/N), inappropriate range (Range), excessive 0-values, t-test results between reference and test sites, and excessive variation among reference sites for primary and primary + secondary references. List sorted by S/N. abn = abundance. EPT = Ephemeroptera, Plecoptera, and Trichoptera. ns = metric non-significant and thus was excluded from *MMI* modeling.

Metric name	Description	Type	S/N	Range & excess 0-value	Secondary difference p-value	Secondary difference t-value	Primary difference p-value	Primary difference t-value	Primary CV	Sec. CV
BI	Biotic Index	sensitivity	5.88	no	<0.001	-4.83	<0.001	-4.79	12.30	13.64
clingrich	Clinger richness	habit	4.78	no	<0.001	3.99	<0.001	4.69	21.63	26.67
pswimrich	% Swimmer richness	habit	4.51	no	ns	ns	ns	ns	60.35	63.78
pclingrich	% Clinger richness	habit	3.95	no	0.002	3.45	<0.001	4.17	15.33	18.78
NCBI	NC Biotic Index	sensitivity	3.81	no	<0.001	-4.53	<0.001	-4.91	12.27	12.78
nanopsectCHI	*Nanocladius* and *Psectrocladius*/Chironomidae abn	composition	3.76	yes	ns	ns	ns	ns	188.95	198.49
per_ortho	% Orthocladiinae	composition	3.66	no	ns	ns	ns	ns	76.23	69.81
pchi	% Chironomidae	composition	3.46	no	ns	ns	ns	ns	24.14	27.07
per_nanopsect	% *Nanocladius* and *Psectrocladius*	composition	3.36	yes	ns	ns	ns	ns	171.38	170.91
per_corynon	% *Corynoneura*	composition	3.33	no	ns	ns	ns	ns	116.70	108.80
swimrich	Swimmer richness	habit	3.06	no	ns	ns	ns	ns	59.27	64.85
burrowers	% Burrowers	habit	3.02	no	ns	ns	ns	ns	56.80	55.03
filtrich	Filterer richness	FFG	2.94	no	0.071	1.88	0.014	2.63	20.25	23.57
climbers	% Climbers	habit	2.92	no	ns	ns	ns	ns	78.04	88.55
perept	% EPT	composition	2.91	no	0.001	3.51	0.001	3.56	50.46	57.92
sdi	Shannon Diversity Index	diversity	2.88	no	ns	ns	ns	ns	8.90	10.13
gatrich	Gathering collector richness	FFG	2.75	no	0.045	2.12	0.054	2.02	17.67	22.24
rich	Total taxa richness	rich	2.72	no	0.009	2.79	0.005	3.06	13.26	15.44
shred	% Shredders	FFG	2.71	no	0.001	3.78	0.001	3.84	52.45	62.94
ppredrich	% Predator richness	FFG	2.62	no	0.005	-3.12	0.022	-2.44	10.74	11.26
filt	% Filtering collectors	FFG	2.54	no	ns	ns	ns	ns	23.26	29.52
psprawlrich	% Sprawler richness	habit	2.44	no	ns	ns	ns	ns	10.14	10.46
orthoCHI	Orthocladiinae/Chironomidae abn	composition	2.37	no	ns	ns	ns	ns	63.61	58.50
scrape	% Scrapers	FFG	2.36	no	ns	ns	ns	ns	79.78	83.78
perHarnComplx	% *Harnischia* complex chironomids	composition	2.27	no	ns	ns	ns	ns	163.64	133.40

Table 38. continued.

swimmers	% Swimmers	habit	2.18	no	ns	ns	ns	ns	159.58	137.65
T	Trichoptera richness	rich	2.16	no	0.014	2.62	0.006	2.95	34.68	38.10
D	Simpson Diversity Index	diversity	2.16	no	ns	ns	ns	ns	43.98	56.20
clingers	% Clingers	habit	2.15	no	ns	ns	ns	ns	42.04	47.54
predrich	Predator richness	FFG	2.13	no	ns	ns	ns	ns	16.88	18.03
HComplxCHI	*Harnischia* complex chironomids/Chironomidae abn	composition	2.09	no	ns	ns	ns	ns	166.28	139.13
orthchirich	Orthocladiinae richness/Chironomidae richness	composition	2.09	no	0.043	2.18	0.029	2.36	15.18	15.61
climbrich	Clibmer richness	habit	2.07	no	0.041	2.19	0.018	2.54	32.30	32.19
pshredrich	% Shredder richness	FFG	2.06	no	0.001	3.83	0.002	3.50	21.41	22.95
Hrich	*Harnischia* complex richness	rich	2.02	no	ns	ns	ns	ns	61.58	60.83
dom1	% Dominant taxon	sensitivity	1.95	no	ns	ns	ns	ns	43.55	47.05
sprawlrich	Sprawler richness	habit	1.94	no	ns	ns	ns	ns	15.83	17.08
chiminiCHI	Chironomini/Chironomidae abn	composition	1.94	no	0.040	-2.19	ns	ns	53.34	54.51
J	Evenness	diversity	1.90	no	ns	ns	ns	ns	7.64	8.35
per_chimini	% Chironomini	composition	1.89	no	ns	ns	ns	ns	61.90	67.53
Tpodchirich	Tanypodinae richness/Chironomidae richness	composition	1.88	no	0.024	-2.50	0.025	-2.47	20.87	20.86
per_procladApsec	% Procladiinae + *Apsectrotanypus*	composition	1.85	no	ns	ns	ns	ns	168.87	159.27
pfiltrich	% Filtering collector richness	habit	1.84	no	ns	ns	ns	ns	14.67	16.88
per_chiminae	% Chironominae	composition	1.83	no	ns	ns	ns	ns	35.45	40.87
P	Plecoptera richness	rich	1.81	no	<0.001	6.21	<0.001	6.05	36.67	38.90
procladApsecCHI	(Procladiinae + *Apsectrotanypus*)/Chironomidae abn	composition	1.79	no	ns	ns	ns	ns	156.93	146.82
eptrich	EPT richness	rich	1.78	no	<0.001	4.08	<0.001	4.39	23.29	27.09
dom2	% 2 Dominant taxa	sensitivity	1.77	no	ns	ns	ns	ns	33.63	29.96
dom3	% 3 Dominant taxa	sensitivity	1.77	no	ns	ns	0.076	-1.84	25.87	25.69
pred	% Predators	FFG	1.76	no	ns	ns	ns	ns	27.93	32.81
Pentanrich	Pentaneurini richness	rich	1.72	no	ns	ns	ns	ns	26.04	25.80
Tpodrich	Tanypodinae richness	rich	1.66	no	ns	ns	ns	ns	27.94	26.48
corynonCHI	*Corynoneura*/Chironomidae abn	composition	1.65	no	ns	ns	ns	ns	94.95	87.46
shredrich	Shredder richness	FFG	1.58	no	<0.001	4.50	<0.001	4.20	27.46	26.63
sprwlers	% Sprawlers	habit	1.57	no	0.047	2.10	0.036	2.22	20.36	22.29

Table 38. continued.

per_Tpod	% Tanypodinae	composition	1.57	no	ns	ns	ns	ns	48.25	49.15
chiminirich	Chironomini richness	rich	1.50	no	ns	ns	ns	ns	20.02	22.95
chiminaeCHI	Chironominae/Chironomidae abn	composition	1.50	no	ns	ns	ns	ns	27.28	29.35
TpodCHI	Tanypodinae/Chironomidae abn	composition	1.42	no	ns	ns	ns	ns	43.70	43.33
Hchirich	*Harnischia* complex richness/Chironomidae richness	composition	1.41	no	ns	ns	ns	ns	54.96	53.17
pclimbrich	% Climber richness	habit	1.40	no	ns	ns	ns	ns	33.74	31.57
chirich	Chironomidae richness	rich	1.36	no	0.034	2.25	0.022	2.44	13.63	16.35
gath	% Gathering collectors	FFG	1.33	no	ns	ns	ns	ns	38.20	38.91
Penchirich	Pentaneurini richness/Chironomidae richness	composition	1.32	no	ns	ns	0.048	-2.12	20.69	23.16
pentanCHI	Pentaneurini/Chironomidae abn	composition	1.32	no	ns	ns	ns	ns	45.77	45.27
per_pentan	% Pentaneurini	composition	1.26	no	ns	ns	ns	ns	48.43	50.32
E	Ephemeroptera richness	rich	1.25	no	0.028	2.31	0.047	2.07	34.31	38.26
orthrich	Orthocladiinae richness	rich	1.24	no	0.009	2.86	0.004	3.17	18.19	22.96
chiminichirich	Chironomini richness/Chironomidae richness	composition	1.15	no	ns	ns	ns	ns	12.17	15.53
pgatrich	% Gathering collector richness	FFG	1.10	no	ns	ns	ns	ns	10.33	12.71
per_tanysini	% Tanytarsini	composition	1.09	no	ns	ns	ns	ns	39.28	46.87
phyept	Hydropsychidae/EPT abn	composition	1.08	no	ns	ns	ns	ns	71.94	82.42
per_rheosmit	% *Rheosmittia*	composition	1.01	no	0.008	2.78	ns	ns	166.59	162.78
tanysinirich	Tanytarsini richness	rich	0.96	no	0.026	2.44	0.030	2.35	14.55	16.02
rheosmitCHI	*Rheosmittia*/Chironomidae abn	composition	0.94	no	0.006	2.88	ns	ns	137.30	149.16
dom5	% 5 Dominant taxa	sensitivity	0.92	no	ns	ns	ns	ns	21.18	23.48
tanysiniCHI	Tanytarsini/Chironomindae abn	composition	0.90	no	ns	ns	ns	ns	32.72	37.20
dom4	% 4 Dominant taxa	sensitivity	0.89	no	ns	ns	ns	ns	29.81	26.46
burrowrich	Burrower richness	habit	0.84	no	ns	ns	ns	ns	25.57	27.17
chminaechirich	Chironominae richness/Chironomidae richness	composition	0.83	no	ns	ns	ns	ns	9.23	9.90
chminaerich	Chironominae richness	rich	0.81	no	0.032	2.29	0.042	2.14	17.19	18.02
pernoinsect	% Non-Insects and non-chironomid Diptera	composition	0.81	no	0.009	-2.92	0.022	-2.47	46.10	42.09
scraprich	Scraper richness	FFG	0.67	no	ns	ns	ns	ns	32.23	31.97
pburrowrich	% Burrower richness	habit	0.47	no	ns	ns	ns	ns	17.77	20.25
pscraprich	% Scraper richness	FFG	0.39	no	ns	ns	ns	ns	28.05	28.62
tanychirich	Tanytarsini richness/Chironomidae richness	composition	0.27	no	ns	ns	ns	ns	13.68	18.06

samples from following years (Kaufmann et al. 1999; Stoddard et al., 2008). Metrics with S/N scores >2 were identified as stable and further considered for model construction. We used t-tests to compare mean metric scores for reference and test sites to determine which metrics were most responsive to disturbance (Barbour et al., 1999; Flotemersch et al., 2006; Stoddard et al. 2008). Metrics with an absolute t-value >2.5 were considered for aggregation into the primary and primary + secondary *MMI*s. Where a metric type was represented by S/N <2, but differed between reference and test sites, they were reconsidered, if S/N >1.5, so that a metric of that type was represented. The subset of metrics passing each of these selection criteria was screened for redundancy to ensure the final set did not display high correlation (Pearson's $r > |0.71|$), thus ensuring each measure contributed separate information about the community.

Metric scoring – The final set of metrics for the primary and primary + secondary *MMI*s were aggregated into indices by assigning scoring criteria to each metric so that the maximum index score was 100. Ceiling and floor values for metric scoring criteria were set at the 95th percentile of reference sites and 5th percentile of all sites, respectively, for metrics that decrease with increased stress (Barbour, et al., 1999). For metrics that increase with increased stress, the 5th percentile of reference sites and the 95th percentile of all sites were used to determine floor and ceiling values, respectively.

5.3.2.3 Evaluation of Models

All *MMI* and RIVPACS-type model scores were tested for differences among training data (reference sites), test sites (sites representing a range of disturbances), and validation data (reference sites that passed screening criteria but were not used in model construction) with an unbalanced ANOVA. Models that detected a significant difference between training versus test sites and validation versus test sites, while also showing no significant difference between training versus validation sites, were considered capable of discriminating disturbed from reference quality sites. Log_{10} transformations were used for scores that were not normally distributed.

The preceding models were then screened against two gradients of habitat quality. The first gradient was a Principal Component Analysis (PCA) axis of land use/land cover (LULC) that included vegetative cover in addition to developed cover and roads in the catchment within 1 km^2 of the sample site; this area was considered a sub-catchment LULC gradient. The second gradient was based on the single measure of stream water dissolved oxygen (Table 39), which showed considerable variation among sites during the study. These variables were chosen for evaluating model performance instead of variables used to delineate reference grades to avoid circular logic. Furthermore, the 1 km^2 sub-catchment buffers represented land cover closest to the sample sites, which possibly was more influential to benthic invertebrate assemblages than land cover of the entire catchment. Each model was plotted against each gradient for a subset of macroinvertebrate sample sites, representing those not used in model construction or evaluation. Subsets consisted of secondary and tertiary reference sites and repeat samplings for models built with primary reference sites and tertiary references and repeat samplings for models built with primary + secondary reference sites.

After rigorous testing of *MMI* and RIVPACS-type models, a final assessment score was generated that could be used as an evaluation tool. The 25th percentile and minimum possible score (zero) of all primary reference quality sites (training, validation, but not repeat samples) was used for trisecting ranges for final assessment of biological integrity for measures that decreased with disturbance (OE, OE_{50}, OE_{CAP}, and *MMI*). The same step was done for remaining secondary models with the addition of secondary reference sites. For models showing increasing scores with increasing disturbance (BC and BC_{50}), the 75th percentiles were used for primary and primary + secondary models with all respective reference quality samples. The trisected model ranges represented "Good", "Fair", "Poor", and "Very Poor" quality.

Table 39. Environmental variables used for disturbance gradient assessment of macroinvertebrate models. LULC PCA = axis-1 from land use land cover Principal Component analysis.

Variable	Definition	Use in Disturbance gradient
00_Drainage_Area__m2_	Total area of catchment (m2) within a 1 km sub-basin buffer above sample site	LULC PCA
_00_HIGH_INTEN	Percentage of 1 km sub-catchment buffer (m2) above sample site consisting of high intensity development	LULC PCA
_00_DECIDUOUS_	Percentage of 1 km sub-catchment buffer (m2) above sample site consisting of deciduous trees	LULC PCA
_00_EVERGREEN_	Percentage of 1 km sub-catchment buffer (m2) above sample site consisting of coniferous trees	LULC PCA
_00_MIXED_FORE	Percentage of 1 km sub-catchment buffer (m2) above sample site consisting of mixed coniferous and deciduous trees	LULC PCA
_00_MEDIUM_INT	Percentage of 1 km sub-catchment buffer (m2) above sample site consisting of medium intensity development	LULC PCA
_00_LOW_INTENS	Percentage of 1 km sub-catchment buffer (m2) above sample site consisting of low intensity development	LULC PCA
_00_DEVELOPED_	Percentage of 1 km sub-catchment buffer (m2) above sample site consisting of open-space development	LULC PCA
_00_Paved_Rd_Dens	Percentage of paved roads within a 1 km sub-catchment buffer (m2) above sample site	LULC PCA
_00_Unpaved_Rd_Dens	Percentage of unpaved roads within a 1 km sub-catchment buffer (m2) above sample site	LULC PCA
domg	Dissolved oxygen (mg/L) measured during time of sampling	Single measure

A final evaluation was implemented by comparing each model to 1) the number of reference sites not evaluated as the "Good"; 2) the number of disturbed (test) sites evaluated as "Good"; and 3) the number of test sites evaluated as "Poor" or "Very Poor" quality. The best primary and primary + secondary predictive reference models, respectively, were selected based on their overall performance. *MMI* models also were evaluated and considered as separate stand alone or complementary evaluation tools.

Unless noted otherwise, analyses and calculations were conducting in SAS® version 9.2 (SAS Institute Inc. 2004).

5.3.3 *Model Results*

5.3.3.1 Predictive Models (RIVPACS)

Classification – Three groups of sites were delineated for primary reference models (Figure 43), representing a large northern group, a small southern group, and a small group from miscellaneous regions (Figure 44). Four classes with primary + secondary reference sites were identified (Figure 45), representing regions associated with the Savannah River Site, Fort Benning, Fort Bragg/Fort Gordon, and a miscellaneous class (Figure 46). Cluster trees were pruned and group memberships were assigned to each of the reference sites for the primary and primary + secondary reference sets, separately, for use in further analysis.

Prediction –There were 38 predictor variables that were important for *RF* construction of simple and expanded predictor sets (Table 39). From the simple predictor set, the primary reference groups were most related to geographic position, drainage association, and water temperature during sampling (Figure 47). The same was true for the simple predictor set for primary + secondary reference groups with the addition of elevation (Figure 48). For the expanded predictor reference sets, geographic position and water temperature also were important, although climate variables were more important in identifying primary reference groups than other variable types including land cover and drainage association (Figure 49). In contrast, for primary + secondary reference groups for expanded sets maximum wetted depth, drainage association, and to a lesser extent land cover and elevation were more important (Figure 50).

The out-of-box error rate estimates for reference classes was 19.1% for both primary reference models with most of the error being made up by the miscellaneous class. The out-of-box error rate estimates for reference classes was 11.1% for the primary + secondary reference model with the simple set of predictors mostly represented by misclassification of the Fort Bragg/Fort Gordon cluster and 13.9% with the expanded set of predictors with most of the misclassifications represented by the miscellaneous cluster.

5.3.3.2 Macroinvertebrate Multimetric Index (MMI)

Metric screening and evaluation – The same 14 metrics passed the S/N and excessive 0 and variation evaluations and were found to show significant differences between test and reference sites for both the primary and primary + secondary reference sets (Table 38). However, there were no macroinvertebrate composition metrics that passed the selection process. Although many composition metrics were screened, all showed non-significant differences between reference and non-reference sites, had a low S/N, or displayed inherently high variation. To represent macroinvertebrate composition, which is typically included in multimetric indices, we decided to use the composition metrics *orthchirich* and *Tpodchirich* (% Orthocladiinae richness/Chironomidae richness, and % Tanypodinae richness/Chironomidae richness). Both metrics are generally related since one often decreases as the other increases; however, this is not universal. In general, Tanypodinae are associated with eutrophication (Sæther, 1979) and organic pollution (Lenat, 1993) and are represented in disturbance-based metrics

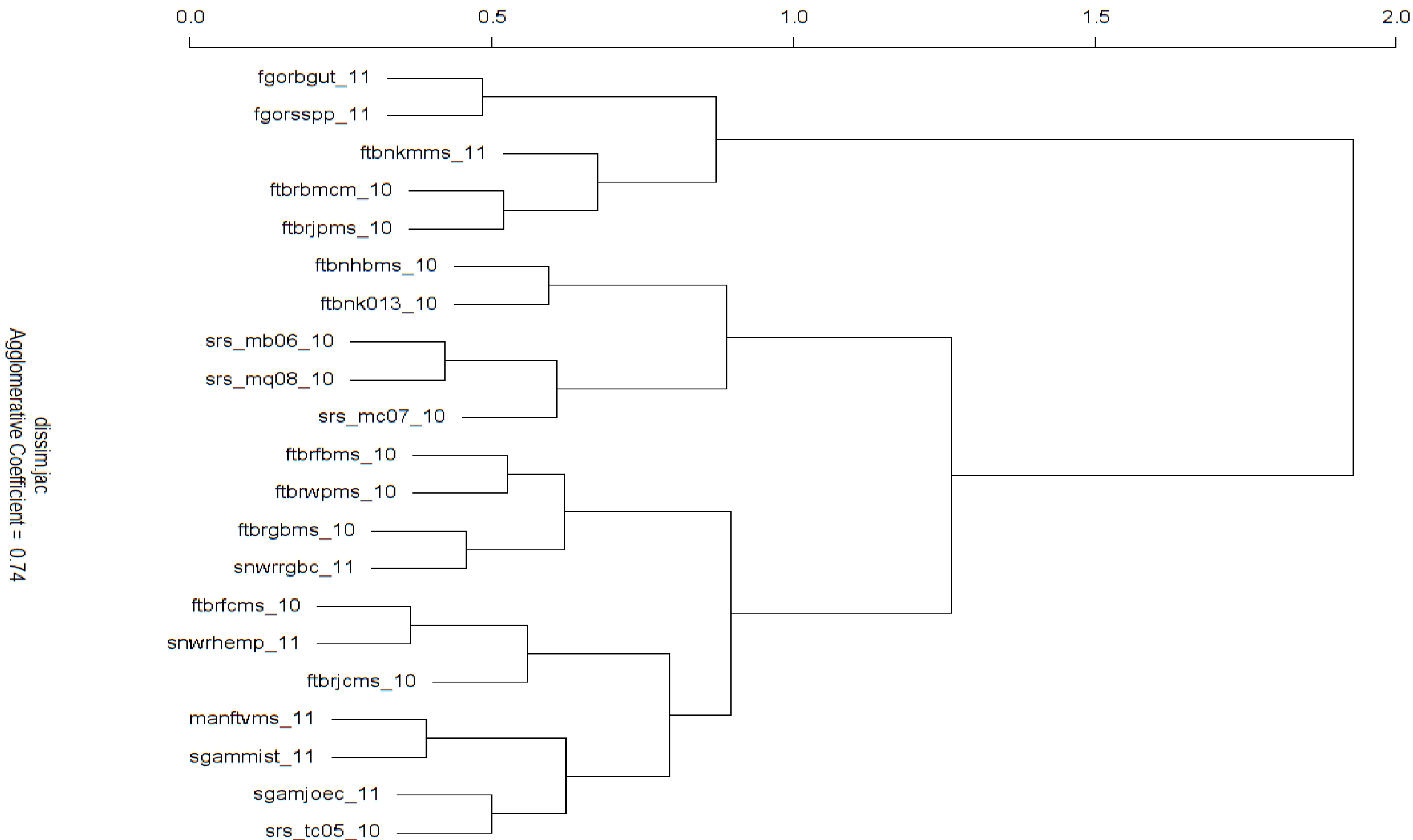


Figure 43. Cluster analysis of macroinvertebrate assemblages without rare (< 10%) and ubiquitous (>90%) taxa for primary reference sites of the sand Hills Ecoregion. The number following the stream abbreviation indicates the year the sample was taken.

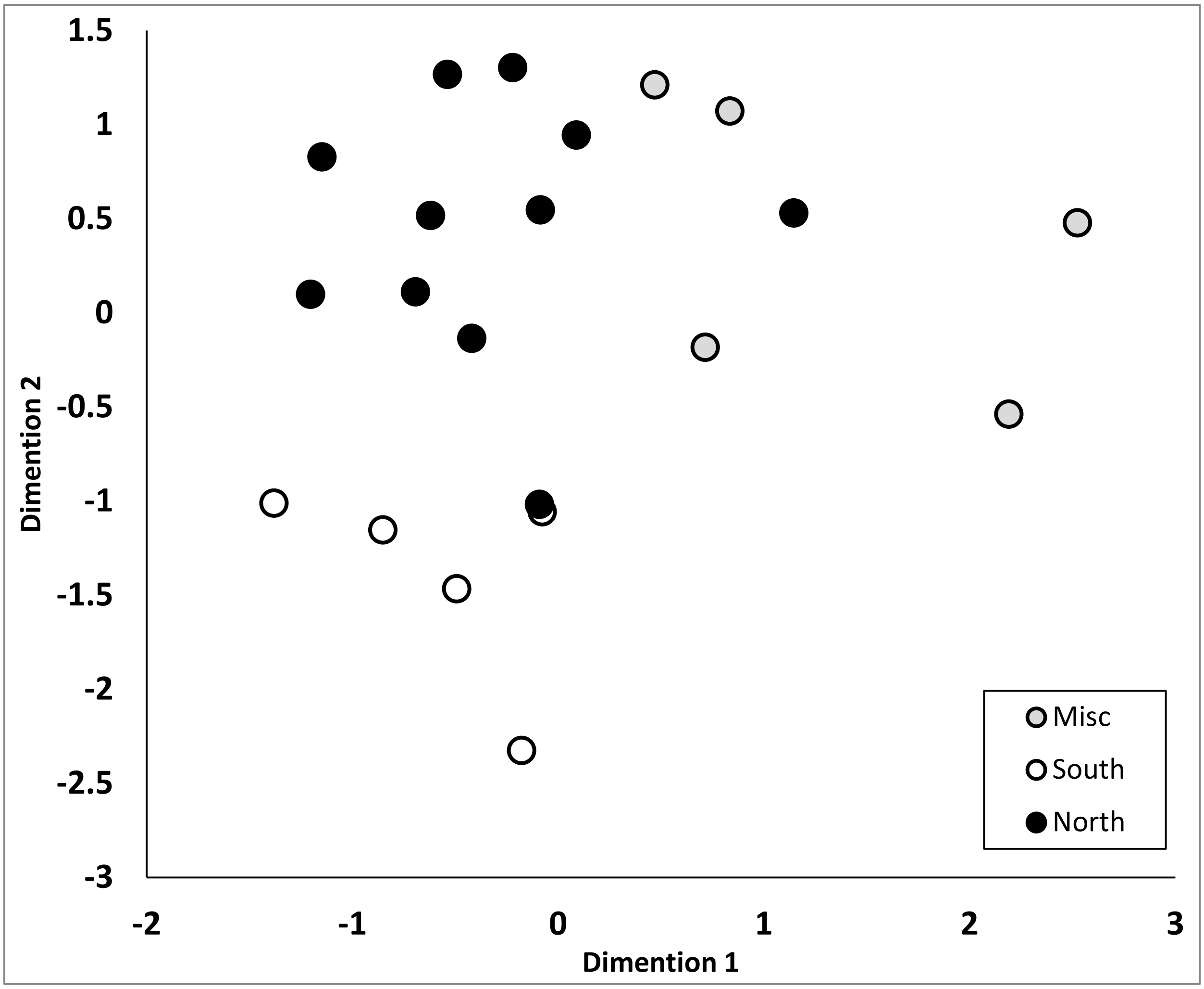


Figure 44. Nonmetric multi-dimensional scaling of primary reference sites across the Sand Hills Ecoregion. Sites with the same color pattern were grouped together from the cluster analysis.

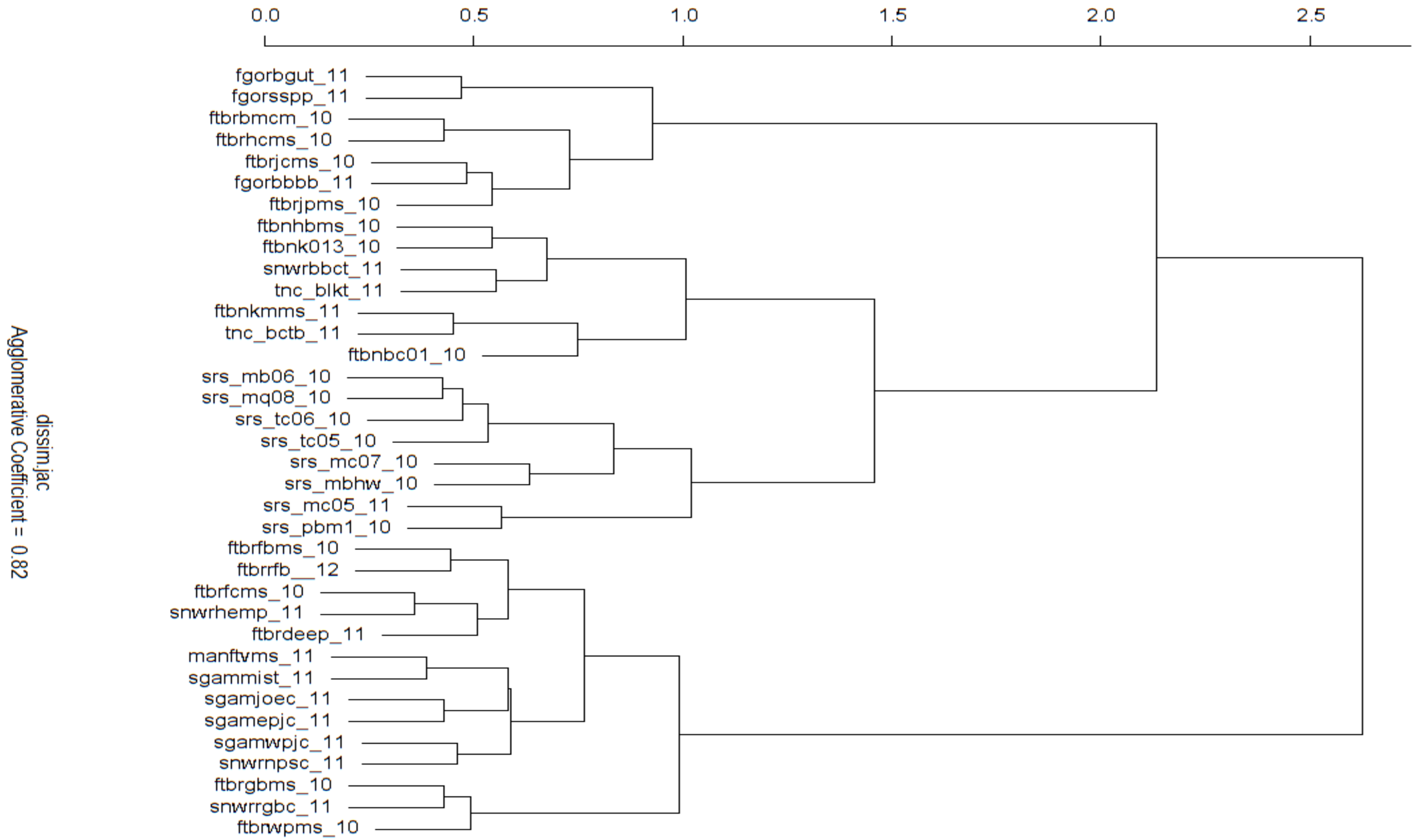


Figure 45. Cluster analysis of macroinvertebrate assemblages without rare (< 10%) and ubiquitous (>90%) taxa for primary and secondary reference sites of the Sand Hills Ecoregion. The number following the stream abbreviation indicates the year the sample was taken.

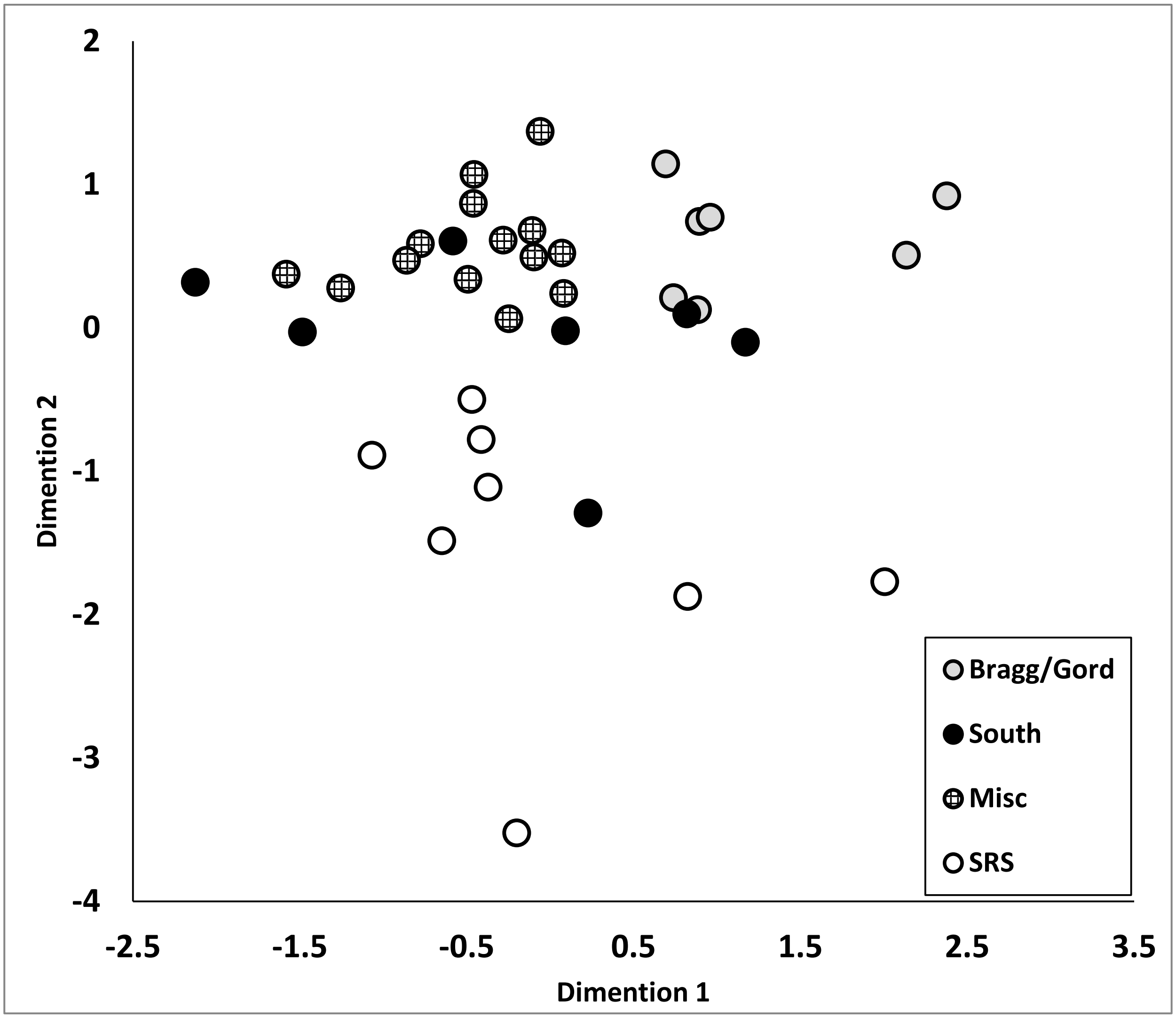


Figure 46. Nonmetric multi-dimensional scaling of primary and secondary reference sites across the Sand Hills Ecoregion. Sites with the same color pattern were grouped together from the cluster analysis.

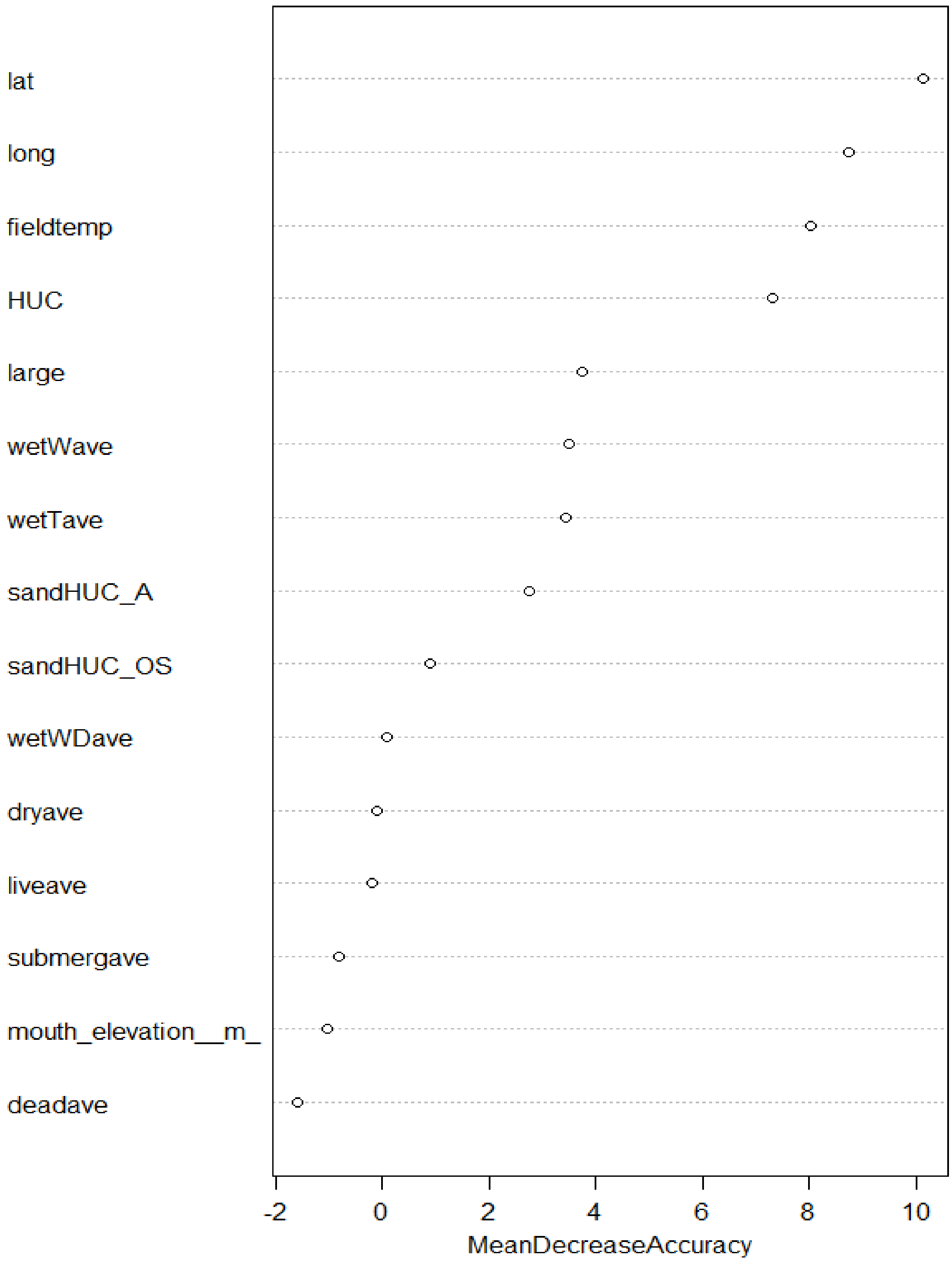


Figure 47. Importance of the simple predictor set for primary reference models based on Random Forests analysis. Variables with the highest mean decrease accuracy were the most important predictors. This set of variables represents a set of predictors that are easily acquired in the field and from basic maps.

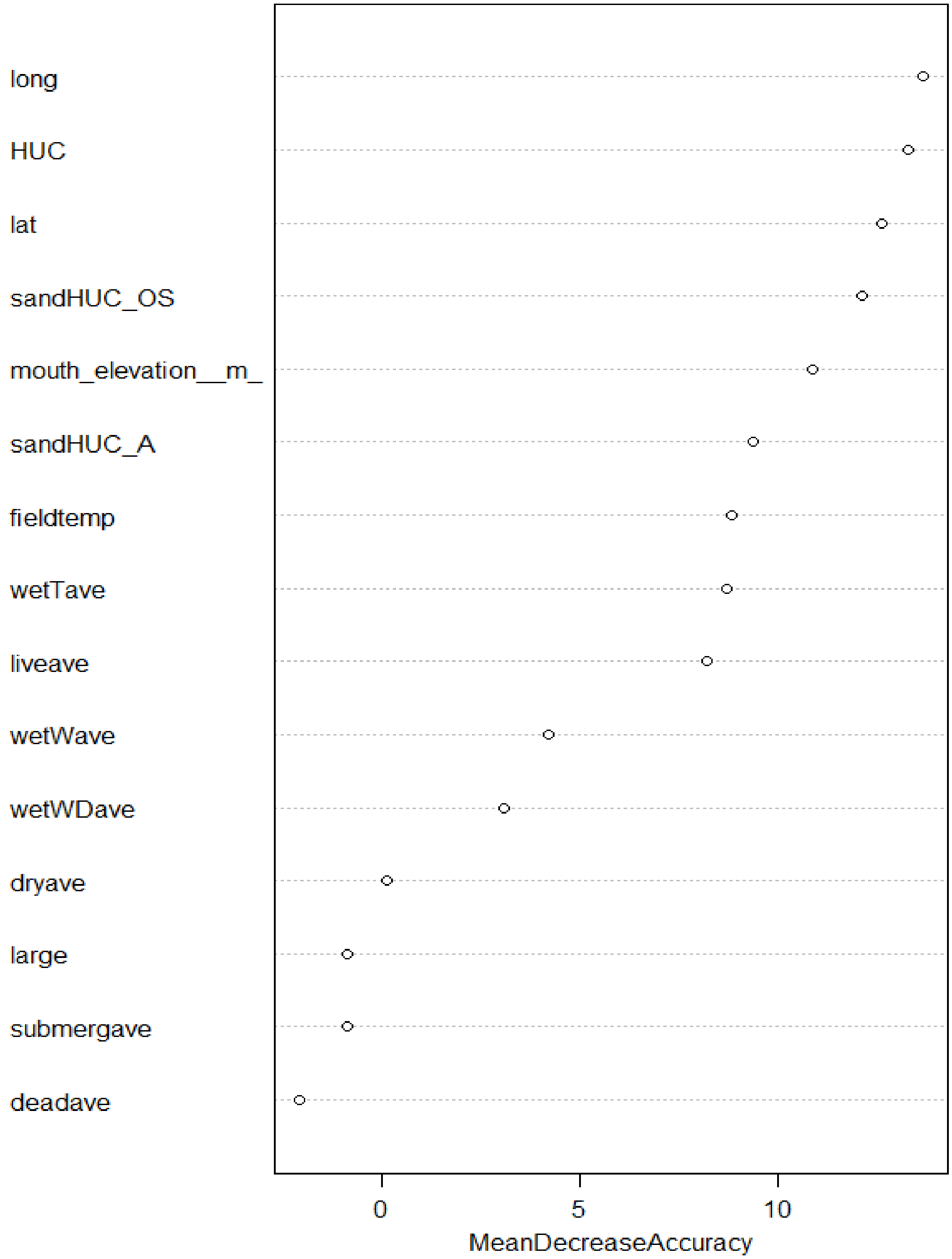


Figure 48. Importance of the simple set of predictors for the primary and secondary reference models based on Random Forests analysis. Variables with the highest mean decrease accuracy were the most important predictors. This set of variables represents a set of predictors that are easily acquired in the field and from basic maps.

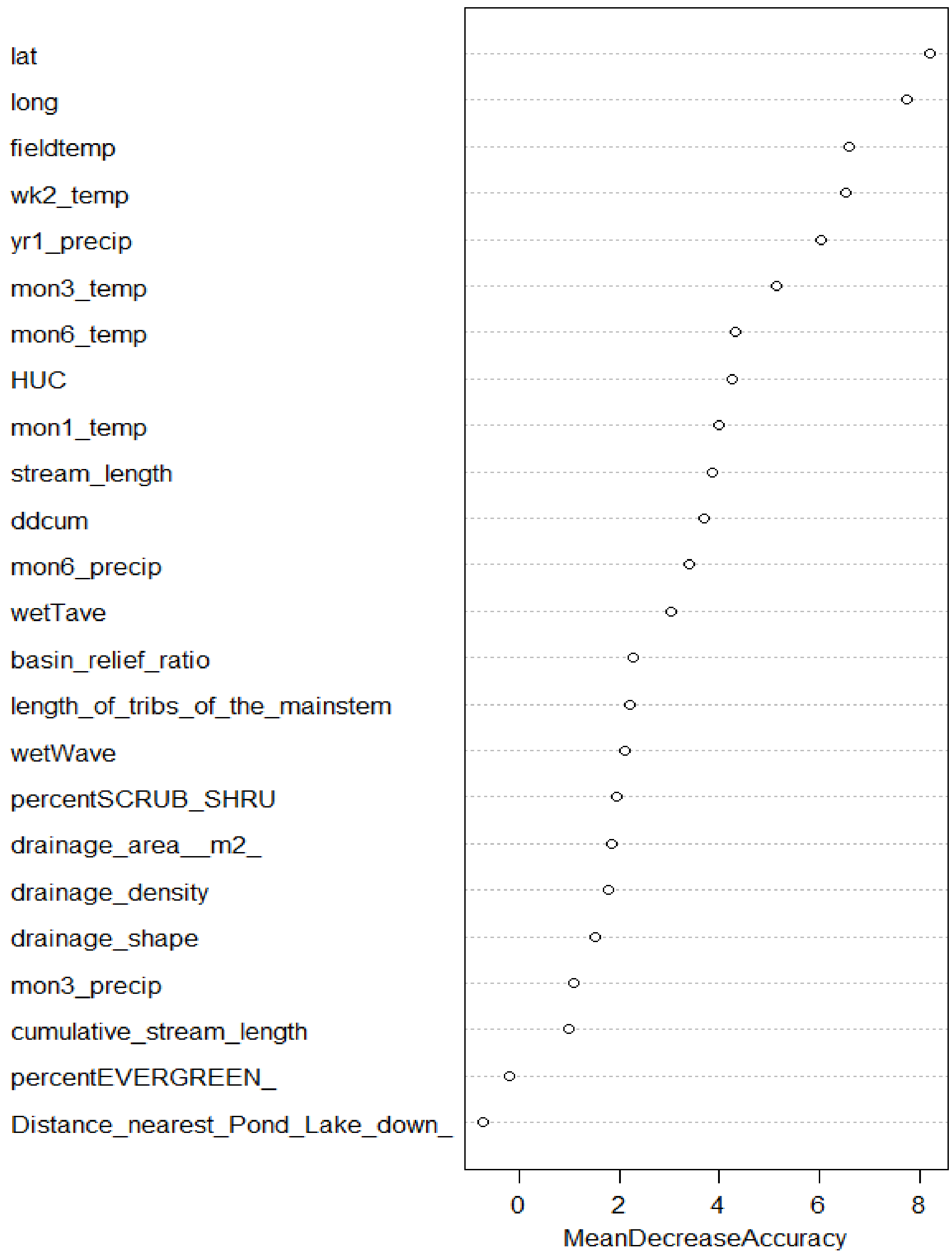


Figure 49. Importance of the expanded set of predictors for the primary reference models based on Random Forests analysis of 38 predictors. Variables with the highest mean decrease accuracy were the most important predictors. This set of variables represents a suite of predictors taken from measurements in the field, GIS, and analysis of climate data.

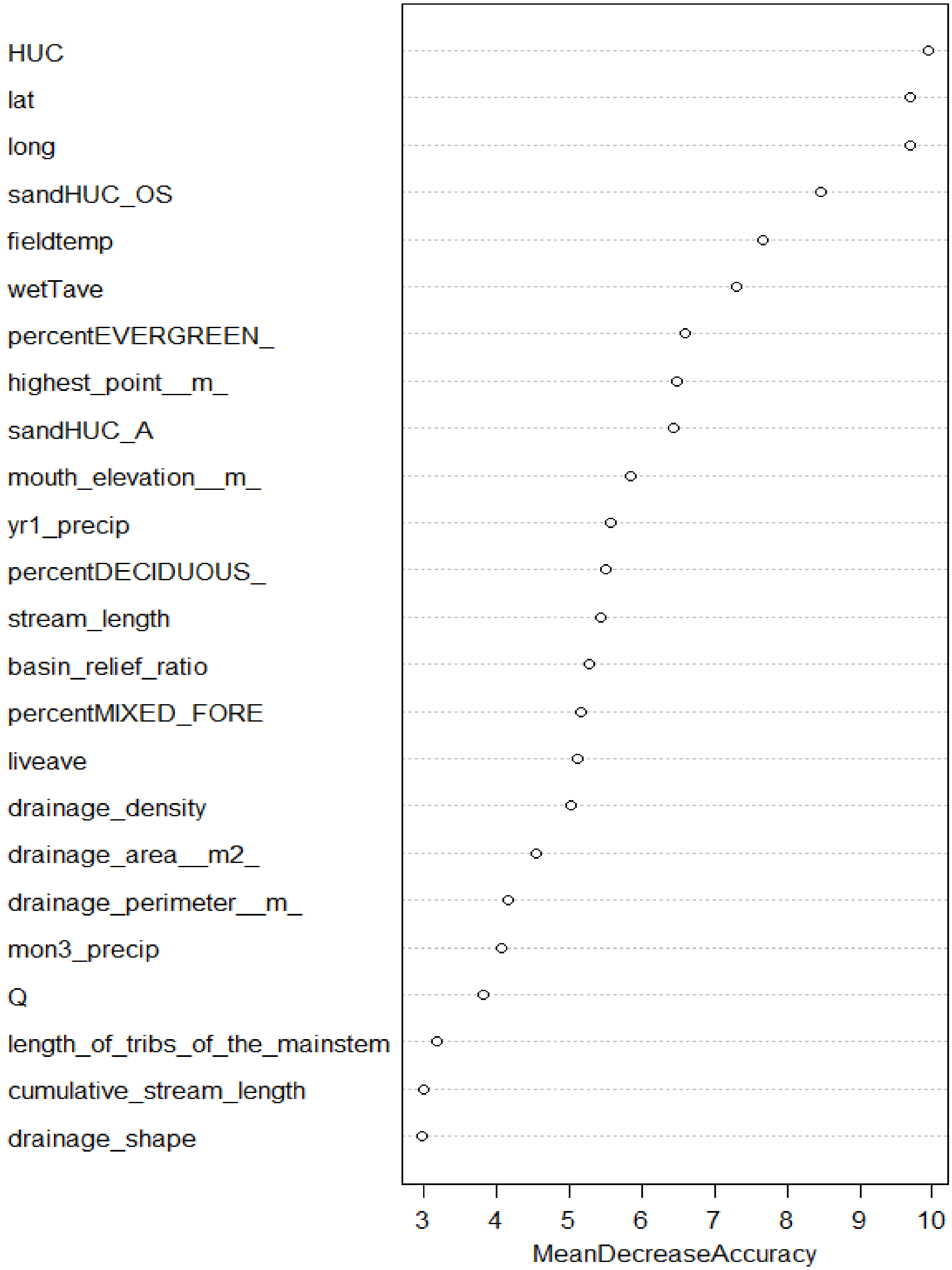


Figure 50. Importance of the expanded set of predictors for the primary and secondary reference models, based on Random Forests analysis of 38 predictors. Variables with the highest mean decrease accuracy were the most important predictors. This set of variables represents a suite of predictors taken from measurements in the field, GIS, and analysis of climate data.

(Kosnicki and Sites, 2007; but see Jessup et al., 2006). Intuitively, the highest quality sites should show higher richness of Orthocladiinae and lower Tanypodinae to receive the highest composition scores. It is also logical that the composition metric should be based on a combination of Chironomidae derived metrics considering about 44% (118) of the 268 OTUs identified from the study sites were from this family (Appendix MM2).

Regarding metric redundancy, a strong correlation existed between *rich* and *clingerich* (Table 38) for primary and primary + secondary data (0.70 and 0.69, respectively) but not between *rich* and *pclingrich* (0.15 and 0.20, respectively). *Rich* was the best choice for richness measures and thus, even though *clingrich* showed higher discrimination between test and reference sites, and displayed a higher S/N, *pclingrich* was incorporated into primary and primary + secondary models. For all other metric types, except diversity measures, no substantial correlations existed, and the best performing metric representative (as *t*-value) was selected for aggregation. None of the diversity measures showed a difference between least disturbed and disturbed sites and thus were excluded from inclusion in the *MMIs*. Final aggregation metric scoring criteria and *MMI* score for primary reference model are as follows:

$$BI_{score} = \frac{20(6.9 - BI)}{(6.9 - 3.8)} \quad (9)$$

$$rich_{score} = \frac{20(rich - 53)}{(86 - 53)} \quad (10)$$

$$pshredrich_{score} = \frac{20(pshredrich - 3.9)}{(11.6 - 3.9)} \quad (11)$$

$$pclingrich_{score} = \frac{20(pclingrich - 11.1)}{(25 - 11.1)} \quad (12)$$

$$orthchirich_{score} = \frac{10(orthchirich - 24)}{(44.4 - 24)} \quad (13)$$

$$Tpodrich_{score} = \frac{10(34.1 - Tpodrich)}{(34.1 - 14.7)} \quad (14)$$

$$MMI_{score} = BI_{score} + rich_{score} + pshredrich_{score} + pclingrich_{score} + orthchirich_{score} + Tpodrich_{score} \quad (15)$$

For the primary + secondary reference metric scoring criteria, final aggregation scoring criteria and *MMI* scores are as follows:

$$BI_{score} = \frac{20(6.9 - BI)}{(6.9 - 3.7)} \quad (16)$$

$$rich_{score} = \frac{20(rich - 53)}{(99 - 53)} \quad (17)$$

$$pshredrich_{score} = \frac{20(pshredrich - 3.9)}{(11.6 - 3.9)} \tag{18}$$

$$pclingrich_{score} = \frac{20(pclingrich - 11.1)}{(25 - 11.1)} \tag{19}$$

$$orthchirich_{score} = \frac{10(orthchirich - 24)}{(44.4 - 24)} \tag{20}$$

$$Tpodrich_{score} = \frac{10(34.1 - Tpodrich)}{(34.1 - 14.3)} \tag{21}$$

$$MMI_{score} = BI_{score} + rich_{score} + pshredrich_{score} + pclingrich_{score} + orthchirich_{score} + Tpodrich_{score} \tag{22}$$

5.3.3.3 Evaluation of Models

Score evaluation – The LULC sub-catchment gradient defined by PCA1 accounted for 34.4% of the variation across all 64 sites and was strongly driven by percentage of developed land and paved road density (Table 40). In general the sub-catchment gradient appeared unbalanced because few sites with large amounts of developed land were sampled. This disparity was especially true for models built with primary references because a large number of secondary reference sites were used to evaluate these models. As a result, models were transformed by 1/x for evaluation with this gradient. After transformation models were significantly related to the sub-catchment gradient in the expected direction (i.e., positive with BC and BC_{50} and negative for all other models) (Figs. 51-56). The dissolved oxygen measures were normally distributed and therefore no transformations were conducted. All models were significantly related to the dissolved oxygen gradient in the expected direction (i.e., negative with BC and BC_{50} and positive for all other models) (Figs. 57-62).

Both *MMI* models passed the unbalanced ANOVA evaluation. RIVPACS-type models built with primary reference sites all passed the unbalanced ANOVA evaluation, except for the *BC* model built with the simple set of predictors. Both *OE* models were less sensitive than all other primary reference models (lower *F*-value, Table 41). Most of the RIVPACS-type models built with primary + secondary reference sites did not pass the unbalanced ANOVA evaluation (Table 42). The BC and BC_{50} built with both simple and expanded predictor sets could not discriminate test from validation sites but did discriminate validation from training sites. This difference suggests these models were made less robust by the addition of secondary reference sites; if true, they are perhaps best when built with only high-quality reference sites. Both *OE* models passed the screening criteria whereas OE_{50} models did not discriminate between test and validation sites. Only the OE_{CAP} built with the expanded predictor set passed the unbalanced ANOVA evaluation. It is possible that the OE_{50} models are more sensitive to disturbances as has been shown elsewhere (Hawkins et al., 2000; Van Sickle et al., 2007) and place a high demand on reference site quality. In this context, use of training sites to build OE_{50} models that are not of the highest quality may hinder their ability to detect impairment. *OE* models allow for inclusion of taxa with occurrence frequencies <0.5 so more taxa are considered in the assessment; and, although not as sensitive in responding to disturbance, their capability to detect differences (at least in this study) with inclusion of secondary reference sites suggest that they may be useful for assessing sites based on a less than optimal benchmark.; e.g., a best attainable model (BAM). However, our OE_{CAP} built from the expanded predictor set may be a better BAM, as it showed better promise in the final screening (see below).

Table 40. Principal Component Analysis (PCA) loadings used for land use and land cover (LULC) variables defining sub-catchments. PCA axis 1 & 2 accounted for 34.4 & 24.1% of the variation, respectively. PCA axes are the rotated factor patterns. LULC variables defined in Table 39.

LULC variable	PCA-1	PCA-2
Drainage area	0.07	0.92
% Deciduous	-0.01	0.58
% Evergreen	0.03	0.75
% Mixed forest	0.04	0.87
% High-intensity developed	0.72	0.06
% Medium intensity developed	0.86	0.18
% Low intensity developed	0.92	0.19
% Developed	0.88	-0.05
Paved road density	0.59	-0.01
Unpaved road density	0.20	<0.00

The 13 models (RIVPACS-type and MMI, Table 43) were finally screened for their ability to discriminate test sites that were disturbed from reference sites. Both secondary reference *OE* models tended to give "Good" ratings to test sites as well as reference sites compared with the OE_{CAP}, which performed somewhat better (Figure 63). Our analysis indicated that the *MMI* built with primary + secondary references included too many test sites as "Good" quality and was not considered as a final reference model. The *BC* and BC_{50} models built with primary references were ineffective at identifying sites representative of "Poor" habitat conditions (Figure 63) and were weak at identifying primary reference sites as representative of "Good" condition. These models tended to clump sites in the middle and evaluate them as "Fair" regardless of their reference type.

The remaining primary reference models were equivalent in evaluating test sites as "Poor" or "Very Poor" but showed some discrepancies in giving a "Good" rating for test and primary reference sites. For simple predictor models, *OE* had a slight tendency to give a "Good" rating to test sites and reference sites over the OE_{CAP} model; and, at our discretion, we favored the later model for its tendency to be more conservative. For the remaining expanded predictor primary reference models, the *OE* was considered too liberal at evaluating test sites as "Good" quality habitats. OE_{CAP} tended to underrepresent the primary reference sites as "Good" quality whereas the OE_{50} evaluated the highest percentage of reference sites as "Good" quality and was retained as the expanded predictor primary reference model.

The primary reference *MMI* was the best model overall at identifying test sites as "Poor" and the only model to designate some test sites as "Very Poor" in habitat quality (Figure 63). It was also the only model to give a primary reference site a "Poor" rating, and many reference sites were not evaluated as highest quality. This pattern could be due to the implementation of the model over the whole region without classification. Streams from some regions may not be equivalent to streams from other regions, and the *MMI* does not compensate for these differences. The construction of *MMI* for a region is usually done *a priori* as we have done here at the level IV ecoregion. However, even within the Sandhills, it is

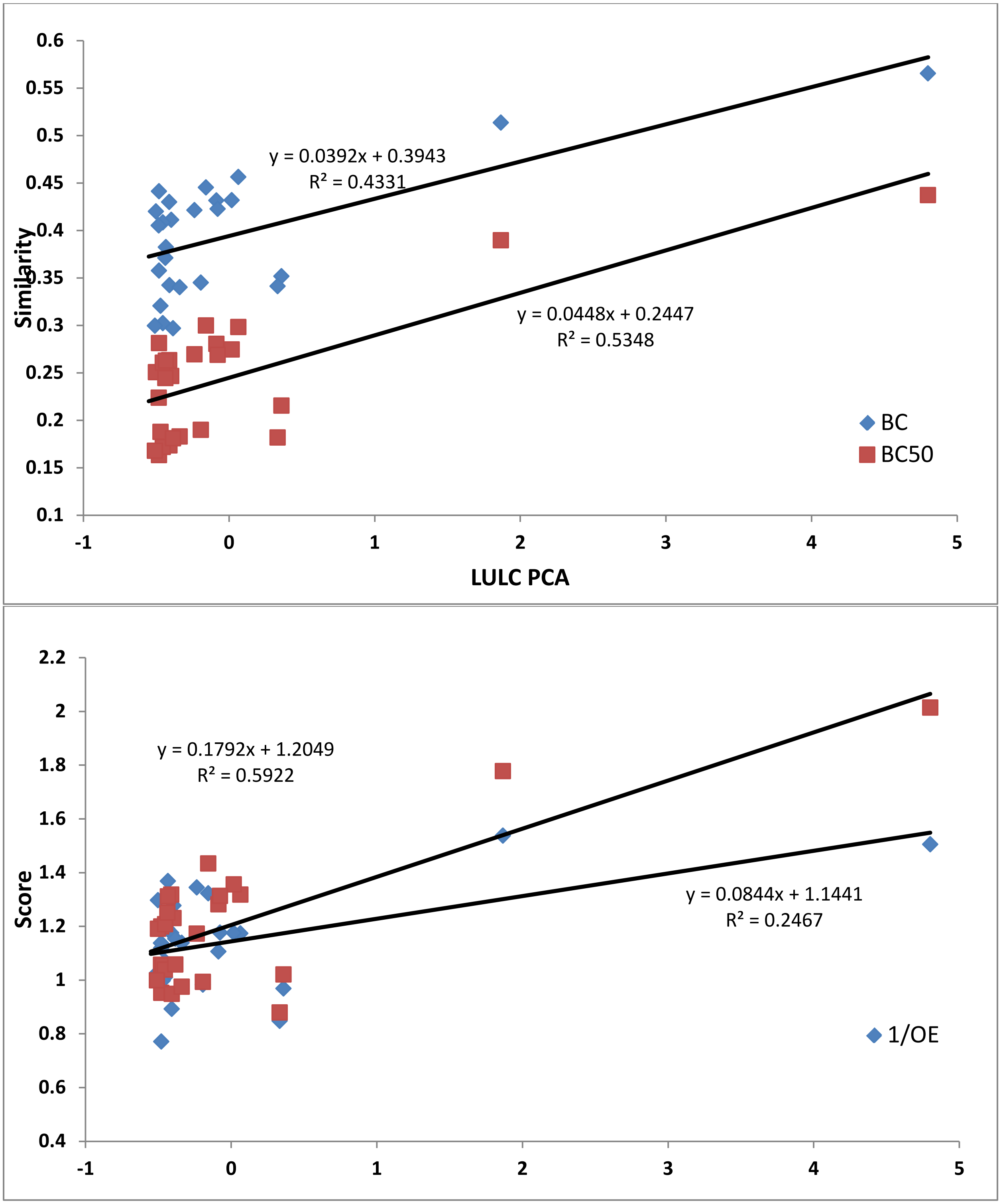
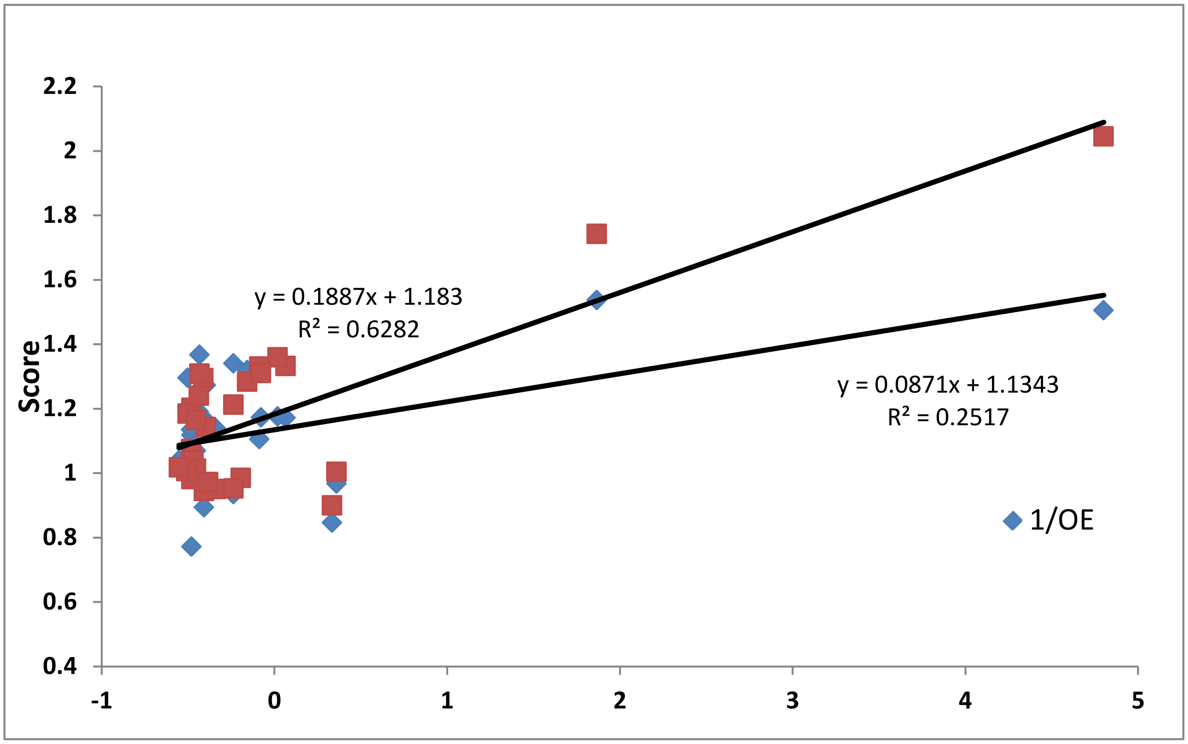


Figure 51. Relationship between land use/land cover (LULC) PCA axis-1 and 1/*x* transformed macroinvertebrate O/E scores (top panel) and LULC and Bray Curtis similarity (bottom panel) including repeat, secondary, and tertiary reference sites. Regression lines with equations and R^2 values indicate significant relationships ($\alpha = 0.05$). All models were built with primary reference samples and the simple predictor set.



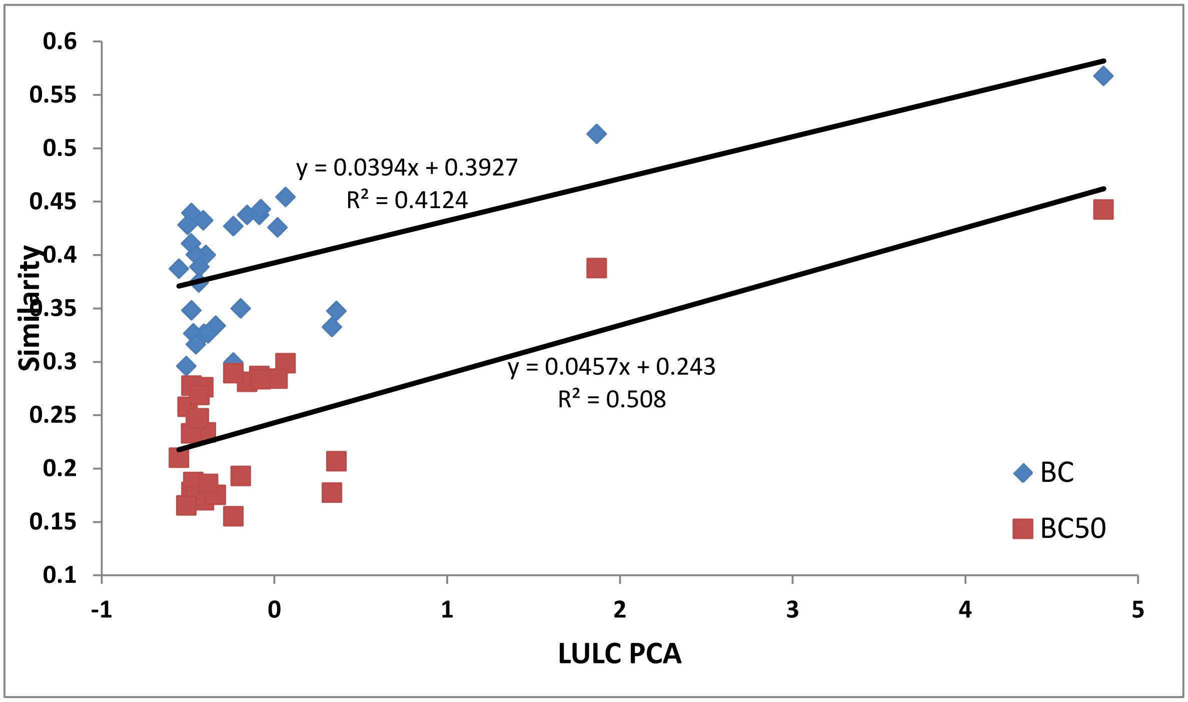


Figure 52. Relationship between land use/land cover (LULC) PCA axis-1 and 1/x transformed macroinvertebrate O/E scores (top panel) and Bray Curtis similarity (bottom panel) including repeat, secondary, and tertiary reference sites. Regression lines with equations and R^2 values indicate significant relationships ($\alpha = 0.05$). All models were built with primary reference samples and the expanded predictor set.

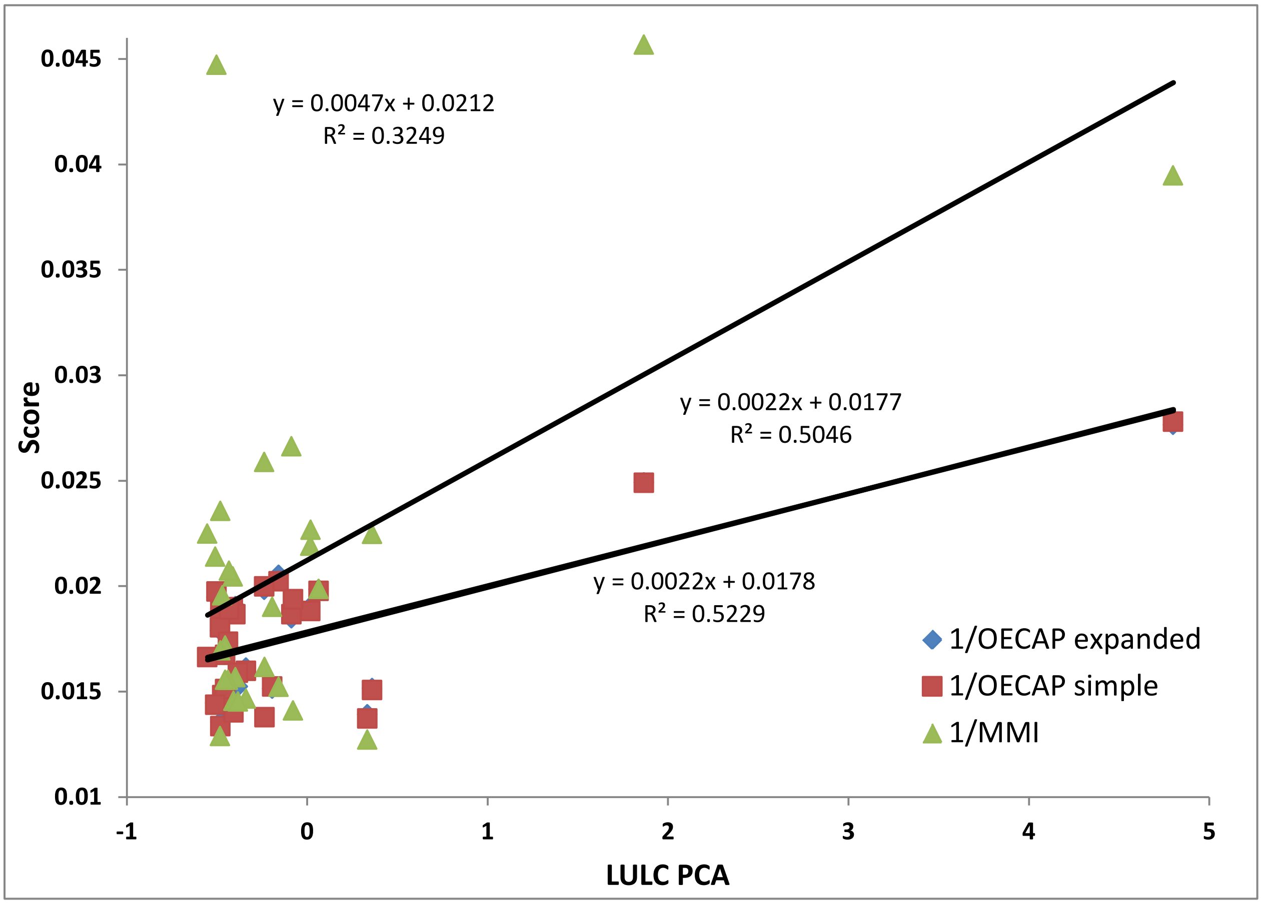


Figure 53. Relationship between land use/land cover (LULC) PCA axis-1 and macroinvertebrate simple and expanded OE_{CAP} and macroinvertebrate multi-metric index (*MMI*) scores including repeat, secondary, and tertiary reference sites. Regression lines with equations and R^2 values indicate significant relationships ($\alpha = 0.05$). All models were built with primary reference samples. All scores were $1/x$ transformed.

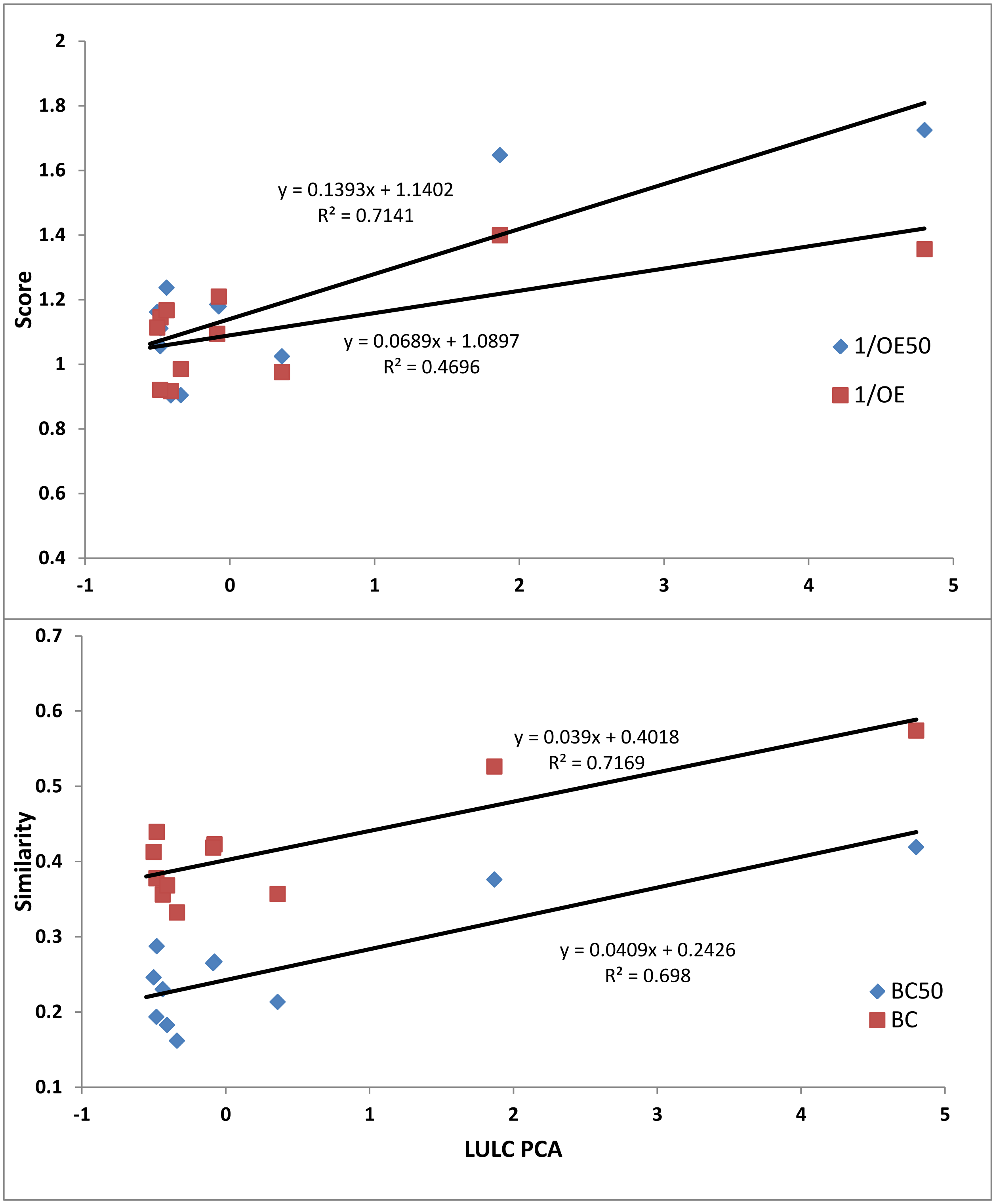


Figure 54. Relationship between land use/land cover (LULC) PCA axis-1 and 1/x transformed macroinvertebrate O/E scores (top panel) and Bray Curtis similarity (bottom panel) including repeat and tertiary reference sites. Regression lines with equations and R^2 values indicate significant relationships (α = 0.05). All models were built with primary and secondary reference samples and the simple predictor set.

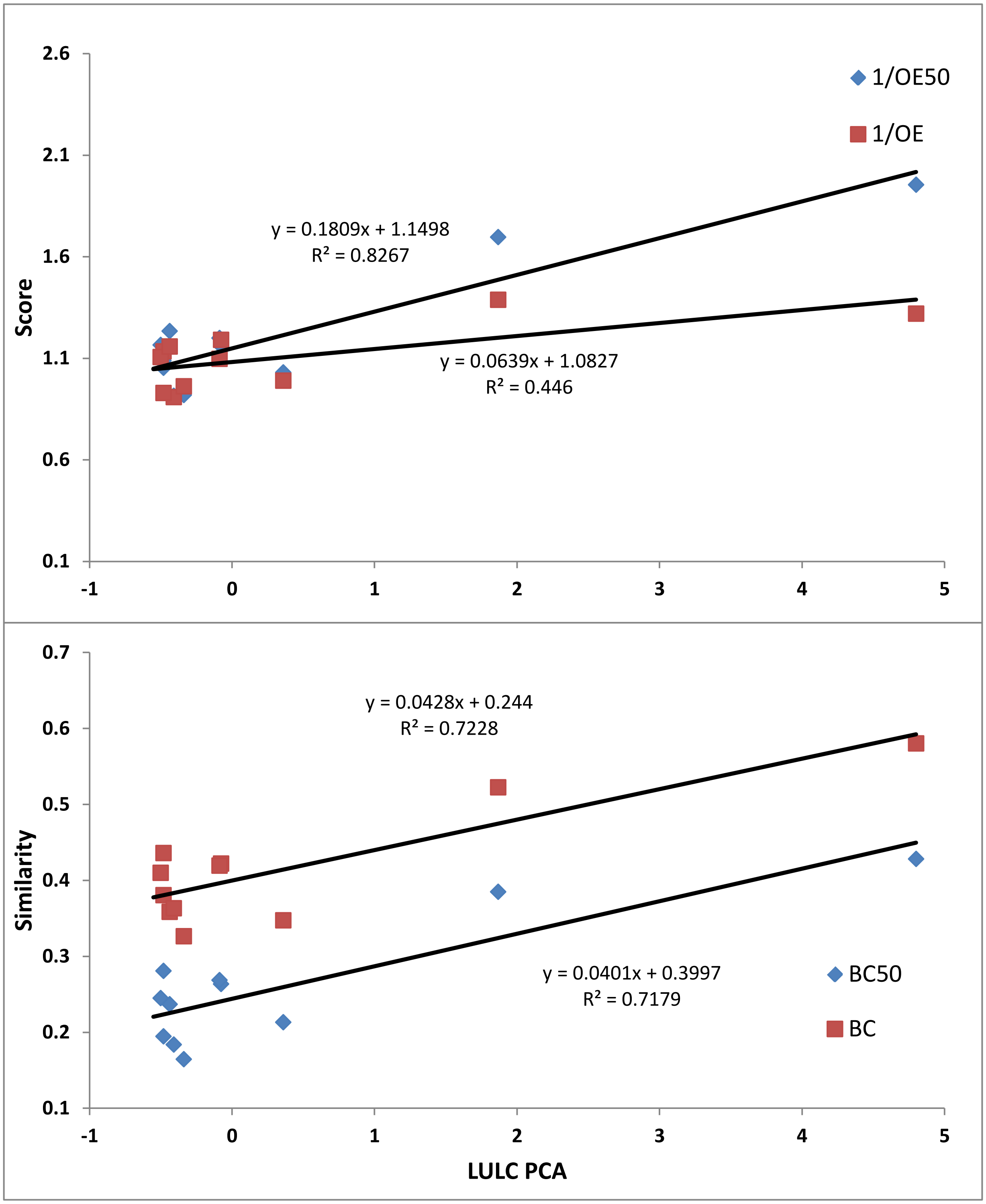


Figure 55. Relationship between land use/land cover (LULC) PCA axis-1 and 1/*x* transformed macroinvertebrate O/E scores (top panel) and Bray Curtis similarity (bottom panel) including repeat and tertiary reference sites. Regression lines with equations and R^2 values indicate significant relationships ($\alpha = 0.05$). All models were built with primary and secondary reference samples and the expanded predictor set.

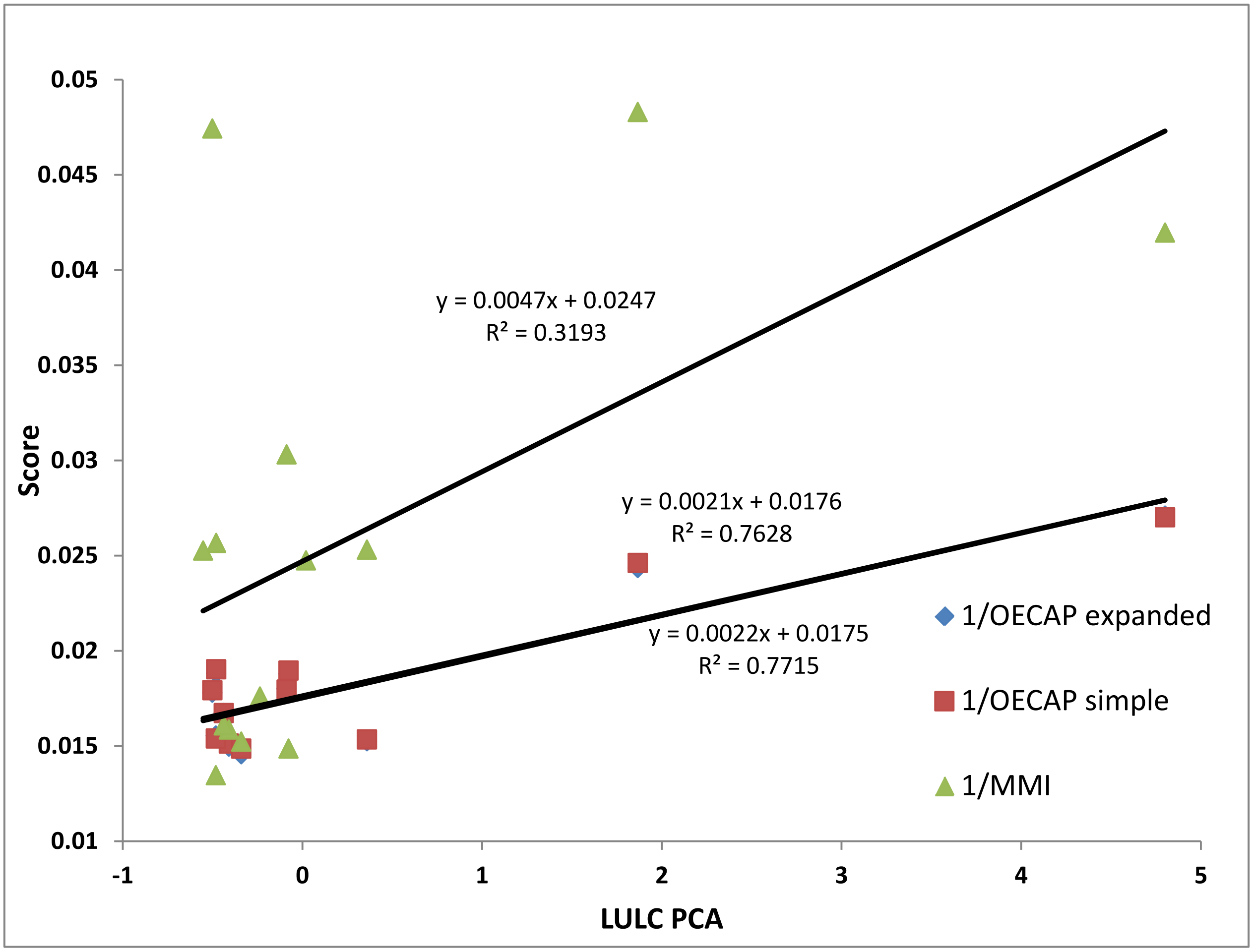


Figure 56. Relationship between land use/land cover (LULC) PCA axis-1 and macroinvertebrate simple and expanded OE_{CAP} and macroinvertebrate multi-metric index (*MMI*) scores including repeat and tertiary reference sites. Regression lines with equations and R^2 values indicate significant relationships ($\alpha = 0.05$). All models were built with primary and secondary reference samples. All scores were 1/*x* transformed.

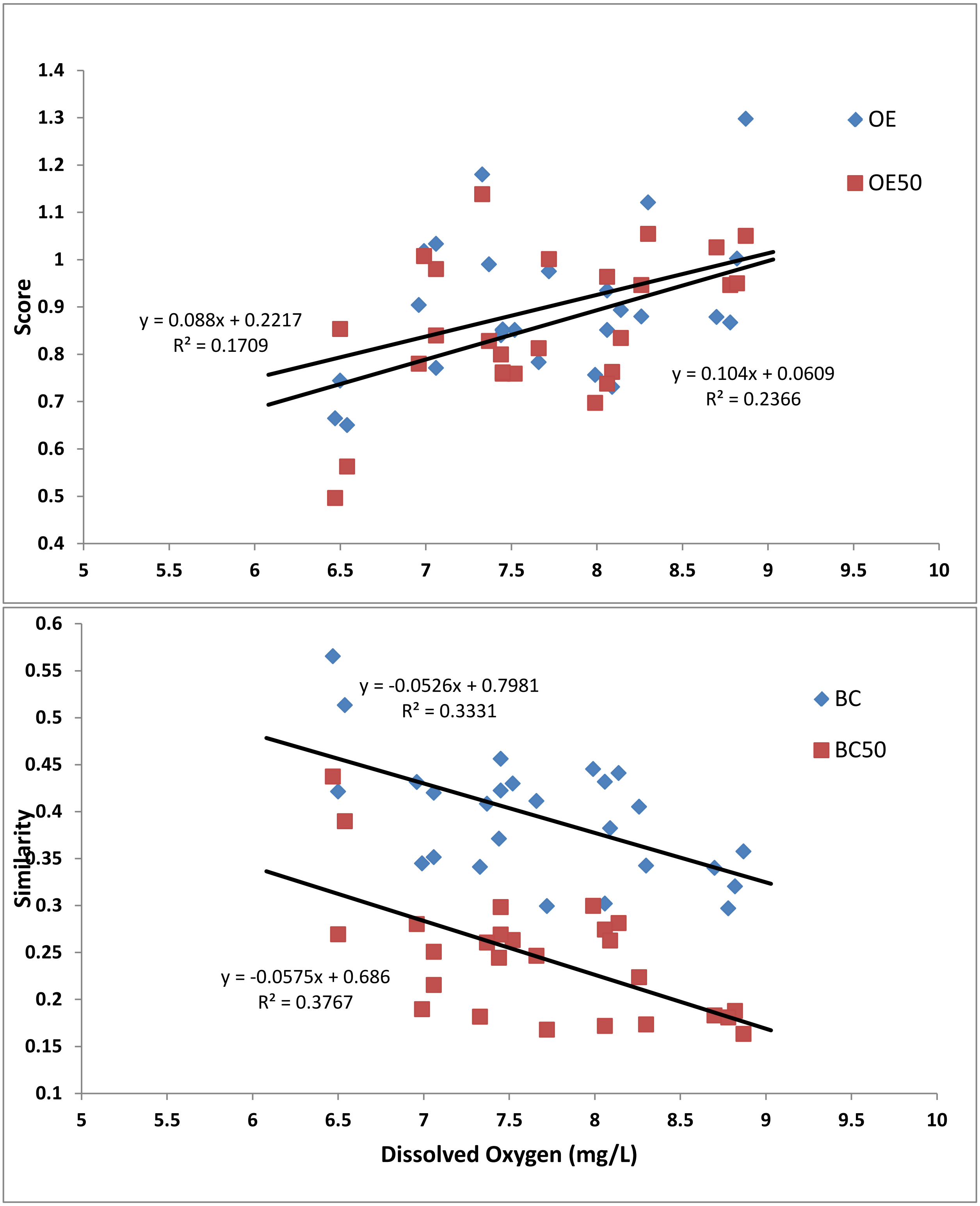


Figure 57. Relationship between site dissolved oxygen and macroinvertebrate O/E scores (top panel) and Bray Curtis similarity (bottom panel) including repeat, secondary, and tertiary reference sites. Regression lines with equations and R^2 values indicate significant relationships ($\alpha = 0.05$). All models were built with primary reference samples and the simple predictor set.

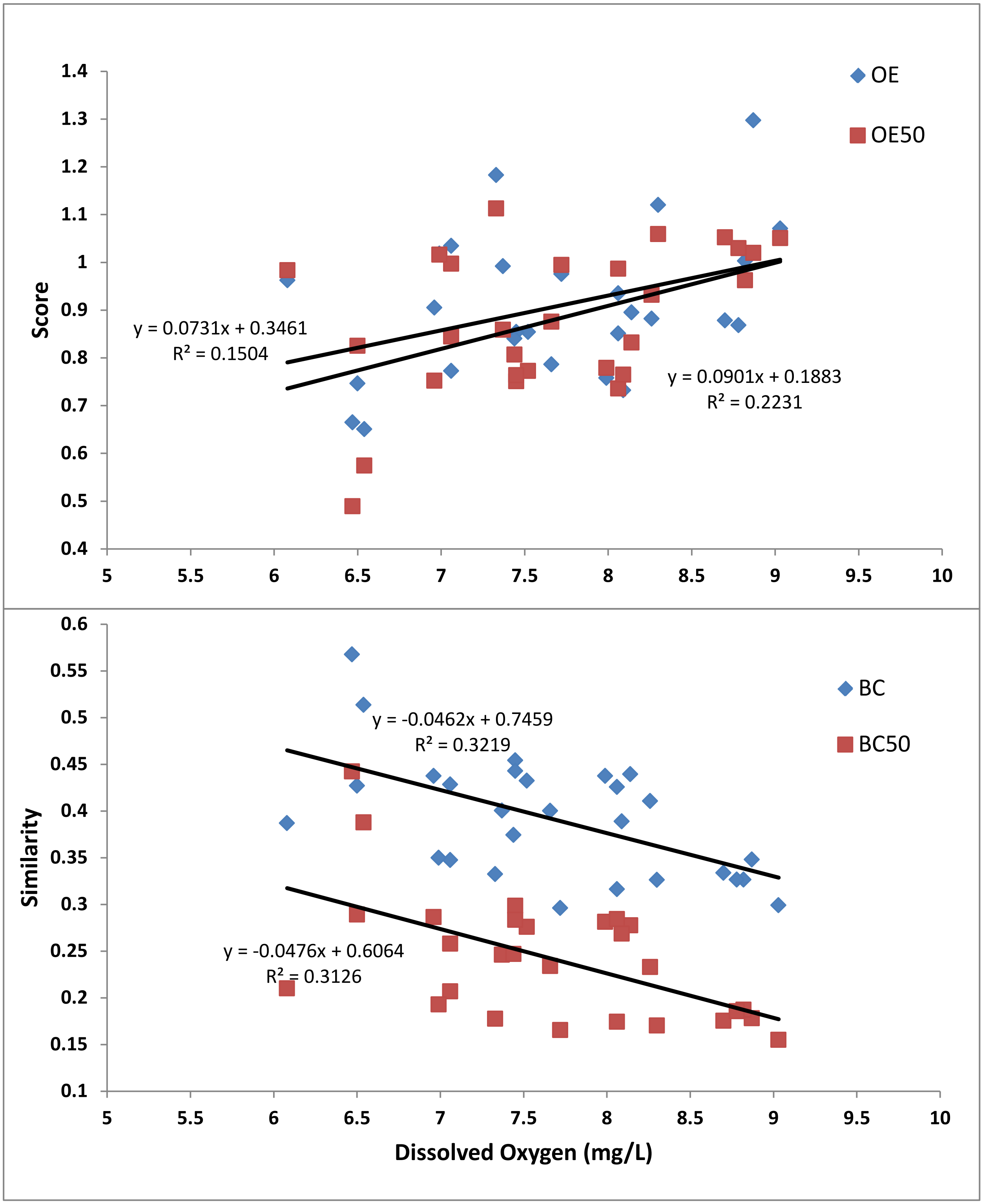


Figure 58. Relationship between site dissolved oxygen and macroinvertebrate O/E scores (top panel) and Bray Curtis similarity (bottom panel) including repeat, secondary, and tertiary reference sites. Regression lines with equations and R^2 values indicate significant relationships ($\alpha = 0.05$). All models were built with primary reference samples and the expanded predictor set.

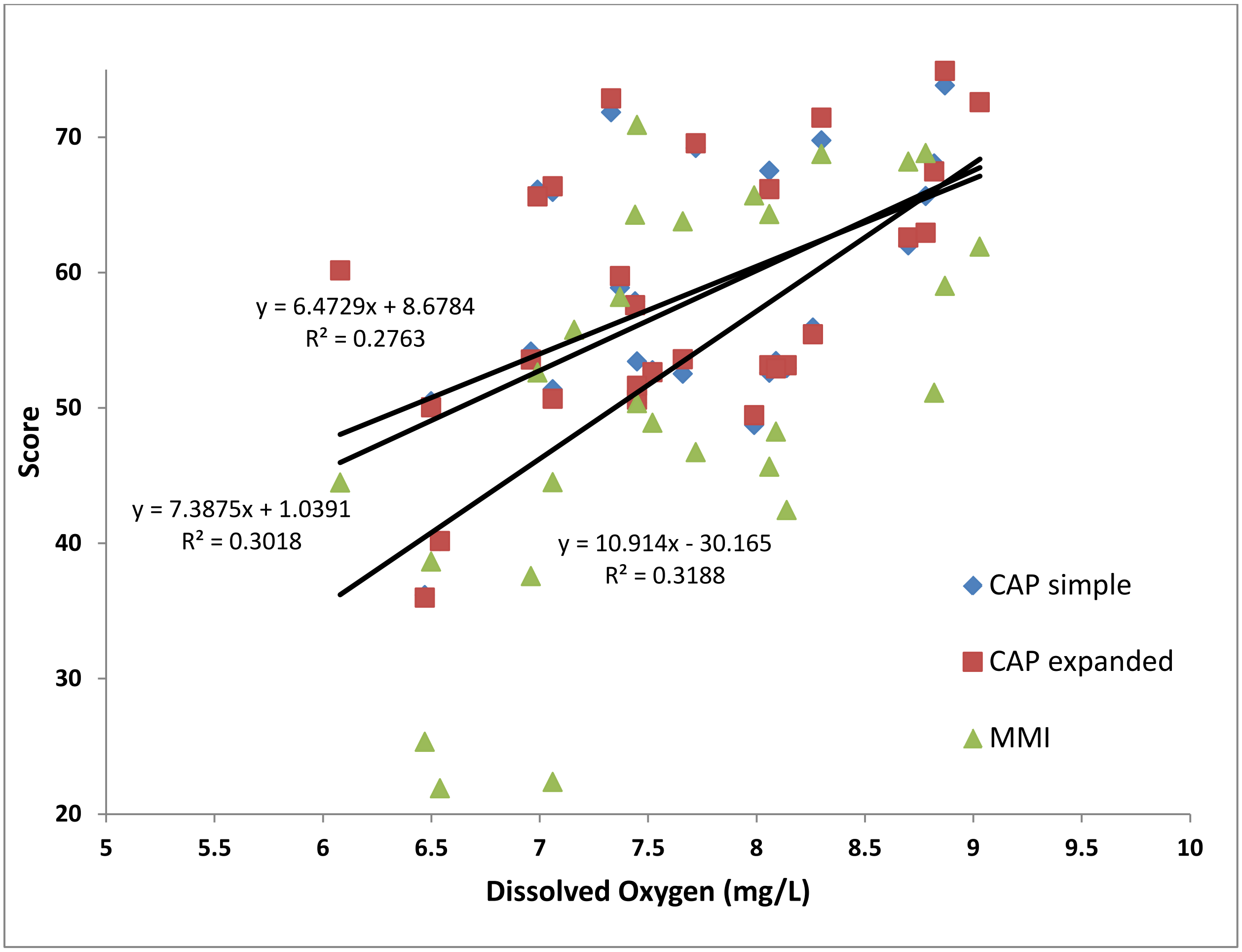


Figure 59. Relationship between site dissolved oxygen and macroinvertebrate simple and expanded OE_{CAP} and multi-metric index model (*MMI*) scores including repeat, secondary, and tertiary reference sites. Regression lines with equations and R^2 values indicate significant relationships ($\alpha = 0.05$). All models were built with primary reference samples.

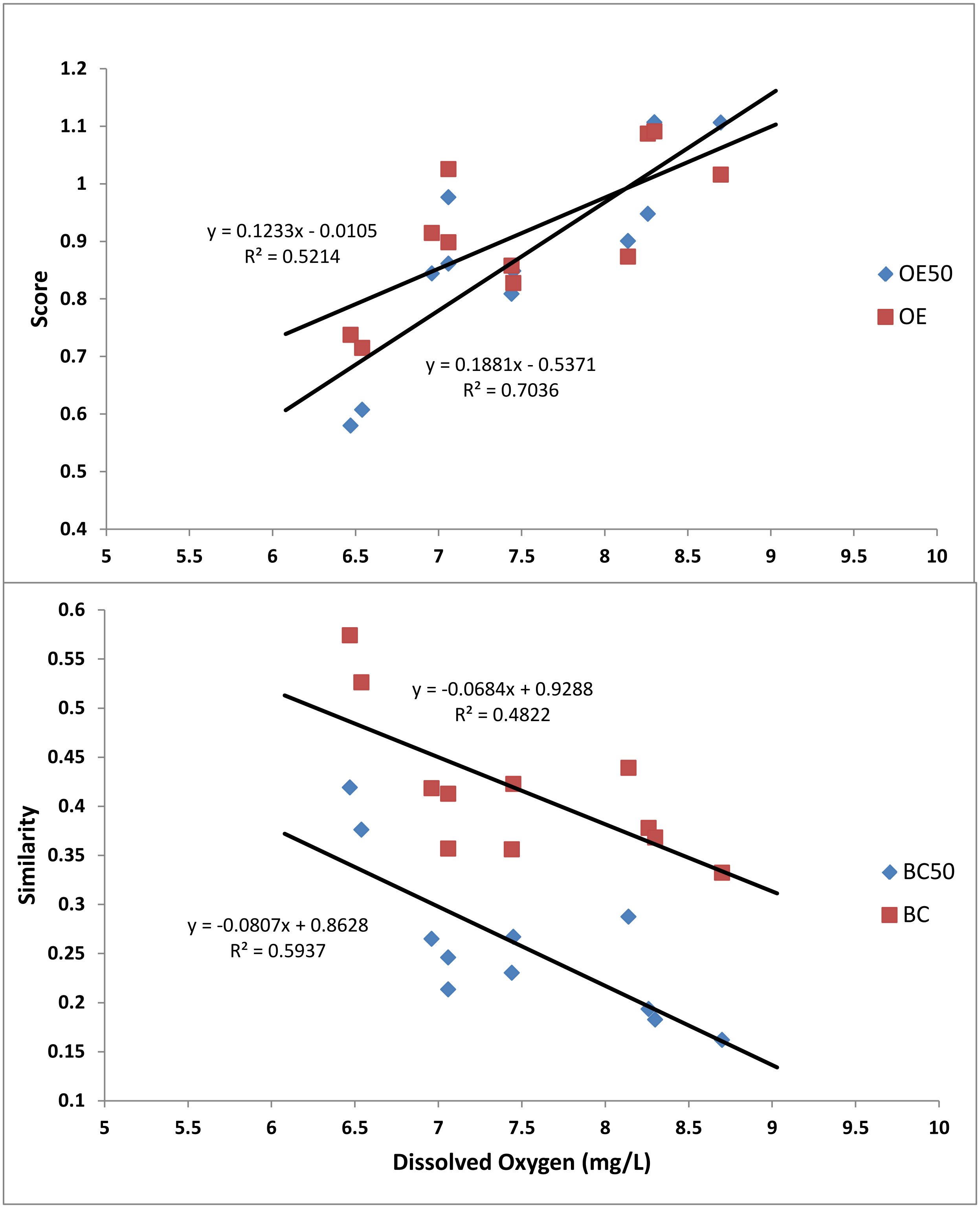


Figure 60. Relationship between site dissolved oxygen and macroinvertebrate O/E scores (top panel) and Bray Curtis similarity (bottom panel) including repeat and tertiary reference sites. Regression lines with equations and R^2 values indicate significant relationships ($\alpha = 0.05$). All models were built with primary and secondary reference samples and the simple predictor set.

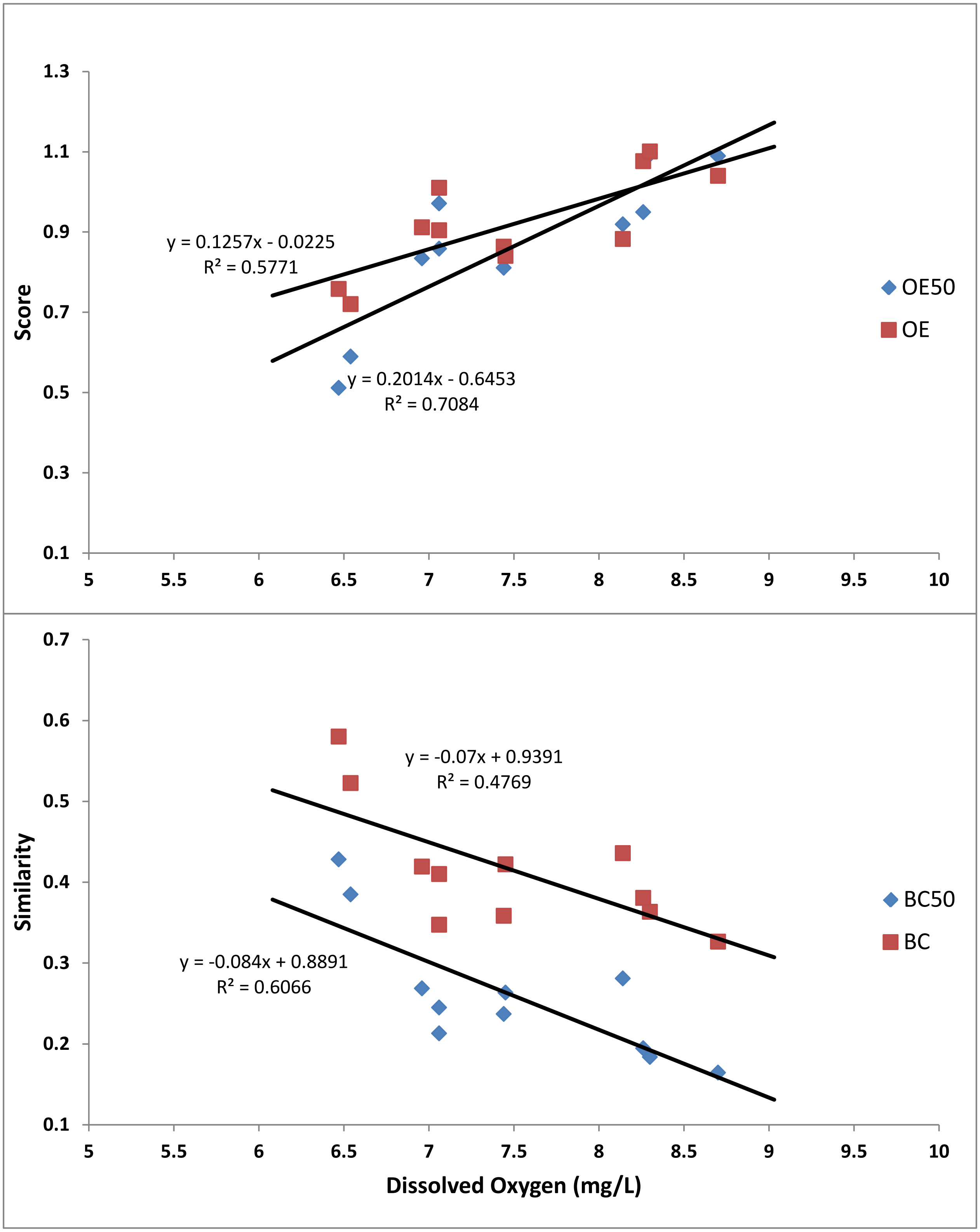


Figure 61. Relationship between site dissolved oxygen and macroinvertebrate O/E scores (top panel) and Bray Curtis similarity (bottom panel) including repeat and tertiary refernece sites. Regression lines with equations and R^2 values indicate significant relationships. All models were build with primary and secondary reference samples and the expanded predictor set.

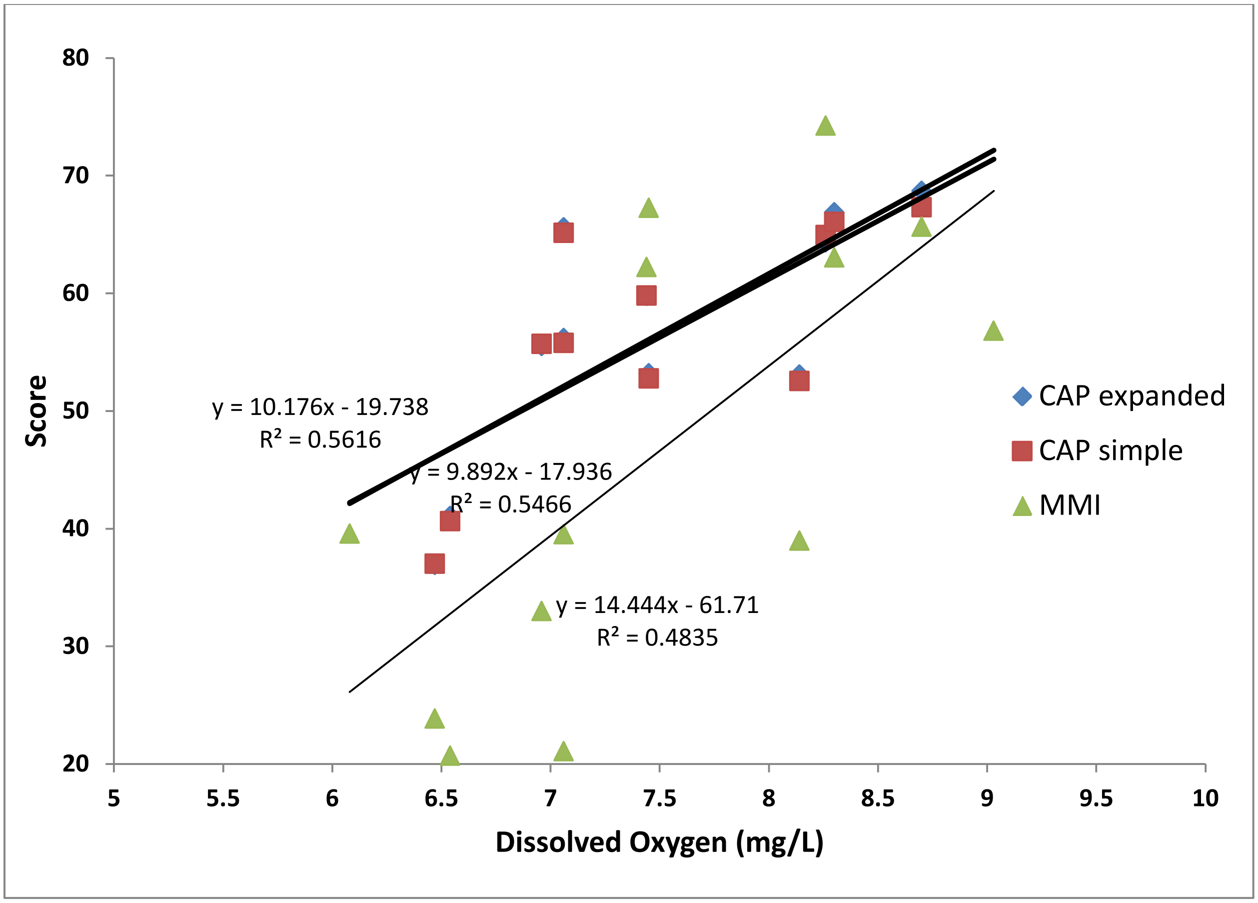


Figure 62. Relationship between site dissolved oxygen and macroinvertebrate simple and expanded OE_{CAP} and multi-metric index (MMI) scores including repeat and tertiary reference sites. Regression lines with equations and R^2 values indicate significant relationships ($\alpha = 0.05$). All models were built with primary and secondary reference samples.

Table 41. Analysis of variance (ANOVA) and gradient analysis results for models built with primary reference sites. Mean separation indicates 'yes' when the mean of training (C) and validation (V) sites are both significantly different than means of test (T) sites and not significantly different from each other, otherwise indicated. Test gradients indicate significant correlation with land use/land cover PCA axis-1 (LULC) and dissolved oxygen (DO) from Figures 51-62.

	No. of sites			ANOVA results				Test gradients	
Models	Training	Test	Validation	F	R^2	p	Mean separation	LULC	DO (mg/L)
Simple predictor set									
OE	21	13	4	8.30	0.32	0.0011	yes	yes	yes
OE_{50}	21	13	4	24.6	0.58	< 0.0001	yes	yes	yes
BC	21	13	4	34.33	0.66	< 0.0001	V ≠ C	yes	yes
BC_{50}*	21	13	4	31.18	0.64	< 0.0001	yes	yes	yes
OE_{CAP}	21	13	4	25.03	0.59	< 0.0001	yes	yes	yes
Expanded predictor set									
OE	21	14	4	9.16	0.34	0.0006	yes	yes	yes
OE_{50}	21	14	4	23.31	0.56	< 0.0001	yes	yes	yes
BC	21	14	4	44.53	0.71	< 0.0001	yes	yes	yes
BC_{50}*	21	14	4	30.83	0.63	< 0.0001	yes	yes	yes
OE_{CAP}	21	14	4	30.05	0.63	< 0.0001	yes	yes	yes
No classification									
MMI	21	15	4	23.55	0.56	< 0.0001	yes	yes	yes

* model mean scores $\log_{10}$ transformed for ANOVA

Table 42. Analysis of variance (ANOVA) and gradient analysis results for models built with primary and secondary reference sites. Mean separation indicates 'yes' when the mean of training (C) and validation (V) sites are both significantly different than means of test (T) sites and not significantly different from each other, otherwise indicated. Test gradients indicate significant correlation with land use/land cover PCA axis-1 (LULC) and dissolved oxygen (DO) from Figures 51-62.

	No. of sites			ANOVA results				Test gradients	
Models	Training	Test	Validation	F	R^2	p	Mean separation	LULC	DO (mg/L)
Simple predictor set									
OE	36	13	6	5.72	0.183	0.0057	yes	yes	yes
OE50*	36	13	6	13.20	0.37	< 0.0001	T=V & V ≠ C	yes	yes
BC	36	13	6	33.63	0.56	< 0.0001	T=V & V ≠ C	yes	yes
BC50*	36	13	6	25.00	0.49	< 0.0001	T=V & V ≠ C	yes	yes
OECAP	36	13	6	21.77	0.46	< 0.0001	T=V	yes	yes
Expanded predictor set									
OE	36	13	6	6.86	0.21	0.0023	yes	yes	yes
OE50	36	13	6	13.05	0.33	< 0.0001	T=V	yes	yes
BC	36	13	6	34.19	0.57	< 0.0001	T=V & V ≠ C	yes	yes
BC50*	36	13	6	22.16	0.46	< 0.0001	T=V & V ≠ C	yes	yes
OECAP	36	13	6	22.69	0.47	< 0.0001	yes	yes	yes
No classification									
MMI	37	15	6	23.40	0.46	< 0.0001	yes	yes	yes

* model mean scores $\log_{10}$ transformed for ANOVA

Table 43. Abbreviations of models selected for final evaluation based on score performance (see text). na=not applicable.

Model abbreviation	Type	Reference grade	Predictor set
BC_P3cE	BC	primary	expanded
BC50_P3cE	BC50	primary	expanded
BC50_PS	BC50	primary	simple
OE_P3cE	OE	primary	expanded
OE_P3cS	OE	primary	simple
OE_S4cE	OE	primary + secondary	expanded
OE_S4cS	OE	primary + secondary	simple
OE50_P3cE	OE50	primary	expanded
CAP_P3cE	OECAP	primary	expanded
CAP_P3cS	OECAP	primary	simple
CAP_secS	OECAP	primary + secondary	expanded
MMIprime	MMI	primary	na
MMIsec [b]	MMI	primary + secondary	na

[b] - did not show a response to LULC gradient

likely that there are differences between habitats of different regions based on a number of factors including geography, temporal effects close to the time of sampling, management differences, and site selection bias. The best models and their suggested scoring ranges for evaluating the biological integrity of streams in the SH Ecoregion are given in Table 44.

5.3.3.4 Summary for Macroinvertebrate Models

Use of the RIVPACS-type model ($OE_{CAP}E$) constructed with primary (highest quality) + secondary (moderate to high quality) reference sites appears useful for evaluating streams of the SH Ecoregion that are still undergoing recovery; however, this tool should be used with caution and should be identified as a BAM in its use. We consider the primary reference sites to represent “best of the best” conditions available in the SH, and that assessments based on these models are likely to give more stringent evaluations. Nevertheless, we feel that delineation of secondary reference sites shows promise as they display ecological conditions and qualities similar to primary reference sites. In many cases, macroinvertebrate assemblages were shown to be statistically similar for both reference sets based on metrics used in the North Carolina and Georgia biological monitoring programs (Kosnicki et al., submitted). The use of secondary reference sites may be appropriate when prevailing land uses or legacy effects preclude the attainment of conditions represented by primary reference sites. In such cases, secondary reference sites may be useful for the development of BAMs, and in this way BAMs may be useful, and perhaps better, at evaluating sites under recovery. Sustainable streams should be represented by a "balanced, integrated, adaptive community of organisms" (Karr, 1991) and we feel that our secondary reference condition sites achieve this goal, although, admittedly more testing of the efficacy of the predictive modeling approach should be done.

We recommend managers use these macroinvertebrate assessment tools as is suitable for their purposes. If resources are limited in obtaining habitat and landscape data, or if data from previous macroinvertebrate collections are available without habitat information, we recommend use of the *MMI* with the understanding that it may be the most stringent (conservative) in its evaluations. In this sense, an evaluation score of "Good" from our *MMI* is a useful indicator of sites representing the healthiest stream conditions in the SH Ecoregion. If tenable, managers also can potentially improve the efficacy of their evaluations by supplementing the *MMI* with one of the predictive models. As a minimum in this regard, use of primary reference OE_{CAP} built with the simple set of predictors offers an easy solution for managers who do not have the time or resources to obtain climate and land use data. Managers who can obtain data for all variables considered in this report will have access to the best performing macroinvertebrate models from this study, specifically, the primary reference OE_{50} built with the expanded predictor set. Moreover, use of the OE_{CAP} model built with primary + secondary reference sites will be most useful in evaluating attainment of sites in areas subjected to heavy uses such as military training, or for evaluating streams that are undergoing restoration. The highest quality rating from this model is an indication that the community has attained sustainable conditions, though possibly not of the highest quality compared to the region, and it should be reported that the results are based on a BAM assessment.

Obviously, these models are not equivalent and may result in different assessment results. We recommend that the *MMI* constructed with primary reference sites is the strictest and most appropriate model identifying the best quality streams comparable within the SH Ecoregion. However, as it may be restrictive, the OE_{50} built from the expanded predictor set appears the best approach at assigning the highest quality rating to our primary reference sites, although, it may not be best at identifying impairment. We therefore suggest the combined use of the *MMI* and OE_{50}. When assessments of these models agree, confidence can be given to that rating. When they disagree, further investigation involving other assessment tools developed in this study is warranted.

Table 44. Scoring ranges and suggested assessment evaluations for the macroinvertebrate multimetric index built with primary reference sites (*MMI*), predictive model built with primary reference sites and the simple set of predictors ($OE_{CAP}S$), predictive model built with primary reference sites and the expanded set of predictors (OE_{50}), and predictive model built with primary and secondary reference sites and the expanded set of predictors ($OE_{CAP}E$).

Model	Good	Fair	Poor	Very Poor
MMI	Score ≥ 60.6	40.4 ≤ Score < 60.6	20.2 ≤ Score < 40.4	Score < 20.2
$OE_{CAP}S$	Score ≥ 64.2	42.8 ≤ Score < 64.6	21.4 ≤ Score < 42.8	Score < 21.4
OE_{50}	Score ≥ 0.95	0.63 ≤ Score < 0.95	0.31 ≤ Score < 0.63	Score < 0.31
$OE_{CAP}E$	Score ≥ 60.7	40.5 ≤ Score < 60.7	20.2 ≤ Score < 40.5	Score < 20.2

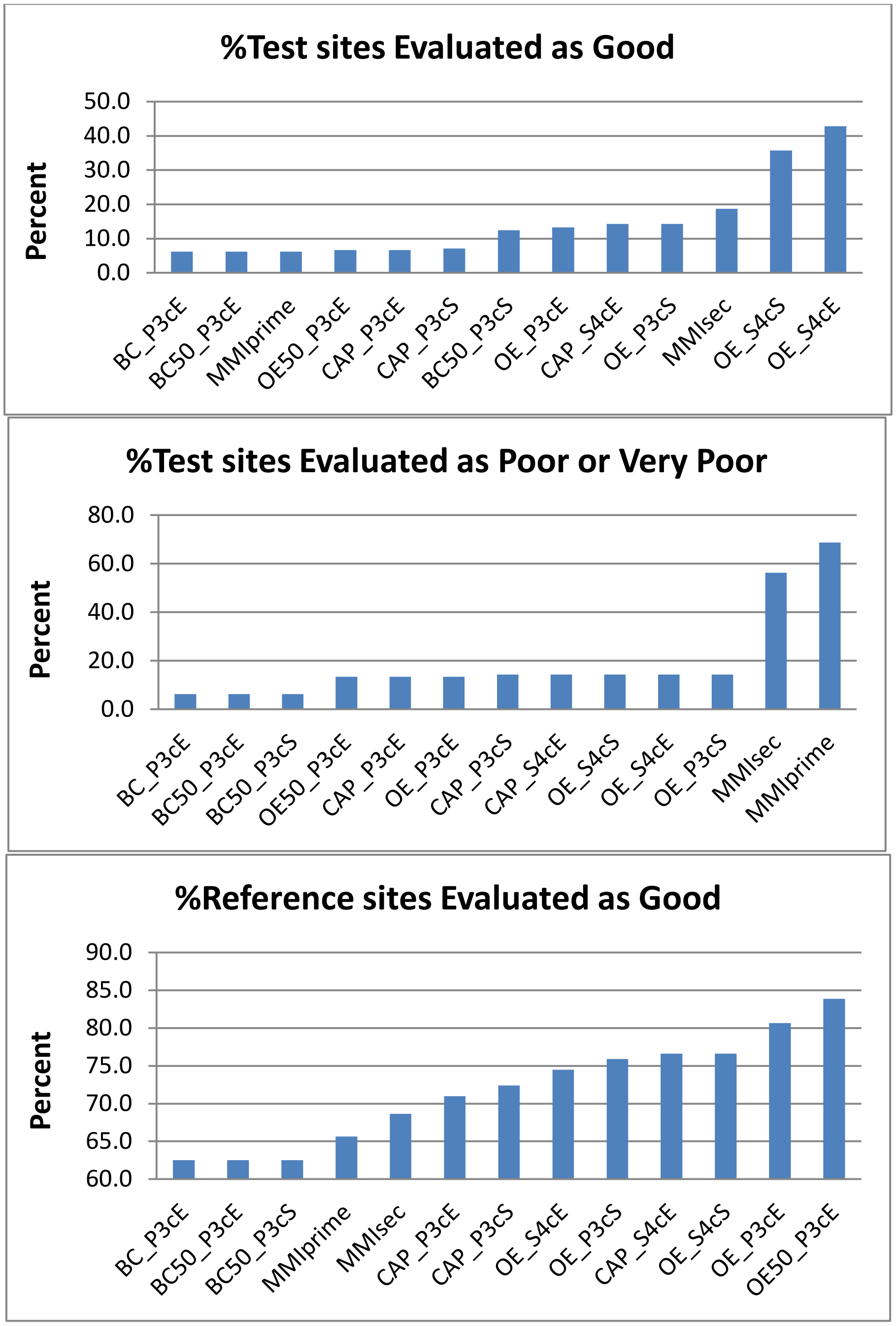


Figure 63. Final set of candidate models showing percent of test sites that were evaluated as "Good" quality (higher value indicates model inability to separate disturbed from high quality conditions; top panel), percentage of test sites that were evaluated as "Poor" or "Very Poor" quality (higher value indicates model ability to detect disturbed conditions; middle panel), and reference sites that were evaluated as "Good" quality (low values indicate model inability to detect high quality conditions; bottom panel). Model abbreviations are given in Table 43.

5.4 Stream Hydrogeomorphology

Hydrology has been identified as the major influence in stream ecosystems (Power et al. 1995, Hart and Finelli 1999, Lake 2000, Bunn and Arthington 2002). It follows that changes in stream hydrology associated with land use can be a major force in shaping the hydraulic geometry of channels, as streams are in dynamic equilibrium with water and sediment supplies delivered from their watersheds (Leopold 1994). Increased watershed impervious cover increases water delivery to the stream during storm events, which alters the hydrograph (i.e., flashiness, duration, etc.) and causes incision (Hammer 1972).
In this context, deviations from predicted relationships between flow regime and channel structure can be used to discriminate disturbed from non-disturbed watersheds. For example, Doll et al. (2002) used the ratio of bank height to bankfull height (bank height ratio, Rosgen and Silvey (1996) in North Carolina Piedmont streams to identify channel incision, and Hammer (1972) used an enlargement ratio (i.e., bankfull area) in northeastern streams to describe altered channels in human-affected watersheds. Both of the preceding ratios are empirically determined in the field with expected values derived from reference streams of a given watershed size. Hydraulic geometry relationships for sand bed, reference streams have been developed for many of the states in the Southeastern Plains Level-III Ecoregion (Omernik 1987) including North Carolina (Doll et al. 2003), Florida (Metcalf et al. 2009), and Maryland (McCandless 2003) and Virginia (Krstolic and Chaplin 2007).

Changes in channel geomorphology associated with altered hydrology can, in turn, lead to altered instream habitats and biota. Southwood (1977) was the first to propose that habitat suitability reflected the “templet” for biotic communities, and Townsend and Hildrew (1994) modified the templet concept to make specific predictions for streams. Keddy (1992) viewed the process of organism-habitat associations in the context of community assembly, where organisms are filtered from the regional species pool based on the interaction of their traits and environmental conditions. Thus, organisms at a site “pass” the environmental filter and are suited to that environment because on their life history traits (Keddy 1992), a concept that also applies to stream communities (Poff 1997, Sokol et al. 2011). Trait-based approaches designed to assess the match between communities and habitats involve use of individualized measures of trait states to produce a community-aggregated mean trait (*mT*, Garnier et al. 2004). Trait values weighted in this way should thus reflect ecosystem-level structure and function as individuals with the highest relative proportions should maximize resource use (Grime 1998, Shipley et al. 2006). Moreover, mean trait values should show strong relationships with the environmental gradients where such traits are advantageous (Garnier et al. 2004) and, thus, also useful in indicating variation in local environmental conditions.

Within the SE Plains, the Sandhills ecoregion (Griffith et al. 2001) is characterized by low gradient, sand-bed streams, which are geomorphically and biologically distinct from upland, gravel/cobble bed streams. Human population in the southeastern US will have the largest increase of any region in the nation (US Census Bureau 2005), largely manifested as increased forestland conversion and expansion of urban land (reviewed by Nagy et al. 2011). In receiving streams, such encroachment underscores the need to 1) develop approaches to define reference conditions for instream habitat and biota, and 2) evaluate the degree to which altered hydrology and geomorphology signal predictable change in biotic communities (Hawkins et al. 2010).

5.4.1 *Methods*

We describe and validate simple empirical models describing the hydrogeomorphic reference condition in the Sandhills ecoregion and its association with trait-based benthic macroinvertebrate assemblages. The hydrogeomorphic reference condition should reflect known variation in the physical channel structure and, in turn, provide useful information about community composition and function in streams spanning a wide range of environmental conditions.

5.4.1.1 Study Area and Landscape Variables

The study area consisted of 62 streams in the SH Ecoregion of Georgia, South Carolina, and North Carolina (Figure 64). Study watersheds spanned a range of sizes (0.64-54.6 km^2; mean=8.5 km^2) and orders (1-3). Stream channels were low-gradient and sandy, and associated watersheds were generally forested. Study watersheds were located on two DoD (Fort Bragg, Fort Benning) installations, one DoE (Savannah River Site) installation, and various state and private lands (Table MM1). Watershed areas were derived from 10m digital elevation models above the downstream sample reach terminus. Landscape conditions and variables were quantified as described earlier in this report.

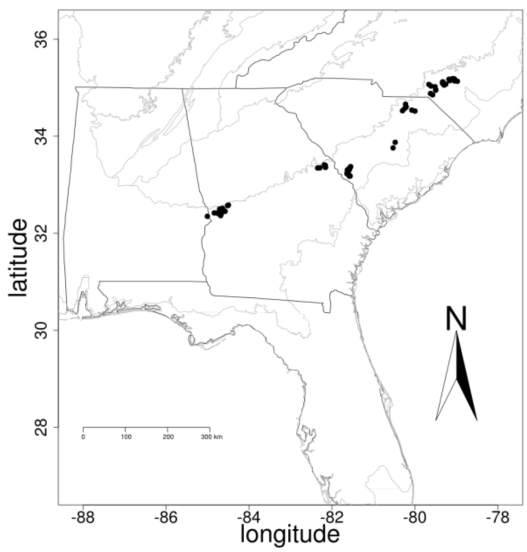


Figure 64. Map of SE United States showing study sites (closed circles) in GA, SC, and NC. Sites were in the SH Ecoregion below the Piedmont and above the Coastal Plain (light grey lines).

5.4.1.2 Stream Hydrology, Geomorphology, and Instream Habitat

Stream stage was used as a measure of temporal variation in stream hydrology, estimated from Solinst Levelogger Junior® pressure transducers (Model 3001, Solinst Canada Ltd., Canada), installed at the downstream end of each study reach. Leveloggers were adjusted for ambient atmospheric pressure using Solinst Barologger Gold® pressure transducers or barometric pressure data obtained from local airport weather stations. Level and barometric pressure data were measured every 15 min for the duration of logger deployment at each stream reach. Temporary stilling wells were used to house leveloggers, constructed from schedule-40 PVC (3.81 cm ID) and perforated on the downstream side for water circulation. Stilling wells have been shown to produce stable water surface elevations in other SE streams (Schoonover et al. 2006). Continuous stream level was summarized using a suite of metrics reflecting variation in flow regime (McMahon et al. 2003, Helms et al. 2009).

Channel geomorphology was quantified by surveying four to six channel cross sections at each study reach. Transects were established in runs dividing the reach into approximately equidistant sections, with reach length fixed regardless of four or six transects, thus allowing comparison. Cross sections were established by staking rebar at determined top of bank height on either side of the channel perpendicular to the direction of flow. Top of bank height was determined as the point where water breaches the lowest of the 2 stream banks (Leopold 1994). A line level was used to establish the relative top of bank datum for each survey and top of bank depths and water depths were recorded every 20 cm along each transect. Top of bank area, depth, and width were summarized as a reach-specific median.

Measures of instream habitat consisted of estimates of substrate size, amount of coarse woody debris (CWD, wood > 2.5 cm diameter) and benthic organic matter (BOM, organic matter material ≤1.6 cm diameter) at each transect (Wallace and Benke 1984). Substrate size and BOM were estimated from PVC cores inserted to a depth of 10 cm (7.62 cm inner diameter diameter, 455.8 cm^3 sample volume for substrate size; 2.5 cm ID, 49.1 cm^3 sample volume for BOM) near the center of the channel directly above each transect (substrate size) or at two locations (midstream and stream margin) along each transect. Samples were dried and combusted to remove BOM (below), and then dry sieved for representative particle sizes (i.e., *d5*, *d50*, *d95*, etc., phi scale -4-5, Lane (1947) and summarized as a reach-specific median. Geometric mean and standard deviation of particle size were calculated from reach medians after equations in Table 2.8 in Bunte and Apt (2001). For removed BOM, samples were oven-dried at 80°C for 24 to 48 h, weighed, and ashed in a muffle furnace at 550°C for 3 h. Samples were then cooled in a desiccator and reweighed; % BOM was determined as the difference between dry and ashed masses divided by total dry mass (Wallace and Grubaugh 1996). For CWD, length and width of each piece of CWD was measured to obtain an areal estimate, and then summed across transects. BOM was processed in the laboratory according to standard methods. Last, specific conductivity (SC), pH, dissolved oxygen, were quantified at the downstream terminus of each reach during one or more sampling dates (below). Multiple water chemistry measures were summarized as the median for a stream reach, whereas temperature was recorded every 15 min from stage leveloggers.

5.4.1.3 Macroinvertebrate Assemblages and Hydrogeomorphology

Benthic macroinvertebrate assemblages were sampled as described earlier in this report. Macroinvertebrates were identified to the lowest practical taxonomic level, usually genus or species. We used a variety of single benthic macroinvertebrate metrics selected from standard USEPA rapid bioassessment protocols (Barbour et al. 1999), including Ephemeroptera, Plecoptera, and Trichoptera (EPT) richness, abundance, and the North Carolina Biotic Index (NCBI, Lenat 1993). Macroinvertebrate trait values were taken from the USGS database http://pubs.usgs.gov/ds/ds187/ (Vieira et al. 2006) and coded so that trait states summed to one. Mean trait values (mT) were calculated as

$$mT = \sum_{i=1}^{n} p_i * trait_i \qquad (24)$$

where p_i was taxa relative abundance and $trait_i$ was value of trait (Garnier et al. 2004) (functcomp; FD package). We predicted that community-weighted *mT* would be useful in relating assemblage trait structure to instream hydrogeomorphic conditions.

5.4.1.3 Classification and Analysis of Sites

Discrimination between reference and non-reference streams as two populations was based on the assumption that predictable relationships exist between channel morphology and watershed area (Leopold 1994), and that streams deviating in morphology measures from the expected (reference) value given watershed area defined a non-reference population. To quantify deviation from this expectation, regression models relating watershed area (x= *Aws*) and 3 geomorphic variables (y= area at top of bank, *Atob*; depth at top of bank, *Dtob*; width at top of bank, *Wtob*) were fit with iteratively re-weighted least squares regression (MASS package, Venables and Ripley 1994) and then tested for significance with a Wald test (robftest; sfsmisc package). Centroid distance from 0, Cd_o, (i.e., distance from the channel morphology reference condition in multivariate space) was calculated to describe variability in the interaction of residual channel morphology, as:

$$Cd_0=((0-(A_e/3))^2)+(0-(D_e/3))^2)+(0-(W_e/3))^2))^{0.5} \quad (25)$$

where A_e = *Atob residuals*, D_e = *Dtob residuals*, and W_e =*Wtob residuals.*

Use of the above three groups of residuals, rather than a single measure of channel morphology, ensured that multivariate differences in channels were accounted for in stream clustering (i.e., adjustments of channels to changes in geomorphic variables were not constrained by a single measure). Channel morphology measures were standardized (mean=0, SD=1) and clustered with Partitioning Around Medoids (PAM), a technique that minimizes the sum of distances from group mediods (actual data used as cluster centers in lieu of centroids that is robust to statistical outliers, Van der Laan et al. 2003), with the number of groups determined objectively by maximizing mean width between clusters (=silhouette width, Borcard et al. 2011). In this context, the maximum mean silhouette width value can be visualized as amplifying the "tightness" of clusters, with the greater the maximum mean silhouette width the greater the clustering, see also Rousseeuw 1987). A classification tree was then used to assess which unstandardized geomorphic variables classified sites into groups (RPART package). To test for putative differences in macroinvertebrate *mT* measures, streams between geomorphic groups were paired randomly with those of similar *Aws* determined from the cluster analysis. We used *t* tests with equal variance (Levene's test for homogeneity of variance and Shapiro's test for normality P>>0.05) to test for significance between reference and non-reference streams for EPT and NCBI. *P* values were adjusted for multiple testing with the Benjamini-Hochberg false discovery rate correction (Benjamini and Hochberg 1995).

Partial Least Squares Regression (PLS), a useful method in ecological research (Carrascal et al. 2009), was used to relate environmental variable and macroinvertebrate traits (hereafter "trait") between reference and non-reference streams. PLS is useful when collinearity is high or when there are many more predictor variables than sites. These two issues often occur in ecological surveys and precludes use of standard statistical techniques (e.g., multiple linear regression) when investigating multivariate data sets (Carrascal et al. 2009). PLS is similar to Principal Component Analysis (PCA) as they both reduce data dimensionality by projecting linear combinations onto fewer derived orthogonal variables (i.e., PCA components or PLS latent variables). PCA projects the X (environmental) matrix onto orthogonal components with the criteria of maximizing the variation in that direction. PCA scores/loadings are intuitive as the score (ordinate) of a site is affected by the original variables (i.e., loading) in the X matrix. In contrast, PLS projects the X matrix onto a latent variable by maximizing the covariance (or correlation, when the data has been standardized) between the Y and X matrix. This procedure results in scores and loadings similar to PCA, but also incorporates information contained in Y (Carrascal et al. 2009).

PLS modeling (also called PLS discriminant analysis, PLS-DA) was used to relate a reference versus. non-reference condition to predictor variables. This procedure maximizes the X matrix's (i.e., environmental variables or trait matrices) ability to predict the Y matrix (i.e., group membership). Using the retained latent variables in a linear discriminant to predict group membership has been shown to increase classification accuracy (Boulesteix 2004, Boulesteix and Strimmer 2007, Turkmen and Billor 2013). PLS-Linear Discriminant Analysis (LDA) is used instead of classical LDA because of predictor variable multicolinearity (i.e., after PLS the latent components are not collinear, making LDA possible). PLS models were fit with the de Jong's (1993) SIMPLS algorithm (plsr; PLS package, Mevik and Wehrens 2007). The general modeling process was as follows: 1) the number of latent variables was chosen to minimize root mean square error of prediction based on leave one out cross-validation (LOOCV), 2) PLS model was fit with LOOCV number of latent variables, 3) variable importance in the projection (vip) scores were calculated and variables with values <1 were removed, which was used as a variable selection step (1 is bounded by values that Chong and June (2005) reported and is commonly reported as a very important variable threshold (see Sonesten 2003), 4) steps 1 and 2 were repeated on the reduced dataset, 5) the retained latent variables were then used in linear discriminate analysis (LDA; lda; package MASS (Venables and Ripley 1994)) to develop classification rules (Turkmen and Billor 2013). The method of Turkmen and Billor (2013) was modified to use classical LDA instead of a robust version. The number misclassified (NMC) by the models were assessed with LOOCV to access likely predictive ability of the model (Turkmen and Billor 2013), and overall model significance was accessed with permuted class labels to access the significance of the classification (number of permutations=10,000, Szymanska et al. 2012). This permutation test was used to determine how often NMC of permuted models was $\leq$ NMC values from the original model. In this way, a probability for accessing the models significance in classification was found. All statistical analyses were done in the *R* language (v. 3.0.1; R Core Team 2013; Ihaka and Gentleman 1996).

5.4.2 Results and Discussion

5.4.2.1 Classification of Reference and Non-Reference sites

Regression analysis showed that *Aws* was significantly related to *Atob* (n=62, F=43.14, P<0.0001, equation *Atob*=0.7381+ 0.0839 *Aws*). Two groups of study sites displaying contrasting hydrogeomorphic conditions were identified from the clustering analysis (Figure 65), and these groups form the basis for the results presented in this report.

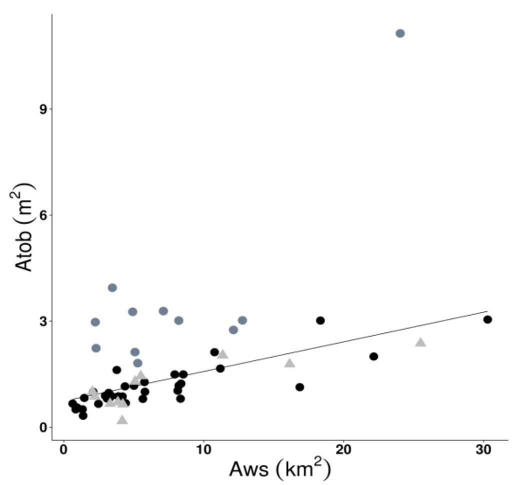


Figure 65. Watershed area (*Aws*) plotted against area at the top of bank (*Atob*) for the 62 study sites. Colors and shapes are indicative of geomorphic group based on the clustering analysis and iteratively re-weighted least squares (see text). ● and ▲are reference sites as indicated by Partitioning Around Mediods (PAM) clustering. ▲ are reference sites that were randomly paired with similar watershed areas to non-reference sites (●) for site comparisons using invertebrate and environmental data.

Classification tree analysis indicated that a residual (unexplained variation) value of $0.6m^2$ A_{tob} from the preceding regression equation separated reference from non-reference streams with a misclassification rate of <2% (1/62, Figure 66). Those streams with *Atob* residuals $>0.6m^2$ were considered non-reference streams, whereas all other sites were classified as reference streams. An example of how the *Atob residuals* criterion was used to classify one of the study sites is presented later.

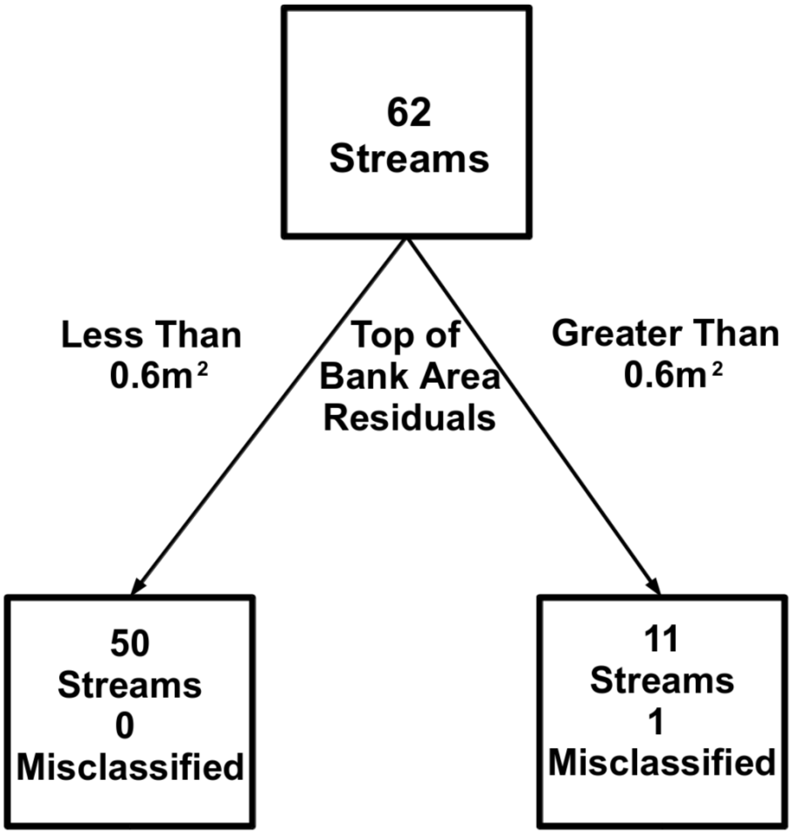


Figure 66. Classification tree of unstandardized residual top of bank area. Misclassification rate was <2%. This figure along with regression model of watershed area and top of bank area (Figure 65) can be used to classify new streams into hydrogeomorphic groups.

These results suggest existence of a hydrogeomorphologic reference condition for SH streams, derived from a multivariate cluster analysis of commonly used and easily measured hydraulic channel geometry variables. In this context, difference in reference and non-reference conditions can be discriminated with independent data from those empirical data that were used to derive to site clusters (i.e., excluding all channel morphology and *Aws*). The following section describes the degree to which macroinvertebrate assemblages, both as standard bioassessment metrics and traits corresponded with these two contrasting stream groups.

5.4.2.2 Macroinvertebrate Assemblages and Hydrogeomorphology

Results of *t*-tests showed that EPT richness was significantly higher in reference sites (t=-2.68; P=0.031; Figure 67A) and NCBI was significantly lower in reference sites at α=0.1 (t=1.95; P=0.067; Figure 69B). Both of these differences occurred in the predicted direction, corroborating the linkage between channel structure and biotic condition.

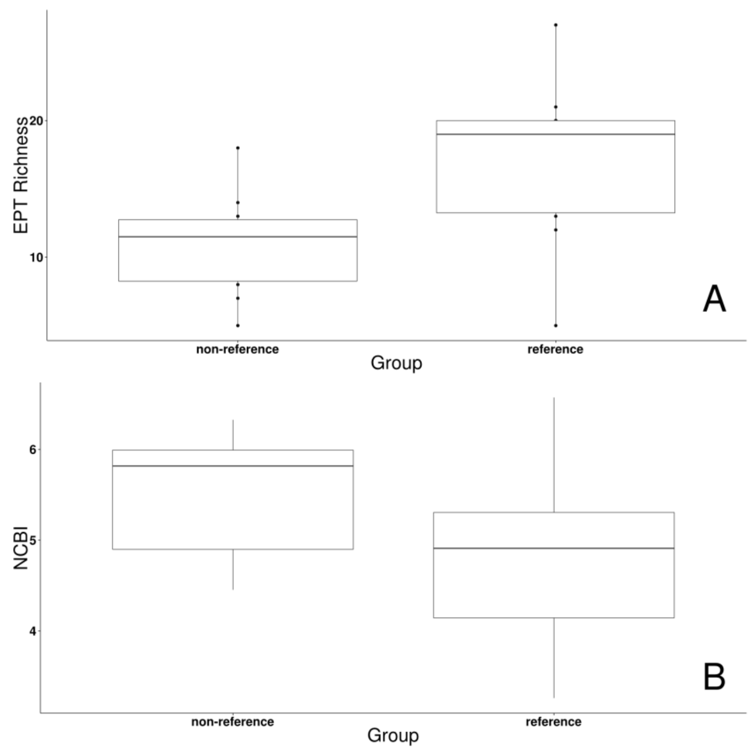


Figure 67. Boxplots comparing Ephemeroptera, Plecoptera and Trichoptera (EPT) richness (A) and North Carolina Biotic Index (NCBI) values (B) between non-reference and reference streams defined by contrasting site hydrogeomorphology (see text). Difference between groups for each measure was significant.

PLS-LDA revealed that environmental variables (env_{PLS}) and macroinvertebrate traits ($trait_{PLS}$) significantly discriminated the two stream groups (permutation values: P=0.02 and <0.001, respectively). Variable importance in projection (VIP) values were used in a variable-selection step before the final models were constructed (see methods). All of the variables retained for the final model showed VIP≥1 in the full models. Thus, VIP values for the final models were used to access relative importance of variables to the model similar to Sonesten (2003) (i.e., vip≥1, strong importance; 1>vip≥0.8, moderately strong importance; and vip<0.8, weak importance). Generally, loadings indicate the importance of a variable in deriving site scores and, thus, discriminating between reference and non-reference streams on the component for which the loadings are presented. The VIP score indicates how important the variable was to predicting the Y values for the model under discussion.

Env_{PLS} identified one latent variable in the non-reference (disturbed) direction and only organic matter in the reference direction (35% explained X variance; 68% explained Y variance, Figure 70). The LDA standardized coefficient derived from the Env_{PLS} scores was 7.52 (included for completeness as there was only one latent variable). Stream water specific conductance and pH both were important in discriminating the two stream groups (highest loadings and high VIP scores), and were higher in the non-reference group. Hydrologic measures of high-magnitude stage duration (i.e., maximums above specified quartiles 75th and 95th; high loadings and moderate high VIP scores). Flashiness (i.e., falling limb differences) also was higher in non-reference group. In contrast, low-magnitude stage duration (i.e., maximums below specified quartiles 10th and 25th) was higher in the reference group. Substrate size standard deviation and mean (Sed SD and Sed Mean, respectively) and % developed land were higher in non-reference group, whereas streambed organic matter (both as CWD and OM) was higher in the reference stream group (Figure 68). NMC for env_{PLS} was 10% (i.e., 90% classified correctly).

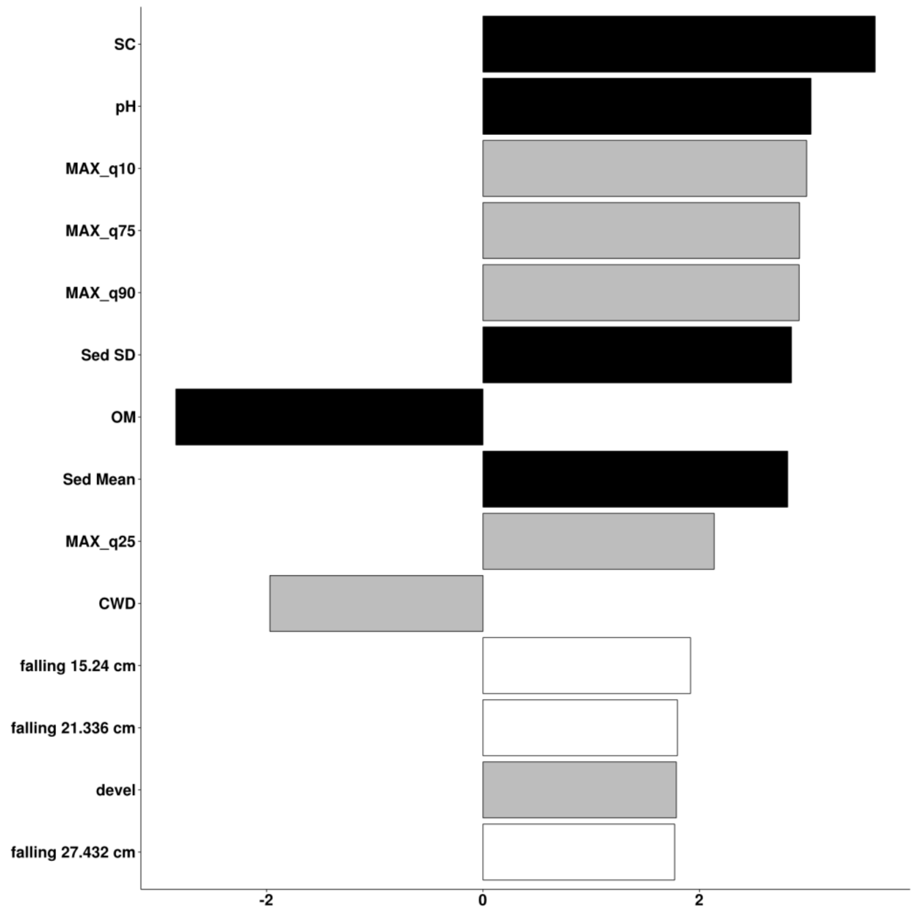


Figure 68. Partial Least Squares (PLS) loadings of original environmental variables onto PLS latent variable 1. Variables are ordered by the absolute magnitude of their loading on latent variable 1 and are colored according to the variable importance in projection value (vip) score), reflecting the importance of the variables in determining site scores along latent variable 1 (vip≥1: black [strong importance], 1>vip≥0.8; grey [moderate], and vip<0.8; white [weak]. Reference stream scores were negative values whereas non-reference stream scores were positive. Variables defined in Appendix 8.

For macroinvertebrates, LDA standardized coefficients derived from $trait_{PLS}$ scores are presented in Table 45. The LDA coefficient for latent component (Comp) 1 was close to or greater than twice Comp 2 and Comp 3 (Table 45). Given the paramount importance of Comp 1 as a latent variable in discriminating sites (36 and 49% of X and Y variance, respectively), interpretation was centered on this component, and data visualization was presented as a biplot of Comps 1 and 2 (Comps 2 and 3 were qualitatively similar, Table 45). Several traits loaded highly on Comp 1 in the non-reference direction (right side of Figure 69), and included taxa in the collector-gathering functional feeding group, having multiple generations per year (multivoltine), preference for fast current, and bluff (bricklike) body shapes. In contrast, traits loading highly on Comp 1 in the reference direction (left side of Figure 69) included taxa with dorsoventrally flattened bodies, in the shredder functional feeding group, having a moderate adult life span, hemimetabolous development (only 2 aquatic stages), one generation per year (univoltine) and fast seasonal cycles, and habitat preferences for the stream bed (Figure 69). Figure 70 shows the full array of PLS loadings of original macroinvertebrate trait variables onto PLS latent variable 1. The NMC for $trait_{PLS}$ was 15% (i.e., 85% classified correctly).

Table 45. Standardized linear discriminant analysis (LDA) coefficients utilizing partial least squares latent variables (Components [Comp] 1-6 from $trait_{PLS}$) to predict reference non-reference streams from mean macroinvertebrate trait values. See text for explanation.

	Standardized Coefficient
Comp 1	9.9
Comp 2	5.5
Comp 3	6.0
Comp 4	2.1
Comp 5	2.5
Comp 6	1.9

Patterns in trait data suggest the presence of two different assemblages of macroinvertebrates in reference and non-reference stream groups. These results support the assertion that assembly of macroinvertebrate communities occurs through filtering from the regional species pool based on the traits that best suit their occurrence in these contrasting stream types. Specifically, the results also suggest that hydrologic disturbance resulting from altered land use in non-reference watersheds affects instream habitats and benthic organisms with traits suiting them to a harsher physical environment. In particular, macroinvertebrate taxa showing fast life cycles, multiple generations per year, and preference for fast current velocity, traits of which are favored in hydrologically disturbed habitats, were strongly associated with non-reference streams. This pattern is consistent with the supposition that flow regimes influence communities in non-reference streams. Similarly, organic matter abundance was lower in non-reference streams, likely resulting from increased high-flow events (flashiness) eroding instream benthic organism matter and CWD (Maloney et al. 2005). In turn, decreased organic matter likely reduced OM resources for shredders in non-reference streams, hence reducing their abundance relative to reference sites. Last, community richness (as EPT) and tolerance (as NCBI), which strongly differed between non-reference and reference streams, also suggested a general difference in the importance of sensitive taxa within non-reference sites.

5.4.2.3 Hydrogeomorphic (HGM) Status as an Assessment Framework

Below (Table 46) is an example application showing determination of HGM status for a study site at Fort Benning (ftbnbc01). Watershed area (*Aws*, m^2) is estimated using GPS (as UTM coordinates at the downstream terminus of the sample reach) and GIS algorithms (e.g., r.stream.basins in GRASS GIS 6.4; comparable software found in ArcGIS). Observed area at top of bank ($Atob_o$, m^2) is measured in the field following methods discussed above. Predicted area at top of bank ($Atob_p$, m^2) is estimated using the derived equation $Atob_p$=0.7381+ 0.0839 *Aws*, which, when subtracted from $Atob_o$ yields *Atob residual* for the site. If *Atob residual* is <0.6m^2 then the site is considered reference; if this value is >0.6m^2 then the site is considered hydrogeomorphically disturbed. In this example, estimated *Atob residual* was $<0.6\ m^2$, resulting in a site classification of reference. Instream environmental and biological variables with the highest classification potential (determined from PLS analyses, above) can then be used as predictors of conditions likely to occur based on HGM classification state.

Table 46. Determination of hydrogeomorphic (HGM) status for site ftbnbc01 at Fort Benning.

Variable	**Value**
UTM Coordinate (Zone 16)	3588705N 0710570 E
Aws (from GIS)	$0.636\ km^2$
$Atob_p = 0.7381 + 0.0839\ Aws$	$0.792\ m^2$
$Atob_o$	$0.668\ m^2$
$Atob\ residual = Atob_o - Atob_p$	$-0.124\ m^2$
Atob residual $<0.6\ m^2$ - REFERENCE	YES
Atob residual $>0.6\ m^2$ - NON-REFERENCE	NO

The preceding results support the existence of a landcover cascade (Burcher et al. 2007), which predicts that changes in watershed land use are linked with altered hydrology and geomorphology, which, in turn, alter stream biotic communities and their habitats. Use of this approach in an assessment framework to define a biologically relevant hydrogeomorphic reference condition appears useful in determining the likelihood that a stream is in the reference condition. A simple empirical measure of $Atob_o$ and its deviation from $Atob_p$ given watershed area (*Aws)* can be used as a basis for evaluating the non-reference condition using the classification criterion $Atob_p > 0.6m^2$. In addition, this field-based approach appears useful in providing expectations for instream environmental and macroinvertebrate trait conditions as well as other key response variables used in assessment. We suggest that hydrogeomorphic approach be expanded and evaluated in the SH and in other Atlantic Coastal Plain ecoregions where low-gradient sand-bed streams predominate.

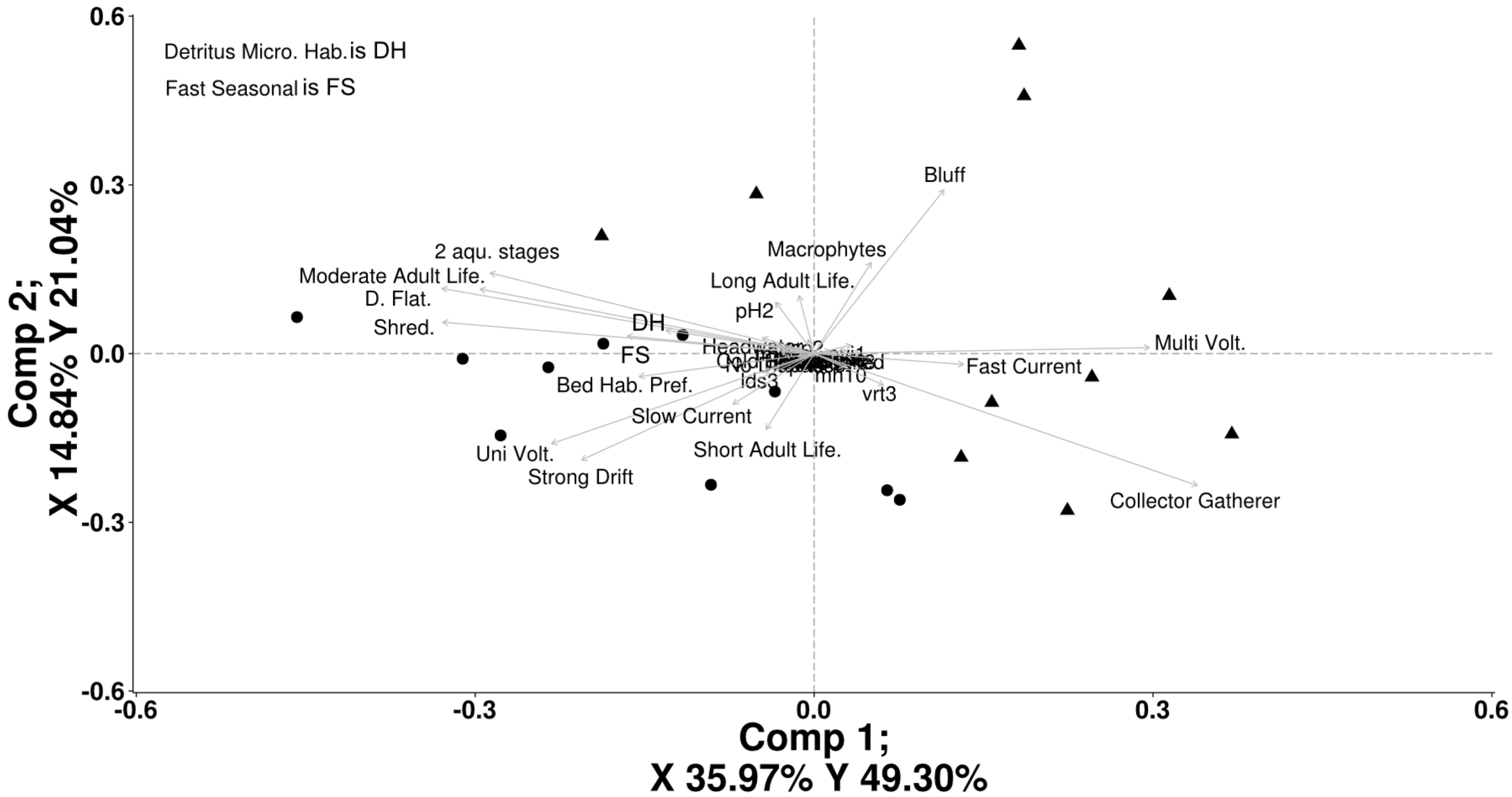


Figure 69. Partial Least Squares (PLS) score and loadings biplot displaying PLS latent components (comp) 1 and 2 for the macroinvertebrate assemblage trait data from the study sites. ● are reference sites and ▲ are non-reference sites as indicated by Partitioning Around Mediods (PAM) clustering. Variables furthest from the origin exerted the greatest influence on the resulting site scores (see HGM Figure 70 for variable loading plots). Variance explained for X and Y matrices are shown on the axes labels. Variables defined in Appendix 8.

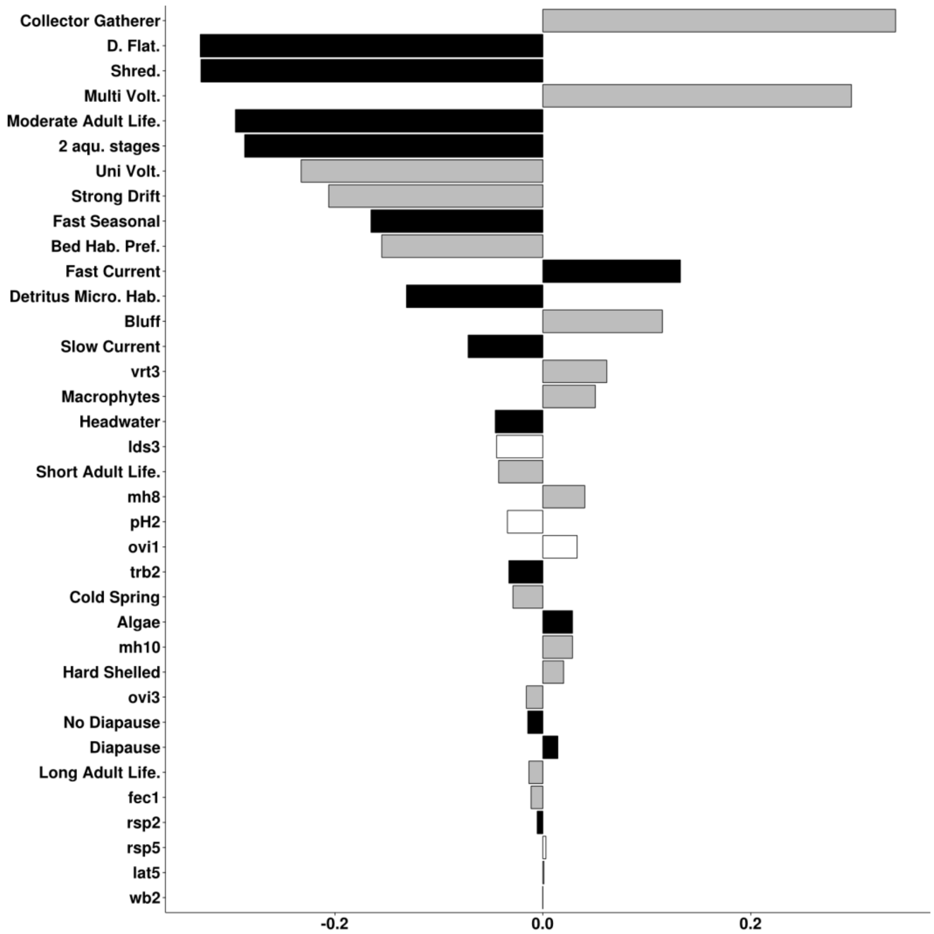


Figure 70. Partial Least Squares (PLS) loadings of original macroinvertebrate trait variables onto PLS latent variable 1. Variables are ordered by the absolute magnitude of their loading on latent variable 1 and are colored according to the variable importance in projection value (vip, vip≥1: black [strong], 1>vip≥0.8; grey [moderate], and vip<0.8; white [weak]), reflecting the importance of the variables in determining site scores along latent variable 1. Reference stream scores were negative values whereas non-reference stream scores were positive. Loadings ordered in terms of decreasing loading magnitude. Variables defined in Figure 69 caption. Variables defined in Appendix 8.

5.5 Performance Assessment

We assessed the performance of three fish and two macroinvertebrate bioassessment indices described in Sections 5.2.9 and 5.3.3: the ANNA, CQI, MMI, BC50_PS, and CAP_secS methods. The performance of bioassessment indices is affected by their variability, sensitivity, accuracy, responsiveness, and statistical power. These measures were calculated as follows:

1) The variability of the indices was assessed by computing coefficients of variation (CVs) for the reference sites. Greater variability among reference sites results in reduced ability to distinguish disturbed sites from reference sites. The lower 10% of the reference site scores for each index were excluded from these computations for reasons explained in Section 5.2.9.1.

2) Sensitivity can be defined as the ability of an index to measure disturbance (as indicated by the magnitude of the difference between disturbed and reference sites) in relation to its variability. Diamond et al. (1996) calculated sensitivity as follows:

$$(M-X)/S \qquad 23$$

where M is the mean index value for the reference sites (excluding the lower 10% of the reference site scores), X is the index value for a site that is being assessed, and S is the standard deviation for the reference sites. Sensitivity was computed for each of the sites ranked as disturbed by the PCA gradient approach. The resulting sensitivity scores were averaged for each index and compared using one-way analysis of variance.

3) Accuracy, the ability of an index to successfully distinguish disturbed sites from reference sites, was computed following Diamond et al. (1996):

$$(M-X)/S > t \qquad 24$$

where M, X, and S are as previously described and t is the statistic for a one-tailed test with degrees of freedom equal to the number of reference sites minus one. A one-tailed t statistic was used because only test site scores lower than the reference scores were of interest for the ANNA, CQI, MMI, and CAP_secS methods, and only test site scores higher than the reference scores were of interest for the BC50_PS method. Accuracy can be placed in the context of Type I and Type II error, with the null hypothesis being no difference between the reference sites and the site being tested, Type I error being the false labeling of an undisturbed site as disturbed, and Type II error being failure to distinguish a disturbed site from reference conditions.

4) Responsiveness is the degree to which an index changes monotonically with disturbance (Diamond et al. 2012). It can be assessed by examining plots of index values versus disturbance measures or by calculating the coefficient of determination (R^2) and significance of the slope between index values and a disturbance measure. The measure of disturbance used in this analysis was the score on axis 1 of the PCA (Section 5.1.2), which was strongly correlated with a number of key abiotic disturbance related variables.

5) Statistical power (1-β) is the probability of rejecting the null hypothesis when it is really false. It is a function of alpha (α, the probability below which the null hypothesis is rejected), effect size (*D*, the magnitude of the difference that is being tested for, e.g., difference between means), number of samples in each group (n), and the sample standard deviation (σ). The analysis of power can focus on number of samples required to detect a specified difference (effect size) between two population means:

$$n = \frac{2\sigma^2(t_\beta + t_{\alpha/2})^2}{D^2} \qquad 25$$

where t_α represents the desired level of statistical significance and t_β represents the desired power (typically 80%). The other variables are as described above.

Calculation of the standard deviation as required for computation of statistical power necessitates the collection of two or more replicate samples from the same site or from closely adjacent sites, which is seldom done for logistical and economic reasons (Diamond et al. 2012). Because this was not done in RC-1694, power was calculated from historical data collected on the SRS in the 1990s by one of the authors (M.H. Paller). These data consisted of samples collected from pairs of closely adjacent sites using methods very similar to those used in RC-1694. There were 10 pairs of sites located in eight second through third order streams.

5.5.1 *Responsiveness*

R^2 values for the ANNA, CQI, MMI, BC50_PS, and CAP_secS methods were 0.09, 0.39, 0.30, 0.37, and 0.41, respectively. All regressions were significant at $P < 0.001$ except for the ANNA regression, which was significant at $P=0.01$. Diamond et al. (2012) observed that R^2 values were under 0.30 for most indices developed for states in the Southeast. Only the responsiveness of the ANNA model fell below this level.

Responsiveness was also evaluated by comparing the distributions of disturbed and reference sites for the five indices (Figure 71). The distribution of scores for each index was divided into ranges representing five levels of quality: very good, good, fair, poor, and very poor. This was done by designating the 25th percentile of the reference site distribution as the threshold between "good" and "fair" and equally sectioning the range between the 25th percentile and the bottom of the scoring range. The bottom of the scoring range for the MMI and ANNA was the lowest score mathematically possible. For indices without a minimum score (e.g., the CQI), the bottom of the scoring range was indicated by a score representative of severe degradation. Although somewhat arbitrary, narrative thresholds are often used to describe levels of relative impairment (Barbour et al. 1999). Box plots of these distributions showed that "good" and "very good" scores" were largely excluded from the distribution of disturbed sites for all indices except the ANNA model (Figure 71).

The preceding evaluations of responsiveness were based on data collected under RC-1694, some of which were used to build the indices. For the CQI and MMI methods, sensitivity was also assessed using the completely independent data shown in Figures 35 and 38; R^2 values were 0.33 for the CQI and 0.26 for the MMI ($P<0.001$ for both). Figure 38 shows that the CQI was particularly successful at distinguishing different levels of disturbance.

5.5.2 *Variability, Sensitivity, and Accuracy*

The variability, sensitivity, and accuracy analyses were conducted only on the 58 RC-1694 sites with sufficient data to compute all five indices. This was done to directly compare the indices and point out advantages of using fish and macroinvertebrate indices concomitantly. The sites were ranked on a gradient of physical habitat disturbance using the PCA approach (Section 5.1.2), with rank being inversely proportional to disturbance. An inflection point on the curve representing the gradient occurred between the ranks of 11 and 12 (Section 5.1.2). On this basis, the first 11 sites were classified as disturbed and the remaining sites as undisturbed.

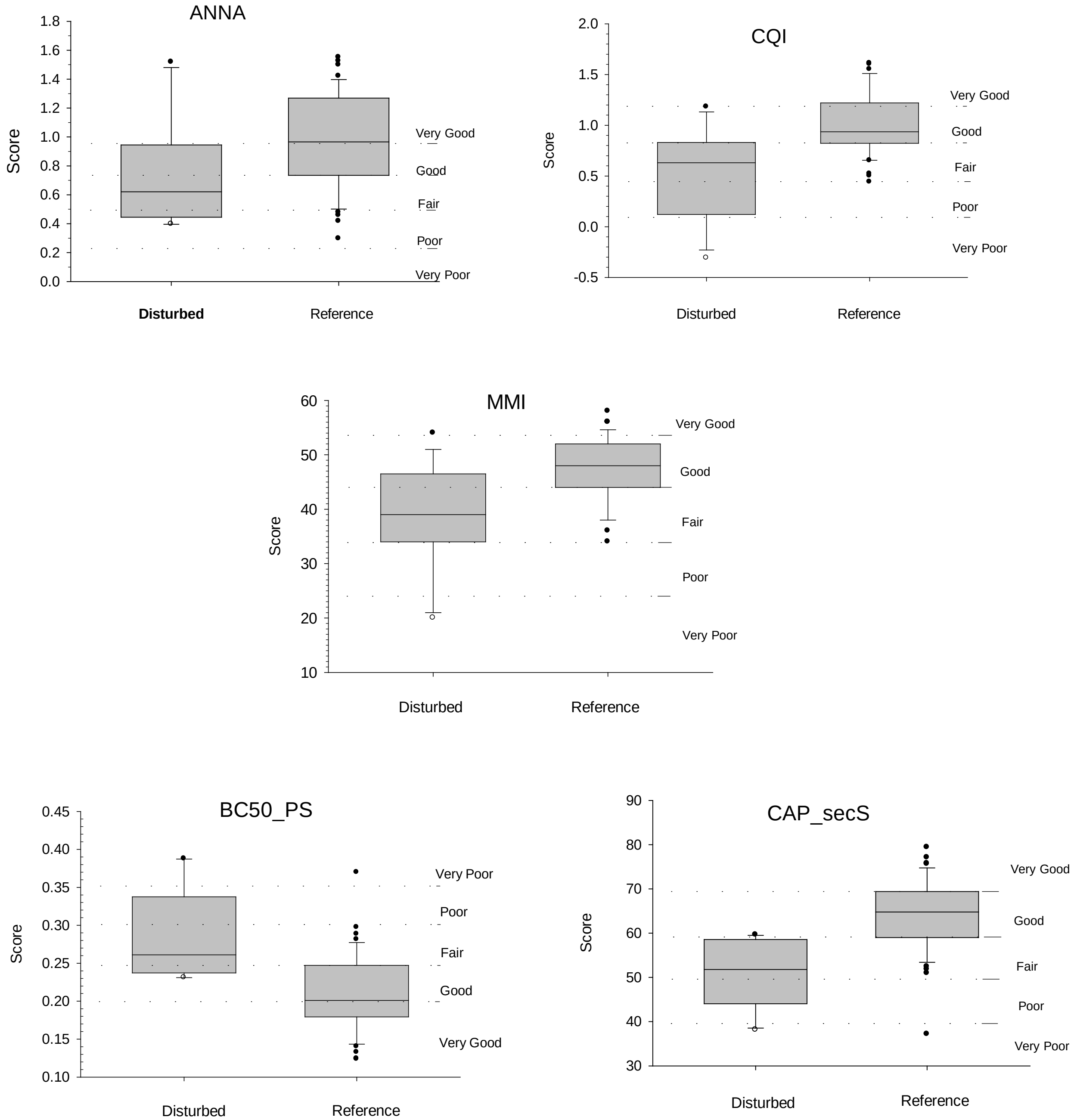


Figure 71. Distribution of bioassessment scores for disturbed and reference sites. The distribution of scores for each index was divided into ranges representing five levels of quality by designating the 25th percentile of the reference site distribution as the threshold between “good” and “fair” and equally sectioning the range between the 25th percentile and the bottom of the scoring range.

Reference site variability, as indicated by CVs, was 25% for ANNA, 22% for the CQI, 12% for the MMI, 20% for BC50_PS, and 9% for CAP_secS. The somewhat lower values for the MMI and CAP_secS indices suggest that they have the potential to detect smaller departures from the reference condition than the other indices. Reference site CVs for the MMI and CAP_secS were similar to the CVs reported by Paller (2001) for fish- and macroinvertebrate-based multimetric indices developed for the SRS (7% and 12%, respectively) and to the CV of 12% and 13% reported for the Florida Stream Condition Index in the Florida Peninsula and Florida Panhandle bioregions, respectively (Gerritsen et al. 2000).

The average sensitivity (difference between disturbed and reference sites in relation to index variability) for the 11 disturbed sites was 1.03 for ANNA, 1.84 for the CQI, 2.75 for the MMI, 2.22 for BC50_PS, and 2.40 for CAP_secS. Although this analysis suggests that the ANNA method was less sensitive, differences among methods were not statistically significant at $P<0.05$ (one-way ANOVA).

From the perspective of resource protection, Type 2 error (failure to identify a disturbed site as disturbed) is more serious than Type 1 error because it may result in insufficient protection or restoration effort. None of the individual methods accurately distinguished all of the 11 disturbed sites from the reference sites. The most accurate method was the MMI, which identified 9 of the 11 sites followed by the CAP_secS, which identified 7 (Table 47). However, all of the disturbed sites were distinguished from the reference sites by at least one method. The seven most disturbed sites were accurately identified by multiple methods including at least one fish and one macroinvertebrate-based method (Figure 72). Indication of disturbance by both fish and macroinvertebrate indices is strong evidence of ecological degradation. These results indicate that the MMI was most the most effective index at avoiding Type 2 error, that the concurrent use of all five indices was 100% effective at avoiding Type 2 error (as indicated by a significant finding by at least one index), and that the most disturbed sites were accurately identified by at least one fish-based and one macroinvertebrate-based method.

Accuracy was calculated for the reference sites as a potential measure of Type 1 error – the false labeling of undisturbed sites as disturbed. Type 1 error suggests a minimally disturbed site may be impacted, thus indicating the need for more protection than may be necessary and/or collection of more data to verify the degree of impact. It is less serious than Type 1 error from a conservation perspective but may result in unnecessary expenditures. The highest rate of Type 1 error was exhibited by the MMI, which identified 10 of the 47 reference sites as disturbed (Table 47). Twenty-one of the reference sites were identified as disturbed by a least one method. However, only one of the reference sites (ranked 14 on the disturbance gradient) was identified as disturbed by both a fish- and a macroinvertebrate-based index. These results suggest a substantial incidence of Type 1 error when a site is considered to be disturbed on the basis of a significant finding by one or more methods but a low rate of Type 1 error when a site is classified as disturbed by a combination of a least one fish and one macroinvertebrate method.

5.5.3 *Statistical Power*

Data were available to calculate power only for the MMI. The computations simulated a question likely to be asked by a biologist, “How many samples are needed to detect a difference between a reference site representing the lower threshold of the reference condition (i.e., 42 for the MMI, Section 5.2.9.2) and a moderately impaired test site with an MMI value 1 to 4 points lower?” The average standard deviation of the MMI for the paired historical SRS sample sites was 1.6, thereby supplying the measure of

Table 47. Accuracy of five assessment frameworks. An "x" indicates the site was distinguished from the reference site mean (calculated following Diamond et al. 2000). Sites are ranked by decreasing disturbance based on an abiotic disturbance gradient calculated by principal component analysis.

Site	ANNA	CQI	MMI	BC50_PS	CAP_secS	Rank	Class
Br-bct		x	x	x	x	1	d
Gr-mrbtr		x	x		x	2	d
Sr-mqh				x	x	3	d
Gr-mcoy	x	x	x	x	x	4	d
Bn-wolf				x	x	5	d
Sr-mb61	x	x	x		x	6	d
Br-mcp	x	x	x	x	x	7	d
Br-cab			x			8	d
Sn-bbct	x		x			9	d
Sh-mill			x			10	d
Bn-lpk			x			11	d
Sr-mbhw				x	x	12	r
Sr-pb4				x	x	13	r
Bn-d12		x			x	14	r
Br-field			x			15	r
Bn-d13F				x		16	r
Nc-bct						17	r
Bn-holl		x	x			18	r
Br-hect						19	r
Sr-mc5				x		20	r
Br-lrt						21	r
Sr-mbm				x	x	22	r
Gr-prong			x			23	r
Sh-cedr						24	r
Nc-pmt	x	x				25	r
Br-cyp				x	x	26	r
Sg-joes						27	r
Sr-tc6						28	r
Sr-mq8						29	r
Sr-pbm						30	r
Sn-rogr			x			31	r
Mn-tav	x	x				32	r
Nc-bjc			x			33	r
Mn-mcra	x					34	r
Gr-bath						35	r
Br-bm						36	r
Sn-hemp			x			37	r
Bn-k11			x			38	r
Br-wp						39	r
Sr-mc6						40	r
Sr-pbhw				x	x	41	r
Br-deep			x			42	r
Sr-tcm						43	r
Sg-bone						44	r
Sr-tc5						45	r
Sr-mcm	x					46	r
Bn-k13						47	r
Bn-ljt						48	r
Br-flat			x			49	r
Br-jen						50	r
Sg-mlst	x	x	x			51	r
Br-lrf						52	r
Sr-mc7						53	r
Sr-mb6						54	r
Br-jun						55	r
Br-rfb						56	r
Br-gum						57	r
Gr-bog						58	r

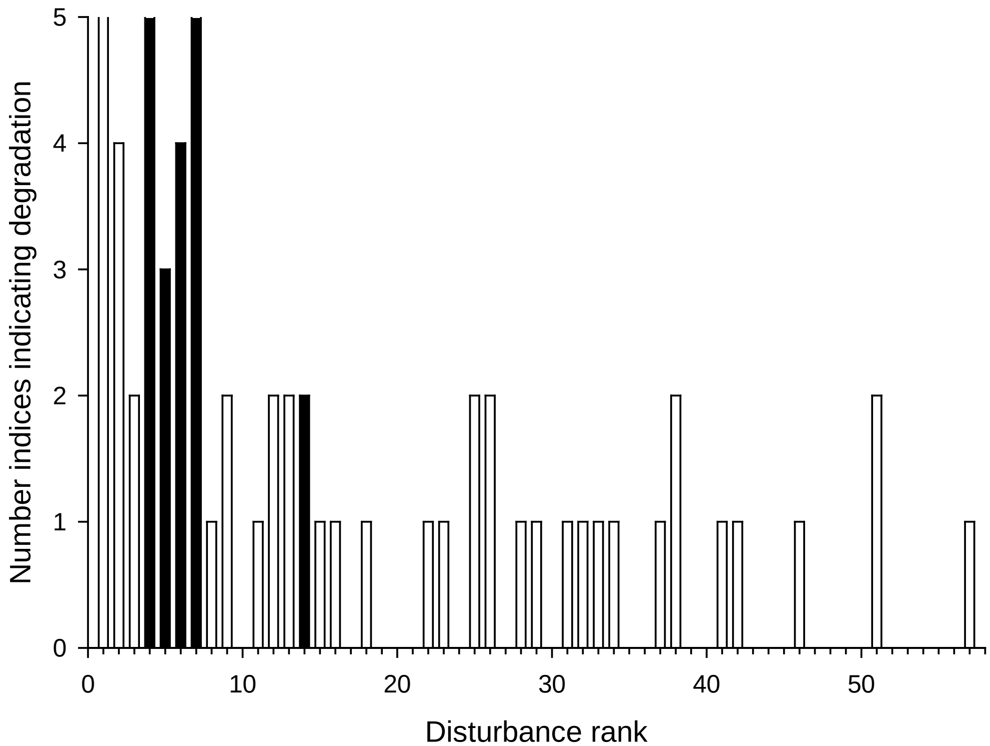


Figure 72. Number of bioassessment indices (out of five) indicating that a site differed from the reference condition. The 58 sites are ranked from most disturbed (1) to least disturbed (58). Solid bars indicate that a site was classified as disturbed by both fish and macroinvertebrate based indices.

Table 48. Number of samples required to detect differences (D) between two means for the fish multimetric index (MMI).

Statistic	D=1	D=2	D=3	D=4
Mean MMI value	42	42	42	42
Mean expected test value	41	40	39	38
Standard deviation	1.6	1.6	1.6	1.6
Power	0.80	0.80	0.80	0.80
Alpha	0.05	0.05	0.05	0.05
Required n for one-sided test	16	4	2	1
Required n for two-sided test	21	6	3	2

measure of inter-replicate variability needed for the computations (Table 48). The number of samples needed in each group, assuming a two-tailed test, ranged from 21 samples for a difference between means of 1 to 2 samples for a difference between means of 4. For a one tailed test, sample size requirements decreased to 16 and 1, respectively. These results indicate that only moderate to small numbers of samples are needed to detect comparatively small differences in scores and that the MMI is capable of distinguishing between levels of site quality (e.g., “good,” “ very good,” etc. as shown in Figure 71). The high level of similarity between the paired fish assemblage samples used in this analysis suggests the likelihood of comparable levels of statistical power for the ANNA and CQI methods.

5.5.4 *Performance Summary*

When interpreting the preceding results, it is important to consider all of the factors that affected the performance evaluations. These include the characteristics of the bioassessment indices plus the following:

1) The severity of disturbance at the disturbed sites: The inclusion of highly disturbed sites in a performance evaluation is likely to decrease the prevalence of Type 2 error because such sites are easier than moderately disturbed sites to discriminate from reference sites. There were relatively few highly disturbed sites in this study, which focused primarily on sampling reference sites. The inclusion of such sites, such as depauperate stream sites in densely populated areas, would have resulted in stronger performance evaluations by all methods.

2) The number of disturbed sites included in the evaluation: Related to number 1, the inclusion of more highly disturbed sites would likely enhance the performance evaluations for all methods.

3) The accuracy of the classification of sites as disturbed or undisturbed based on abiotic factors: The sites in this study were ranked on a PCA gradient that summarized a number of disturbance related abiotic factors but may not have included all salient factors or may have weighted the included factors suboptimally. Accurately measuring, evaluating, and summarizing the many abiotic factors that potentially affect biotic integrity is extremely difficult – a factor that has led to the widespread use of biotic indices rather than abiotic indicators to evaluate ecological health.

4) Sampling issues: Biological samples may occasionally lack representativeness because of unrecognized factors that result in inappropriately high or low bioassessment scores.

The performance evaluation indicated that most methods performed well at highly disturbed sites but sometimes evaluated reference sites as disturbed. They also suggest that the inclusion of both fish and macroinvertebrate methods can result in a more accurate assessment with greater potential to detect different types of environmental degradation. This is because different types of organisms differ in their responses and sensitivities to environmental degradation. Fish may be more sensitive than macroinvertebrates to metal pollution, and macroinvertebrates may be more sensitive than fish to organic pollution (Mount et al. 1984). Fish and macroinvertebrates may also differ in their rates of recovery from effects of disturbance (Yoder and Rankin 1995). Bioassessment accuracy can also be increased by using more than one assessment method for each taxonomic group since methods may differ in relative ability to detect different types of community responses to degradation. The application of multiple methods produces a “weight of evidence” that contributes to accuracy and a better evaluation of degradation. For example, the information on occurrence of different classes of species provided by the CQI and the relative values of different metrics provided by the MMI can contribute insights regarding possible causes of degradation; e.g., low scores for darter and benthic fluvial specialist metrics suggest degradation of the benthic environment by siltation or other factors.

6.0 CONCLUSIONS AND IMPLICATIONS FOR FUTURE RESEARCH/IMPLEMENTATION

The findings of RC-1694 are based on four years of field work in wadeable streams in the Sand Hills Ecoregion in NC, SC, and GA. The work was conducted in major DoD and DOE installations plus state, federal, and private holdings that protect large and relatively intact regional ecosystems including blackwater streams, a distinctive stream type of the southeastern coastal plain. Over seventy streams were sampled with an emphasis on sites that represented the least disturbed conditions remaining within the region. The objectives were to 1) develop reference models for ecosystem recovery based on the characteristics of fish and macroinvertebrate assemblages from least disturbed streams, 2) identify and evaluate habitat variables associated with least disturbed conditions, and 3) develop assessment frameworks for measuring progress towards benchmarks specified by the reference models. GIS, instream habitat, water quality, hydrogeomorphological, macroinvertebrate, and fish assemblage data were collected to meet these goals.

The following are the main conclusions of this study:

- Three independent methods were developed for reference site selection: method 1 involved use of a multivariate disturbance gradient derived from several stressors, method 2was based on variation in channel morphology, and method 3was based on passing 6 of 7 environmental criteria. Sites selected as reference by all 3 methods were considered primary reference, whereas those selected by 2 or 1 method were considered secondary or tertiary reference.
- Primary reference sites represented least disturbed conditions suitable for the development of reference models for full recovery, and secondary reference sites represented lesser quality but sustainable conditions for the development of best attainable models.
- The fish and benthic macroinvertbrate data collected from the reference sites provided a basis for developing reference models specifically designed for DoD installations and surrounding areas located in the Sand Hills.
- Fish and benthic macroinvertbrate faunas of the Sand Hills are heterogeneous, with river basin, stream size, and connection with larger streams being especially influential. This indicates a need for multiple reference models in which landscape or basin-scale classifications are supplemented with information concerning key environmental gradients.
- Hydrogeomorphic condition is biologically relevant and useful in determining if a stream is in the reference condition. A simple measure of the stream area at the top of the bank and its relation to the depth at the top of the bank (given watershed area) can be used to identify non-reference conditions.
- Least disturbed sites do not necessarily possess the highest fish species richness. Moderately disturbed sites may exhibit greater richness, likely because of faunal homogenization. Few bioassessment protocols consider this, which can lead to overrating of moderately disturbed and underrating of minimally disturbed sites.
- Fish and BMI assemblage structure responded predictable to disturbance. This, together with reference site data, provided the basis for several assessment frameworks.
- Assessment frameworks developed for the Sand Hills included multimetric indices that summarize scores from selected metrics, predictive models that relate observed taxonomic composition to the taxonomic composition expected under reference conditions, and a community quality index, specifically designed for Sand Hills fish assemblages.
- The predictive models developed under RC-1694 include RIVPACS type model, a ANNA type model, index of compositional dissimilarity (Bray-Curtis) models, and a novel predictive model termed O/E "capped."
- Data requirements for the assessment frameworks developed in this study range from minimal to moderate and ease of implementation from simple to more complex. This provides a range of options for managers with different resources.

- Most assessment frameworks performed well at highly disturbed sites but sometimes evaluated reference sites as disturbed.
- The use of both fish and macroinvertebrate assesment methods can result in more accurate assessments with greater potential to detect degradation.
- Bioassessment accuracy can be increased by using more than one assessment method for each taxonomic group since methods may differ in relative ability to detect different responses to degradation.

Bioassessment methods developed under RC-1694 provide an array of tools that can be used by DoD resource managers with varying resources and type of information to monitor the ecological health of lotic resources and assess their recovery to reference conditions characteristic of the highest expectations for the ecoregion or, where necessary, best attainable conditions imposed by factors that limit full recovery. Because the assessment frameworks represent different biotic communities as well as abiotic features, they provide a basis for comprehensive evaluations that include key communities of stream organisms with different environmental requirements as well as the habitat needed to support them. These tools can be used together to produce a weight-of-evidence assessment that provides a more accurate evaluation, greater understanding of the full range of ecosystem responses to recovery programs, and a better understanding of obstacles to recovery.

The methods developed in this study provide the basis for economical monitoring and assessment programs that provide information about the ecological health of wadeable streams. Their implementation necessitates acceptance by DoD resource managers, which can be facilitated by providing information and education concerning the methods. The usefulness of the methods can be further expanded by addressing the following issues:

- Adapting the methods to larger streams. The protocols developed in this study are designed for streams under about 5 m in width. Adapting the reference models and sampling strategies to larger streams would permit the assessment of a wider range of habitats.
- Comparing the methods developed under RC-1694 with existing bioassessment methods. Comparisons of this type could illustrate the relative advantages and disadvantages of different methods as well as way to combine methods for optimal results.
- Sampling additional disturbed sites. Sampling additional disturbed sites would permit a better understanding of disturbance gradients, which could help fine-tune metrics and better assess method accuracy.
- Designing effective monitoring programs. The optimal use of bioassessment frameworks for monitoring ecological health requires the selection of sampling sites and sampling frequencies that provide maximum information at minimum cost.

7.0 LITERATURE CITED

Allen, T.H.F. and T.B. Starr. 1982. Hierarchy: Perspectives for ecological complexity. University of Chicago Press, Chicago, IL.

Bailey, R.C., R.H. Norris, and T.B. Reynoldson. 2004. Bioassessment of freshwater ecosystems: using the reference condition approach. Kluwer Academic Publishers, New York, NY, USA.

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

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 15:185-211.

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling. 1999. Rapid Bioassessment Protocols for Use in Wadeable Streams and Rivers: Periphyton, Benthic Macroinvertebrates and fish, 2nd ed. EPA 841-B-99-002. Office of Water, United States Environmental Protection Agency, Washington, DC.

Beaty, S.R. 2011a. The Ephemeroptera of North Carolina: A biologist's handbook with standard taxonomic effort levels. North Carolina Department of Environment and Natural Resources, Division of Water Quality, Biological Assessment Unit.

Beaty, S.R. 2011b. The Plecoptera of North Carolina: A biologist's handbook with standard taxonomic effort levels. North Carolina Department of Environment and Natural Resources, Division of Water Quality, Biological Assessment Unit.

Beaty, S.R. 2011c. The aquatic Coleoptera of North Carolina: A biologist's handbook with standard taxonomic effort levels. North Carolina Department of Environment and Natural Resources, Division of Water Quality, Biological Assessment Unit.

Benjamini, Y. and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B (Methodological) 57:289-300.

Benke, A.C. and J.L. Meyer. 1988. Structure and function of a blackwater river in the southeastern U.S.A. Proceedings International Association of Theoretical and Applied Limnology 23:1209-1218.

Benke, A.C., R.L. Henry III, D.M. Gillespie, and R.J. Hunter. 1985. Importance of snag habitat for annual production in southeastern streams. Fisheries 10:8-13.

Borcard, D., F. Gillet, and P. Legendre. 2011. Numerical ecology with R. Springer, NY.

Boulesteix, A.L. and K. Strimmer. 2007. Partial least squares: a versatile tool for the analysis of high-dimensional genomic data. Briefings in Bioinformatics 8:32-44.

Boulesteix, A.L. 2004. PLS dimension reduction for classification with microarray data. Statistical Applications in Genetics and Molecular Biology 3:1-32.

Bunn, S.E. and A.H. Arthington. 2002. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. Environmental Management 30:492-507.

Bunte, K. and S.R. Abt. 2001. Sampling surface and subsurface particle-size distributions in wadeable gravel-and cobble-bed streams for analyses in sediment transport, hydraulics, and streambed monitoring. US Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, Colorado.

Burcher, C.L., H.M. Valett, and E.F. Benfield. 2007. The Land-Cover Cascade: relationships coupling land and water. Ecology 88:228-242.

Cam, E., J.D. Nichols, J.E. Hines, J.R. Sauer, R. Alpizar-Jara, and C.H. Flather. 2002. Disentangling sampling and ecological explanations underlying species–area relationships. Ecology 83:1118–1130

Cao, Y., D.D. Williams, and N.E. Williams. 1998. How important are rare species in aquatic community ecology and bioassessment? Limnology and Oceanography. 43:1402-1409.

Carlough, L.A. 1994. Origins, structure, and trophic significance of amorphous seston in a blackwater river. Freshwater Biology 31:227–237.

Carrascal, L.M., I. Galvan, and O. Gordo. 2009. Partial least squares regression as an alternative to current regression methods used in ecology. Oikos 118:681-690.

Chong, I.G. and C.H. Jun. 2005. Performance of some variable selection methods when multicollinearity is present. Chemometrics and Intelligent Laboratory Systems 78:103-112.

Chutter, F.M. 1972. An empirical biotic index of the quality of water in South African stream and rivers. Water Research 6:19-30.

Clarke, K.R. and R.M. Warwick. 2001. Change in marine communities: An approach to statistical analysis and interpretation, 2nd edition. PRIMER-E Ltd., Plymouth, UK.

Clarke, R.T., and J.F. Murphy. 2006. Effects of locally rare taxa on the precision and sensitivity of RIVPACS bioassessment of freshwaters. Freshwater Biology 51:1924-1940.

Clarke, R.T., J.F. Wright, and M.T.Furse. 2003. RIVPACS models for predicting the expected macroinvertebrate fauna and assessing the ecological quality of rivers. Ecological Modeling 160, 219–233.

Cohn, J.P. 1996. New defenders of wildlife. BioScience 46:11-14

Connell, J.H. 1978. Diversity in tropical rain forests and coral reefs. Science 199:1302-1310.

Cutler, D.R., T.C. Edwards, Jr., K.H. Beard, A. Cutler, K.T. Hess, J. Gibson, and J.J. Lawler. 2007. Random forests for classification in ecology. Ecology. 88:2783-2792.

Davies, P.E. 2000. Development of a national rivers bioassessment system (AUSRIVAS). In J.F. Wright, D.W. Sutcliffe, and M.T. Furse (eds), Assessing the biological quality of fresh waters; RIVPACS and other techniques. Freshwater Biological Association, Ambleside, UK. pp 113-124.

Davis, W.S. and J. Scott. 2000. Mid-Atlantic highlands streams assessment: Technical support document. EPA-903-B-00-004. U.S. Environmental Protection Agency, Region 3, Office of Research and Development, Ft. Meade, MD.

de Jong, S. 1993. SIMPLS: an alternative approach to partial least squares regression. Chemometrics and Intelligent Laboratory Systems 18:251-263.

Diamond, J.M., Barbour, M.T., and J.B. Stribling. 1996. Characterizing and comparing bioassessment methods and their results: a perspective. Journal of the North American Benthological Society 15:713-727.

Dillon, R. T., Jr., B.T. Watson, T. W. Stewart, and W. K. Reeves 2006. The freshwater gastropods of North America. Internet address: http://www.fwgna.org

Dingman, S.L. 2009. Fluvial hydraulics. Oxford University Press, New York, NY.

Doll, B.A., A.D. Dobbins, J. Spooner, D.R. Clinton, and D.A. Bidelspach. 2003. Hydraulic geometry relationships for rural North Carolina coastal plain streams Pages appendix A, in

Doll, B.A., Grabow G.L., Hall, K.R., Halley J., Harman, W.A., Jennings, G.D., and Wise, D.E. Stream restoration: a natural channel design handbook. NC Stream Restoration Institute, NC State University.

Doll, B.A., D.E. Wise-Frederick, C.M. Buckner, S.D. Wilkerson, W.A. Harman, R.E. Smith, and J. Spooner. 2002. Hydraulic geometry relationships for urban streams throughout the piedmont of North Carolina. JAWRA Journal of the American Water Resources Association 38:641-651.

Dunne, T. and L.B. Leopold. 1978. Water in environmental planning. W.H. Freeman Co. San Francisco, CA.

EPA (United States Environmental Protection Agency). 2006. Best practices for identifying reference condition in Mid-Atlantic streams. EPA-260-F-06-002. Office of Environmental Information, Washington, DC.

Epler, J.H. 2001. Identification Manual for the Larval Chironomidae (Diptera) of North and South Carolina. North Carolina Department and Natural Resources, Division of Water Quality.

Fausch, K.D., J.R. Karr, and P.R. Yant. 1984. Regional application of an index of biotic integrity based on stream fish communities. Transactions of the American Fisheries Society 113:39-55.

Filzmoser, P., R. Maronna, and M. Werner. 2008. Outlier identification in high dimensions. Computational Statistics and Data Analysis 52:1694-1711.

Flint, Jr., O.S. 1984. The genus *Brachycentrus* in North America, with a proposed phylogeny of the genera of Brachycentridae (Trichoptera). Smithsonian contributions to zoology, Smithsonian Institution Press. No. 398.

Flotemersch, J.E., J.B. Stribling, and M.J. Paul. 2006. Concepts and approaches for the bioassessment of non-wadeable Streams and Rivers. EPA 600-R-06-127. US Environmental Protection Agency, Cincinnati, OH.

Floyd, M.A., J.C. Morse, and J.V. McArthur. 1993. Aquatic insects of Upper Three Runs Creek, Savannah River Site, South Carolina. Part IV: Caddisflies (Trichoptera) of the lower reaches. Journal of the Georgia Entomological Society 28:85-95.

Fore, L, J.R. Karr, and R. Wisseman. 1996. Assessing invertebrate responses to human activities: evaluating alternative approaches. Journal of the North American Benthological Society 15:212-231.

Freeman, M.C. 2010. UGA Odum School of Ecology, personal communication.

Frissell, C.A., W.L. Liss, C.E. Warren, and M.D. Hurley. 1986. A hierarchical framework for stream habitat classification: Viewing streams in a watershed context. Environmental Management 19:199-214.

GDNR (Georgia Department of Natural Resources). 2005. Part III: Scoring criteria for the Index of Biotic Integrity and the Index of Well-Being to monitor fish communities in wadeable streams in the Apalachicola and Atlantic Slope drainage basins of the Southeastern Plains ecoregion of Georgia. Wildlife Resources Division, Fisheries Management Section. http://www.georgiawildlife.com/node/913

Garnier E., J. Cortez, G. Billes, M.L. Navas, C. Roumet, M. Debussche, G. Laurent, A. Blanchard, D. Aubry, A. Bellmann, C. Neill, and J.P. Toussaint. 2004. Plant functional markers capture ecosystem properties during secondary succession. Ecology 85:2630-2637.

GDNR (Georgia Department of Natural Resources). 2007. Macroinvertebrate biological assessment of wadeable stream in Georgia: standard operating procedures. Environmental Protection Division.

Gerritsen J., B .Jessup, E.W. Leppo, and J. White. 2000. Development of lake condition indexes (LCI) for Florida. Tetra Tech, Inc., Owings Mills, MD.

Gordon, N.D., T.A. McMahon, and B.L Finlayson. 1993. Stream hydrology: An introduction for ecologists. John Wiley and Sons, NY.

GRASS Development Team. 2008. Geographic Resources Analysis Support System (GRASS) software. Open Source Geospatial Foundation Project. http://grass.osgeo.org.

Griffith, G.E., J.M. Omernik, J.A. Comstock, S. Lawrence, G. Martin, A. Goddard, V.J. Hulcher, and T. Foster. 2001. Ecoregions of Alabama and Georgia, (color poster with map, descriptive text, summary tables, and photographs). US Geological Survey (map scale 1:1,700,000), Reston, Virginia.

Grime J. 1998. Benefits of plant diversity to ecosystems: immediate, filter and founder effects. Journal of Ecology 86:902-910.

Gschwandtner, M. and P. Filzmoser. 2010. Multivariate outlier detection based on robust methods. R package version 1.5. http://CRAN.R-project.org/package=mvoutlier.

Hammer, T.R. 1972. Stream channel enlargement due to urbanization. Water Resources Research 8:1530-1540.

Hart, D.D. and C.M. Finelli. 1999. Physical-biological coupling in streams: the pervasive effects of flow on benthic organisms. Annual Review of Ecology and Systematics 30:363-395.

Hassan-Williams, C.T., H. Bonner, and C. Thomas. 2010. Texas freshwater fishes, Texas State University-San Marcos Biology Department/Aquatic Station.

http://www.bio.txstate.edu/~tbonner/txfishes/index.htm.

Hawkins, C.P., J.R. Olson, and R.A. Hill. 2010. The reference condition: predicting benchmarks for ecological and water-quality assessments. Journal of the North American Benthological Society 29:312-343.

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.

Hawkins, C.P., Y. Cao, and B. Roper. 2010. Method of predicting reference condition biota affects the performance and interpretation of ecological indices. Freshwater Biology 55:1066–1085.

Helms, B.S., J.E. Schoonover, and J.W. Feminella. 2009. Assessing influences of hydrology, physiocochemistry, and habitat on stream fish assemblages across a changing landscape. Journal of the American Water Resources Association 45**:**157-169.

Herlihy, A.T., S.G. Paulsen, J. Van Sickle, J.L. Stoddard, C.P. Hawkins, and L.L. Yuan. 2008. Striving for consistency in a national assessment: the challenges of applying a reference-condition approach at a continental scale. Journal of the North American Benthological Society 27:860-877.

Hill, J., and G.D. Grossman. 1987. Home range estimates for three North American stream fishes. Copeia 1987:376-380.

Hilsenhoff, W.L. 1977. Use of arthropods to evaluate water equality of streams. Wisconsin Department of Natural Resources Technical Bulletin 100.

Hitt, N.P. and P.L. Angermeier. 2011. Fish community and bioassessment response to stream network position. Journal of the North American Benthological Society 30:296-309.

Hobbs, H.H. and D.M. Lodge. 2010. Chapter 22: Decapoda. *In*: Ecology and classification of North American freshwater invertebrates. 3rd ed. Thorp, J.H. and A.P. Covich (editors). Academic Press.

Horwitz, R.J. 1978. Temporal variability patterns and the distributional patterns of stream fishes. Ecological Monographs 48:307-321.

Hughes, R.M. 1995. Defining acceptable biological status by comparing with reference conditions. pp. 31-47. In W.S. Davis and T.P. Simon (eds.), Biological assessment and criteria: Tools for water resource planning and decision making for rivers and streams. CRC Press, Boca Raton, FL.

HydroGeoLogic. 2007. Proceedings from the Southeast Region Threatened, Endangered, and At-Risk Species Workshop, 27 February – 1 March 2007, Cocoa Beach, Florida. Prepared for the Legacy Resource Management Program, Strategic Environmental Research and Development Program, and Environmental Security Technology Certification Program by HydroGeoLogic Inc., Reston, VA.

Ihaka, R. and R. Gentleman. 1996. R: A language for data analysis and graphics. Journal of Computational and Graphical Statistics 5:299-314.

Jessup, B., C. Hawkins, and J. Stribling. 2006. Biological Indicators of stream condition in Montana using benthic macroinvertebrates. Montana Department of Environmental Quality.

Jongman R.H.G., C.J.F. Ter Braak, and O.F.R. van Tongeren (eds). 1995. Data analysis in community and landscape ecology. Cambridge University Press, Cambridge, USA.

Jordan R.A., K.S. Wheaton, W.M. Weiher, and T.J. Hayden. 1997. Integrated endangered species management recommendations for army installations in the southeastern United States. USACERL Special Report 97/94, U.S. Army Construction Engineering Research Laboratories, Champaign, IL.

Joy, M.K. and R.G. Death. 2003. Assessing biological integrity using freshwater fish and decapods habitat selection functions. Environmental Management 32:747-759.

Karr, J.R. 1981. Assessment of biotic integrity using fish communities. Fisheries 6:21-27.

Karr, J. R. 1991. Biological integrity: a long-neglected aspect of water resource management. Ecological Applications. 1: 66-84.

Karr, J.R., K.D. Fausch, P.L. Angermeier, P.R. Yant, and I.J. Schlosser. 1986. Assessing biological integrity in running waters: A method and its rationale, special publication 5. Illinois Natural History Survey, Champaign, IL.

Karr, J.R. and E.W. Chu. 1999. Restoring life in running waters, better biological modeling. Island Press, Washington, DC.

Kaufman, L. 2010. A base for war training, and species preservation. New York Times, Feb 22, A1

Kaufmann, P.R., P. Levine, E.G. Robinson, C. Seeliger, and D. Peck. 1999. Quantifying physical habitat in wadeable streams. EPA/620/R-99/003. Office of research and Development, US Environmental Protection Agency, Washington, DC.

Keddy, P.A. 1992. Assembly and response rules: two goals for predictive community ecology. Journal of Vegetation Science 3:157-164.

Kennard, M.J., B.J. Pusey, A.H. Arthington, B.D. Harch, and S.J. Machay. 2006. Development and application of a predictive model of freshwater fish assemblage composition to evaluate river health in eastern Australia. Hydrobiologia 572:33-57.

Kerans, B.L. and J.R. Karr. 1994. A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. Ecological Applications 4:768-785.

Klemm, D.J., K.A. Blocksom, F.A. Fulk, A.T. Herlihy, R.M. Hughes, P.R. Kaufmann, D.V. Peck, J.L. Stoddard, W.T. Thoeny, M.B. Griffith, and W.S. Davis. 2003. Development and evaluation of a Macroinvertebrate Biotic Integrity Index (MBII) for regionally assessing Mid-Atlantic Highlands streams. Environmental Management 31:656-669.

Knighton, D. 1998. Fluvial forms and processes. Rutledge, Chapman, and Hall, Inc., New York, NY.

Kosnicki, E., S.A. Sefick, M.H. Paller, M.S. Jarrell, B.A. Prusha, S.C. Sterrett, T.D. Tuberville, and J.W. Feminella. (submitted). Defining the reference condition for wadeable streams in the Sand Hills subdivision of the Southeastern Plains ecoregion, USA. Environmental Management.

Kosnicki, E. and R.W. Sites. 2007. Least-Desired Index for assessing the effectiveness of grass riparian filter strips in improving water quality in an agricultural region. Environmental Entomology 36:713-724.

Kosnicki, E. and R.W. Sites. 2011. Seasonal predictability of benthic macroinvertebrate metrics and community structure with maturity-weighted abundances in a Missouri Ozark stream, USA. Ecological Indicators 11:704-714.

Krstolic, J.L. and J.J. Chaplin. 2007. Bankfull regional curves for streams in the non-urban, non-tidal Coastal Plain physiographic province, Virginia and Maryland. USGS Scientific Investigation Report 2007-5162. U.S. Department of the Interior; U.S. Geological Survey, Reston, Virginia.

Lake, P.S. 2000. Disturbance, patchiness, and diversity in streams. Journal of the North American Benthological Society 19:573-592.

Lane, E. 1947. Report of the subcommittee on sediment terminology. Transactions, American Geophysical Union 28:936-938.

Lenat, D.R. 1993. A biotic index for the southeastern United States: Derivation and list of tolerance values, with criteria for assigning water-quality ratings. Journal of the North American Benthological Society 12:279-290.

Leopold, L.B. 1994. A view of the river. Harvard University Press, Cambridge, MA.

Leopold, L.B. and T. Maddock Jr. 1953. The hydraulic geometry of stream channels and some physiographic implications. U.S. Geological Survey Professional Paper 252.

Lepš, J. and P. Šmilauer. 2007. Multivariate analysis of ecological data using CANOCO. Cambridge University Press, New York, USA.

Linke, S., R.H. Norris, D.P. Faith, and D. Stockwell. 2005. ANNA: a new prediction method for bioassessment programs. Freshwater Biology 50:147-158.

Loehle, C., Wigley T.B., Schilling E., Tatum V., Beebe J., Vance E., Van Deusen P., Weatherford . 2009. Achieving conservation goals in managed forests of the Southeastern Coastal Plain. Environmental Management 44:1136-1148.

Lohr, S.C., and K.D. Fausch. 1997. Multiscale analysis of natural variability in stream fish assemblages of a western Great Plains watershed. Copeia 1997:706-724.

Maloney, K.O., P. J. Mulholland, and J.W. Feminella. 2005. Influence of catchment-scale military land use on stream physical and organic matter variables in small Southeastern Plains streams (USA). Environmental Management 35:677–691.

Maloney, K.O. and J.W. Feminella. 2006. Evaluation of single and multi-metric benthic macroinvertebrate indicators of catchment disturbance over time at the Fort Benning Military Installation, Georgia, USA. Ecological Indicators 6:469-484.

Marcinek, P. 2010. Georgia DNR Stream Survey Team, personal communication.

Marcy, B.C. Jr., D.E. Fletcher, F. D. Martin, M.H. Paller, and M.J.M. Reichert. 2005. Fishes of the middle Savannah River Basin. pp. 462. University of Georgia Press, Athens, GA.

Matthews, W.J. 1986. Fish faunal structure in an Ozark stream: stability, persistence and a catastrophic flood. Copeia 1986:388-397.

Matthews, W.J., R.C. Cashner, and F.P. Gelwick. 1988. Stability and persistence of fish faunas and assemblages in three midwestern streams. Copeia 1988:945-955.

Maxted, J.R., M.T. 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 North American Benthological Society 19:128-144.

McCoy, E.D. and H.R. Mushinsky. 2002. Measuring the success of wildlife community restoration. Ecological Applications 12:1861-1871.

McCune, B. and J.B. Grace. 2002. Analysis of ecological communities. MjM Software Design, Gleneden Beach, OR.

McCune, B. and M.J. Mefford. 1999. PC-ORD. Multivariate analysis of ecological data, version 4. MjM Software Design, Gleneden Beach, OR.

McMahon, G., J.D. Bales, J.F. Coles, E.M.P. Giddings, and H. Zappia. 2003. Use of stage data to characterize hydrologic conditions in an urbanizing environment. Journal of the American Water Resources Association 39:1529-1546.

Meffe, G.K. and T.M. Berra. 1988. Temporal characteristics of fish assemblage structure in an Ohio stream. Copeia 1988: 684-690.

Merritt R.W., K.W. Cummins, M.B. Berg (eds.). 2008. An introduction to the aquatic insects of North America. 4th ed. Kendall Hunt, Iowa.

Metcalf, C.K., S.D. Wilkerson, and W. A. Harman. 2009. Bankfull regional curves for north and northwest Florida streams. Journal of the American Water Resources Association 45:1260-1272.

Mevik, B.H. and R. Wehrens. 2007. The pls package: principal component and partial least squares regression in R. Journal of Statistical Software 18:1-24.

Meyers, J.L., A.C. Benke, R.T. Edwards, and J.B. Wallace. 1997. Organic matter dynamics in the Ogeechee River, a blackwater river in Georgia, USA**.** Journal of the North American Benthological Society 16:82-87.

Morse, J.C., J.W. Chapin, D.D. Herlong, and R.S. Harvey. 1980. Aquatic insects of Upper Three Runs, Savannah River Plant, South Carolina, Part I: Orders other than Diptera. Journal Georgia Entomological Society 15:73-101.

Morse, J.C., J.W. Chapin, D.D. Herlong, and R.S. Harvey. 1983. Aquatic insects of Upper Three Runs, Savannah River Plant, South Carolina, Part I: Diptera. Journal Georgia Entomological Society 18:303-316.

Nagy, R.C., B.G. Lockaby, B. Helms, L. Kalin, and D. Stoeckel. 2011. Water resources and land use and cover in a humid region: the southeastern United States. Journal of Environmental Quality 40: 867-878.

NCDENR (North Carolina Department of Environmental and Natural Resources). 2006a. Standard operating procedure: Biological monitoring: Stream fish community assessment program. Division of Water Quality, Environmental Sciences Section, Biological Assessment Unit. http://portal.ncdenr.org/web/wq/ess/bau.

NCDENR (North Carolina Department of Environment and Natural Resources). 2006b. Standard operating procedures for benthic macroinvertebrates biological assessment unit. Division of Water Quality, Environmental Science Section.

Omernik, J.M. 1987. Ecoregions of the conterminous United States. Annals of the Association of America Geographers 77:118-125.

Osborne, L.L. and M.J. Wiley. 1992. Influence of tributary spatial position on the structure of warmwater fish communities. Canadian Journal of Fisheries and Aquatic Sciences 49: 671-681.

Osborne, L.L., S.L. Kohler, P.B. Bailey, D.M. Day, W.A. Bertrand, M.J. Wiley, and R. Sauer. 1992. Influence of stream location in a drainage network on the index of biotic integrity. Transactions of the American Fisheries Society 121:635-643.

Paller, M.H. 1994. Relationship between fish assemblage structure and stream order in South Carolina coastal plain streams. Transactions of the American Fisheries Society 123:150-161.

Paller, M.H. 1995. Relationships among number of fish species sampled, reach length surveyed, and sampling effort in South Carolina coastal plain streams. North American Journal of Fisheries Management 15:110-120.

Paller, M.H. 2001. Comparison of fish and macroinvertebrate bioassessments from South Carolina coastal plain streams. Aquatic Ecosystem Health and Management 4: 175-186.

Paller, M.H., M.J.M. Reichert, and J.M. Dean. 1996. The use of fish communities to assess environmental impacts in South Carolina coastal plain streams. Transactions of the American Fisheries Society 125:633-644.

Paller, M.H., J. Dean, M.J.M. Reichert, C. Trettin, and W.L. Specht. 2000. Use of fish community data to evaluate restoration success in Pen Branch. Ecological Engineering 15: S171-S187.

Paller, M.H., F.D. Martin, L.D. Wike, and W.L. Specht. 2007. Factors influencing the accuracy of a macroinvertebrate bioassessment protocol in South Carolina coastal plain streams. Journal of Freshwater Ecology 22:23-32.

Patrick, R. 1996. Rivers of the United States, Volume 3, The Eastern and Southeastern States. John Wiley & Sons Inc., Somerset, NJ.

Plafkin, J.L., M.T. Barbour, K.D. Porter, S.K. Gross, and R.M. Hughes. 1989. Rapid bioassessment protocols for use in streams and rivers. Benthic macroinvertebrates and fish. EPA 440-4-89-001. Office of Water Regulations and Standards, U.S. Environmental Protection Agency, Washington, D.C.

Poff, N.L. 1997. Landscape filters and species traits: Towards mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological Society 16:391-409.

Power, M.E., A. Sun, G. Parker, W.E. Dietrich, and J.T. Wotton. 1995. Hydraulic food-chain models. BioScience 45:159-167.

Prasad, A.M., L.R. Iverson, and A. Liaw. 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems 9:181-199.

Quist, M.C., P.A. Fay, C.S. Guy, A.K. Knapp, and B.N. Rubenstein. 2003. Military training effects on terrestrial and aquatic communities on a grassland military installation. Ecological Applications 13:432-442.

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

R Development Core Team (R-core). 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org/.

Rader, R.B., J.V. McArthur, and J.M. Aho. 1994. Relative importance of mechanisms determining decomposition in a southeastern blackwater stream. American Midland Naturalist 132:19-31.

Rahel, F.J. 2000. Homogenization of fish faunas across the United States. Science 288:854-856.

Rahel, F.J. and W.A. Hubert. 1991. Fish assemblages and habitat gradients in a Rocky Mountain-Great Plains stream: biotic zonation and additive patterns of community change. Transactions of the American Fisheries Society 120:319-332.

Rhode, F.C., R.G Arndt, J.W. Foltz, J.M. Quattro. 2009. Freshwater fishes of South Carolina. The University of South Carolina Press. Columbia, SC.

Rodríguez, M.A. 2002. Restricted movement in stream fish: the paradigm is incomplete, not lost. Ecology, 83, 1-13.

Rose, S. and N.E. Peters. 2001. Effects of urbanization on streamflow in the Atlanta area (Georgia, USA): a comparative hydrological approach. Hydrological Processes 15:1441-1457.

Rosenzweig, M.L. 1995. Species Diversity in Space and Time. Cambridge: Cambridge University Press.

Rosenzweig, M.L., W.R. Turner, J.G. Cox, and T.H. Ricketts. 2003. Estimating diversity in unsampled habitats of a biogeographical province. Conservation Biology 17:864–874.

Rosgen D.L. and H.L. Silvey. 1996. Applied river morphology. Wildland Hydrology Pagosa Springs, Colorado.

Rousseeuw P.J. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics 20:53-65.

Sabater, F., J.L. Meyer, and R.T. Edwards. 1993. Longitudinal patterns of dissolved organic carbon concentration and suspended bacterial density along a blackwater river. Biogeochemistry 21:73-93.

Sæther, O. A. 1979. Chironomid communities as water quality indicators. Holarctic Ecology 2:65-74.

SAS Institute Inc. (2004) SAS/STAT® 9.1 User's Guide. Cary, NC: SAS Institute Inc.

SCDHEC (South Carolina Department of Health and Environmental Control). 1998. Standard operating and quality control procedures for macroinvertebrate sampling. Technical Report No. 004-98. Bureau of Water, Division of Water Monitoring, Assessment and Protection, Aquatic Biology Section.

Schlosser, I.J. 1990. Environmental variation, life history attributes, and community structure in stream fishes: implications for environmental management and assessment. Environmental Management 14:621-628.

Schoonover, J.E., B.G. Lockaby, and B.S. Helms. 2006. Impacts of stream hydrology in the west Georgia piedmont, USA. Journal of Environmental Quality 35:2123:2131.

Scott, M.C. and G.S. Helfman. 2001. Native invasions, homogenization, and the mismeasure of integrity of fish assemblages. Fisheries 26: 6-15.

Sheldon, A.L. 1968. Species diversity and longitudinal succession in stream fishes. Ecology 49:193-198.

Shipley B., D. Vile, and E. Garnier. 2006. From plant traits to plant communities: a statistical mechanistic approach to biodiversity. Science 314:812-814.

Simon, T.P. and J. Lyons. 1994. Application of the index of biotic integrity to evaluate water resource integrity in freshwater ecosystems. pp. 245-262. In W.S. Davis and T.P. Simon (eds.), Biological assessment and criteria: Tools for water resource planning and decision making in rivers and streams. Lewis Publishers, Boca Raton, FL.

Simpson, J.C. and R.H. Norris. 2000. Biological assessment of river quality: Development of the AUSRIVAS models and outputs. pp. 125-142. In Wright, J.F., D.W. Sutcliffe, M.T. Furse, (eds.), Assessing the biological quality of freshwaters: RIVPACS and other techniques. Freshwater Biological Association, Ambleside, Cumbria, UK.

Sokol, E.R., E. Benfield, L.K. Belden, and H.M. Valett. 2011. The assembly of ecological communities inferred from taxonomic and functional composition. American Naturalist 177: 630-644.

Sonesten, L. 2003. Catchment area composition and water chemistry heavily affects mercury levels in perch (*Perca fluviatilis* L.) in circumneutral lakes. Water, Air, and Soil Pollution 144:117-139.

Southerland, M.T. and J.B. Stribling. 1995. Status of biological criteria development and implementation. pp. 81-96. In W.S. Davis and T.P. Simons (eds.), Biological assessment and criteria: Tools for water resource planning and decision making in rivers and streams. CRC Press, Boca Raton, FL.

Southwood, T.R.E. 1977. Habitat, the templet for ecological strategies? Journal of Animal Ecology 46:337-365.

Stoddard, J.L., D.V. Peck, A.R. Olsen, D.P. Larsen, J. Van Sickle, C.P. Hawkins, R.M. Hughes, T.R. Whittier, G. Lomnicky, A.T. Herlihy, P.R. Kaufmann, S.A. Peterson, P.L. Ringold, S.G. Paulsen, and R. Blair. 2005. Environmental Monitoring and Assessment Program (EMAP): western streams and rivers statistical summary. EPA 620/R-05/006. Office of Research and Development, US Environmental Protection Agency, Washington, DC.

Stoddard, J.L., D.P. Larsen, C.P. Hawkins, R.K. Johnson, and R.H. Norris. 2006. Setting expectations for the ecological condition of streams: The concept of reference condition. Ecological Applications 16:1267-1276.

Stoddard, J.L., A.T. Herlihy, D.V. Peck, R.M. Hughes, T.R. Whittier, and E. Tarquinio. 2008. A process for creating multimetric indices for large-scale aquatic surveys. Journal of the North American Benthological Society. 27:878-891.

Straight, C.A., B. Albanese, and B.J. Freeman. 2009. Fishes of Georgia Website, Georgia Museum of Natural History. http://fishesofgeorgia.uga.edu.

Swift, C.C., C.R. Gilbert, S.A. Bortone, G.H. Burgess, and R.W. Yerger. 1986. Zoogeography of the freshwater fishes of the southeastern United States: Savannah River to Lake Pontchartrain. pp 213-265 In C.H. Hocutt and E.O. Wiley (eds.), The zoogeography of North American freshwater fishes. John Wiley and Sons, New York.

Szymanska, E, E. Saccenti, A.K. Smilde, and J.A. Westerhuis. 2012. Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies. Metabolomics 8:3-16.

Taylor, C.M., M.R. Winston, and W.J. Matthews. 1996. Temporal variation in tributary and mainstem fish assemblages in a Great Plains stream system. Copeia 1996:280-289.

Thornbrugh, D.J. and K.B. Gido. 2010. Influence of spatial positioning within stream networks on fish assemblage structure in the Kansas River basin. Canadian Journal of Fisheries and Aquatic Sciences 67: 143-156.

Tonn, W.M. 1990. Climate change and fish communities: A conceptual model approach. Transactions of the American Fisheries Society 119:337-352.

Townsend, C.R. and A.G. Hildrew. 1994. Species traits in relation to a habitat templet for river systems. Freshwater Biology 31:265-275.

Tsang, Y.P., G.K. Felton, G.E. Moglen, and M. Paul. 2011. Region of influence method improves macroinvertebrate predictive models in Maryland. Ecological Modeling. 222:3473-3485.

Turkmen, A. and N. Billor. 2013. Partial least squares classification for high dimensional data using the PCOUT algorithm. Computational Statistics 28:771-778.

U.S. Census Bureau PD. 2005. Interim state population projections. Washington, D.C.

USEPA (US Environmental Protection Agency). 2006. Wadeable streams assessment: the states assess the nation's streams. EPA 841/B/06/002. Office of Research and Development, US Environmental Protection Agency, Washington, DC.

USFWS (US Fish and Wildlife Service). 2003. Red-Cockaded Woodpecker (*Picoides borealis*) Recovery, second ed. Atlanta, GA, p. 296.

Van der Laan, M., K. Pollard, and J. Bryan. 2003. A new partitioning around medoids algorithm.

Journal of Statistical Computation and Simulation 73:575-584.

Van Sickle, J, D.D. Huff, and C.P. Hawkins. 2006. Selecting discriminant function models for predicting the expected richness of aquatic macroinvertebrates. Freshwater Biology 51:359-372.

Van Sickle, J. 2008. An index of compositional dissimilarity between observed and expected assemblages. Journal of the North American Benthological Society 27:227-235.

Van Sickle, J., D. P. Larsen, and C. P. Hawkins. 2007. Exclusion of rare taxa affects performance of the O/E index in bioassessments. Journal of the North American Benthological Association 26:319-331.

Vannote, R.L., G.W. Minshall, K.W. Cummins, J.R. Sedell, and C.E. Cushing. 1980. The river continuum concept. Canadian Journal of Fisheries and Aquatic Sciences 37:130-137.

Venables, W.N. and B.D. Ripley. 1994. Modern applied statistics with S-PLUS. 4th edition. Springer-Verlag, NY.

Vieira, N.M.K., N.L.M. Poff, D.M. Carlisle, S.R. Moulton II, M.L. Koski, and B.C. Kondratieff. 2006. A database of lotic invertebrate traits for North America. Available: http://pubs.usgs.gov/ds/ds187/.

Vinson, M.R. and C.P. Hawkins. 1996. Effects of sampling area on subsampling procedures on composition of taxa richness among streams. Journal of the North American Benthological Society 15:392-399.

Waite, I.R., A.T. Herlihy, D.P. Larsen, and D.J. Klemm. 2000. Comparing strengths of geographic and nongeographic classifications of stream benthic macroinvertebrates in the Mid-Atlantic Highlands, USA. Journal of the North American Benthological Society 19:429-441.

Wallace, J.B. and A.C. Benke. 1984. Quantification of wood habitat in subtropical Coastal Plain streams. Canadian Journal of Fisheries and Aquatic Sciences 41:1643–1652.

Wallace, J.B. and J. Grubaugh. 1996. Transport and storage of FPOM Pages 191-215, in Hauer R and Lamberti GA. Methods in stream ecology. Academic Press, San Diego, CA.

Wallace, J.B. and J.R. Webster. 1996. The role of macroinvertebrates in stream ecosystem function. Annual Review of Entomology 41:115-139.

Wallace, J.B., J.W. Grubaugh, and M.R. Whiles. 1996. Biotic indices and stream ecosystem processes: Results from an experimental study. Ecological Applications 6:140-151.

Wallace, J.B., J.J. Hutchens, and J.W. Grubaugh. 2006. Transport and storage of FPOM. *In* Hauer, F.R. and G.A. Lamberti (eds.), Methods in stream ecology. 2nd edition. Elsevier, Amsterdam.

Warren, M. L. Jr., B.M. Burr, S.J. Walsh, H.L. Bart Jr., R.C. Cashner, D.A. Etnier, B.J. Freeman, B.R. Kuhajda, R.L. Mayden, H.W. Robinson, S.T. Ross, and W.C. Starnes. 2000. Diversity, distribution, and conservation status of the native freshwater fishes of the Southern United States. Fisheries 25:10:7-29.

Washington, H.G. 1984. Diversity, biotic and similarity indices: A review with special relevance to aquatic ecosystems. Water Resources 18:653-694.

Whitney, E., D.B. Means, and A. Rudloe. 2004. Priceless Florida. Pineapple Press Inc., Sarasota, FL.

Whittier T.R., J.L Stoddard., D.P.Larsen, and A.T. Herlihy. 2007. Selecting reference sites for stream biological assessments: best professional judgment or objective criteria. Journal of the North American Benthological Society 26:349-360.

Wiggins, G.B. 1996. Larvae of the North American caddisfly genera (Trichoptera). 2nd ed. University of Toronto Press, Toronto.

Wolman, M.G. and J.P. Miller. 1960. Magnitude and frequency of forces in geomorphic processes. Journal of Geology 68:54-74.

Wright, J.F. 1995. Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. Australian Journal of Ecology 20:181-197.

Wright, J.F., D. Moss, P.D. Armitage, and M.T. Furse. 1984. A preliminary classification of running-water sites in Great Britain based on macro-invertebrate species and the prediction of community type using environmental data. Freshwater Biology 14:221-256.

Wright, J.F., D.W. Sutcliffe, and M.T. Furse (eds). 2000. Assessing the biological quality of freshwaters: RIVPACS and other techniques. Freshwater Biological Association. Ambleside, Cumbria, UK.

8.0 APPENDICES

A. Supporting Data

A.1. R code for calculating O/E models for the Sand Hills subdivision of the Southeastern Plains Ecoregion

```
## Required components:
## 1. A csv file of predictor data given in Table 37; the file should be named "pred_new2"; the first
## column must be called "Sample" and contains the sample sites
## 2. A csv file of taxonomic data and counts called "mytaxa". The OTUs in Table 35 must
## be the column headings for each taxon. The first column must be called "SAMPLE" and
## contains the sample sites identical' to the names or codes given in the predictor data set.

## Once you have saved your model(s) you can use this program to load the model and use
## VanSickle's function to predict Expected and calculate the O/E, Prob of group, and Pc scores
## Put all data files, models (R Workspace), functions, etc.. into working directory

## First load relevant packages
library(Hmisc);
library(randomForest);
library(reshape2);
library(reshape);

## set your working directory
setwd("")

## Load your data files; in this example I use the original RC-1694 data
## Also include the original calibration data from the RC-1694 dataset
## NOTE: secondary and primary references will be different files from each other
## Rename or select for specific classes
predall<-read.csv("pred_new2.csv",row.names="Sample",header=T);
bugall<-read.csv("mytaxa.csv",row.names="SAMPLE",header=T)

bugall.pa<-bugall;
bugall.pa[bugall.pa>0]<-1;  #This is the presence absence matrix

predxport<-predall
bugxport.pa<-bugall.pa

## Make sure row names of your input bug and variable datasets match
bugxport.pa<-bugxport.pa[row.names(predxport),]         ## Fix row names so they match
row.names(bugxport.pa)==row.names(predxport)  ## See if row names match

## Load VanSickle function and your RF predictive model
## Pay close attention to the formatting
source("model.predict.RanFor.4.2.r");
load('.rdat');

## Note sample sites will be dropped if they do not have all the model predictors
predxport<-predxport[complete.cases(predxport[,preds.final]),];
```

```
bugxport.pa<-bugxport.pa[row.names(predxport),];

## Make predictions... Note that OOB=FALSE
OE.results<-model.predict.RanFor.4.2(bugcal.pa,grps.final,preds.final,ranfor.mod=rf.mod,
prednew=predxport,bugnew=bugxport.pa,Pc=0.000000005,Cal.OOB=FALSE);
OE50.results<-model.predict.RanFor.4.2(bugcal.pa,grps.final,preds.final,ranfor.mod=rf.mod,
prednew=predxport,bugnew=bugxport.pa,Pc=0.5,Cal.OOB=FALSE);
OE.results$OE.scores[,c(3,7)]; #data frame of O/E scores;
OE50.results$OE.scores[,c(3,7)];  #data frame of O/E50 scores;
OE.results$Capture.Probs; #predicted capture probabilties;
OE.results$Group.Occurrence.Probs; #predicted group occurrence probabilities;

## OK THE NEXT STEP IS TO MERGE RESULTS, SCORES ETC... DONE BY MAKING NEW DF
WITH ifelse()
## first check that row and colnames are the same
row.names(bugxport.pa)==row.names(OE.results$Capture.Probs)
colnames(bugxport.pa)==colnames(OE.results$Capture.Probs)

OE.bug<-data.frame(ifelse(bugxport.pa>=OE.results$Capture.Probs, OE.results$Capture.Probs, 0))
cap<-data.frame(rowSums(OE.bug))
cap<-rename(cap, c(rowSums.OE.bug.="OEcap"))
OE.scores<-data.frame(OE.results$OE.scores[,c(3,7)])
OE.scores<-rename(OE.scores, c(OoverE="OE"))
OE50.scores<-data.frame(OE50.results$OE.scores[,c(3,7)])
OE50.scores<-rename(OE50.scores, c(OoverE="OE50", BC="BC50"))
OEcapit<-data.frame(cap, OE.results$OE.scores[,c(2)])
OEcapit<-rename(OEcapit, c(OE.results.OE.scores...c.2..="E"))
OEcap<-data.frame(OEcapit, 100*OEcapit$OEcap/OEcapit$E)
OEcap<-rename(OEcap, c(X100...OEcapit.OEcap.OEcapit.E="CAP"))
myOEscores<-cbind(OE.scores, OE50.scores, OEcap)

## Name your Exported CSV file
```

A.2. Definitions of environmental and stream benthic macroinvertebrate variables used in Partial Least Squares (PLS) regression modeling. Variables highlighted grey were those retained in PLS modeling after the variable importance in the projection (vip), variable selection step (see text). GIS LULC = Land use/Land Cover from GIS data.

Variable Category	Variable Class	Variable Condition	Variable Code	Variable Definition
ENVIRONMENTAL	***LANDSCAPE***		elevation	Stream Elevation
			slope	GIS stream slope
			devel	GIS LULC Low+Medium+High Development
			forest	GIS LULC mixed+evergreen+deciduous forest+woody wetlands
	HYDROLOGY (STAGE)	**Duration of Low Stage**	MAX_q10	Maximum duration below 10th quartile
			MAX_q25	Maximum duration below 25th quartile
			MAX_q5	Maximum duration below 5th quartile
			MED._q10	Median duration below 10th quartile
			MED._q25	Median duration below 25th quartile
			MED._q5	Median duration below 5th quartile
		Duration of High Stage	MAX_q75	Maximum duration above 75th quartile
			MAX_q90	Maximum duration above 90th quartile
			MAX_q95	Maximum duration above 95th quartile
			MED._q75	Median duration above 75th quartile
			MED._q90	Median duration above 90th quartile
			MED._q95	Median duration above 95th quartile
		Frequency of Stage Change (Flashiness)	rising 15.24 cm	Number of hours stage rises by at least 15.24cm
			rising 21.336 cm	Number of hours stage rises by at least 21.336cm
			rising 27.432 cm	Number of hours stage rises by at least 27.432m
			rising 3.048 cm	Number of hours stage rises by at least 3.048cm
			rising 9.144 cm	Number of hours stage rises by at least 9.144cm
			falling 15.24 cm	Number of hours stage falls by at least 15.24cm
			falling 21.336 cm	Number of hours stage falls by at least 21.336cm
			falling 27.432 cm	Number of hours stage falls by at least 27.432cm
			falling 3.048 cm	Number of hours stage falls by at least 3.048cm
			falling 9.144 cm	Number of hours stage falls by at least 9.144cm
			CV	Coefficient of Variation for the Period of Record of stage
	HYDROLOGY (Discharge)		max_Q	Maximum Discharge
			median_Q	Median Discharge

	HYDRAULICS		max_power	Maximum Stream Power
			median_power	Median Stream Power
			max_Fr	Maximum Froude Number
			median_Fr	Median Froude number
			max_t	Maximum Tractive Force
			median_t	Median Tractive Force
			prop_sed_move	Proportion of the time median particle entrained
			max_U_shear	Maximum Shear Velocity
			median_U_shear	Median Shear Velocity
	INSTREAM HABITAT	**Organic Matter**	CWD	Coarse woody debris (wood >0.25m diam.)
			OM	Benthic organic matter deposited in channel
		Stream Chemistry	pH	pH
			SC	Specific Conductance
			DO %	Dissolved oxygen %
			Temperature	Water Temperature
		Insolation	prop_NC	Proportion stream canopy
		Bed Sediment	Sed Mean	Mean diameter of substrate particles in stream bed
			SED SD	Substrate size Standard Deviation
MACROINVERTEBRATES		**Reproductive Preference**	ovi1	oviposition (egg laying) on algal mats
			ovi2	oviposition (egg laying) on bank soil
			ovi3	oviposition (egg laying) on bed substrate
			ovi4	oviposition (egg laying) on floating debris
			ovi5	oviposition (egg laying) on moss/ submerged macrophytes
			ovi6	oviposition (egg laying) on wet wood
			ovi7	oviposition (egg laying) on/under stones
			ovi8	oviposition (egg laying) on overhanging substrate dry
			eggc1	Laying cemented eggs
			eggc2	Laying Non-cemented eggs
	HABITAT	**Waterbody Preference**	wb1	Lentic
			wb2	Warm spring
			Cold Spring	Cold spring
			Headwater	Headwater
			wb5	2nd to 4th order
			wb6	River
			wb7	Temporary habitat
		Current Preference	vel1	Quiet current

		Slow Current	Slow current
		vel3	Fast laminar current
		Fast Current	Fast turbulent current
	Current Adaptation	mflo1	Adaptations to flow
		mflo2	No flow adaptations
	Microhabitat Preference	mh1	Sand
		mh2	Silt
		mh3	Gravel
		mh4	Rocks
		mh5	Boulder
		mh6	Large woody debris
		Detritus Micro. Hab.	Detritus
		mh8	Phytoplankton
		Algae	Algae
		mh10	Pelagic
	Lateral Preference	lat1	Lotic margin
		lat2	Lentic shore
		lat3	Pools
		lat4	Riffle
		lat5	Hyporheic (subsurface)
	Vertical Preferences	vrt1	Water surface
		Macrophytes	Aquatic plants
		vrt3	Pelagic
		Bed Hab. Pref.	Stream bed
		vrt5	Hyporheic (subsurface)
PHYSIOLOGICAL	**Oxygen Tolerance**	oxy1	normal
		oxy2	low
	Chemical Tolerance	pH1	Acidic (<6.0)
		pH2	Circumneutral (~7.0)
		pH3	Alkaline (>8.0)
		sal1	Fresh water
		sal2	Brackish
		sal3	Salt water
	Thermal Preference	thrm1	Cold water
		thrm2	No temperature preference
		thrm3	Warm water

	Turbidity Preference	trb1	Low turbidity
		trb2	No turbidity preference
		trb3	High turbidity
ECOLOGY	**Functional Feeding Groups**	ffg1	Collector-filterer
		Collector Gatherer	Collector-gatherer
		ffg3	Parasite
		ffg4	Predator
		ffg5	Scraper/grazer
		Shred.	shredder
	Habit	hab1	Burrower
		hab2	Climber
		hab3	Clinger
		hab4	Sprawler
		hab5	Swimmer
LIFE HISTORY		2 aqu. stages	egg and nymph
		ast3	egg, larvae, and pupae
		ast4	egg, larvae, pupae, and adult
	Voltinism	vlt1	Semivoltine (<1 generation per year)
		Uni Volt.	Univoltine (1 generation per year)
		Multi Volt.	Multivoltine (>1 generation per year)
	Development Speed	dsp1	Slow seasonal cycle
		Fast Seasonal	Fast seasonal cycle
		dsp3	Non-seasonal
	Adult Life Span	Short Adult Life.	Hours
		Moderate Adult Life.	Weeks
		Long Adult Life.	Months
	Fecundity	fec1	<100 eggs
		fec2	100-1000
		fec3	1000-10000
		Diapause	Diapause (resting stage)
		No Diapause	No Diapause (resting stage)
MOBILITY	**Drift**	drf1	Weak drifter
		drf2	Passive/occasional drifter
		Strong Drift	active/frequent drifter
	Larval Dispersal	lds1	Dispersal distance <1m
		lds2	Dispersal distance 1-10m

		lds3	Dispersal distance 11-100m
	Adult Dispersal	dis1	Dispersal distance 10m
		dis2	Dispersal distance 1km
		dis3	Dispersal distance 10km
		dis4	Dispersal distance 100km
		ext1	Ability to temporarily exit water
		ext2	Inbility to temporarily exit water
MORPHOLOGY	**Larval size**	siz1	length <9mm
		siz2	length 9-16mm
		siz3	length >16mm
	Body shape	Bluff	Bluff (bricklike)
		shp2	Round (humped)
		shp3	Tubular
		shp4	Streamlined/fusiform
		D. Flat.	Dorsoventrally flattened
	Sclerotization (body armor)	arm1	Soft (unhardened) body
		arm2	Partially sclerotized (hardened)
		Hard Shelled	Hard Shelled
		arm4	All sclerotized (completely hardened)
	Respiration mode	rsp1	Cutaneous respiration (through cuticle)
		rsp2	Spiracular gills (water breather)
		rsp3	Hemolymph with hemoglobin
		rsp4	Tracheal gills (water breather)
		rsp5	Atmospheric respiration (air breather)
		rsp6	Plastron (water breather)
		rsp7	Temporary air store (air breather)

B. List of Scientific/Technical Publications

1. Articles in peer-reviewed journals

Kosnicki, E., S.A. Sefick, M.S. Jarrell, M.H. Paller, B.A. Prusha, S.C. Sterrett, T.D. Tuberville, and J.W. Feminella. 2014. Defining the reference condition for wadeable streams in the Sand Hills subdivision of the Southeastern Plains ecoregion, USA. Environmental Management. (DOI) 10.1007/s00267-014-0320-0.

Kosnicki, E., S.A. Sefick, M.S. Jerrell, S.C. Sterrett, M.H. Paller, B.A. Prusha, T.D. Tuberville, and J.W. Feminella. Development of best attainable and least disturbed macroinvertebrate reference models for wadeable streams in the Sand Hills of the Southeastern Plain ecoregion, USA. Formatted for Freshwater Science. (in prep).

Paller, M.H., B.A. Prusha, Kosnicki, E., S.A. Sefick, M.S. Jarrell, S.C. Sterrett, T.D. Tuberville, and J.W. Feminella. Factors influencing reference site fish assemblages in the Sand Hills ecoregion of the southeastern United States. Submitted to Transactions of the American Fisheries Society.

Paller, M.H., B.A. Prusha, Kosnicki, E., S.A. Sefick, M.S. Jarrell, S.C. Sterrett, T.D. Tuberville, and J.W. Feminella. A comparison of bioassessment methods for streams in the Sand Hills ecoregion of the southeastern United States. (in prep).

Paller, M.H., S.C. Sterrett, T. D. Tuberville, D. E. Fletcher, and A. M. Grosse. 2014. Effects of disturbance at two spatial scales on macroinvertebrate and fish metrics of stream health, Journal of Freshwater Ecology, 29:1, 83-100.

Sefick, S.A., L. Kalin, E. Kosnicki, B.P. Schneid, M.S. Jarrel, C.J. Anderson, M.H. Paller, and J.W. Feminella. Estimating discharge in low-gradient, sand-bed streams of the Southeastern Plains. Journal of the American Water Resources Association. (accepted).

Sefick, S.A., L. Kalin, E. Kosnicki, B.P. Schneid, M.S. Jarrel, C.J. Anderson, M.H. Paller, and J.W. Feminella. Identifying the hydrogeomorphic reference condition and its relationship with benthic macroinvertebrate assemblages of Sand Hills Streams of the southeastern US. (in prep).

2. Technical reports

3. Conference or symposium proceedings

Paller, M., J. Feminella, E. Kosnicki, E., S. Sefick, M. Jarrell, D. Fletcher, T. Tuberville, S. Sterrett, and B. Prusha. 2010. Ecological Reference Models for Blackwater Streams: a Prerequisite for Successful Ecosystem Recovery and Management. Proceedings of the 2010 South Carolina Water Resources Conference, held October 13-14, 2010, at the Columbia Metropolitan Convention Center.

4. Conference or symposium abstracts

Fletcher, D.E., G.K. Stillings, M.H. Paller, and C.D. Barton. 2011. Legacy disturbances and restoration potential of coastal plain streams. Proceedings of the 2011 Georgia Water Resources Conference, April 11, 12, and 13, 2011, Athens, Georgia.

Helms, B., S. Sefick, S. Reithel, E. Kosnicki, D. Werneke, B. Schneid, J. Zink, J. Feminella, and G. Jennings. 2014. Geomorphic assessments and instream ecological endpoints: integration for restoration and management. Society for Freshwater Science, Portland Oregon.

Kosnicki, E., S. Sefick, M. Jarrell, A. Grosse, S. Sterrett, T. Tuberville, M. Paller, and J. Feminella. 2013. Development of a macroinvertebrate O/E model framework for assessing stream biological integrity in the Sand Hills of the Southeastern Coastal Plains, USA. Society for Freshwater Science. Session S10:8087

Paller, M., B. Prusha, E. Kosnicki, S. Sefick, M. Jarrell, J. Feminella, D. Flectcher, T. Tuberville, A. Grosse, and S. Sterrett. 2013. Fish assemblage structure in minimally disturbed streams on the upper southeastern coastal plain. Society for Freshwater Science. Session S10:7617

Prusha, B.A., M.H. Paller, C.S. Farrow, M.C. Freeman, D.E. Fletcher, T.D. Tuberville, A.M. Grosse, S.C. Sterrett, E. Kosnicki, and J.W. Feminella. 2011. Selecting scientifically rigorous fish IBI metrics: creating a Sand Hills stream assessment program. North American Benthological Society, Annual Meeting. Session T01E: 9756

Sefick, S.A., M.S. Jarrell, E. Kosnicki, M.H. Paller, and J.W. Feminella. 2013. Channel morphology-benthic macroinvertebrate trait relationships in southeastern coastal plains streams. Society for Freshwater Science. Session S10:8067

Sefick, S. A., M.S. Jerrell, E. Kosnicki, B.P. Schneid, M.H. Paller, and J.W. Feminella. 2012. Is the hydrogeomorphic reference condition related to benthic macroinvertebrate assemblages in coastal plains streams? Society for Freshwater Science. Session T24:7196

Kosnicki, E., B.A. Prusha, S.C. Sterrett, A.M. Grosse, S.A. Sefick, M.S. Jerrell, T.D. Tuberville, M.H. Paller, and J.W. Feminella. 2011. Preliminary fish observed to expected (O/E) models for the sand hills subecoregion of the southeastern United States. North American Benthological Society, Annual Meeting. Session T01:37

Jarrell, M.S., E. Kosnicki, B.A. Prusha, S.A. Sefick, J.A. Feminella, and M.H. Paller. 2011. Riparian-aquatic response to timber harvest in a coastal plains stream food web. North American Benthological Society, Annual Meeting. Session T18B:95

Sefick, S.A., M.S. Jarrell, E. Kosnicki, M.H. Paller, and J.W. Feminella. 2011. Preliminary hydrogeomorphic classification of stream reference condition for the Atlantic Coastal Plain and Sand Hills ecoregions. North American Benthological Society, Annual Meeting. Session T14:82

Organized Conference Session: Coastal plain streams of the southeastern US: advances in biological reference modeling and monitoring. Society for Freshwater Science. 2013. Session S10

5. Text books or book chapters