Environmental Management (2014) 54:494–504
DOI 10.1007/s00267-014-0320-0

Defining the Reference Condition for Wadeable Streams in the Sand Hills Subdivision of the Southeastern Plains Ecoregion, USA

Ely Kosnicki · Stephen A. Sefick · Michael H. Paller · Miller S. Jarrell · Blair A. Prusha · Sean C. Sterrett · Tracey D. Tuberville · Jack W. Feminella

Received: 4 September 2013 / Accepted: 17 June 2014 / Published online: 11 July 2014

Abstract The Sand Hills subdivision of the Southeastern Plains ecoregion has been impacted by historical land uses over the past two centuries and, with the additive effects of contemporary land use, determining reference condition for streams in this region is a challenge. We identified reference condition based on the combined use of 3 independent selection methods. Method 1 involved use of a multivariate disturbance gradient derived from several stressors, method 2 was based on variation in channel morphology, and method 3 was based on passing 6 of 7 environmental criteria. Sites selected as reference from all 3 methods were considered primary reference, whereas those selected by 2 or 1 methods were considered secondary or tertiary reference, respectively. Sites not selected by any of the methods were considered non-reference. In addition, best professional judgment (BPJ) was used to exclude some sites from any reference class, and comparisons were made to examine the utility of BPJ. Non-metric multidimensional scaling indicated that use of BPJ may help designate non-reference sites when unidentified stressors are present. The macroinvertebrate community measures Ephemeroptera, Plecoptera, Trichoptera richness and North Carolina Biotic Index showed no differences between primary and secondary reference sites when BPJ was ignored. However, there was no significant difference among primary, secondary, and tertiary reference sites when BPJ was used. We underscore the importance of classifying reference conditions, especially in regions that have endured significant anthropogenic activity. We suggest that the use of secondary reference sites may enable construction of models that target a broader set of management interests.

Keywords Reference condition · Sand Hills ecoregion · Stream management · Historical land use · Macroinvertebrates · Stressors

Introduction

Biological reference models are fundamental tools used in assessing biological integrity, which, in turn, can be used to measure ecosystem health, identify degraded conditions, monitor anthropogenic activities, and determine restoration effectiveness (Karr and Dudley 1981; Karr et al. 1986; Barbour et al. 1999). Model performance is critically dependent on the condition of the reference sites used to construct them. A major tenet of the reference condition concept is that sites reflect minimally disturbed conditions for the region (Reynoldson et al. 1997). However, many geographic regions have undergone transformations or are influenced by human activities such that comparable reference quality sites are at best represented by least-disturbed conditions (Stoddard et al. 2006). Unfortunately, the use of low-quality streams as references can result in misleading assessments (Kosnicki and Sites 2007; Herlihy et al. 2008).

E. Kosnicki (✉) · S. A. Sefick · M. S. Jarrell · J. W. Feminella
Department of Biological Sciences, 331 Funchess Hall, Auburn University, Auburn, AL 36849, USA
e-mail: ekosnick@brockport.edu

M. H. Paller
Savannah River National Laboratory, SRS, Aiken, SC 29808, USA

T. D. Tuberville
Savannah River Ecology Laboratory, Drawer E, Aiken, SC 29802, USA

B. A. Prusha
Midwest Biodiversity Institute, 4673 Northwest Parkway, Hilliard, OH 43235, USA

S. C. Sterrett
Warnell School of Forestry and Natural Resources, University of Georgia, Athens, GA, USA

The Southeastern Plains (SP) level III ecoregion of the United States is a low-gradient region that extends from the Chesapeake Bay, west along the Piedmont, to the Mississippi Valley Loess Plains (Omernik 1987). Once dominated in places by longleaf pine (*Pinus palustris*) forests, the SP has undergone significant anthropogenic modification, reducing forest cover to ~3 % of pre-settlement levels (Landers et al. 1995). Legacy effects from historical agriculture contribute to present-day impacts for many SP streams (Loehle et al. 2009). More recently, however, this region has undergone moderate succession, restoration, and improved management. SP streams are typically characterized as acidic, low-gradient, nutrient-poor systems with sandy substrate and high biotic diversity, including many imperiled species (Voelz and McArthur 2000).

Contemporary land use in SP is a mosaic of agriculture, private lands, urban development, public natural areas, and military training installations. These coupled with extensive land use modifications since pre-settlement times (Loehle et al. 2009) render the pristine condition (Hughes 1995; Stoddard et al. 2006; Whittier et al. 2007) virtually unattainable. For instance, some streams considered reference quality by regulatory agencies occur within well-forested watersheds; however, these same watersheds were historically converted from forest to agriculture, and likewise many streams have been impounded by low-head dams (the word "Mill" appears in names of many streams).

The Sand Hills level IV ecoregion (SH) is a subsection of the SP, adjoining the eastern portion of the Piedmont along the fall line from west central Georgia to south central North Carolina (Griffith et al. 2001). The SH spans ~20,600 km^2, consisting of extensive quartz sand deposits, formed from the late Cretaceous to the Holocene with natural vegetation represented by longleaf pine, turkey oak (*Quercus laevis*), and wire grass (*Aristida beyrichiana*), historically maintained by recurring low-intensity fires (Markewich and Markewich 1994; Schmidt 2013).

Identifying reference sites in the SH is not only a challenge because of difficulty in finding least-disturbed conditions, but also because a large part of the land is privately owned, making accessibility difficult. It is still important to seek reference sites from within the ecoregion because models developed with reference sites outside of the ecoregion may represent different communities that have limited comparability (Whittier et al. 2007). 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 have been used to identify sites as reference quality (Waite et al. 2000; Whittier et al. 2007; Herlihy et al. 2008). However, other approaches for establishing reference condition could be developed and combined with criteria-driven methods and BPJ to determine a more rigorous classification. Our objective was to combine 3 different selection methods and BPJ to create a tiered system for classifying reference quality streams and evaluate their efficacy for developing biological assessment tools for the SH ecoregion. First, we classified reference condition by evaluating streams based on 3 selection methods: (1) a generalized disturbance gradient, (2) geomorphologic disturbance, and (3) passing a series of environmental stress criteria. Second, we placed sites into reference classes based on the number of methods that selected them as reference quality. Last, after the selection process, BPJ was used to reclassify streams as non-reference for instances where sites passed at least 1 of our methodologies, but were initially chosen for sampling as disturbed sites. The resulting classes of this process were then compared with key macroinvertebrate metrics to test for significant differences of biotic integrity.

Methods

Study sites were located on lands managed by the Department of Defense (DoD, Fort Benning GA; Fort Gordon, GA; Fort Bragg, NC) and Department of Energy (DOE, Savannah River Site, SC) installations and public (Manchester State Forest, SC; Sand Hills State Forest, SC; Sandhills National Wildlife Refuge, SC; Sandhills Gamelands, NC) and non-profit (The Nature Conservancy, GA) parcels, as they represented ideal locations for finding least-disturbed habitats across the SH (Fig. 1). 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, particularly the red-cockaded woodpecker (*Picoides borealis*) (Jordan et al. 1997; USFWS 2003). We preliminarily selected 72 wadeable streams as potential reference (52) or disturbed (20), based on GIS evaluations of watershed land cover, previous investigations (e.g., Maloney et al. 2005), land manager perception, or from our onsite reconnaissance.

We sampled benthic macroinvertebrates from 64 of the 72 sites during summers of 2010–2012 (Fig. 1). Two multihabitat samples were taken with a D-frame dipnet (244 µm mesh) over a ~150 m reach in each stream. Samples represented ~1 m^2 of combined depositional, coarse woody debris, root mat, and macrophyte microhabitats from within-available pools, runs, and riffles. Samples were preserved in the field with 95 % ethanol and transported on ice to the laboratory. We used a 2-step sample processing method modified after Feminella (1996). First, we visually removed all large organisms (>2 cm). Second, we elutriated organic materials using a

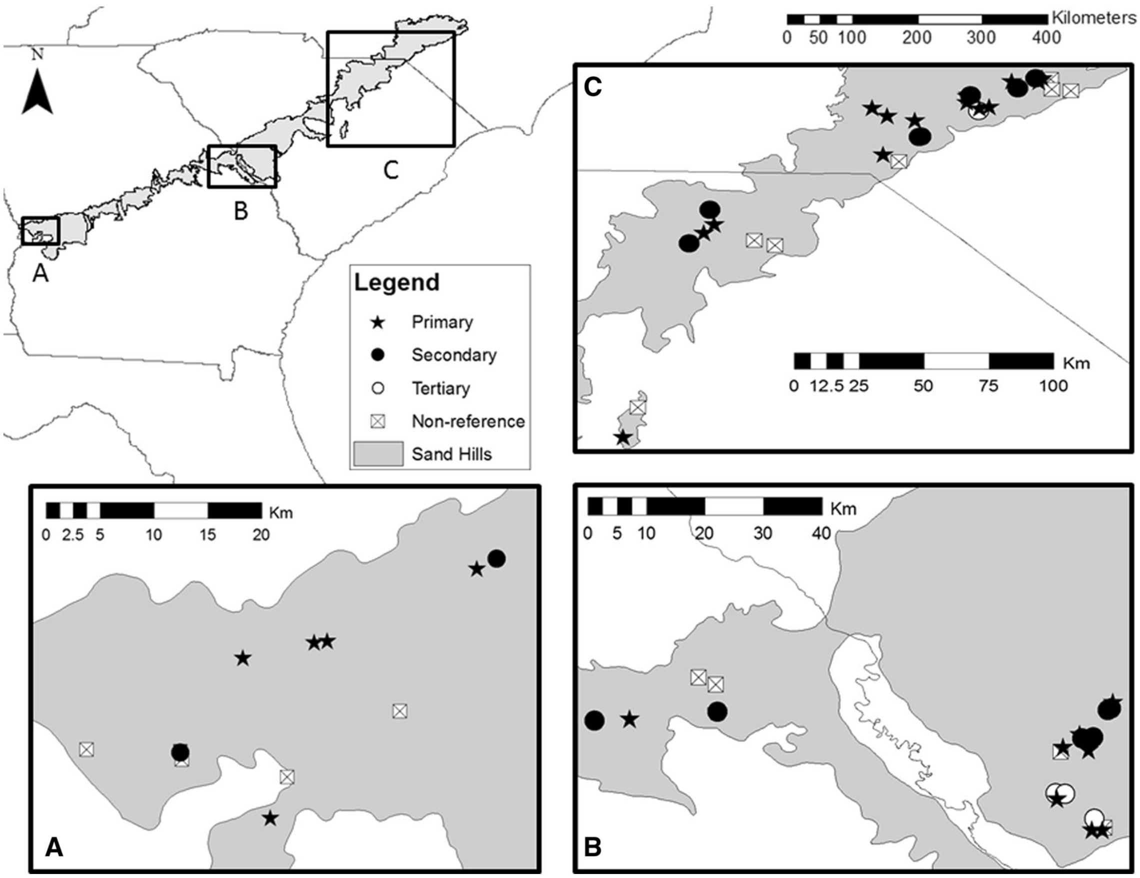

Fig. 1 The Sand Hills ecoregion and study sites in the southeast (**a**), central (**b**), and northeast (**c**) regions selected as reference or non-reference within the region

brine solution and then volumetrically subsampled by removing 5 % of the homogenized material. Subsamples were then microscopically sorted at ~7×, and all organisms were removed. Additional subsamples were processed until ≥300 organisms were counted (excluding large organisms removed earlier). Taxa were identified to lowest practical taxonomic level (usually genus or species group) except for Oligochaeta (left at Class) using available keys (Brigham et al. 1982; Kowalyk 1985; Epler 2001; Epler 2010; Merritt et al. 2008; Thorp and Covich 2010). After resolving ambiguous taxa (Cuffney et al. 2007), we adjusted counts by the number of subsamples to estimate the total individuals per sample plus larger organisms; then, both samples were combined, and densities (individuals/m^2) were estimated based on number of subsamples and sample area for each reach.

We measured streamwater-specific conductance (YSI 56 MPS, Yellow Springs, OH, USA) and pH (Orion 290A, Thermo Electron Corporation, Waltham, MA, USA) at the time of macroinvertebrate sampling. Channel morphology was surveyed from cross sections at the top of bank over 20-cm intervals with a stadia rod and leveled surveyor's tape across 4–6 runs that were at least one pool/run sequence and ≥20 m apart. We conducted surveys from the top of bank because it was a consistently discernible feature, whereas indicators of bankfull such as vegetation breaks or scour lines (e.g., Williams 1978; Johnson and Heil 1996) were less evident. We used a modified bank height ratio ($\mathrm{BHR_{mod}}$) as a measure of incision, calculated by:

$$\mathrm{BHR_{mod}} = \left| \left(\frac{D}{Y} \right) - 1 \right|$$

where D is the top of bank mean depth measured in situ and Y is the mean bankfull predicted from watershed area based on regional curves for the Coastal Plain hydrologic region (McCandless 2003). Y was predicted from watershed area, so we subtracted 1 from the ratio to account for stream beds showing obvious aggradation. We conducted habitat quality surveys onsite after methods from SCDHEC (1998) and Barbour et al. (1999). Erosion and channel modification were each scored from 0 (none) to 4 (extreme) based on visual instream examination. Last, we collected water samples from each stream between November 9 and December 8, 2012 and analyzed for total N and total P to assess potential nutrient enrichment.

We obtained land use land cover (LULC) data from the 2006 USGS National Land Cover Dataset. Watersheds were delineated from LIDAR and 10-m-resolution digital elevation data. Paved and unpaved roads (including firebreaks) were delineated from shapefiles and aerial photographs by digitizing along each roadside edge to create a polygon. We estimated road density (percent cover) and number of stream-road crossings for each study watershed. A disturbance index was generated by finding the percentage of watershed that included low, medium, and high-intensity development plus cultivated, pasture, and bare land coverage, using ArcGIS 9.2® (ESRI 2009).

We developed three methods for characterizing reference condition based on data available for each independent screening method. Method 1 involved defining a multivariate disturbance gradient based on a principal components analysis (PCA) derived from the habitat quality score, erosion score, channel modification score, bank depth, paved road area, unpaved road area, number of stream-road crossings, percentage watershed disturbance, and percentage bare ground for 70 of 72 stream reaches, representing 61 of the 64 macroinvertebrate sample sites. A cutoff point was determined by using the value of 0 as guidance and finding rates of change (inflection points) within the gradient based on the second derivative of a smooth spline function of ranked PCA axis values. PCA was conducted with PC-ORD (McCune and Mefford 1999), and the smooth spline function and second derivative were calculated with R software version 2.14.0 (Ihaka and Gentleman 1996, R Development Core Team 2011).

Method 2 was based on variation in channel morphology cross-sectional surveys of 62 of 72 stream reaches and 62 of the macroinvertebrate sample sites. Residuals were generated from iteratively re-weighted least squares regression of top of bank width, top of bank mean depth, and top of bank area, separately, with watershed area. We used iteratively re-weighted least squares regression (i.e., robust regression) because it performs well with outliers (i.e., non-reference streams in our case, Maindonald and Braun 2007). We reasoned that non-reference streams would deviate significantly (i.e., show larger residuals) from hydraulic geometry relationships of reference streams (Leopold and Maddock 1953; Dunne and Leopold 1978; Leopold 1994). We calculated the Euclidian distance of the centroid of channel morphology residuals for each site from 0 to account for variability in the interaction of channel morphology variables as follows:

$$\mathrm{Cd_o} = \sqrt{[}2](0-(A_e/3))^2 + (0-(D_e/3))^2 + (0-(W_e/3))^2$$

where A_e is top of bank area residuals, D_e is top of bank mean depth residuals, and W_e is top of bank width residuals. Cluster analysis using partitioning around medoids (*pam*) (Borcard et al. 2011) with a Euclidean distance matrix of the residuals and $\mathrm{Cd_o}$ was performed to find the average silhouette width for all possible numbers of clusters (*k*). This procedure compares within and between cluster dissimilarities for each site ("silhouette") for each *k*; the optimum *k* has the highest average silhouette width (Rousseeuw 1987), and we assumed hydraulically non-disturbed streams would cluster separately from disturbed at this optimum. Robust regression was performed using 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 2013).

Method 3 involved use of 7 environmental measures of stress as screening criteria developed for 66 of 72 stream reaches and all 64 of the macroinvertebrate sample sites, including streamwater specific conductance, streamwater total N, streamwater total P, streamwater pH, watershed road density, stream habitat quality scores, and $\mathrm{BHR_{mod}}$. Cutoff points for total N and P were from Herlihy et al. (2008). We determined remaining variable cutoffs by identifying inflection points based on the second derivative of smooth spline functions of ranked values (as in method 1) and guidance for potential breaks from published and unpublished sources. Guidance for specific conductance came from Alabama SP reference streams (L. Huff, AL Department of Environmental Management, personal communication), where the median range among level IV ecoregions is 20.4–129.7 µS/cm and the median value for the Fall Line Hills which is a western continuance of the SH is 25.8 µS/cm. Guidance for pH came from the values reported for Georgia state reference sites in the SP with a mean and standard deviation of 6.03 ± 0.94 and a maximum of 7.51 ($n = 32$) (Hughes 2005). Guidance for $\mathrm{BHR_{mod}}$ followed the US EPA stability ratings, based on measured bank height ratios where moderately to deeply incised streams have values of 0.3–0.5 (Rosgen 2007). We used a habitat quality score of 75 % as guidance for identifying a criteria threshold for streams in excellent condition (after Rankin 1995). Road density has been noted as a significant source of stream disturbance (Forman and Alexander 1998; Jones et al. 2000), influencing biota (Smith and Kaster 1983; Benton et al. 2008). However, we could not find any published or defined criteria regarding road density limits for reference condition; therefore, we visually considered the inflection points of our dataset as a means of guidance. Using these measures, we considered a stream reference if it passed 6 out of 7 environmental criteria; if a site only had 6 measured habitat criteria, it was only considered reference if it passed all 6.

The reference designation of each site was cross-referenced among the 3 selection methods for the 64 macroinvertebrate sample sites. In this way, reference groups depicting "classes" or tiers of quality were delineated.

Sites that could not be evaluated with at least 2 methods were excluded. We classified sites as *primary reference* if they passed all 3 selection methods, whereas we classified sites as *secondary reference* if they failed 1 of the screening methods, or passed 2 methods and were only evaluated with 2 methods. We classified sites as *tertiary reference* if evaluated with all 3 methods but were selected as reference by 1 method. Sites that did not pass any method were classified as non-reference sites. After reference selection, those sites showing clear signs of disturbance (e.g., presence of household waste, extreme scouring, excessive sedimentation determined from previous studies) were reclassified as non-reference by BPJ, irrespective of their outcome of each selection method. The later sites were flagged and considered for analyses as their reference classification before and after BPJ reclassification.

We used non-metric multidimensional scaling (NMS) with Bray-Curtis dissimilarities of macroinvertebrate presence absence assemblages to examine the relationships of reference and non-reference classes, with and without the use of BPJ. *T* test was used to test for significant differences of mean axis values between sites reclassified as non-reference with BPJ and remaining *primary reference* sites. A general linear model was used to conduct an unbalanced ANOVA with Tukey's HSD to identify significance across all 3 reference and non-reference classes, before and after BPJ, separately, with the macroinvertebrate community measures Ephemeroptera, Plecoptera, Trichoptera (EPT) richness and the North Carolina Biotic Index (NCBI), both representing North Carolina aquatic bioassessment tools (NCDENR 2006), and the Georgia SH Macroinvertebrate Multi-Metric Index (GAMMI) (GDNR 2007). We chose these measures as response variables because they represent existing state biological assessments in the SH. All analyses were performed with SAS® version 9.2 (SAS Institute Inc. 2004), except where noted.

Table 1 Variables and their source for principal component analysis (PCA) used for reference selection method 1

Variable	Data source	PC-1
Habitat quality score	EPA Habitat quality	-0.786
Erosion score	On-site evaluation	0.779
Channel modification score	On-site evaluation	0.701
Bank depth (m)	On-site evaluation	0.694
Paved road area (m^2)	GIS	0.865
Unpaved road area (m^2)	GIS	-0.080
Number of road crossings	GIS	-0.055
Watershed disturbance (%)	GIS	0.881
Percentage bare ground (%)	GIS	-0.189

Standardized eigen vectors (loadings) of PCA axis 1 (PC-1) are also given. PC-1 explained 41.86 % of the variance

GIS geographical information systems

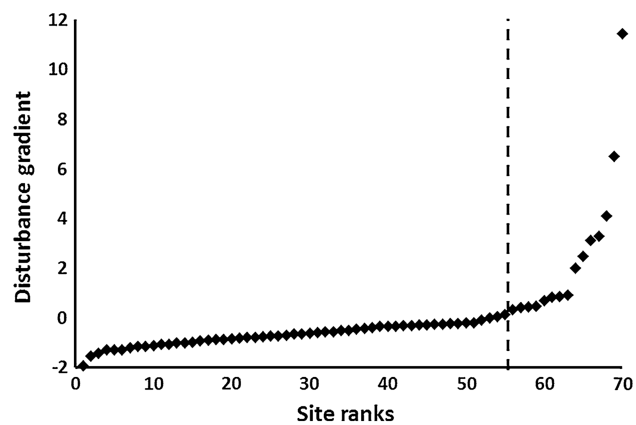

Fig. 2 Plot of stream sites along a disturbance gradient interpreted from the rank of PCA axis-1 scores derived from the variables in Table 1. The *vertical dotted line* indicates the cutoff between candidate reference (*left side*) and non-reference (*right side*) determined from a 2nd derivative of a smooth spline function. Seventy sites were used to establish reference criteria, but only 64 sites were evaluated with macroinvertebrate community measures

Results

The PCA used in Method 1 showed that 6 of the 9 stress-related variables loaded strongly on PCA axis-1 (PC-1) which we interpreted as a disturbance gradient (Table 1). PC-1 accounted for 41.9 % of the variation among sites and was the only significant PCA axis (randomization test, $P = 0.001$). Loadings from PC-1 indicated that some variables increased with disturbance while others decreased; therefore, we considered PC-1 value of 0 as a target to determine the inflection point that best represented the threshold between reference and non-reference sites. Ranking of sites along PC-1 identified 55 sites as reference and 15 as non-reference based on an inflection point derived from second derivatives and guidance from the PC-1 loadings (Fig. 2). Robust regression of Method 2 indicated that all 3 channel morphology variables were significantly related to watershed area; however, top of bank area had the highest *F* value (Table 2). The *k* with the highest average silhouette width was 2 which we interpreted as a demarcation between disturbed and non-disturbed channels; *pam* clustered 51 and 11 streams into reference and non-reference classes, respectively. Method 3 identified 47 sites as reference and 19 as non-reference based on criteria developed from 7 environmental stressors (Fig. 3; Table 3). Of the 64 sites sampled for macroinvertebrates, 55 were evaluated with all 3 methods and 9 were evaluated with a combination of 2 methods. Cross-comparison of each stream reference designation resulted in 34 *primary*, 17 *secondary*, and 8 *tertiary reference* sites.

Table 2 Equation coefficients from robust regressions generating residuals used in cluster analysis, separating reference from non-reference sites on the basis of channel morphology

Variable	Coefficient	y-intercept	F	P
Top of bank width (m)	2.52	9.5e−08	43.14	<0.0001
Top of bank mean depth (m)	0.35	1.2e−06	25.21	<0.0001
Top of bank area (m^2)	0.74	8.4e−08	75.69	<0.0001

Each variable was modeled with watershed area (m^2)

Table 3 Variables and criteria used to delineate reference from non-reference. Sites had to pass at least 6 criteria to be considered reference

Variable	Criteria
Streamwater-specific conductance (µS/cm)	<36
Watershed road density (%)	<6.3
Streamwater pH	<6.34
Habitat quality scores	>153
BHR_{mod}	<0.46
Streamwater total N (µg/L)	<1,000[a]
Streamwater total P (µg/L)	<30[a]

[a] Based on Herlihy et al. 2008

Five sites were considered as non-reference because of failure to pass any reference condition selection methods. Eleven sites were reclassified as non-reference because of BPJ even if they were selected as reference by any of the 3 methods. Reclassification refined the reference set to 29 *primary*, 15 *secondary*, 4 *tertiary*, and 16 non-reference sites. *Primary reference* sites were distributed throughout the study area; however, only 5 of these occurred within southwestern portion of the ecoregion, indicating that there is a disparity of our highest quality classifications for some areas of the SH (Fig. 1).

A total of 268 macroinvertebrate taxa were identified for the 64 sampled reaches. NMS fully converged for 3 dimensions with stress = 0.13, indicating that the data fit the model well. Unbalanced t-test results comparing NMS dimensions of sites removed by BPJ with *primary reference* sites after BPJ resulted in significant differences for dimensions 1 and 3 (Table 4). Plots of these axes showed a general progression from disturbed to less disturbed from upper left to lower right, with same sites designated as non-reference with BPJ occurring to the upper left (Fig. 4).

EPT richness differed significantly between reference and non-reference classes both with ($df = 3$, $F = 9.07$, $P < 0.0001$) and without ($df = 3$, $F = 9.06$, $P < 0.0001$) BPJ. Sites designated as *primary*, *secondary*, and *tertiary reference* were statistically indistinguishable from each other when BPJ was used; however, when BPJ was not considered, *tertiary references* were significantly different from *primary* and *secondary,* indicating that a number of non-reference sites were able to pass at least 1 of the selection methods (Fig. 5a). NCBI significantly differed between reference and non-reference groups both with ($df = 3$, $F = 9.98$, $P < 0.0001$) and without ($df = 3$, $F = 6.04$, $P = 0.0012$) BPJ (Fig. 5b); sites designated as

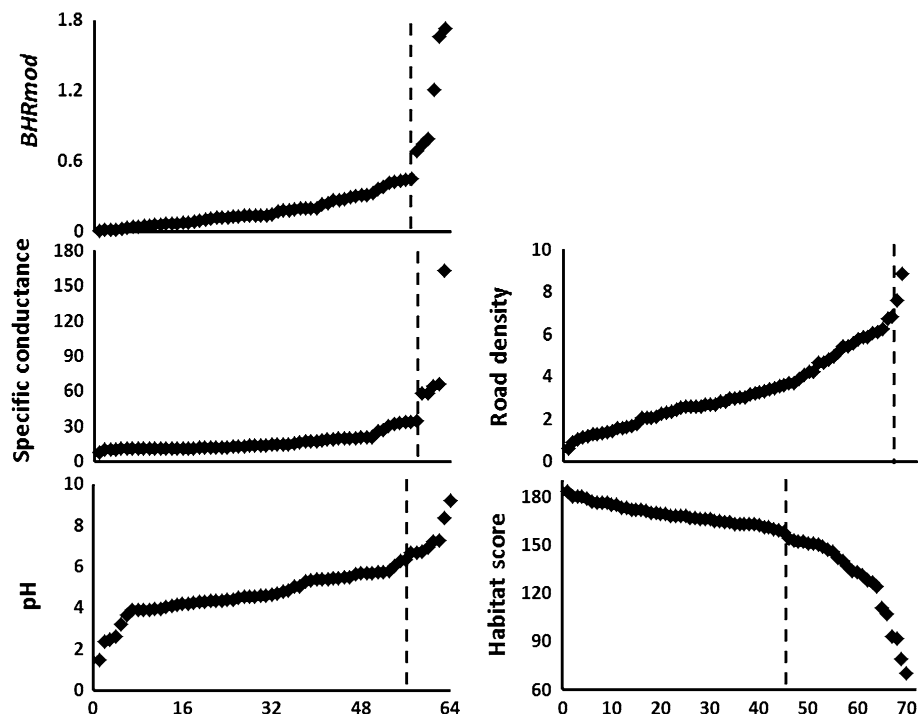

Fig. 3 Reference criteria derived by method 3. The *vertical dotted lines* indicate the cutoff between candidate reference (*left side*) and non-reference (*right side*) determined from 2nd derivatives of smooth spline functions. Road density and habitat scores were plotted in association with 70 sites to establish reference criteria, but only 64 sites were evaluated with macroinvertebrate community measures

Table 4 Unbalanced t-test results for non-metric multidimensional scaling of sites reclassified as non-reference by best professional judgment (BPJ, Non-ref-BPJ) and *primary reference* sites (*Primary*) classified after the method selection process and BPJ

n	Dimension 1		Dimension 2		Dimension 3	
	Mean	SD	Mean	SD	Mean	SD
Non-ref BPJ						
11	−0.952	1.170	0.348	1.048	0.823	1.115
Primary						
29	0.356	0.832	0.116	0.768	−0.347	0.703

Test statistics	Dimension 1	Dimension 2	Dimension 3
t	−3.96	0.77	3.97
P	<0.001	0.514	<0.001

SD Standard deviation

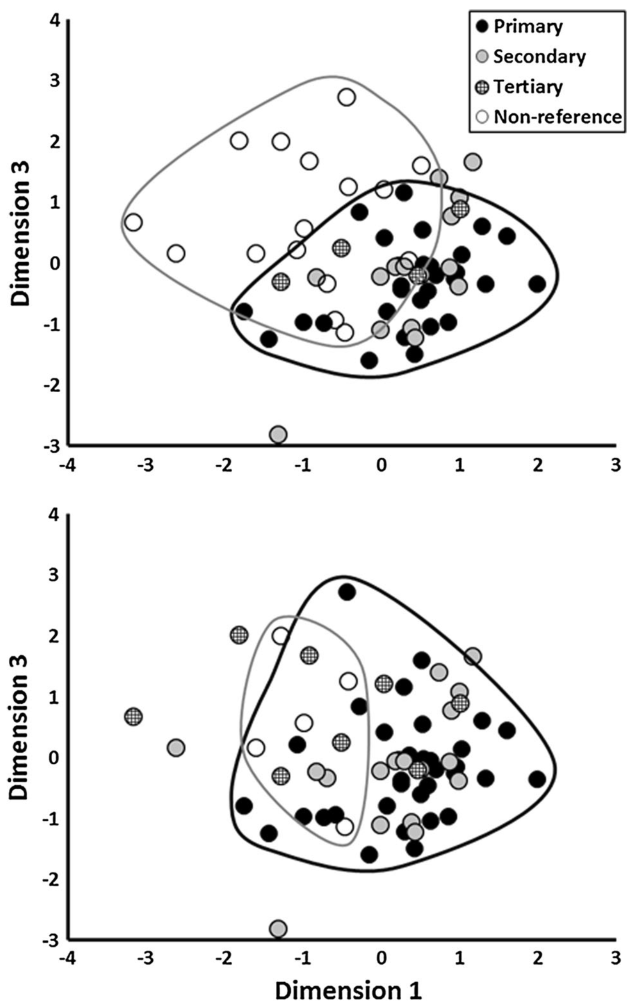

Fig. 4 Non-metric multidimensional scaling ordination of axis 1 and 3 for streams determined as *primary*, *secondary*, and *tertiary reference*, with non-reference streams determined from the reference selection approach with best professional judgment (*top*) and without best professional judgment (*bottom*). *Primary reference* sites are *corralled in black* and non-reference sites are *corralled in gray*

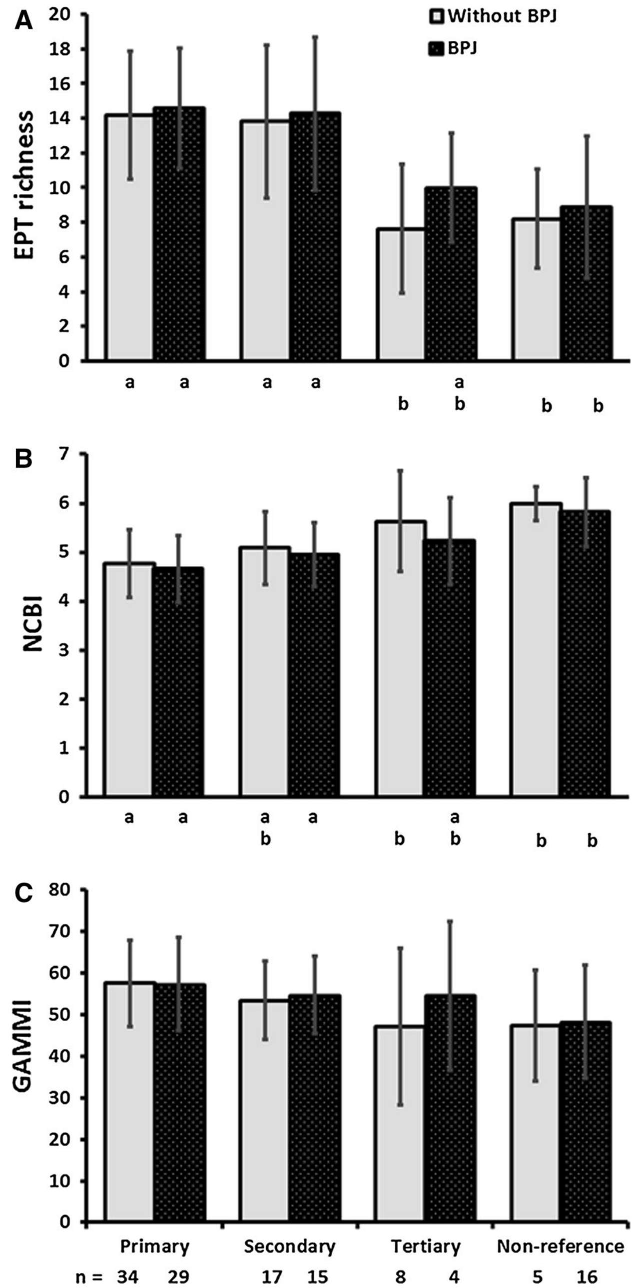

Fig. 5 Means and standard deviations of sites determined as *primary*, *secondary*, or *tertiary reference*, and non-reference groups for EPT richness (**A**), North Carolina Biotic Index (NCBI, **B**), and the Georgia Macroinvertebrate Multimetric Index (GAMMI, **C**). Unbalanced ANOVA was performed on reference sets with and without the use of best professional judgment (BPJ), separately. *Bars* with the *same letters* indicate means are not significantly different

primary, *secondary*, and *tertiary reference* were not significantly different from each other when BPJ was used; however, without BPJ, *secondary*, *tertiary*, and non-reference sites were indistinguishable. GAMMI was highly variable and did not differ between reference and non-reference sites either with ($df = 3$, $F = 1.99$, $P = 0.1248$) or without ($df = 3$, $F = 2.48$, $P = 0.0698$) BPJ (Fig. 5c).

Discussion

The Sand Hills ecoregion of the Southeastern Plains has undergone many land use changes in recent history, and virtually all streams within this region have been influenced by human activities since pre-settlement to present day. Many areas are currently managed for differing land uses and stakeholder interests; therefore, natural areas in the SH reflect variation in human legacies and contemporary management practices, which are likely to vary among managers depending on their goals. Identifying pristine streams within this landscape is impractical if not impossible, and thus using least-disturbed sites as reference condition is warranted, even though they are likely to differ from each other given the combination of historical land use and contemporary management.

Results from the NMS ordinations indicated that our use of best professional judgment eliminated some sites that were biologically "on the fringe" of what we would expect from least-disturbed conditions. We note that remaining *primary reference* sites appearing "on the fringe" may represent natural variation or may have been impaired by unaccounted disturbance that eluded our BPJ. The macroinvertebrate community metrics we used indicated that the *primary reference* streams delineated by our 3-method reference selection process likely represent least-disturbed conditions of the region, with or without the use of BPJ to reclassify them. However, we included the use of BPJ in the selection process because it may help eliminate sites that are impacted by unaccounted sources of disturbance as indicated by the improved metric values for *secondary* (Fig. 5b) and *tertiary reference* sites (Fig. 5a). We concede that some sites with healthy biotic communities may be excluded by BPJ, but this potential source of error is more conservative and more desirable than including sites that are disturbed by unknown impacts or classifying sites as reference by BPJ alone. Our results suggest that sites classified as *tertiary reference* should probably not be used as references for assessment or management purposes. The biotic communities of these sites were not significantly different from non-reference sites and therefore will not indicate streams of low quality unless impaired conditions are extreme.

Developing final assessments based on conditions of varying quality is a serious problem and there is a strong need to estimate differences in reference quality (Hawkins et al. 2010). We believe that our tiered approach in combining independent methods provides a means of weighting reference condition. Our classification of *secondary reference* sites may represent varying quality, and the use of these sites for establishing reference criteria should be clearly indicated. Although these sites were indistinguishable from *primary reference* sites and were different from non-reference sites for EPT richness and the NCBI with the use of BPJ, they did not pass all 3 selection methods. This result would indicate these sites experienced a level of disturbance, though possibly negligible in terms of least-disturbed conditions. However, there may be utility in using *secondary reference* sites in combination with *primary reference* sites if a particular study region has limited sites of *primary reference* condition, as in the southeastern portion of the SH. Furthermore, using *secondary reference* sites for assessment purposes in heavily impacted regions may help guide managers toward a best attainable condition (see below).

Ecological assessments often are based on reference conditions indicative of the natural conditions in a region (Hughes 1995; Davies and Jackson 2006); however, since there is no general consensus with regard to defining reference condition (Hawkins et al. 2010), it should be, at least, region-specific (Stoddard et al. 2006). BPJ has long been used by management programs. Our aim was not to censure the usefulness of BPJ, but rather to apply it more conservatively in excluding streams from being considered reference quality. Our reasoning is not to discount our methods in being rigorous enough to measure all watershed disturbances, but instead to assume that there are some disturbance sources or their manifestations that may not be accountable. Our results confirm that there are differences between non-reference sites and sites of the highest reference quality and also confirm a conservative use of BPJ in helping make better decisions about reference site quality.

Defining reference conditions as a basis for differentiating "levels of quality" allows for the use of contrasting benchmarks to be established that represent a holistic and integrated characterization of least-disturbed conditions. We acknowledge that using *secondary reference* conditions to construct reference models may lower the assessment benchmark, and therefore, models that include these conditions should be indicated as such. However, adopting different sets of reference conditions encourages the development of different models, which, when combined, may provide a more useful benchmark (Van Sickle et al. 2006). Furthermore, reference sites of different overall quality are useful with regard to developing models based on specific stressors (Bailey et al. 2004).

Our use of a tiered reference selection (i.e., *primary*, *secondary*, *tertiary*) allows for the construction of different sets of models and can be applied to target a broad range of management questions. For instance, a strict model that evaluates sites with a high standard of quality in the SH would be based on models that only include our *primary* *reference* sites. However, a more flexible model could also include our *secondary reference* sites as a means of accounting for more regional variation and may be more relevant for evaluating specific management goals (e.g., restoration success). Some areas of the SH, such as DoD installations, are managed for multiple land use activities, including military training, timber or wildlife management, and conservation. It is well documented that training in areas of military property can create considerable impacts to terrestrial and aquatic ecosystems (Quist et al. 2003; Maloney et al. 2005); however, these installations also represent some of the largest refuges for imperiled species (Cohn 1996; Kaufman 2010). Land managers on DoD installations must consider the dual and sometimes conflicting missions of military readiness and adherence to legal mandates such as the Endangered Species Act. Therefore, managers may need a broad set of tools for evaluating and assessing habitat conditions with a reasonably achievable set of standards, particularly those relevant to their region. We recommend the use of *primary* plus *secondary reference* sites to develop a Best Attainable Model (BAM) for evaluating streams that are in the process of slow recovery, such as those subjected to intense military training and/or legacy impacts from historical agriculture (Maloney et al. 2008). In developing assessment tools, we argue that it is better to identify reference streams of different tiers rather than simply including them as "reference" condition. A BAM approach will provide managers with the capacity to measure improvements, without erroneously determining that conditions have reached the most desirable quality.

We have retained the term "least-disturbed" to define the quality of our *primary reference* sites (Stoddard et al. 2006). It is noted that almost all of the SH surrounding our study area have undergone some form of recent anthropogenic alteration, though present-day conditions at some of these areas are under strict conservation management for imperiled species, and these areas may be in a state of minimal disturbance (Stoddard et al. 2006). Many of our *primary reference* sites occur within military installations that were previously agricultural lands, but are currently managed with controlled burning in addition to being subjected to training exercises and silviculture. Such sites are influenced by an amalgamation of past and present human activities. However, based on our extensive work in the SH, we feel that these lands are least exposed to stressors, managed to maintain naturalness, and thus represent the best reference conditions for this ecoregion.

We should note that our results from the state biological assessment tools may not reflect direct assessments of biological integrity (Karr 1991) because our sampling and processing methodologies differ from protocols used to design them. In this context, our goal was not to evaluate the performance or sensitivity of regional assessment tools, but rather test the efficacy of our reference classifications. We also note that state assessment tools were constructed with the restriction of using streams found within the state of origin alone. Frameworks that use streams across state lines are expected to be more robust and may be useful for identifying regional similarities and differences.

Conclusions

We recommend that reference sites selected using a combination of methods provide an effective means to designate least-disturbed sites within a region, especially one such as the Sand Hills ecoregion that has undergone multiple land use changes. The conservative use of BPJ has utility to eliminate potential non-reference sites from reference classification. Identifying streams as *secondary reference* allows for construction and testing of models, such as BAMs, which may greatly facilitate effective water and land management. We also suggest that combining models developed from different tiers of reference condition may be more robust in integrating regional variation important in accounting for stream community structure and function.

Acknowledgments We thank Hugh Westbury (Fort Benning, Natural Resources Branch), Chuck Bryan (Fort Bragg, Endangered Species Branch), Michael Juhan (Fort Gordon, Natural Resources Branch), Harvey Belser (Manchester State Forest), Michele Brossett (Georgia Department of Natural Resources), Michael Walters (North Carolina Biological Assessment Unit), George Williams (Osage of Virginia, Inc.), Nancy Jordan (Sandhills National Wildlife Refuge), Brian Davis (Sand Hills State Forest), Michele Elmore and Geoff Sorrell (The Nature Conservancy), and special thanks to Brady Beck (Sandhills Gamelands). We also thank Brian Helms at the Auburn Museum of Natural History for field assistance and Andrew Grosse and Bess Harris at Savannah River Ecology Lab for help with GIS analysis. This project was funded by the Strategic Research and Development Program Project RC-1694. Support was provided in part by Award Number DE-FC09-07SR22506 from Department of Energy to the University of Georgia Research Foundation.

References

Bailey RC, Norris RH, Reynoldson TB (2004) Bioassessment of freshwater ecosystems: using the reference condition approach. Kluwer, New York

Barbour MT, Gerritsen J, Snyder BD, Stribling JB (1999) Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates and fish. 2nd edn. EPA 841-B-99-002. U.S. Environmental Protection Agency; Office of Water; Washington

Benton PD, Ensign WE, Freeman BJ (2008) The effect of road crossings on fish movements in small Etowah Basin streams. Southeast Nat 7:301–310

Borcard D, Gillett F, Legendre P (2011) Numerical ecology with R. Springer, New York

Brigham AR, Brigham WU, Gnilka A (eds) (1982) Aquatic insects and oligochaetes of North and South Carolina. Midwest Aquatic Enterprises, Illinois

Cohn JP (1996) New defenders of wildlife. Bioscience 46:11–14

Cuffney TF, Bilger MD, Haigler AM (2007) Ambiguous taxa: effects on the characterization and interpretation of invertebrate assemblages. J N Am Benthol Soc 26:286–307

Davies SP, Jackson SK (2006) The biological condition gradient: a descriptive model for interpreting change in aquatic ecosystems. Ecol Appl 16:1251–1266

Dunne T, Leopold LB (1978) Water in environmental planning. Freeman, San Francisco

Epler JH (2001) Identification manual for the larval Chironomidae (Diptera) of North and South Carolina: a guide to the taxonomy of the midges of the southeastern United States, including Florida. St. Johns River Water Management District

Epler JH (2010) Identification manual for the water beetles of Florida:(Coleoptera: Dryopidae, Dytiscidae, Elmidae, Gyrinidae, Haliplidae, Hydraenidae, Hydrophilidae, Noteridae, Psephenidae, Ptilodactylidae, Scirtidae). State of Florida, Department of Environmental Protection, Division of Water Facilities

ESRI (Environmental Systems Resource Institute) (2009) ArcGIS® 9.2. ESRI, Redlands, California

Feminella JW (1996) Comparison of benthic macroinvertebrate communities in small streams of contrasting hydrologic permanence. J N Am Benthol Soc 15:651–669

Forman RTT, Alexander LE (1998) Roads and their major ecological effects. Annu Rev Ecol Syst 29:207–231

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

Griffith GE, Omernik JM, Comstock JA, Lawrence S, Martin G, Goddard A, Hulcher VJ, Foster T (2001) Ecoregions of Alabama and Georgia. Two-sided color poster with map, descriptive text, summary tables, and photographs). U.S. Geological Survey, Reston, VA, scale 1:1,700,000

Hawkins CP, Olson JR, Hill RA (2010) The reference condition: predicting benchmarks for ecological and water-quality assessments. J N Am Benthol Soc 29:312–343

Herlihy AT, Paulsen SG, Van Sickle J, Stoddard JL, Hawkins CP, Yuan LL (2008) Striving for consistency in a national assessment: the challenges of applying a reference-condition approach at a continental scale. J N Am Benthol Soc 27:860–877

Hughes RM (1995) Defining acceptable biological status by comparing with reference conditions. In: Davis WS, Simon TP (eds) Biological assessment and criteria: tools for water resource planning and decision making for rivers and streams. CRC Press, Boca Raton, pp 31–47

Hughes DL (2005) Development of biological reference conditions of wadeable streams in the major ecoregions and subecoregions of Georgia. Thesis, Columbus State University

Ihaka R, Gentleman R (1996) R: a language for data analysis and graphics. J Comput Graph Stat 5:299–314

Johnson PA, Heil TM (1996) Uncertainty in estimating bankfull conditions. J Am Water Resour Assoc 32:1283–1291

Jones JA, Swanson FJ, Wemple BC, Synder KU (2000) Effect of roads on hydrology, geomorphology, and disturbance patches in stream networks. Conserv Biol 14:76–85

Jordan RA, Wheaton KS, Weiher WM, Hayden TJ (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

Karr JR (1991) Biological integrity: a long-neglected aspect of water resource management. Ecol Appl 1:66–84

Karr JR, Dudley DR (1981) Ecological perspective on water quality goals. Environ Manag 5:55–68

Karr JR, Fausch KD, Angermeier PL, Yant PR, Schlosser IJ (1986) Assessing biological integrity in running waters: a method and its rationale. Illinois Natural History Survey, Champaign, Illinois, Special Publication 5

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

Kosnicki E, Sites RW (2007) Least-Desired Index for assessing the effectiveness of grass riparian filter strips in improving water quality in an agricultural region. Environ Entomol 36:713–724

Kowalyk HE (1985) The larval cephalic setae in the Tanypodinae (Diptera: Chironomidae) and their importance in generic determination. Can Entomol 117:67–106

Landers JL, Van Lear DH, Boyer WD (1995) The longleaf pine forests of the Southeast: requiem or renaissance? J For 93:39–44

Leopold LB (1994) A view of the river. Harvard Univ Press, Cambridge

Leopold LB, Maddock T (1953) The hydraulic geometry of stream channels and some physiographic implications; USGS Professional Paper 252. United States Government Printing Office, Washington

Loehle C, Wigley TB, Schilling E, Tatum V, Beebe J, Vance E, Van Deusen P, Weatherford P (2009) Achieving conservation goals in managed forests of the Southeastern Coastal Plain. Environ Manag 44:1136–1148

Maindonald JH, Braun J (2007) Data analysis and graphics using R: an example-based approach. Cambridge University Press, Cambridge

Maloney KO, Mulholland PJ, Feminella JW (2005) Influence of catchment-scale military land use on stream physical and organic matter variables in small Southeastern Plains catchments (USA). Environ Manag 35:677–691

Maloney KO, Feminella JW, Mitchell RM, Miller SA, Mulholland PJ, Houser JN (2008) Landuse legacies and small streams: identifying relationships between historical land use and contemporary stream conditions. J N Am Benthol Soc 27: 280–294

Markewich HW, Markewich W (1994) An overview of pleistocene and Holocene Inland dunes in Georgia and the Carolinas: morphology, distribution, age, and paleoclimate, U.S. Geological Survey Bulletin no. 2069

McCandless TL (2003) Maryland stream survey: bankfull discharge and channel characteristics of streams in the coastal plain hydrologic region. U.S. Fish and Wildlife Service, Chesapeake Bay Field Office. CBFO-S03-02

McCune B, Mefford MJ (1999) PC-ORD. Multivariate analysis of ecological data, version 4. MjM Software Design, Gleneden Beach, OR

Merritt RW, Cummins KW, Berg MB (eds) (2008) An introduction to the aquatic insects of North America, 4th edn. Kendall-Hunt, Iowa

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

Omernik JM (1987) Ecoregions of the conterminous United States. Map (scale 1:7,500,000). Ann Assoc Am Geogr 77:118–125

Quist MC, Fay PA, Guy CS, Knapp AK, Rubenstein BN (2003) Military training effects on terrestrial and aquatic communities on a grassland military installation. Ecol Appl 13:432–442

Rankin ET (1995) Chapter 13: Habitat indices in water resource quality assessments. In: Davis WS, Simon TP (eds) Biological assessment and criteria: Tools for water resource planning and decision making for rivers and streams. CRC Press, Boca Raton, pp 181–208

Reynoldson TB, Norris RH, Resh VH, Day KE, Rosenberg DM (1997) The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. J N Am Benthol Soc 16:833–852

Rosgen D (2007) Watershed Assessment of River Stability and Sediment Supply (WARSSS). Wildland Hydrology. Fort Collins, CO. http://www.epa.gov/warsss/

Rousseeuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 20:53–65

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

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

Schmidt JP (2013) "Sandhills." New Georgia Encyclopedia. 19 August 2013. http://www.georgiaencyclopedia.org/articles/geography-environment/sandhills

Smith ME, Kaster JL (1983) Effect of rural highway runoff on stream benthic macroinvertebrates. Environ Pollut (Series A) 32:157–170

Stoddard JL, Peck DV, Olsen AR, Larsen DP, Van Sickle J, Hawkins CP, Hughes RM, Whittier TR, Lomnicky G, Herlihy AT, Kaufmann PR, Peterson SA, Ringold PL, Paulsen SG, Blair R (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 JL, Larsen DP, Hawkins CP, Johnson RK, Norris RH (2006) Setting expectations for the ecological condition of streams: the concept of reference condition. Ecol Appl 16:1267–1276

R Development Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Version 3.0.1. http://www.R-project.org/

Thorp JH, Covich AP (eds) (2010) Ecology and classification of North American freshwater invertebrates. 3rd edn. Academic Press, London

USFWS (2003) US Fish and Wildlife Service (2003) Red-Cockaded Woodpecker (*Picoides borealis*) Recovery, 2nd edn. Atlanta, GA

Van Sickle J, Huff DD, Hawkins CP (2006) Selecting discriminant function models for predicting the expected richness of aquatic macroinvertebrates. Freshw Biol 51:359–372

Voelz NJ, McArthur JV (2000) An exploration of factors influencing lotic insect species richness. Biodivers Conserv 9:1543–1570

Waite IR, Herlihy AT, Larsen DP, Klemm DJ (2000) Comparing strengths of geographic and nongeographic classifications of stream benthic macroinvertebrates in the Mid-Atlantic Highlands, USA. J N Am Benthol Soc 19:429–441

Whittier TR, Stoddard JL, Larsen DP, Herlihy AT (2007) Selecting reference sites for stream biological assessments: best professional judgment or objective criteria. J N Am Benthol Soc 26:349–360

Williams GP (1978) Bank-full discharge of rivers. Water Resour Res 14:1141–1154