WETLANDS, Vol. 19, No. 3, September 1999, pp. 560–569

HYDROGEOMORPHIC (HGM) ASSESSMENT—A TEST OF USER CONSISTENCY

Dennis F. Whigham[1], Lyndon C. Lee[2], Mark M. Brinson[3], Richard D. Rheinhardt[3], Mark C. Rains,[2,6]
Jeffrey A. Mason[2,7], Humaira Kahn[1,8], Melanie B. Ruhlman[5], and Wade L. Nutter[4,5]
[1] *Smithsonian Environmental Research Center*
Box 28
Edgewater, Maryland, USA 21027

[2] *National Wetland Science Training Cooperative*
221 First Avenue West, Suite 415
Seattle, Washington, USA 98199

[3] *Biology Department*
East Carolina University
Greenville, North Carolina, USA 27858

[4] *School of Forest Resources*
University of Georgia
Athens, Georgia, USA 30602

[5] *Nutter, Overcash & Associates*
1073 S. Milleage Ave.
Athens, Georgia, USA 30605

[6] *Present Address:*
Hydrologic Sciences
212 Veihmeyer Hall
University of California
Davis, California, USA 95616

[7] *Present address:*
University of California at Santa Barbara
Department of Geography
Santa Barbara, California, USA 93106-4060

[8] *Present address:*
University of Maryland
Department of Geography
College Park, Maryland, USA 20742

Abstract: We describe the first test conducted to determine user consistency in the application of hydrogeomorphic (HGM) functional assessment models. Over a three-week period, two teams of individuals trained in the HGM methodology assessed 44 riverine wetlands on the Coastal Plain of Delaware, Maryland, and Virginia, USA. Results demonstrated a high degree of agreement between the two assessment teams for both Variable Subindices and Functional Capacity Index Scores, indicating that the assessment models were robust and results were repeatable. Analyses of the data demonstrated the importance of only using variables whose measurements are repeatable. When variable measurements are not repeatable, HGM functional capacity scores are detrimentally affected, especially functions that are modeled by only a few variables.

Key Words: hydrogeomorphic approach, HGM, riverine wetlands, reference wetlands, test of user consistency, Maryland, Virginia, Delaware

INTRODUCTION

A theoretical context for the Hydrogeomorphic (HGM) approach to functional assessment of wetlands has been developed (Brinson et al. 1994, 1998, Smith et al. 1995), and examples have been provided for its application (Brinson and Rheinhardt 1996, Rheinhardt et al. 1997). Functional assessment under the HGM approach differs from alternative assessment methods in that (1) wetlands are classified using hydrogeomorphic properties, (2) data from reference wetlands are used in model development, and (3) relative, rather than absolute, indices are used to increase efficiency and consistency of the assessment process. A national guidebook for assessing riverine wetlands for the United States (Brinson et al. 1995) has been used to develop Regional Guidebooks for the riverine HGM subclass (e.g., Ainslie et al. 1999). Regional Guidebooks are under development for other HGM wetland subclasses such as flats in the Atlantic and Gulf Coastal Plains (R. R. Rheinhardt and M. M Brinson, pers.com.). Zentner (1999) reported on an effort to develop a regional guidebook for California's Central Valley streams. The National Wetland Science Training Cooperative (NWSTC), in cooperation with various federal and state agencies, has been developing HGM Guidebooks for riverine and slope wetlands in southeast and interior Alaska, depressional wetlands in the Prairie Pothole region, riverine wetlands in central and southern California, and riverine and depressional wetlands in the Mid-Atlantic region (L. C. Lee, pers. com.). Ongoing development of guidebooks and testing of HGM models has been funded by public (U.S. Environmental Protection Agency, U.S. Army Corps of Engineers, and the National Resource Conservation Service) and private organizations. NWSTC has used the HGM approach as part of the 404 permit process for three approved projects in Washington, one in New York, and four in California. A protocol for the development of regionally specific guidebooks has been proposed, and iterative testing of the regional models is an important component of the process (Federal Register 1996, 1997, Brinson et al. 1999, Wakeley and Smith, In press).

To be effective and widely accepted, the HGM approach should be repeatable to facilitate consistent application of federal, state, and local regulations. Thus, part of the development process should include testing models for consistency of results. In this paper, we present results from a test of user consistency for variables and models for a Regional Guidebook being developed for riverine wetlands in the Mid-Atlantic region. At the time this study was conducted, Regional Guidebook development and testing proceeded in six related steps: (1) establishment of the Reference Domain, (2) selection and sampling of the Reference Wetland System, (3) selection of Reference Standard wetlands, (4) identification of models (e.g., functions) and variables and scaling of variables using the Reference Wetland System, (5) application of the assessment models at Reference Wetlands by two independent teams, and (6) analysis and interpretation of results leading to revision of HGM variables and models. The interdisciplinary team of scientists that conducted this study had expertise in wetland hydrology and biogeochemistry, soil science, plant community ecology, and landscape ecology.

METHODS

Establishment of the Reference Domain

The Reference Domain is defined as the geographic region from which a wetland representing the regional subclass is selected (Smith et al. 1995). The regional wetland subclass used in this study included 1st–3rd order streams (*sensu* Strahler 1957) that were identifiable on USGS 1:24,000 maps. The sites were located within the Coastal Plain of Maryland, Virginia, and Delaware (Figure 1). In Maryland and Virginia, all sampled wetlands were within the Inner Coastal Plain on the Western Shore of the Chesapeake Bay. Wetland sites in Delaware were located on the Coastal Plain.

Within the Reference Domain, forested wetlands associated with 1st–3rd order streams, similar to those selected in this study, have been described by several authors. All of the sites that we studied would be classified as either Seasonally or Temporarily Flooded Swamp Forests in the context of the National Wetlands Inventory (Tiner and Burke 1995). Wetlands similar to those that were selected have also been referred to as non-tidal palustrine forested wetlands (Hupp et al. 1993, Puckett et al. 1993, Walbridge and Struthers 1993), Coastal Plain swamps (Parsons and Ware 1982), forests on small stream bottoms (Glascock and Ware 1979), and inland forested wetlands (Tiner 1987). All of the natural wetlands were on floodplains associated with small streams in narrow, valley-bottom positions that experience seasonal flooding. Brush et al. (1980) characterized forested areas of bottomlands on the Western Shore Coastal Plain in Maryland as the River Birch-Sycamore Association. Tree and shrub species composition varied from site to site, but there was extensive overlap between sites and the dominant species were *Acer rubrum* L., *Fraxinus pennsylvanica* Marshall, *Liquidambar styraciflua* L., *Liriodendron tulipifera* L., and several oak (*Quercus*) species in the tree stratum and *Alnus serrulata* Aiton (Willd.), *Carpinus caroliniana* Walter., *Ulmus* spp., *Betula nigra* L., and *Viburnum* spp. in the shrub stratum (Rheinhardt et al., In press).

Selection and Sampling of Reference Wetland System

Within the Reference Domain, we selected 44 Reference Wetlands (Figure 1) that represented a range of altered and relatively unaltered conditions within the subclass and, presumably, encompassed the range of ecological functioning shown by the subclass (*sensu* Brinson and Rheinhardt 1996). The range of variation included (1) sites that had no sign of recent perturbation (e.g., beaver activity, logging, hydrologic alterations), (2) sites that had been perturbed by natural processes and were recovering (e.g., abandoned beaver impoundments), (3) sites that had been perturbed by anthropogenic activities (e.g., logging) in the past but were recovering, (4) sites that were subject to continuous and ongoing perturbation by anthropogenic activities (e.g., under-fit culverts and subsequent impoundment of water), and (5) sites that were recent attempts to restore or create riverine wetlands.

Each Reference Wetland was sampled during the spring-summer of 1995 for hydrologic (e.g., channel morphology and evidence of flooding depth and frequency), edaphic (e.g., percent cover of leaf litter), biological (e.g., plant species composition), and landscape (e.g., characteristics of upland habitats immediately adjacent to the wetland) attributes and processes using procedures similar to those described in Brinson et al. (1995) and Rheinhardt et al. (1997). Data collected during this phase of the study were used to establish reference standards and scale variables that were used in models, as described in the next section.

Selection of Reference Standard Wetlands

Reference Standard Wetlands are defined as the subset of Reference Wetlands that have been least altered and, thus, represent a sustainable level of ecological functioning across the suite of functions used in the models (Brinson et al. 1995, Brinson et al. 1998). Reference Standard Wetlands were chosen by a subset of the authors (DFW, HK, RDR) that had selected and sampled the 44 Reference Wetlands. Selection criteria included field observations, preliminary analysis of the data collected in the spring-summer of 1995, and best professional judgment. Hydrologic, edaphic, biological, and landscape data that had been collected at the Reference Standard Wetlands were used to scale variables used in the HGM models, as described in the next section.

Functions, Variables and Scaling of Variables Using Data from Reference Wetland System

Hydrologic, biogeochemical, biota, and habitat functions for the subclass were identified using Brinson et al. (1995) as a guide. Models for each function and the variables aggregated into the models were also selected based on guidelines in Brinson et al. (1995). Model variables are attributes of wetlands that can be measured directly (e.g., tree basal area, presence of tree seedlings as evidence of tree regeneration) or indirectly as indicators (e.g., stains on vegetation indicating flooding depth) and are combined in equations to calculate a Functional Capacity Index (FCI) Scores for each function (Smith et al. 1995, Rheinhardt et al. 1997, Brinson et al. 1998). Model variables and functions used in this study are described in Tables 1 and 2.

We assessed 14 variables in three general areas: hydrology (Functions 1–5 in Table 2), biogeochemistry (Functions 6–9 in Table 2), and biota/habitat (Functions 10–14 in Table 2). The 28 variables (Table 1) were used in the model equations from one to six times for calculating FCI Scores (Table 2). Procedures originally described in Smith et al. (1995) and elaborated by Brinson et al. (1995) and Rheinhardt et al. (1997) were used to transform the data for the field variables (i.e., field measurements of the variables) to Variable Subindices scaled from 1.0 to 0.0.

In most instances, the Variable Subindex categories were 1.0, 0.5, 0.1, or 0.0. A Subindex of 1.0 represented conditions at Reference Standard Wetlands, 0.5 represented a significant deflection from Reference Standards, 0.1 was used to indicate complete or near absence of the variable but with potential for recovery (regrowth of vegetation, etc.), and 0.0 indicated absence of the variable without potential for recovery (i.e., construction such as dams, buildings, highways, etc.). When quantitative data were not available for variables, the Variable Subindices were based on qualitative observations and/or best professional judgment of the interdisciplinary team. Functional Capacity (FCI) Scores were computed (see equations in Table 2) as continuous numbers that range from 1.0, characteristic of Reference Standard Wetlands, to 0.0, absence of function.

Application of HGM Models at Reference Wetlands by Two Independent Teams

Two assessment teams were established to independently assess the Reference Wetlands over a three-week period in November 1995. None of the authors (MMB, LCL, WLN, DFW) involved in development of the national guidebook for assessment of riverine wetlands (Brinson et al. 1995) were members of the assessment teams. Each team consisted of an individual who was responsible for making final decisions for scoring each variable at the 44 sites. The team leaders had been trained in the application of HGM but had

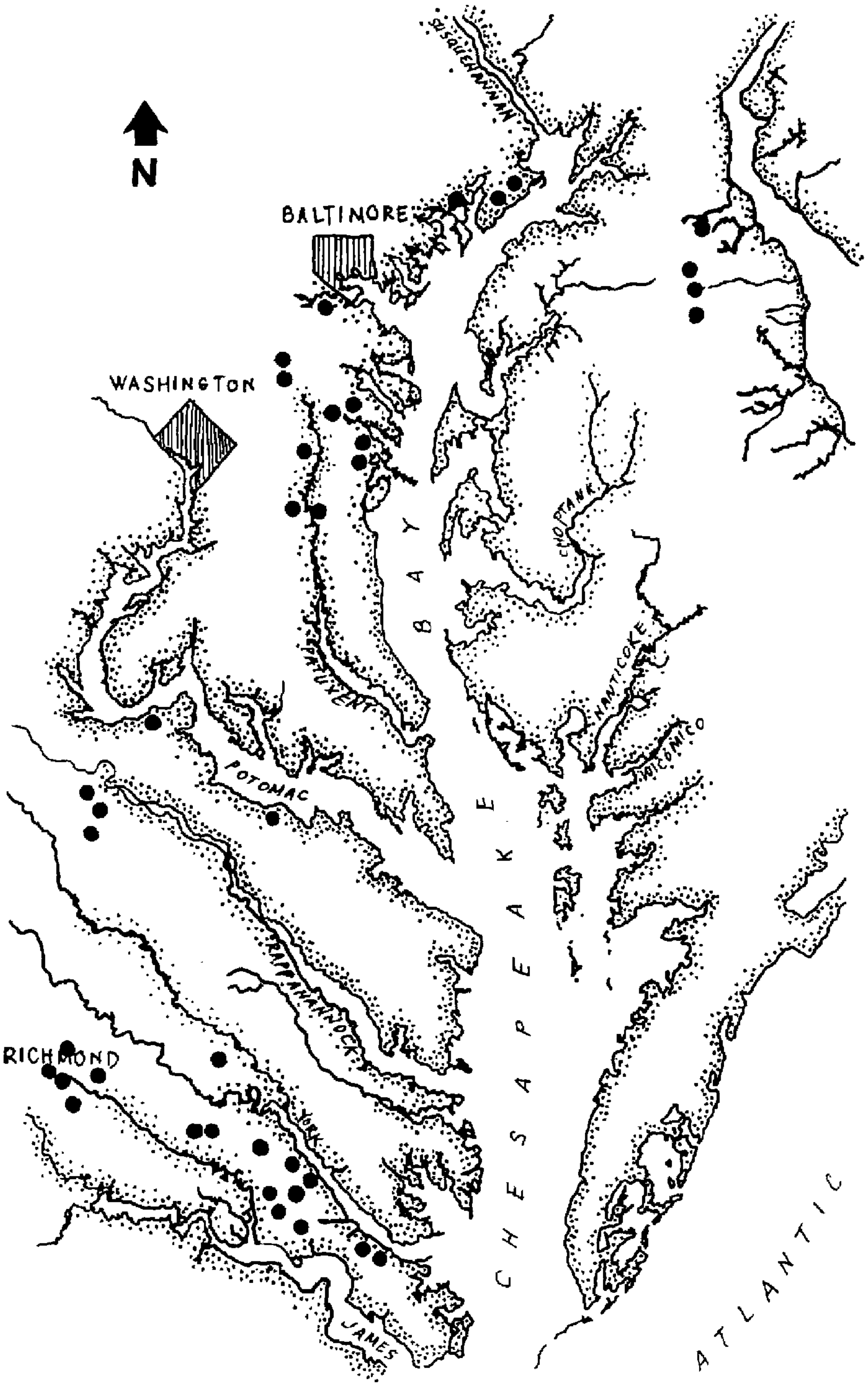

Figure 1. Map of a portion of the Chesapeake Bay region showing the distribution of wetland study sites. The 39 dots on the map represent all 44 sites because, in three instances, one dot covers more than one wetland at the scale used to draft the map.

not previously visited any of the 44 sites. A second member of each team was one of the authors (KH or RR) who was present to assist with data collection for assigning variable subindex scores. Each team also included one or more individuals from federal and state resource agencies who had not previously been to any of the sites but had received training in the HGM approach.

Each assessment team was given a work schedule and instructions that included procedures for conducting the assessment, directions to each wetland, and the location of the Assessment Area within each wetland

Table 1. Field variables used in this study are listed in the first column and a brief description of how the variables were measured in the field is provided in the third column. Numbers in the second column refer to the FUNCTIONS that variables were used in (see Table 2 for a list of functions and equations). Mean Variable Subindices for the two assessment teams are shown in the TEAM 1 and 2 columns, and results of the Wilcoxon signed rank test are shown in the TEST column (NS indicates that Variable Subindices did not differ significantly).

Variable	Models Used In	Description of Method	Team 1	Team 2	Test
V_{BTREE} - Tree Basal Area	1, 7, 8, 10	Basal area of trees	0.83	0.81	NS
V_{CANOPY} - Canopy Cover	10	Estimate coverage of tree canopy	0.72	0.91	0.0015
V_{COMP} - Tree Species Composition	10	Presence of certain tree species	0.61	0.65	NS
V_{CONTIG} - Contiguity	13	Percent of wetland-upland boundary covered by upland forest	0.73	0.78	0.047
V_{DTREE} - Tree Density	3, 8, 10	Density of trees	0.57	0.70	0.005
V_{FREQ} - Frequency of Flooding	1, 3, 7, 8, 9, 13	Evidence of flooding every 1–2 years as shown by the presence of sediment lines on trees, drift and/or wrack lines	0.71	0.62	NS
V_{FWD} - Fine Woody Debris	6, 11	Percent cover of fine woody debris (wood < 10 cm diameter and < 1 meter long)	0.72	0.93	0.0001
V_{GAPS} - Canopy Gaps	12	Estimate percent of tree canopy that has gaps formed by the death of standing or fallen trees	0.70	0.81	0.049
V_{HERB} - Herbaceous Cover	8	Percent of area covered by herbaceous plants	0.45	0.35	NS
V_{INUND} - Flooding Depth	1	Depth of water and sediment stains on vegetation	0.75	0.71	NS
V_{LINVT} - Litter Invertebrates	14	Evidence of invertebrates in leaf litter and logs	0.86	0.95	NS
V_{LOGS} - Log Decomposition	6, 11	Number of stages of log decomposition present	0.73	0.65	NS
V_{MACRO} - Macrotopography	2, 3	Presence of meander scars and/or secondary stream channels	0.38	0.52	0.017
V_{MATUR} - Forest Successional Stage	12	Maturation stage of forest as indicated by presence of logs and standing dead canopy trees	0.95	0.91	NS
V_{MICRO} - Microtopography	1, 2, 7, 8, 13	Percent of area with evidence of surface ponding	0.68	0.81	0.048
V_{ORGAN} - Organic Matter	9	Percent leaf litter cover and presence of an A soil horizon	0.87	0.87	NS
V_{PATCH} - Forest Heterogeneity	12	Changes in vegetation cover along transects	0.61	0.58	NS
V_{PORE} - Soil Mottling	4	Presence of gley or low chroma Munsell colors and soil structure	0.23	0.27	NS
V_{REDVEL} - Impact of Flooding	1, 3	Evidence for the presence or absence of high energy overbank flooding events	0.74	0.58	0.019
V_{REGEN} - Tree Regeneration	10	Species composition and cover of tree seedlings	0.87	0.87	NS
V_{SHRUB} - Shrub Density	1, 8	Shrub density	0.43	0.64	0.0001
V_{SINVT} - Soil Invertebrates	14	Evidence of invertebrates in soil and logs on ground	0.86	0.84	NS
V_{SORP} - Soil Nutrient Sorption	7	Soil texture, chroma, and roots	0.50	0.57	NS
V_{STRATA} - Vegetation Stratification	6, 12	Number of strata from canopy to ground	0.86	0.89	NS
V_{SUBIN} - Subsurface Water Input	5, 7, 9, 13	Evidence of seeps that discharge subsurface water	0.87	0.85	NS
V_{SUBOUT} - Subsurface Discharge	5	Evidence of groundwater discharge to stream	0.75	0.75	NS
V_{SURFIN} - Surface Water Input	7, 8, 9, 13	Evidence of surface water input	0.55	0.68	NS
V_{WTF} - Water Table Fluctuation	4	Evidence of high chroma mottles in the soil or presence of surface water ponding	0.27	0.50	0.0005

Table 2. Comparisons of Functional Capacity Index (FCI) Scores for TEAM 1 and 2. Results of Wilcoxon ranked sign tests are shown in the TEST columns (NS indicates that FCI scores did not differ significantly). Variables used in the equations to calculate FCI scores are described in Table 1, and variables that are in **bold** were different for the two teams when they were statistically compared (Table 1).

Functions	Equation	Team 1	Team 2	Test
1. Dynamic Surface Water Storage	$(V_{FREQ}*((V_{INUND} + \mathbf{V_{MICRO}} + \mathbf{V_{SHRUB}} + V_{BTREE} + \mathbf{V_{REDVEL}})/5))^{1/2}$	0.65	0.61	NS
2. Long-term Surface Water Storage	$(\mathbf{V_{MACRO}} + \mathbf{V_{MICRO}})/2$	0.53	0.67	0.002
3. Energy Dissipation	$(((\mathbf{V_{MACRO}} + \mathbf{V_{DTREE}} + \mathbf{V_{REDVEL}})/3)*V_{FREQ})^{1/2}$	0.57	0.55	NS
4. Substrate Storage of Water	$(V_{PORE} + \mathbf{V_{WTF}})/2$	0.25	0.39	0.002
5. Moderation of Groundwater Flow or Discharge	$(V_{SUBIN} + V_{SUBOUT})/2$	0.81	0.80	NS
6. Nutrient Cycling	$(V_{STRATA} + V_{LOGS} + \mathbf{V_{FWD}})/3$	0.77	0.82	NS
7. Removal of Elements and Compounds	$(V_{FREQ} + V_{SURFIN} + V_{SUBIN} + \mathbf{V_{MICRO}} + V_{SORP} + V_{BTREE})/6$	0.69	0.72	NS
8. Retention of Particulates	$(((V_{FREQ} + V_{SURFIN})/2)*((\mathbf{V_{MICRO}} + \mathbf{V_{SHRUB}} + V_{HERB} + V_{BTREE} + \mathbf{V_{DTREE}})/5))^{1/2}$	0.58	0.62	NS
9. Organic Carbon Export	$(((V_{FREQ} + V_{SURFIN} + V_{SUBIN})/3)*V_{ORGAN})^{1/2}$	0.75	0.75	NS
10. Maintain Characteristic Plant Community	$(V_{COMP} + V_{REGEN} + \mathbf{V_{CANOPY}} + ((\mathbf{V_{DTREE}} + V_{BTREE})/2)/4$	0.72	0.80	0.001
11. Maintain Characteristic Detrital Biomass	$(V_{LOGS} + \mathbf{V_{FWD}})/2$	0.73	0.79	NS
12. Maintain Spatial Structure of Habitat	$(V_{MATUR} + V_{STRATA} + V_{PATCH} + \mathbf{V_{GAPS}})/4$	0.78	0.80	NS
13. Maintain Interspersion and Connectivity	$(V_{FREQ} + \mathbf{V_{MICRO}} + V_{SURFIN} + V_{SUBIN} + \mathbf{V_{CONTIG}})/5$	0.71	0.75	NS
14. Maintain Distribution and Abundance of Invertebrates	$(V_{SINVT} + V_{LINVT})/2$	0.76	0.90	NS

where they would conduct sampling and assign Variable Subindex scores. The first team to assess each site marked the Assessment Area in which they worked so that the second team would conduct their assessment in the same location. One team began its work in Delaware and the other in southern Virginia. The two teams compiled Variable Subindex scores for all variables at each site and also were asked to provide written commentaries and rationale for each assigned score.

Data Analyses and Interpretation

Data for each team and site were compiled, and the results for the two teams for the Variable and FCI scores were compared with the Wilcoxon signed rank test. The Wilcoxon test was used to compare the scores for the two teams because the data were not normally distributed. Variable subindices and FCI scores for the two teams were further compared to determine the probabilities that the two teams would differ by more than a specified level. Normal probabilities were computed using the PROBNORM function of SAS (SAS 1990), and a description of the procedure used to calculate probabilities for the Variable Subindices and FCI Scores follows.

The Variable Scores assigned by each team were multinomial in distribution and were placed in one of four groups that accounted for 16 possible levels of agreement for each site.

Group 1. No difference between the two teams in Variable Subindex scores. Possible combinations of Variable Subindex scores were (0.0, 0.0), (0.1, 0.1), (0.5, 0.5), and (1.0, 1.0).

Group 2. Variable Subindex scores differ by one scalar unit. Possible combinations of Variable Subindex scores were (0.0, 0.1), (0.1, 0.5), (0.5, 1.0), (0.1, 0.0), (0.5, 0.1), and (1.0, 0.5).

Group 3. Variable Subindex scores differ by two scalar units. Possible combinations of Variable Subindex scores were (0.0, 0.5), (0.1, 1.0), (0.5, 0.0), and (1.0, 0.1).

Group 4. Variable Subindex scores differ by three scale units. Possible combinations of Variable Subindex scores were (0.0, 1.0) and (1.0, 0.0).

Using this procedure, each site assessed by the two teams created a 16-cell multinomial distribution where the cells were defined by all possible pairs of Variable Subindices. The probability of each level of disagreement was estimated by the number of sites that had Variable Subindices in each group divided by the total number of sites.

In contrast to Variable Subindices, FCI scores were treated as if they had continuous frequency distributions. The probability of exceeding any fixed level of difference between the two teams was estimated by integrating the tails of a normal density function over values whose absolute value exceeded the fixed level. The normal density function was assumed to have a mean of zero and variance equal to the variance of the differences in FCI Scores between teams. The variance was estimated by the sample variance of the differences between teams. Assuming a mean of zero was the same as assuming that each team produced an unbiased estimate of the true site FCI Score.

RESULTS

The two teams differed statistically in scoring 10 of the 28 (35.7%) variables (Table 1), and they were distributed approximately evenly across the three categories of functions (e.g., hydrology, biogeochemistry, biota/habitat). Three of the ten statistically different variables (V_{MACRO}, V_{REDVEL}, V_{WTF}) were only used in equations used to calculate FCI Scores for hydrology functions (Table 2). Three statistically different variables (V_{CANOPY}, V_{CONTIG},) were only used to calculate FCI Scores for biota/habitat functions, and the remaining four variables (V_{DTREE}, V_{FWD}, V_{MICRO}, V_{SHRUB}) were used to calculate FCI Scores for more than one category of functions. Mean FCI scores differed significantly between teams for two of the five hydrologic functions, for none of the four biogeochemical functions, and one of the five biota/habitat functions (Table 2).

DISCUSSION

Variable Subindex Scores

Variable Subindex scores for the two teams were further analyzed by calculating probabilities based on comparisons between the two teams (Table 3). Variable Subindices for the 44 Reference Wetlands were within 0.1 of each other for 19 of the 28 variables (68%). However, the degree of agreement between the two teams was even higher based upon the probability analysis (Table 3).

Table 3. Summary of the probabilities that the differences in Variable Subindices between the two assessment teams disagree by less than or equal to 0, 1, 2, and 3 scale units.

Variable	0	1	2	3
V_{BTREE}	0.818	1.000	1.000	1.000
V_{CANOPY}	0.591	0.977	1.000	1.000
V_{COMP}	0.727	1.000	1.000	1.000
V_{CONTIG}	0.795	1.000	1.000	1.000
V_{DTREE}	0.523	1.000	1.000	1.000
V_{FREQ}	0.681	0.795	0.955	1.000
V_{FWD}	0.568	0.977	1.000	1.000
V_{GAPS}	0.614	0.977	1.000	1.000
V_{HERB}	0.568	0.909	0.977	1.000
V_{INUND}	0.636	0.955	0.955	1.000
V_{LINVT}	0.864	0.864	0.864	1.000
V_{LOGS}	0.432	0.977	0.977	1.000
V_{MACRO}	0.682	0.795	0.887	1.000
V_{MATUR}	0.955	0.955	0.977	1.000
V_{MICRO}	0.545	0.886	0.955	1.000
V_{ORGAN}	0.818	0.818	0.909	1.000
V_{PATCH}	0.364	0.886	0.955	1.000
V_{PORE}	0.591	0.841	0.977	1.000
V_{REDVEL}	0.591	0.795	0.841	1.000
V_{REGEN}	0.795	0.955	0.977	1.000
V_{SHRUB}	0.500	0.909	1.000	1.000
V_{SINVT}	0.886	0.886	0.886	1.000
V_{SORP}	0.455	0.614	0.841	1.000
V_{STRATA}	0.864	0.977	1.000	1.000
V_{SUBIN}	0.818	0.886	0.955	1.000
V_{SUBOUT}	0.795	1.000	1.000	1.000
V_{SURFIN}	0.727	0.727	0.727	1.000
V_{WTF}	0.477	0.706	0.886	1.000

Most ecological data are continuously distributed, and when data are classified into discrete categories, such as is done in the HGM Approach, borderline cases are difficult to classify. Thus, in cases in which the probability was low that the two teams did not differ in their assessment of a variable (cases where the two teams usually differed), one does not know whether these differences were due to (1) the measurement requiring too much subjective interpretation, (2) too many borderline cases, or (3) improper scaling of the variable. Consider, for example, V_{FREQ} and V_{GAPS} in Table 3. The probability that the two teams did not differ was 0.681 and 0.614, respectively, suggesting that V_{FREQ} was a slightly more reliable variable than V_{GAPS} (e.g., the probability that they did not differ was higher). The probabilities that the average scores differed by two or more scalar units was, however, 1.000–0.795 = 0.205 for V_{FREQ} and 1.000–0.977 = 0.023 for V_{GAPS}. Thus, for V_{GAPS}, the two teams usually only differed by one scalar unit when they did not agree. This result would be expected if many sites were borderline for the value assigned for V_{GAPS}. The higher probability of the score for V_{FREQ} differing by two or more scalar units indicates that the teams had differed fundamentally in how they assessed that variable. Indeed, the team leaders reported, independent of viewing results, difficulty in applying the method for assessing V_{FREQ}, and they each slightly modified the suggested procedure when they applied the sampling protocol in the field. Thus, in spite of the evidence offered by the high degree of probability agreement for the two variables, one might infer that V_{FREQ} was a less reliable variable than V_{GAPS}. From this reasoning, it follows that the best benchmark for comparing the raw Variable Subindices is determining the probability that teams differ by 1 or fewer scalar units. Table 3 shows that the probabilities that the two teams differ by 1 or fewer scalar units ranges from 0.614 for V_{SORP} to 1.000 for five variables (V_{BTREE}, V_{COMP}, V_{CONTIG}, V_{DTREE}, V_{SUBOUT}). There were only three variables (V_{SORP}, V_{SURFIN}, V_{WTF}) for which this level of agreement was less than 80%. A general conclusion would be that teams differed by one or fewer scalar units on ≥ 75% of the sites for the vast majority of Variable Subindices.

The two teams differed statistically in their assessment of four hydrologic variables (V_{MACRO}, V_{MICRO}, V_{REDVEL}, V_{WTF} ; Table 1). Scoring of each of these variables required the use of subjective rather than objective measurements. For example, V_{REDVEL} was scored by qualitatively assessing the presence and magnitude of high energy overbank flooding events such as sediment scour and deposition patterns rather than by directly measuring flow velocities. Both teams also reported that these four variables appeared to be scaled incorrectly (*i.e.*, the scaling did not concur with field observations relative to degree of impact). It is likely that imprecise definitions and methods of measurement, combined with incorrect scaling, were the reasons why there were significantly different Variable Subindices for V_{REDVEL}. Problems of this type can only be solved by continual evaluation, testing, and improvement of models using reference data (Wakeley and Smith In press).

Detailed hydrologic data were available for only one site at the time of the study. Assessment of hydrology variables will almost always be problematic in wetland assessment because there are very few wetlands for which there are long-term hydrology data that can be used to quantify the frequency, depth, and duration of flooding, factors that influence soil chemistry (Ponnamperuma 1972), elemental cycling (Gambrell and Patrick 1978, Sánchez-Pérea and Trémolières 1997), plant community composition (Piégay 1977, Bedinger 1979, Toner and Keddy 1997), plant growth (Will et al. 1995, Jones et al. 1996) and ecological functioning of wetlands (Verhoeven et al. 1994). Assessment of hydrology variables will almost always be based on "indirect evidence" of field indicators, such as channel and floodplain morphology, magnitude of sediment deposition, drift lines, and water staining of trees. Clear directions of how to assess variables must be provided in the Regional Guidebooks to minimize differences in measurement results due to differences in sampling or interpretation (i.e., measurements must be consistent and objective). Measurements based on subjective interpretations or those that rely on the level of experience of the assessor should be avoided.

Five of seven objectively measured variables used in calculating biogeochemical and biota/habitat FCIs differed significantly between assessment teams [variables shown in bold for FUNCTIONS 6–14 in Table 2) and described (V_{CANOPY}, V_{CONTIG}, V_{FWD}, V_{GAPS}, V_{FWD}) in Table 1]. The two team leaders reported that differences between the two teams for those variables were most likely due to slight differences in where data were collected within the Assessment Areas and, in part, to differences in the sizes of the Assessment Areas among sites with respect to the size and scale of the river. Spatial variability of biological and physical features of riverine forested wetlands within the Reference Domain is common (e.g., Parsons and Ware 1982, Puckett et al. 1993) and adequate sampling (*e.g.*, area sampled, sample size, number of samples) would be required to resolve potential differences in Variable Subindex scores for such area-based variables. A similar logic applies to more quantitative variables such as V_{DTREE}. Density-based measurements are subject to spatial variation and appropriate sample sizes, and areas sampled must be determined during the guidebook development phase.

Another problem with objective measurements is the variability inherent to the measurement techniques and the timing of the assessments. For example, V_{CANOPY} was measured in the field in November, so many of the leaves had fallen and tree canopies lacked continuous crown outlines. The lack of continuous crown outlines is an important source of error in canopy cover assessments (Mueller-Dombois and Ellenberg 1974). Thus, measuring V_{CANOPY} in the field in November likely increased sample variability. Problems such as this could be avoided by restricting the assessment to the growing season or by scoring the variable through the use of recent high quality low altitude aerial photographs taken during the growing season.

Functional Capacity Index Scores

Average FCIs for the two teams were significantly different for three of the fourteen (21.4%) functions (Table 2). Probability analyses of the data further suggest that there was a high degree of agreement between the two teams, as mean FCI Scores differed by less than 0.1 for twelve of the fourteen (85.7%) functions (Table 4). The probability analysis also showed that, for all functions but Dynamic Surface Water Storage, between-team differences in FCI Scores were less than 0.4 units at a probability level ≥ 0.990 (Table 4). FCI Scores calculated with equations that had only two variables (Table 2), with the exception of Moderation of Ground-Water Flow or Discharge, had a higher

Table 4. Probabilities that the differences between teams in the Functional Capacity Index (FCI) Scores will be less than or equal to the level (e.g., 0.1, 0.2, 0.3, 0.4) indicated in the columns. Equations used to calculate FCI values are provided in Table 2.

Function	0.1	0.2	0.3	0.4
1. Dynamic Surface Water Storage	0.243	0.465	0.785	0.987
2. Long-term Surface Water Storage	0.281	0.527	0.849	0.996
3. Energy Dissipation	0.260	0.493	0.816	0.992
4. Subsurface Storage of Water	0.252	0.480	0.802	0.990
5. Moderation of Groundwater Flow or Discharge	0.420	0.732	0.973	0.999
6. Nutrient Cycling	0.415	0.725	0.971	0.999
7. Removal of Elements and Compounds	0.411	0.720	0.969	0.999
8. Retention of Particulates	0.361	0.652	0.939	0.999
9. Organic Carbon Export	0.322	0.594	0.903	0.999
10. Maintain Characteristic Plant Community	0.494	0.816	0.992	1.000
11. Maintain Characteristic Detrital Biomass	0.323	0.596	0.905	0.999
12. Maintain Spatial Structure of Habitat	0.503	0.825	0.993	1.000
13. Maintain Interspersion and Connectivity	0.418	0.729	0.972	0.999
14. Maintain Distribution and Abundance of Invertebrates	0.310	0.575	0.890	0.999

probability of differing by more than 0.3. Examination of Table 2, however, suggests that the differences in FCI Scores for functions that used two variables, may be due in part to problems associated with assigning subindex scores to the variables in the field more than the fact that only two variables were used in the models.

Variables that were significantly different in Table 1 have been marked in bold in Table 2 to show how the variables were distributed among the functions when FCI scores were calculated. FCI Scores for the Moderation of Ground-Water Flow or Discharge and Maintain Distribution and Abundance of Invertebrates functions did not differ significantly (Table 4), and the four variables used in models did not differ statistically when applied in the field by the two teams. In contrast, both variables used in the Long-term Surface Water Storage function differed significantly, and the FCI scores were also significantly different (Table 2). Functions that had multiplicative terms (Dynamic Surface Water Storage, Energy Dissipation, Retention of Particulates, Organic Carbon Export) also tended to differ by more than 0.3 units (Table 4) but did not differ statistically between the two teams (Table 2). These results would be expected due to analytical properties of variance, which should decrease as more variables are used in calculating FCI Scores but increase when multiplication is used in the calculation. If the variable or variables used in the multiplications differed significantly, there would be a greater chance that the FCI Scores would have differed between the two teams. That possibility seems to have little influence on the results in this study because none of the multiplication variables differed significantly (Table 1).

CONCLUSIONS

The hydrogeomorphic approach to wetland assessment is still in the developmental stages, and a number of assumptions in model development and application need to be rigorously tested (Wakeley and Smith In press). The objectives of this study were to test one aspect of the HGM approach by evaluating whether or not two teams obtained similar results when they applied the same methodology to a set of wetlands representing a wide range of conditions from relatively unaltered to highly altered sites. We found that highly repeatable results occurred, especially in FCI Scores. Analyses of the data, however, also demonstrated that several issues need to be considered when models are being developed, tested, and refined. The statistical approaches used to evaluate the teams' results indicate the importance of identifying and eliminating variables whose measurements are not repeatable, either because the measurement required too much subjective interpretation or the variable was scaled improperly. Further, we demonstrated that variables whose measurements were not repeatable can detrimentally affect the repeatability of FCI scores derived from those same variables. Not surprisingly, functions that are most affected are thosethat are modeled by only a few variables. We hope that these findings will be useful as the HGM approach to wetlands assessment is further developed and tested.

ACKNOWLEDGMENTS

We thank the members of the sampling teams (Jerry Quesenberry, Janet Norman, Greg Moser, Chris Victoria, Kim Marbaine, Mitch Keiler, Brian Yanchik, Anne Lynn, Al Rizzo, Sally Griffith-Keper, Rebecca Hicks, Pete Stokley, and Steve Harmon). Diane Eckles helped guide the project from its inception, and we thank her for the great effort and support that she provided. Funding was provided through a Fish and Wildlife Service Cooperative Agreement (14-48-0005-94-9075), a Cooperative Agreement (68-3B19-5-2) with the Natural Resources Conservation Service, the Smithsonian Institution, and the National Wetlands Science Training Cooperative. Statistical consultation for the probability analysis was provided by Elgene S. Perry. We thank R. Daniel Smith and two anonymous reviewers for comments on a draft version of the manuscript. Doug Kuzmiak assisted Humaira Khan in sampling the Reference Wetlands.

LITERATURE CITED

Ainslie, W. B., R. D. Smith, B. A. Pruitt, T. H. Roberts, E. J. Sparks, L. West, G. L. Godshalk, and M. V. Miller. 1999. A Regional Guidebook for Assessing the Functions of Low Gradient, Riverine Wetlands in Western Kentucky. Army Corps of Engineers, Waterways Experiment Station, Vicksburg, MS, USA. Technical Report WRP-DE-17.

Bedinger, M. S. 1979. Relation between forest species and flooding. p. 427–435. *In* P. E. Greeson, J. R. Clark, and J. E. Clark (eds.) Wetland Functions and Values: the State of Our Understanding. American Water Research Association, Minneapolis, MN, USA.

Brinson, M. M. and R. D. Rheinhardt. 1996. The role of reference wetlands in functional assessment and mitigation. Ecological Applications 6:69–76.

Brinson, M. M., W. Kruczynski, L. C. Lee, W. L. Nutter, R. D. Smith, and D. F. Whigham. 1994. Developing an approach for assessing the functions of wetlands. p. 615–624. *In* W. J. Mitsch (ed.) Global Wetlands: Old World and New. Elsevier Science B.V., Amsterdam, The Netherlands.

Brinson, M. M., F. R. Hauer, L. C. Lee, W. L. Nutter, R. D. Rheinhardt, R. D. Smith, and D. F. Whigham. 1995. Guidebook for Application of Hydrogeomorphic Assessments to Riverine Wetlands. Army Corps of Engineers, Waterways Experiment Station, Vicksburg, MS, USA. Technical Report TR-WRP-DE-11.

Brinson, M. M., R. D. Smith, D. F. Whigham, L. C. Lee, R. D. Rheinhardt, and W. L. Nutter. 1998. Progress in development of the hydrogeomorphic approach for assessing the functioning of wetlands. p. 383–406. *In* A. J. McComb and J. A. Davis (eds.). Wetlands for the Future. Gleneagles Publishing, Adelaide, Australia.

Brush, G., C. Lenk, and J. Smith. 1980. The natural forests of Maryland: an explanation of the vegetation map of Maryland. Ecological Monographs 50:77–92.

Federal Register. 1996. August 16, 1996. 61(160):42593–42603.

Federal Register. 1997. June 20, 1997. 62(114):33607–33620.

Gambrell, R. P. and W. H. Patrick. 1978. Chemical and microbiological properties of anaerobic soils and sediments. p. 375–423. *In* D. D. Hook and R. M. M. Crawford (eds.) Plant Life in Anaerobic Environments. Ann Arbor Science Publications, Inc., Ann Arbor, MI, USA.

Glascock, G. and S. Ware. 1979. Forests of small stream bottoms in the Peninsula of Virginia. Virginia Journal of Science 30:17–21.

Hupp, C. R., M. D. Woodside, and T. M. Yanosky. 1993. Sediment and trace element trapping in a forested wetland, Chickahominy River, Virginia. Wetlands 13:95–104.

Jones, R. H., B. G. Lockaby, and G. L. Somers. 1997. Effects of microtopography and disturbance on fine-root dynamics in wetland forests of low-order stream floodplains. American Midland Naturalist 136:57–71.

Mueller-Dombois, D. and H. Ellenberg. 1974. Aims and Methods of Vegetation Ecology. John Wiley and Sons, New York, NY, USA.

Parsons, S. E. and S. Ware. 1982. Edaphic factors and vegetation in Virginia Coastal Plain swamps. Bulletin of the Torrey Botanical Club 109:365–370.

Piégay, H. 1997. Interactions between floodplain forests and overbank flows: data from three piedmont rivers of southeastern France. Global Ecology and Biogeography Letters 6:187–196.

Ponnamperuma, F. N. 1972. The chemistry of submerged soils. Advances in Agronomy 24:29–96.

Puckett, L. J., M. D. Woodsite, B. Libby, and M. R. Schening. 1993. Sinks for trace metals, nutrients, and sediments in wetlands of the Chickahominy River near Richmond, Virginia. Wetlands 13:105–114.

Rheinhardt, R. D., M. M. Brinson, and P. M. Farley. Applying wetland reference data to functional assessment, mitigation, and restoration. Wetlands 17:195–215.

Rheinhardt, R. R., D. Whigham, H. Kahn, and M. Brinson. In Press. Vegetation of headwater wetlands in the Inner Coastal Plain of Virginia and Maryland. Castanea.

Sánchez-Pérez, J. -M. and M. Trémolières. 1997. Variation in nutrient levels of the groundwater in the Upper Rhine alluvial forests as a consequence of hydrological regime and soil texture. Global Ecology and Biogeography Letter 6:211–217.

SAS Institute Inc. 1990. SAS Language: Reference. Version 6, First Edition. SAS Institute, Inc., Cary, NC, USA.

Smith, R. D., A. Ammann, C. Bartoldus, and M. M. Brinson. 1995. An Approach for Assessing Wetland Functions Using Hydrogeomorphic Classification, Reference Wetlands, and Functional Indices. U.S. Army Corps of Engineers, Waterways Experiment Station, Vicksburg, MS, USA. Technical Report TR-WRP-DE-10.

Strahler, A. N. 1957. Quantitative analysis of watershed geomorphology. American Geophysical Union Transactions 38:913–920.

Tiner, R. W. 1987. Mid-Atlantic Wetlands. A Disappearing Natural Treasure. Cooperative Publication, U.S. Fish and Wildlife Service, Region 5, Newton Corner, MA, USA, and U.S. Environmental Protection Agency, Region Three, Philadelphia, PA, USA.

Tiner, R. W. and D. G. Burke. 1995. Wetlands of Maryland. Cooperative Publication, U.S. Fish and Wildlife Service, Region 5, Hadley, MA, USA, and Maryland Department of Natural Resources, Annapolis, MD, USA.

Toner, M. and P. Keddy. 1997. River hydrology and riparian wetlands: a predictive model for ecological assembly. Ecological Applications 7:236–246.

Verhoeven, J. T. A., D. F. Whigham, K. Van Kerkhoven, J. O'Neill, and E. Maltby. 1994. Comparative study of nutrient-related processes in geographically separated wetlands: towards a science base for functional assessment procedures. p. 91–106. *In* W. J. Mitsch (ed.) Global Wetlands: Old World and New. Elsevier Science B.V., Amsterdam, The Netherlands.

Walbridge, M. R. and J. P. Struthers. 1993. Phosphorus retention in non-tidal palustrine forested wetlands of the Mid-Atlantic Region. Wetlands 13:84–94.

Wakeley, J. S. and R. D. Smith. In press. Verifying, field testing, and validating assessment models. *In* Hydrogeomorphic Approach to Assessing Wetland Functions: Guidelines for Developing Regional Guidebooks. Army Corps of Engineers, Waterways Experiment Station, Vicksburg, MS, USA.

Will, R. E., J. R. Seiler, P. P. Feret, and W. M. Aust. 1995. Effects of rhizosphere inundation on the growth and physiology of wet and dry-site *Acer rubrum* (Red Maple) populations. American Midland Naturalist 134:127–139.

Zentner, J. 1999. HGM: A practitioner's perspective. SWS Bulletin 16:25–26.

Manuscript received 6 April 1998; revision received 11 January 1999; accepted 20 April 1999.